THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE GRENOBLE ALPES

Spécialité : Physique pour les Sciences du Vivant

Arrêté ministériel : 25 mai 2016

Présentée par

Aline CISSE

Thèse dirigée par **Judith PETERS, Professeur des Universités, Université Grenoble Alpes**

préparée au sein du **Laboratoire Interdisciplinaire de Physique** dans **l'École Doctorale de Physique**

L'étude de la lipoprotéine à basse densité par microscopie électronique et diffusion neutronique

Studies of Low-Density Lipoproteins by Cryo-Electron Microscopy and Neutron Scattering

Thèse soutenue publiquement le **9 Décembre 2021**, devant le jury composé de :

Madame Elisabeth CHARLAIX
PROFESSEUR DES UNIVERSITES, Université Grenoble Alpes, Présidente du jury
Madame Aurélie BERTIN
DIRECTEUR DE RECHERCHE, CNRS, Rapportrice
Monsieur Jörg PIEPER
PROFESSEUR, University of Tartu, Rapporteur
Madame Ruth PRASSL
PROFESSEUR ASSOCIE, Medical University of Graz, Examinatrice
Madame Giovanna FRAGNETO
DOCTEUR EN SCIENCES, Institut Laue-Langevin, Examinatrice
Monsieur Marc JAMIN
PROFESSEUR DES UNIVERSITES, Université Grenoble-Alpes, Examinateur
Monsieur Ambroise DESFOSSES
DOCTEUR EN SCIENCES, CNRS, Invité

Remerciements

Je tiens tout d'abord à remercier ma directrice de thèse, Judith Peters, qui fut également par deux fois ma responsable de stage, mais aussi mon enseignante et reponsable durant le Master. Merci encore pour m'avoir aiguillée vers ce parcours de biophysique, acceptée en stage, puis en thèse, et m'avoir guidée et tant appris en recherche durant ces quatre dernières années. J'ai eu plusieurs fois l'opportunité de réaliser à quel point c'était une chance de travailler dans ton groupe, merci pour ta patience, tes précieux conseils, ton soutien, ainsi que tous les projets passionnants que tu m'as fait découvrir et fait participer.

Je remercie également mes rapporteurs de thèse, Jörg Pieper et Aurélie Bertin, ainsi que les membres du jury, Elisabeth Charlaix, Ruth Prassl, Giovanna Fragneto, Marc Jamin et Ambroise Desfosses, pour avoir accepté d'examiner mes travaux de thèse, et d'y porter un intérêt. J'espère vivement que cela donnera lieu à des discussions enrichissantes.

Je tiens à exprimer ma reconnaissance envers la Fondation CFM, et son président, Jean-Philippe Bouchaud, qui a soutenu mes travaux de thèse au travers d'une bourse Jean-Pierre Aguilar durant ces trois années, ainsi que deux mois supplémentaires. Merci également à notre interlocutrice, Nathalie Bilimoff, et son accueil lors de la Journée de la Fondation en 2019.

Au même titre, je souhaite adresser mes remerciements à l'Université Grenoble-Alpes et l'Ecole Doctorale de Physique, ainsi que le personnel impliqué dans les inscriptions et les formations, notamment Isabelle Reverdy-Médélice. Merci également à la Vice-Présidence Recherche pour les deux mois additionnels alloués.

De la même manière, tous mes remerciements à mon laboratoire, le LiPhy, et à sa direction, pour les conditions dans lesquelles j'ai effectué ma thèse. Je tiens à saluer les gestionnaires pour leur sérieux et leur travail auprès des étudiants; Kristina Slavcheva, Sabine Gustave, Sandrine Nadau, Nadine d'Andréa, Michel Legoubin, Philippe Beys, Jean-Marc Sache.

Évidemment, ce projet de thèse n'existerait pas sans l'Université médicale de Graz (Autriche) et mes plus vifs remerciements vont à Ruth Prassl et Karin Kornmueller. Vielen Dank für diese wunderbare Zusammenarbeit. Vous m'avez enseigné énormément sur les LDLs, sur la préparation des échantillons, et je reste impressionnée par votre travail, spécialement après en avoir été témoin moi-même lors de mon échange à Graz. Merci encore pour ces trois semaines et l'accueil que vous m'avez tous fait; Anna Schachner, Ivan Vidakovic, Hans, Gerhard Ledinski et Gerd Leitinger.

Tout autant, un immense merci à l'ILL et au groupe spectroscopie pour m'accueillir en tant que visiteure long-terme. Merci à Laurence Tellier, pour l'organisation du groupe, et pour toutes les fois où je suis passée au secrétariat depuis mes stages. Merci au service informatique, en particulier Holger Gebhard et Sylvain Dufour pour l'installation de la machine pour les calculs cryo-EM.

J'ai eu la chance de pouvoir mener des expériences à l'ILL, et je remercie tous les responsables des instruments concernés pour leur assistance et les discussions que nous avons pu avoir; Markus Appel, Tilo Seydel, Jacques Ollivier, Francesca Natali, Marek Koza. Je n'oublie pas les techniciens d'IN13 avec qui j'ai beaucoup appris sur l'instrument, Bastien Gervasoni et Yann Berlemont. Mes remerciements vont de même au PSCM et au laboratoire de chimie, notamment David Hess, Martina Sandroni et Sandrine Verdon, pour toujours avoir été présents et répondu à mes multiples questions. Pour inclure tous les doctorants d'une manière égale, pour les séminaires et les retours enthousiastes, je tiens aussi à remercier Henry Fischer.

Pour les tests, les préparations de grilles et les premiers screenings en cryo-EM, un grand merci à Guy Schoehn à l'IBS (Grenoble), mais surtout au travail fantastique d'Armel Bézault, sur la plateforme de l'IECB (Bordeaux), qui nous a vraiment permis de commencer le projet sur les LDLs. Pour les magnifiques collectes de données et pour avoir accepté de se lancer avec nous, j'adresse mes remerciements à Eaazhisai Kandiah de l'ESRF et Daouda Traoré de l'ILL. Merci également à Irina Gutsche et Maria Bacia de l'IBS pour les tests en tomographie, les belles grilles Lacey, et les conseils sur le projet. Enfin, un merci tout particulier à Ambroise Desfosses, qui m'a accompagnée non seulement pour les tests en tomographie, mais ensuite pour l'analyse des données, avec ses indications sur tous les tests menés sur les LDLs (et cet arbre géant de calculs). Merci pour les conseils et l'enthousiasme que tu manifestes pour le projet.

Une mention spéciale à Geoffrey Woollard, dont l'initiative de lancer un groupe d'études sur la cryo-EM en ligne (Online CryoEM Study Group) m'a énormément aidé dans la compréhension de la cryo-EM.

J'ai beaucoup appris au travers d'autres collaborations, et je remercie les chercheurs concernés. Tout d'abord, Leonardo Chiappisi, pour m'avoir patiemment partagé ton expérience sur le DSC. Je suis heureuse que l'on puisse maintenant le partager avec plus de monde via *pyDSC*. Je remercie Burkhard Bechinger et Arnaud Marquette pour m'avoir fait confiance pour les mesures sur IN13 et par DSC. Un merci particulier à Dominique Bicout, avec qui j'ai pu assister à la création d'un modèle QENS depuis mon tout premier stage, et même eu la chance d'y participer. Merci pour toutes les explications sur la théorie, pour avoir écouté mes propositions, et m'avoir confié les tests du modèle Matryoshka.

Je tiens à remercier tout spécialement Tatsuhito Matsuo, qui nous a rejoint dans l'élaboration du modèle Matryoshka dés son arrivée dans notre groupe, et partage avec nous sa grande expérience en QENS sur plein de projets différents. Merci également pour répondre aux questions à tout moment de la journée au bureau, et toutes les discussions le midi (et la découverte des Manju).

Pour les trois comités de thèse, pour avoir pris le temps de lire chaque rapport avec minutie et leurs retours instructifs, merci encore à Marc Joyeux et Hervé Guillou. Merci également à Hervé qui nous a accueilli dans son laboratoire pour des mesures DSC.

À tous les doctorants, anciens, nouveaux ou maintenant docteurs, avec qui on a partagé cette grande aventure de thèse, mille mercis pour tous les midis, les cafés, les soirées, les conférences, les beamtimes, en particulier, pour ne citer qu'eux: Dominik Zeller, Loreto Misuraca, Marta Salvador-Castell, Tatiana Renzi, Mélanie Léger, Stéphanie Bolik, Cyrielle Métais, Daniele Di Bari, Antonio Caliò, Louise Colin, Myriam Rahalia, Agathe Nidriche, Elizabeth Barthelemy (merci encore pour la relecture de l'anglais, et merci à toi et Pierrick Morin pour la photo de l'aquarelle qui est maintenant en couverture du JPC Letters de Décembre). Une autre avalanche de mercis aux post-doctorants qui ont survécu avant nous: Marco Grimaldo, Sylvain Mézil, à tous les amis randonneurs pour changer d'air: Olivier Lecarme, Arthur Rönisch, Joris Martinez, Lucie Perez, Christian Beck, ceux qui me suivent depuis le master: Marine Deleu, Adrien Carron, et celle qui me suit depuis encore plus longtemps: Justine Hauchard.

Enfin, pour tout leur soutien, je tiens à remercier ma famille, mon oncle Pham et nos thés sur la physique, mon petit frère Antoine pour les memes quotidiens (sic), ma petite soeur Apolline avec qui nous avons maintenant des conversations passionnantes sur la pédagogie, et mes parents, qui ont toujours cru en moi, et m'ont toujours poussé à faire de mon mieux.

Contents

Abbreviations

Table 1: Abbreviations of the thesis.

3DVA	3D Variability Analysis
apo B-48	apolipoprotein B-48
apo B-100	apolipoprotein B-100
cryo-EM	cryo-Electron Microscopy
cryoDRGN	Deep Reconstructing Generative Networks
cryoSPARC	cryo-EM Single Particle Ab-initio Reconstruction and Classification
CCD	Charged Coupled Devices
CVDs	Cardiovascular diseases
CTF	Contrast Transfer Function
DSC	Differential Scanning Calorimetry
DED	Direct Electron Detectors
EINS	Elastic Incoherent Neutron Scattering
EISF	Elastic Incoherent Structure Factor
EM	Electron Microscopy
ET	Electron Tomography
ESR	Electron Spin Resonance
ESRF	European Synchroton Radiation Facility
FSC	Fourier Shell Correlation
FEG	Field-Emission Gun
FFT	Fast-Fourier Transform
HDL	High-Density Lipoprotein
HHP	Hydrostatic High-Pressure
HWHM	Half-Width at Half-Maximum
IBS	Institut de Biologie Structurale
IDL	Intermediate-Density Lipoprotein
IECB	Institut Européen de Chimie et Biologie
ILL	Institut Laue-Langevin
LAMP	Large-Array Manipulation Program

LDL	Low-Density Lipoprotein
LDLr	Low-Density Lipoprotein receptor
N-LDL	Normolipidemic Low-Density Lipoprotein
Mantid	Manipulation and Analysis Toolkit for Instrumentation Data
MD	Molecular Dynamics
MSDs	Mean-Square Displacements
NP40	Nonidet P-40
PCA	Principal Component Analysis
QENS	Quasi-Elastic incoherent Neutron Scattering
RELION	REgularized LIkelihood OptimizatioN
SANS	Small-Angle Neutron Scattering
SAS	Small-Angle Scattering
SAXS	Small-Angle X-ray Scattering
SDS-PAGE	Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis
SEC	Size-Exclusion Chromatography
SNR	Signal-over-Noise Ratio
SPA	Single-Particle Analysis
STA	Sub-Tomogram Averaging
TG-LDL	Triglyceride-rich Low-Density Lipoprotein
TEM	Transmission Electron Microscopy
UAc	Uranyl Acetate
VLDL	Very-Low-Density Lipoprotein

Abstract

English version (EN)

Low-density lipoproteins (LDL) play a crucial role in the metabolism of cholesterol in the blood. Responsible for its transport from the liver to the organs, their accumulation in the arteries is the cause of cardiovascular diseases, such as atherosclerosis. Consisting of a core composed of cholesterol, in its free or esterified form, as well as triglycerides, LDL is surrounded by a membrane of phospholipids, as well as a huge protein: apolipoprotein B-100 (apo B-100). On LDL, the protein contains parts exposed on the surface, and others partially incorporated into the membrane. Apo B-100 is involved in many functionalities of LDL, such as the reception of LDL by organs, or the conversion of VLDL (very low density lipoproteins) into LDL.

In this context, this thesis work focused on several fundamental questions, which are still scarcely studied: What is the structure of LDL and where is the apo B-100 protein located? What are the molecular dynamics of apo B-100?

In a first part, the dynamics of apo B-100 was studied by elastic and quasi-elastic incoherent neutron scattering (EINS, QENS), techniques that allow to access the picosecond to the nanosecond time scale. Due to the nature of apo B-100, and the consequences of its extraction, the measurements gave rise to new questions. Indeed, the final sample includes both apo B-100, but also the detergent used for its solubilization; Nonidet P-40 (NP40). However, incoherent scattering measurements integrate all the contributions of hydrogen atoms, both from the protein and the detergent. To separate the contributions, and to solve a problem that is not only found in the case of apo B-100, but more generally for membrane proteins, we built a new model to analyze the QENS data. From there, we highlighted the acceleration of the internal dynamics of NP40 in the presence of apo B-100, compared to measurements on pure NP40. Furthermore, we were able to quantify the molecular dynamics of apo B-100. With this methodology, we hope to pave the way for more dynamical studies with neutron scattering on systems such as membrane proteins, where the presence of detergent must be taken into account.

In parallel to the dynamical investigation, we explored the structure of whole LDL by cryo-electron microscopy (cryo-EM), with these objectives in mind: to improve the quality of 3D maps of LDL in view of the recent technical developments, and to localize the apo B-100 protein on LDL's surface. In cryo-EM, LDL images are used as a basis to reconstruct a 3D map. As the orientations of LDL are not known a priori, the 3D reconstruction relies on different algorithms, implemented in software packages, such as RELION and cryoSPARC. Their application first allowed us to quantify more precisely the esterified cholesterol layers observed in the LDL core at low temperature, and we propose a revised bilayer model. Regarding the localization of apo B-100, our map supports the current «belt» model in which apo B-100 encircles LDL. Nevertheless, the map did not provide details about the protein, such as alpha helices or beta sheets, indicating that apo B-100 should be extremely flexible on LDL, a property that was not a consensus in the LDL community until now. The presence of lipids, coupled with significant flexibility of apo B-100, makes high-resolution reconstruction of LDL and apo B-100 a real challenge, which cannot yet be solved with current tools. By making our raw data available online, and by presenting the state of our current work, we hope to initiate new discussions in cryo-EM, and the emergence of new algorithms to reconstruct such complexes.

Version française (FR)

Les lipoprotéines à basse densité (LDL) jouent un rôle crucial dans le métabolisme du cholestérol. Responsables de son transport du foie vers les organes, leur accumulation dans les artères est à l'origine de maladies cardiovasculaires, telles que l'athérosclérose. Constituées d'un cœur composé de cholestérol, dans sa forme libre ou estérifié, ainsi que de triglycérides, les LDL sont entourées d'une membrane de phospholipides, ainsi que d'une protéine immense : l'apolipoprotéine B-100 (apo B-100). Sur la LDL, la protéine contient des parties exposées en surface, et d'autres partiellement incorporées dans la membrane. Apo B-100 est impliquée dans de nombreuses fonctionnalités de la LDL, comme la réception des LDL par les organes, ou la conversion des VLDL (lipoprotéines à très basse densité) en LDL.

Dans ce contexte, ce travail de thèse s'est concentré sur plusieurs questions fondamentales, qui sont encore peu étudiées : Quelle est la structure d'une LDL et où se trouve la protéine apo B-100 ? Quelle est la dynamique moléculaire d'apo B-100 ?

Dans un premier volet, la dynamique d'apo B-100 a été étudiée par diffusion incohérente élastique et quasi-élastique de neutrons (EINS, QENS), des techniques qui permettent d'accéder à des échelles de temps de la picoseconde à la nanoseconde. De par la nature de apo B-100, et des conséquences de son extraction, les mesures ont

donné lieu à de nouvelles problématiques. En effet, l'échantillon final inclut à la fois apo B-100, mais aussi le détergent utilisé pour sa solubilisation ; Nonidet P-40 (NP40). Hors, les mesures de diffusion incohérente intègrent toutes les contributions des atomes d'hydrogène, à la fois la protéine ou le détergent. Pour séparer les contributions, et apporter une solution à un problème qui ne se rencontre pas seulement dans le cas d'apo B-100, mais plus généralement avec les protéines membranaires, nous avons construit un nouveau modèle pour analyser les données QENS. Nous avons ainsi mis en évidence l'accélération de la dynamique interne de NP40 en présence de apo B-100, par comparaison aux mesures sur NP40 pur. En outre, nous avons pu quantifier la dynamique moléculaire d'apo B-100. Avec cette méthodologie, nous espérons ouvrir la voie à plus d'études dynamiques en diffusion de neutrons sur des systèmes comme les protéines membranaires, où la présence de détergent doit être prise en compte.

En parallèle des travaux sur la dynamique, nous avons exploré la structure des LDL entières par cryo-microscopie électronique (cryo-EM), avec ces objectifs en tête : améliorer la qualité des cartes 3D de LDL au vu des récentes évolutions techniques, et localiser la protéine apo B-100 à la surface de la LDL. En cryo-EM, les images des LDL servent de base pour reconstruire une carte 3D. Comme les orientations des LDL ne sont pas connues a priori, la reconstruction 3D s'appuie sur différents logiciels de reconstruction, tels que RELION et cryoSPARC. Leur application a d'abord permis de quantifier plus précisément les couches de cholestérol estérifié observées dans le cœur des LDL à basse température, et nous proposons un modèle révisé de bicouches. Concernant la localisation de apo B-100, notre carte soutient le modèle actuel dans lequel apo B-100 « ceinture » la LDL. Néanmoins, elle n'a pas permis d'obtenir des détails sur la protéine, comme des hélices alpha ou feuillets bêta, ce qui indique que apo B-100 serait extrêmement flexible sur la LDL, une propriété qui n'était pas un consensus dans la communauté des LDL jusqu'ici. La présence de lipides, couplée à une importante flexibilité de apo B-100, font de la reconstruction à haute résolution de la LDL et de apo B-100 un véritable défi, qui ne peut pas encore être résolu avec les outils actuels. En mettant en ligne nos données brutes, et en présentant l'état de nos travaux actuels, nous espérons amorcer de nouvelles discussions en cryo-EM, et l'émergence de nouveaux algorithmes pour reconstruire de tels complexes.

Introduction

English version (EN)

The thesis work presented here was conducted under the supervision of Pr. Judith Peters, in the Laboratoire Interdisciplinaire de Physique (LiPhy, Grenoble, France), at the University Grenoble-Alpes (France), and at the Institut Laue-Langevin (ILL, Grenoble, France) with a long-term visitor status, where all our neutron scattering experiments have been carried out. The thesis is part of a strong collaboration with the group of Assoc. Pr. Ruth Prassl at the Medical University of Graz (Austria), whose expertise on lipoproteins dates back to their very first molecular characterizations. In particular, we worked closely with Assoc. Pr. Ruth Prassl and Dr. Karin Kornmueller. First, to pursue investigation on the molecular dynamics, using incoherent neutron scattering, a technique on which our group has expertise. Second, to explore the Low-Density Lipoprotein (LDL) structure using cryo-electron microscopy (cryo-EM). For the latter, we built collaborations with cryo-EM experts, namely Dr. Eaazhisai Kandiah (European Synchroton Radiation Facility, ESRF, Grenoble, France) and Dr. Daouda Traoré (ILL, Grenoble, France), both in charge of the Titan Krios at the ESRF, but also with the group of Dr. Irina Gutsche (Institut de Biologie Structurale, IBS, Grenoble, France), in particular with Dr. Ambroise Desfosses. My contribution to this joint collaboration and the related results are the subject of this thesis, under the title « Studies of Low-Density Lipoproteins by Cryo-Electron Microscopy and Neutron Scattering ».

Low-Density Lipoproteins (LDLs) act in the blood circulation as the main cholesterol transporters. Their hydrophilic shell is composed of a monolayer of phospholipids and the apolipoprotein B-100 (apo B-100) protein, whose molecular weight reaches about 550 kDa [1, 2]. LDLs encompass various types of hydrophobic elements, starting with the cholesterol in its free or esterified form, as well as small amount of triglycerides [3–5]. They are often misleadingly termed as the «bad cholesterol», in opposition to the High-Density Lipoproteins (HDLs), designated as the «good cholesterol». The metabolic pathways are obviously much more complex than this simple picture, and LDLs remain essential in our body. However, it was recognized

that a high concentration of LDLs in the plasma is a risk factor in cardiovascular diseases, while reversely a high concentration of HDLs was shown to reduce this risk [6]. Indeed, the onset of many of these diseases, such as atherosclerosis, is the uncontrolled accumulation of LDLs in the arteries, which leads to the formation of foam cells, and ultimately to thrombosis or stroke [7]. Often, modified forms of LDLs are involved in this process, like triglyceride-rich (TG-rich) or oxidized LDLs [8–14]. Nevertheless, the molecular properties of the LDL, which are inherently related to its binding properties or recognition, still contain open questions, starting with the exact structure of the LDL, or its protein moiety apo B-100. In fact, their molecular properties started to be investigated quite early, and with cutting-edge techniques all along the years, like small-angle scattering techniques [15–23], X-ray crystallography [24, 25], electron microscopy [26–34], or incoherent neutron scattering [35–37]. However, the huge size of the LDL, around 250 Å of diameter, as well as the apo B-100 protein, which remains non-exchangeable by nature, and the diversity in the composing lipids, make the molecular investigation of the LDL highly challenging.

Following the characterization of the LDL and apo B-100 at the molecular scale, this thesis addresses several fundamental questions, which are still scarcely studied: What is the structure of the LDL and where is the apo B-100 protein located? What is the molecular dynamics of the apo B-100 and how it compares to that of the whole LDL?

In a first chapter, the LDL and its protein apo B-100 will be presented in details, their properties and related metabolism, along with the molecular characterizations performed thorough the last years. A focus will be made on the structural and dynamical investigation performed so far. Furthermore, a section will be devoted to the sample preparation of both LDL and apo B-100.

Then, a second chapter will expose the theoretical framework behind the core techniques of this thesis, namely the incoherent neutron scattering and cryo-EM. In particular, the fundamentals of scattering theory, from the Schrödinger's equation, to the underlying approximations and hypotheses, will be described. An overview of the image formation, whose bases also reside in scattering, will be given as well.

The third chapter will draw a general overview from the production to the detection of the probe particles, neutrons and electrons, at the root of the techniques employed. The related instruments, spectrometers and microscopes are outlined, and an overall view of the image processing in single-particle cryo-EM is given.

Investigation of the molecular dynamics of detergent-solubilized apo B-100 will be covered in Chapter 4. Due to the non-exchangeable nature of apo B-100, and

despite extensive dialysis and trials, the final sample contains both apo B-100 protein and the detergent used for its solubilization; Nonidet P-40 (NP40). To separate the contributions in the context of incoherent neutron scattering, a new approach to analyze quasi-elastic neutron scattering (QENS) data will be introduced in this chapter. Application of this approach to our apo B-100 / NP40 sample, as well as on pure NP40, will be presented, and results exposed.

The last chapter will deal with the exploration of the LDL structure with single-particle cryo-EM. 3D reconstructions from different software packages will be described, and an actualized model of the LDL organization will be proposed.

Finally, a last overview of this work, along with perspectives, will be provided in the conclusion.

Version française (FR)

Ces travaux de thèses ont été menés sous la direction de Pr. Judith Peters, au Laboratoire Interdisciplinaire de Physique (LiPhy, Grenoble, France), à l'Université Grenoble-Alpes (France), et en qualité de visiteure long-terme à l'Institut Laue-Langevin (ILL, Grenoble, France), où toutes nos expériences de diffusion neutroniques ont été effectuées. Cette thèse s'inclut dans une collaboration avec le groupe de Assoc. Pr. Ruth Prassl de l'Université Médicale de Graz (Autriche), dont l'expertise dans le domaine des lipoprotéines remonte à leurs toutes premières caractérisations à l'échelle moléculaire. En particulier, nous avons travaillé en étroite collaboration avec Assoc. Pr. Ruth Prassl et Dr. Karin Kornmueller. Tout d'abord, dans le but de poursuivre les études de dynamique moléculaire, à l'aide de la diffusion incohérente de neutrons, une technique dans laquelle notre groupe possède l'expertise. Ensuite, pour explorer la structure de la lipoprotéine à basse-densité (LDL) avec la cryo-microscopie électronique (cryo-EM). Pour ce faire, nous avons bâti de nouvelles collaborations avec des experts de cryo-EM, à savoir Dr. Eaazhisai Kandiah (European Synchroton Radiation Facility, ESRF, Grenoble, France) et Dr. Daouda Traoré (ILL, Grenoble, France), tous deux en charge du Titan Krios de l'ESRF, mais aussi avec le groupe de Dr. Irina Gutsche (Institut de Biologie Structurale, IBS, Grenoble, France), en particulier Dr. Ambroise Desfosses. Ma contribution à cette large collaboration et les résultats associés sont le sujet de cette thèse, dont le titre est « L'étude de la lipoprotéine à basse densité par microscopie électronique et diffusion neutronique ».

Les lipoprotéines à basse-densité (LDLs) agissent dans la circulation sanguine comme le principal transporteur du cholestérol. Leur enveloppe hydrophile se compose d'une monocouche de phospholipides et de l'apolipoprotéine B-100 (apo B-100), dont

la masse moléculaire atteint jusque 550 kDa [1, 2]. Les LDLs englobent divers types d'éléments hydrophobes, à commencer par le cholestérol, dans sa version libre ou estérifiée, ainsi que des petites quantités de triglycérides [3–5]. Elles sont souvent dénommées de manière trompeuse comme le «mauvais cholestérol», en opposition aux lipoprotéines de haute-densité (HDLs), désignées comme le «bon cholestérol». Les chemins métaboliques sont évidemment beaucoup plus compliqués que cette image simplifiée, et les LDLs demeurent essentiels dans notre corps. Cependant, plusieurs consensus d'experts ont reconnu qu'une concentration élevée de LDLs dans le plasma est un facteur de risque dans les maladies cardiovasculaires, alors qu'inversement une concentration élevée de HDLs réduit ce risque [6]. En effet, le début d'une large majorité de ces maladies, telle que l'athérosclérose, réside dans l'accumulation incontrôlée des LDLs dans les artères, ce qui conduit à la formation de plaques d'athérome, et finalement à une thrombose ou un arrêt cardiaque [7]. Souvent, des formes modifiées des LDLs sont impliquées dans ce processus, tels que les LDLs riches en triglycérides (TG) ou oxydées [8–14]. Cependant, les propriétés moléculaires de la LDL, qui sont intimement liées à leurs propriétés de liaison ou reconnaissance, comportent encore des questions ouvertes, à commencer par la structure exacte de la LDL, ou de sa protéine apo B-100. En réalité, leurs propriétés moléculaires ont commencé à être étudiées assez tôt, et avec des techniques de pointe tout aux long des années, telles que les techniques de diffusion aux petits angles [15–23], la cristallographie aux rayons X [24, 25], la microscopie électronique [26–34], ou la diffusion incohérente de neutrons [35–37]. Néanmoins, la taille considérable des LDLs, autour de 250 Å, ou encore de la protéine apo B-100, qui demeure non-échangeable par nature, ainsi que la diversité des lipides qui les composent, rendent l'étude moléculaire des LDLs extrêmement ambitieuse.

Poursuivant la caractérisation des LDLs et de l'apo B-100 à l'échelle moléculaire, cette thèse aborde plusieurs questions fondamentales, encore peu étudiées : Quelle est la structure des LDLs et où se trouve la protéine apo B-100 ? Quelle est la dynamique moléculaire de l'apo B-100 et comment se compare-t-elle à celle de la LDL entière ?

Dans un premier chapitre, la LDL et sa protéine apo B-100 seront présentées en détail, notamment leurs propriétés et le métabolisme associé, ainsi que les caractérisations moléculaires réalisées au cours des dernières années. L'accent sera mis sur les études structurelles et dynamiques réalisées jusqu'à présent. En outre, une section sera consacrée à la préparation des échantillons de LDL et d'apo B-100.

Ensuite, un deuxième chapitre exposera le cadre théorique qui sous-tend les techniques centrales de cette thèse, à savoir la diffusion incohérente des neutrons et la cryo-EM. En particulier, les principes fondamentaux de la théorie de la diffusion, de l'équation de Schrödinger aux approximations et hypothèses sous-jacentes, seront

décrits. Un aperçu de la formation d'images, dont les bases résident également dans la diffusion, sera donné.

Le troisième chapitre donnera un aperçu général de la production jusqu'à la détection des particules sondes employées, à savoir les neutrons et les électrons, qui sont à la base des techniques utilisées. Les instruments connexes, les spectromètres et les microscopes, seront décrits, ainsi qu'une vue d'ensemble du traitement des images dans la cryo-EM en «particule seule».

L'étude de la dynamique moléculaire de l'apo B-100 solubilisée par détergent sera abordée au chapitre 4. En raison de la nature non échangeable de l'apo B-100, et malgré une dialyse et des essais approfondis, l'échantillon final contient à la fois la protéine apo B-100 et le détergent utilisé pour sa solubilisation, Nonidet P-40 (NP40). Pour séparer les contributions dans le contexte de la diffusion incohérente des neutrons, une nouvelle approche pour analyser les données de diffusion quasi-élastique des neutrons (QENS) sera présentée dans ce chapitre. L'application de cette approche à notre échantillon d'apo B-100 / NP40, ainsi qu'au NP40 pur, sera présentée, et les résultats exposés.

Le dernier chapitre traitera de l'exploration de la structure des LDL à l'aide de la cryo-EM en «particule seule». Des reconstructions 3D issues de différents logiciels seront décrites, et un modèle actualisé de l'organisation des LDL sera proposé.

Enfin, une dernière synthèse de ces travaux, ainsi que leurs perspectives, seront fournies dans la conclusion.

Chapter 1

Low-density lipoproteins

Lipids and fatty acids are essential components of the cell membranes [38]. Among them, cholesterol holds many physiological roles [39]. To cite a few, it modulates the membrane fluidity and its permeability, but is also a precursor for bile acids, steroid hormones and vitamin D. However, like other lipids, it is hydrophobic, thus insoluble in water. As a consequence, cholesterol transport is ensured by an important class of transporters; the lipoproteins. Found in the serum plasma, lipoproteins carry major lipids in the blood circulation, such as the cholesterol, but also triglycerides.

1.1 Lipoproteins

1.1.1 Classification

In a general picture, lipoproteins are composed of a core, comprising the various fatty acids to transport (free cholesterol, cholesterol esters, triglycerides), and a phospholipid membrane, associated to one or several apolipoproteins [1, 2]. They vary in size, role, lipid constituents, major apolipoproteins, and are classified by their densities, as summarized in Figure 1.1 and Table 1.1.

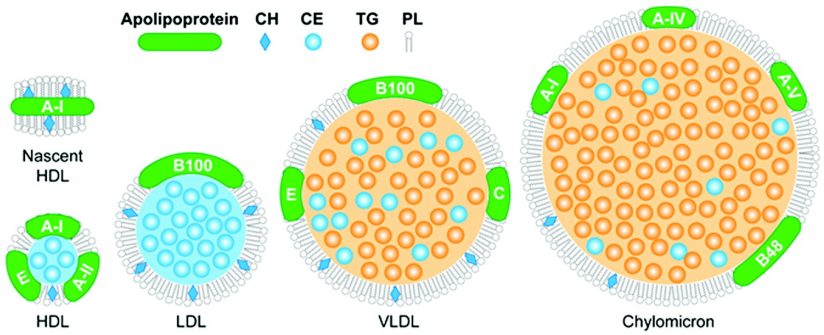


Figure 1.1: Scheme of the different types of lipoproteins. CH stands for cholesterol, CE to its esterified form, TG for triglycerides, and PL for phospholipids. Image taken from Bricarello et al. [40].

Lipoproteins	Abbreviation	Diameter (nm)	Density (g/cm^3)	Main apolipoproteins
Chylomicrons	Chylo, CM	75 - 1200	0.93	Apo B-48, Apo E, Apo C, Apo A-I, Apo A-II, Apo A-IV
Very-low-density lipoproteins	VLDL	30 - 80	0.93 - 1.006	Apo B-100, Apo E, Apo C
Intermediate density lipoproteins	IDL	25 - 35	1.006 - 1.019	Apo B-100, Apo E, Apo C
Low-density lipoproteins	LDL	18 - 25	1.019 - 1.063	Apo B-100
High-density lipoproteins	HDL	5 - 12	1.063 - 1.21	Apo A-I, Apo A-II, Apo C

Table 1.1: Properties of the lipoproteins. Adapted from Segrest et al. [2].

Chylomicrons and Very-Low-Density Lipoproteins (VLDLs) are the largest lipoproteins, and are enriched in triglycerides, whereas Low-Density Lipoproteins (LDLs) and High-Density Lipoproteins (HDLs) are enriched in cholesterol. As the range of density in each class of lipoproteins is quite large, as remarked in Table 1.1, each class can still be separated in subspecies, whose density and size will slightly differ between them, as studied for LDLs in Chapman et al. [41].

Apolipoproteins, abbreviated as apo, are also classified in several classes, with different letters [2, 42]. They can be separated in two groups; the non-exchangeable apolipoproteins (apo B-48 and apo B-100), which are insoluble, and the exchangeable apolipoproteins (all the others), which are soluble in water [1]. The non-exchangeable proteins are tightly associated to the lipoproteins from their synthesis [43, 44], to the different reactions occurring subsequently, until their degradation or absorption. Due to their high insolubility, the study of the non-exchangeable proteins is quite challenging, and their molecular properties are much less known, contrary to exchangeable proteins.

1.1.2 Metabolic pathways

As represented in Figure 1.2, the metabolic path related to lipoproteins is complex and involved several pathways [2, 42, 45].

The first one is called the exogeneous pathway, because it corresponds to the lipids coming from the dietary. This pathway is related to the intestine, from where the lipids are transported by chylomicrons. In the blood, they are reduced in chylomicrons remnants by the lipoprotein lipase, which hydrolyses triglycerides into fatty acids and glycerol. The fatty acids are then either stored in adipocytes, or oxidized in skeletal muscle tissues, where they act as an energy source. The chylomicrons remnants are

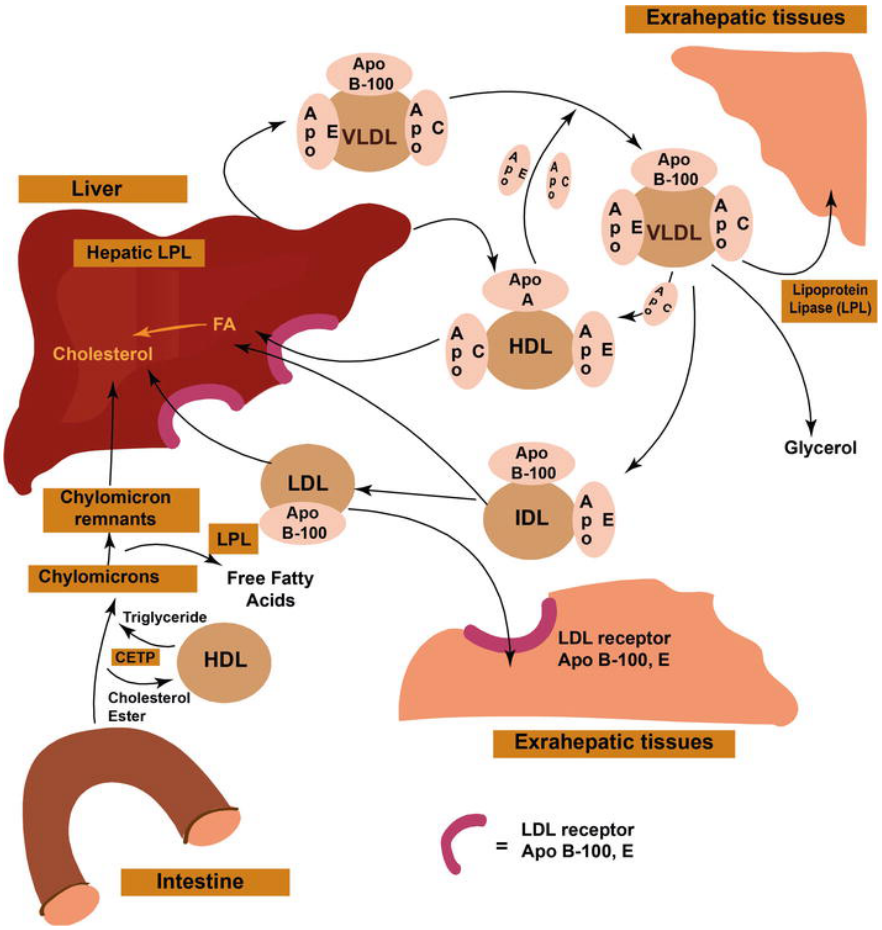


Figure 1.2: Representation of the lipoproteins' metabolism. Image taken from Zhyvotovska et al. [45].

finally cleared in the liver, where their residuals, triglycerides and cholesterol, are then re-used, either for the synthesis of VLDLs, or for conversion into bile acids.

In parallel, the endogeneous pathway concerns lipids of the body, and starts in the liver, with the formation of VLDLs. Similarly to chylomicrons, VLDLs are then migrating in the blood, where they will be reduced by the lipoprotein lipase, losing triglycerides, and becoming smaller. Their remnants are called Intermediate-Density Lipoproteins (IDLs), which are either cleared in the liver, or further reduced by the lipase, leading then to LDLs, where few triglycerides molecules remain. At each hydrolysis step, residuals of VLDLs or IDLs are recovered by HDLs, either the apolipoproteins, or the lipidic content.

The major role of LDLs is to transport cholesterol to the peripheral organs. It will bind with the LDL receptor of cells, through apolipoprotein B-100 (apo B-100), and will be taken by endocytosis [46, 47]. In the cells, LDLs are broken up, each component is recovered, and used, for instance, in the composition of the cell membrane. Remaining LDLs are supposed to be recaptured in the liver. The uncontrolled accumulation of LDLs in the arteries marks the onset of cardiovascular

diseases (CVDs), and high levels of LDLs are associated with higher risk of CVDs [6].

Finally, excess of cholesterol in the tissues is removed by HDLs, through the reverse transport of cholesterol, whose end point is again in the liver. In general, a high level of HDL is associated with a lower risk for CVDs [6, 48].

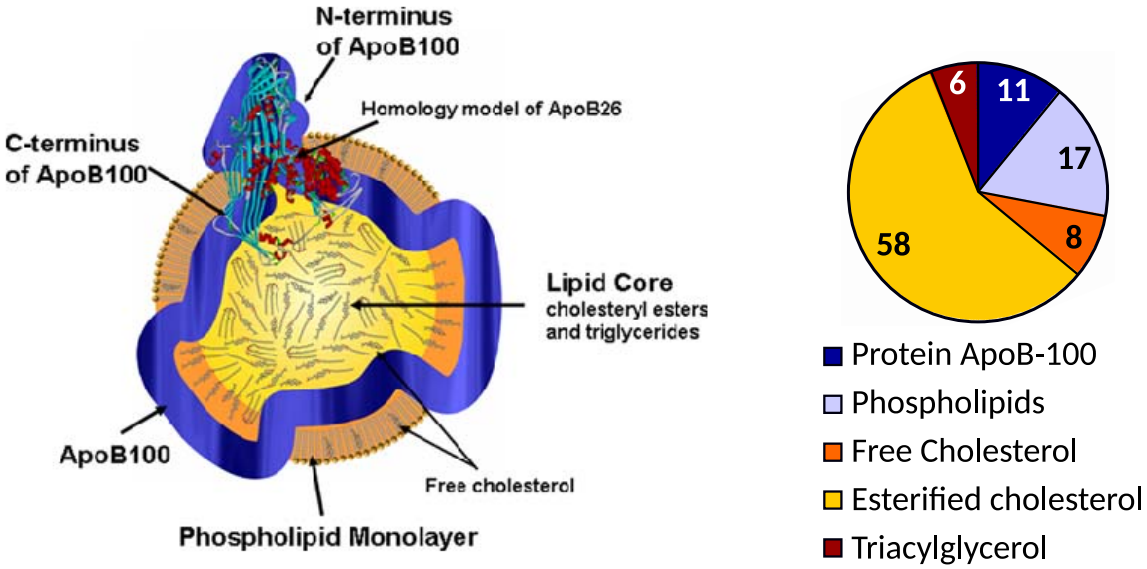
As mentioned previously, lipoproteins' metabolism is often linked to CVDs, especially atherosclerosis. This disease starts with the retention of LDLs in the sub-endothelial cells of the arteries [49, 50]. Then, oxidation of LDLs, and subsequent aggregation, will trigger an inflammatory reaction, with the coming of macrophages. The uptake of the oxidized LDLs by macrophages leads to the formation of foam cells, whose accumulation generates fibrous plaques in the arteries. These plaques are characteristic of atherosclerosis, and promote thrombosis [7, 51, 52]. Several research axes are followed to better understand this pathology, and to help in the development of therapeutics. Numerous studies investigate the possible genetic factors, as reviewed for example by Hegele [53]. Understanding the many familial diseases related to the metabolism of lipoproteins is crucial, with the isolation of possible involved genes, the potential mutations and how they affect the phenotypes through the proteins. In parallel, atherosclerosis also involves modified forms of LDLs. The smaller subspecies of LDLs were shown to be more atherogenic through molecular and epidemiological studies [12, 13, 54], as well as triglyceride-rich (TG) LDLs [9–11], while oxidized LDLs are strongly implicated in the foam cell formation [8, 14]. The reasons of such atherogenic properties can be related to different binding properties with the LDL receptor, or with the proteoglycans lying in the sub-endothelial space, which imply possible structural modifications of the LDL and apo B-100. In addition, apo B-100 misfolding is also advanced as a factor for LDLs accumulation [55, 56].

However, compared to the HDL, the molecular properties of the LDL, which are ultimately linked to its binding and recognition, are much less known, even in its normolipidemic form. Indeed, studies of the LDL at the molecular level remain challenging, due to its large size, high content and variety of lipids, or the non-exchangeable nature of apo B-100 protein. In the following, a short overview of LDL molecular properties will be presented, with an emphasis on structural and dynamical studies.

1.2 LDL composition

1.2.1 Constituents

LDLs can be described by a core containing cholesterol, in its free and esterified forms, and triglycerides, surrounded by a shell, composed of a monolayer of phospholipids, some free cholesterol molecules, as well as a single copy of apo B-100 [3–5]. A scheme is represented in Figure 1.3a, while the Figure 1.3b displays the fraction of each of these components for one LDL batch.



(a) Scheme of a LDL

(b) Composition

Figure 1.3: (a) Representation of a LDL particle (image taken from Prassl and Laggner [5]). (b) Percentage of each LDL component, from concentrations measurements by photometric assays in the Medical University of Graz for one LDL sample.

The diagram in Figure 1.3b shows the high content of lipids, compared to the protein proportion (only 11 %). Table 1.2 gives an estimation of the amount of each molecule, assuming a molecular weight of 2.5 MDa for the LDL [4, 8]. Among the lipids, whose chemical structures are shown in Figure 1.4, the phospholipids are amphipathic, with a hydrophilic polar head group, and one or several hydrophobic tail(s). Due to their nature, phospholipids are the main components of the LDL membrane and separate the other constituents, all hydrophobic, from the aqueous environment. Inside the LDL, triglycerides and cholesteryl esters are densely packed, whereas free (unesterified) cholesterols are supposed to form an intermediate layer between the core and the surface, as reviewed in Hevonoja et al. [4].

LDL part	Molecule	Number of molecules
Core	Triglyceride (TG)	70
	Free (unesterified) cholesterol (UC)	∼200
	Cholesteryl ester (CE)	1600
Surface	Apo B-100	1
	Free (unesterified) cholesterol (UC)	∼400
	Phospholipids	700
Phospholipids	Phosphatidylcholine (PC)	∼450
	Sphingomyelin (SM)	∼185
	Lysophosphatidylcholine (lyso-PC)	∼80
	Phosphatidylethanolamine (PE)	∼10
	Dyacylglycerol (DAG)	∼7
	Ceramide (CER)	∼2

Table 1.2: Main constituent molecules of the LDL, assuming a mean molecular weight of 2.5 MDa. Adapted from Hevonoja et al. [4], Esterbauer et al. [8].

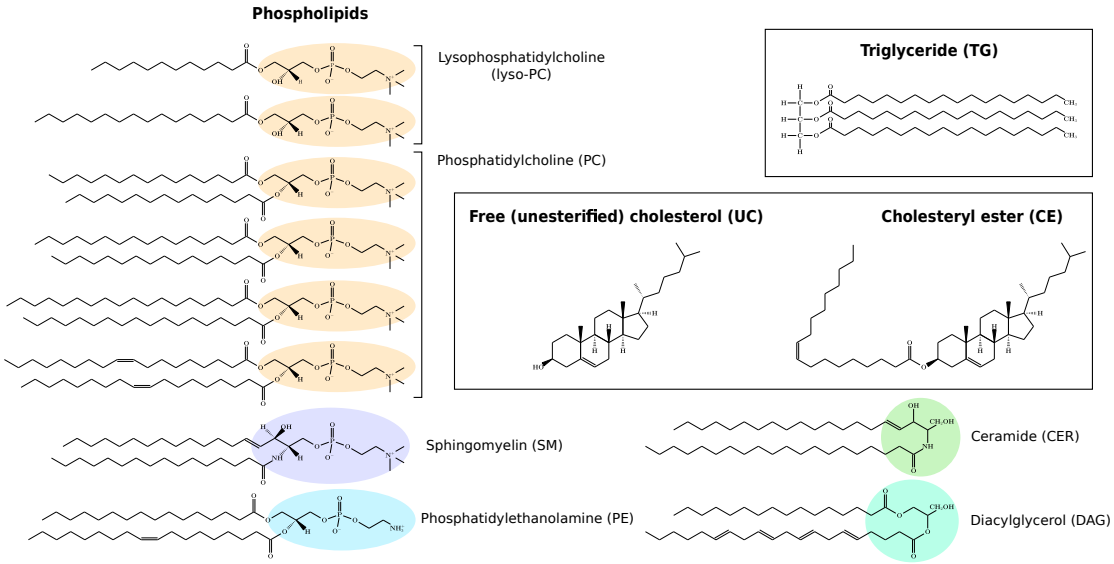


Figure 1.4: Examples of chemical structures for the main constituents of LDLs (chemical structures reproduced from [57–60]). Polar head groups of the phospholipids are represented by colored ellipses, lyso-PC and PC with different chain lengths and saturation are displayed as an example. This list is not exhaustive, and in native LDLs, many various chain lengths and saturation can coexist for the lipids.

1.2.2 Apolipoprotein B-100

Apo B-100 is one of the largest monomeric proteins known, with a molecular weight of about 550 kDa, and 4536 amino acids' residues [2, 61, 62]. Apo B-100 translation occurs in the rough endoplasmic reticulum [62], at the same time as the assembly of

the corresponding lipoprotein, in the liver. From there, the protein will remain bound to the lipoprotein across the different steps of its metabolism [43].

Its sequence can be found in [63, 64], and was determined independently by Yang et al. [65] and Knott et al. [66]. Already from the primary sequence and the complementary studies both groups performed, several properties of apo B-100 were identified and were summarized by Segrest et al. [67] in Figure 1.5:

- Domains accessible for trypsin digestion (supposedly at the LDL surface) and inaccessible ones (possibly buried within the LDL membrane) are alternating. Separation of apo B-100 in several domains was already apparent.

- A cluster of Cysteine residues (filled circles in Figure 1.5), 12 of the total 25, can be found in the NH_2 terminal 500 residues of apo B-100. They are supposed to form disulfide bonds between them, which give more stability to the protein. In addition, and because this region is considered at the surface of the LDL, it led to the hypothesis of a NH_2 domain more globular in appearance compared to the other domains.

- There are 20 possible N-glycosylation sites in the protein, but not all of them are glycosylated (open circles in Figure 1.5). Glycosylation is estimated to represent around 10 % of the total mass of the protein, from the discrepancy between the relative molecular weight of apo B-100 ($\sim$512 kDa) and its measurement from SDS-gel electrophoresis ($\sim$550 kDa). A cluster was identified within the residues $\sim$3000-3500, with all sites containing carbohydrates.

- Identification of the potential receptor-binding site of the LDL with LDL-receptor tending towards a region approximately between residues 3000-3600, before the COOH-terminal region. This region is close to the cluster of glycosylated sites, and to a disulfide bridge. Furthermore, some studies indicate that this site is not accessible when apo B-100 is on the VLDL, but only after conversion to LDL.

The secondary structure of apo B-100 protein was predicted by Segrest et al. [68], with a computer program called LOCATE, and constitutes the so-called «pentapartite» model, where five domains enriched in alpha helices or beta sheets follow one another, and writes as: $NH_2-\beta\alpha_1-\beta_1-\alpha_2-\beta_2-\alpha_3-$COOH. In that model, the $NH_2-\beta\alpha_1$ is supposed to be globular, the amphipathic helices in α_2 and α_3 domains are assumed to be reversibly associated to the lipids, whereas amphipathic β-sheets are considered to form huge domains on the LDL and to be irreversibly linked to the lipids [67]. Infrared spectroscopy measurements of Goormaghtigh et al. [69] showed an orientation of the β-sheets parallel to the phospholipid membrane, which is consistent with the model proposed by Segrest et al. [67]. A direct comparison between

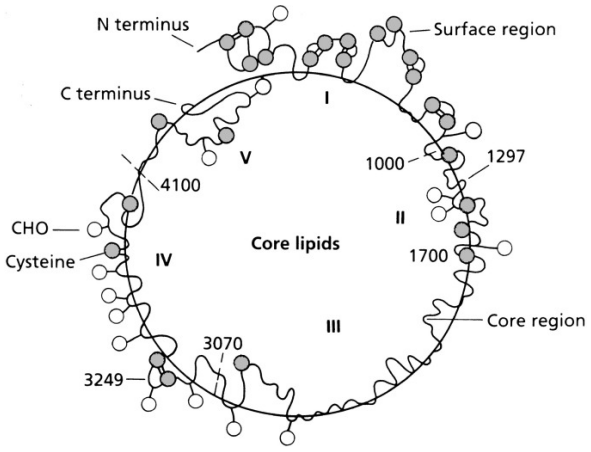


Figure 1.5: Representation of the apo B-100 localization on the LDL. Filled circles correspond to the Cysteine residues, whereas open circles stand for the N-glycosylated sites with carbohydrates. Image taken from Segrest et al. [67], adapted from Yang et al. [65].

the pentapartite model, secondary structure prediction, and the primary sequence is reported for example in Johs et al. [19], and represented in the next section with small-angle scattering studies, in Figure 1.8.

Due to its huge size and high insolubility, and thus difficulties to study apo B-100 in solution, no structure is reported until now. An homology model for the $NH_2 - \beta\alpha_1$ domain with the lamprey lipovitellin [70] was proposed by several groups: Segrest et al. [67], Mann et al. [71], Herscovitz et al. [72], Dashti et al. [73], Jiang et al. [74], Richardson et al. [75], and is represented in Figure 1.6 for $NH_2 - \beta\alpha_1$ 1150 residues.

This homology model enabled Segrest et al. [67], Dashti et al. [73], Richardson et al. [75] among others to formulate hypothesis on the $NH_2 - \beta\alpha_1$ domain function, notably for the co-assembly of apo B-100 with VLDL in the liver.

In a similar way, Kriško and Etchebest [78] built homology models for eight different domains of apo B-100, and proposed a model for the overall shape of the protein, which globally agrees with cryo-electron microscopy images of Gantz et al. [29] (presented below).

However, all these predictions still await for confirmation, and the determination of the structure of full apo B-100, which still remains highly challenging. In parallel, localization of apo B-100 on the native LDL is also of great importance. In the following, a global view of the structural studies on the whole LDL or solubilized apo B-100 will be given.

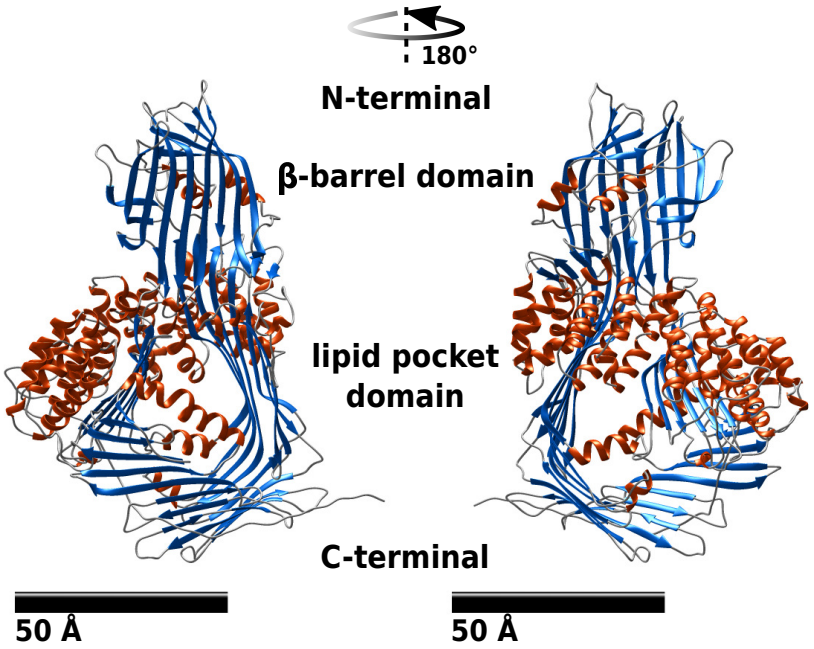


Figure 1.6: Homology model for the $NH_2 - \beta\alpha_1$ 1150 residues of apo B-100, calculated from Modeller V7.7 [76], with input coordinates from the crystal structure of lamprey lipovitellin (PDB-ID: 1LSH [70]). Reproduced from Prassl and Laggner [77]. α-helices are represented in orange, β-sheets in blue. The scale bar corresponds to 50Å.

1.3 Structural studies

1.3.1 X-ray diffraction

Attempts to study the high resolution structure of LDLs with X-ray diffraction were performed by Prassl et al. [24] and Ritter et al. [79], and also reviewed in Prassl and Laggner [77]. In these studies, sub-species of LDLs were crystallized to ensure higher homogeneity. A reconstruction at 2.7 nm resolution was published in Lunin et al. [25], which agreed with small-angle X-rays scattering (SAXS) models from that time [17, 80], as well as cryo-electron microscopy (cryo-EM) images from van Antwerpen et al. [81] or the 3D map of Orlova et al. [28]. Overall, even if the crystals did not diffract at high resolution, these studies showed that LDL crystallization was possible.

1.3.2 Small-angle X-rays and neutron scattering

The advantage of small angle scattering (SAS) techniques is the possibility to perform measurements under native conditions, contrary to X-ray diffraction or electron microscopy, which require specific sample conditions. Small-angle X-rays and neutron scattering (SAXS, SANS) measure an average of all correlations taking place within the sample, and enable to access structural information at low-resolution, like the

general shape of the sample. SAXS and SANS give complementary information, due to the differences in the probe. SAXS is more sensitive to components with higher atomic mass Z, and SANS can sense lighter elements, like Hydrogen or Deuterium.

Very early, studies on native LDLs using SAXS and SANS were carried out, as reported by numerous studies from: Atkinson et al. [15], Deckelbaum et al. [16], Muller et al. [17], Laggner et al. [18], Baumstark et al. [80], Deckelbaum et al. [82], Laggner et al. [83], Laggner and Kostner [84], Laggner et al. [85]. They reported crucial information on the shape of the LDL and its core organization. First, they all agreed with a general spherical shape, of diameter around 22 nm, and the core-shell separation, as presented earlier. Given the number of phospholipids molecules in one LDL, as noted in Table 1.2, Laggner and Kostner [84] expected some parts of the apo B-100 protein to be on the surface of the LDL, which correlate well with trypsin-digestion experiments [65]. Using contrast-matching methods with SANS, and considering a constant distribution of apo B-100, Laggner et al. [18] estimated that the protein should be around 8 Å above the phospholipid membrane, implying a huge interaction with the membrane, but probably also with the lipidic core. Concerning the core, measurements at different temperatures highlighted a transition, occurring between 15°C and 35°C [16, 82], corresponding to the transition of cholesteryl esters from an ordered crystalline phase to a disordered liquid phase, as represented in Figure 1.7.

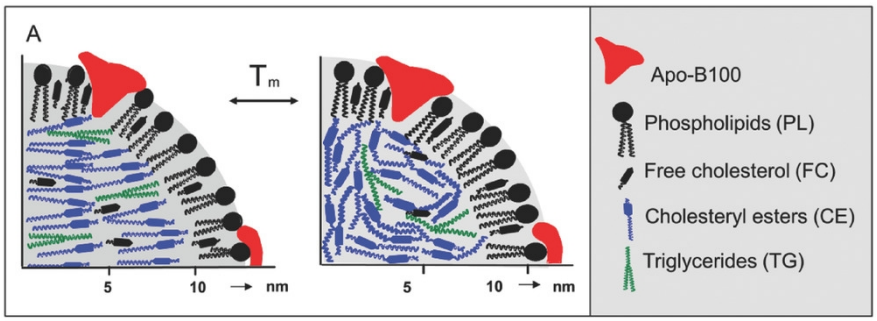


Figure 1.7: Representation of the core transition. A: crystalline phase, where the cholesteryl esters are aligned. B: disordered liquid state. Image taken from Lehofer et al. [23].

The ordering at low temperature is characterized by a high intensity peak around 1.7 nm^{-1} in all mentioned studies, corresponding to a periodic structure of about 3.6 nm, which fits well with the length of a cholesteryl ester molecule. A study by Pregetter et al. [86] using other techniques confirmed the arrangement of the cholesteryl ester lipids, but not the triglycerides molecules. In agreement with these measurements, several models for the core arrangement at low temperature were proposed. Generally, three concentric layers of ordered cholesteryl ester molecules were considered, with the head of cholesteryl pointing towards the shell.

However, some times after, cryo-electron microscopy (cryo-EM) images and reconstruction from Orlova et al. [28] have shown that the core of LDLs has not a radial organization, but a lamellar arrangement, with three flat layers separated by a distance of 3.6 nm, as predicted by the 1.7 nm^{-1} peak from SAXS measurements. In a review from Prassl and Laggner [5], they showed that indeed, SAXS measurements on LDLs can predict periodic distribution and distances from the measured peaks, but cannot distinguish between three concentric or three flat layers.

Even with the rise of cryo-EM and direct reconstructions of the LDL, which we will review in the next section, SAXS and SANS with related modelling, continue to be an essential tool for the LDL investigation. It combines the possibility of working under near-physiological conditions with quick measurements, both in large scale facilities or in-house instruments. During the last years, we can report the modelling of detergent-solubilized apo B-100 structure at low-resolution, through SANS experiments, by Johs et al. [19], whose results are represented in Figure 1.8. They also proposed a possible arrangement of apo B-100 on the LDL, and retrieved a globular $NH_2 - \beta\alpha_1$ domain protruding from the sphere, as predicted by the homology model [75]. In addition, the modelled β_2 domain revealed to be large, and was consistent with proposed loops near the COOH terminal [87, 88], which are hypothesized to modulate the binding to LDL receptor. Overall, this SANS study unveiled structural characteristics of apo B-100, and remains so far the best available model of the protein.

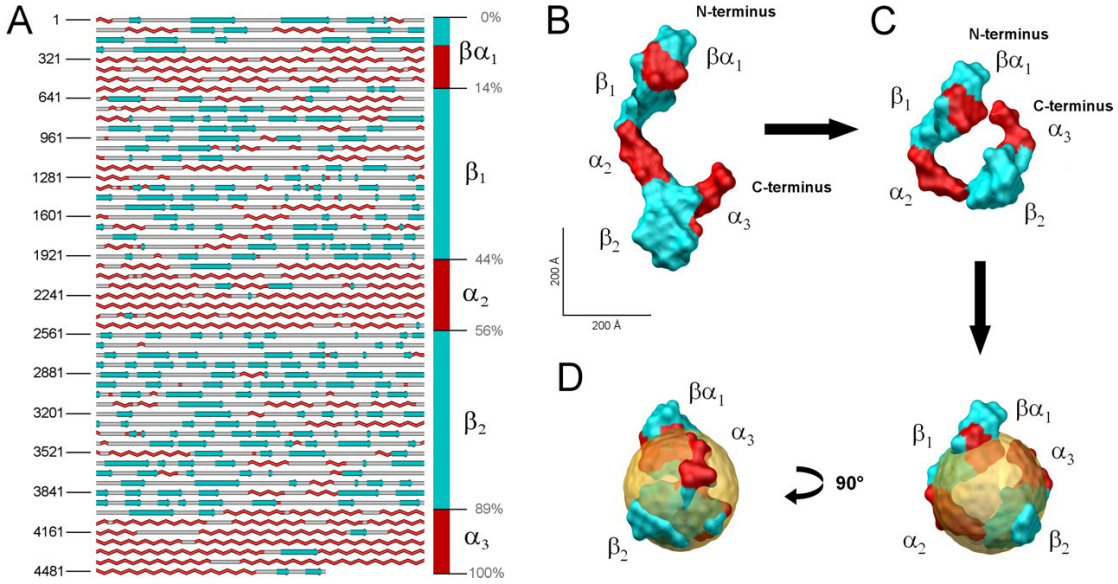


Figure 1.8: Modelling of apo B-100 protein from SANS measurements. A: primary sequence related to the pentapartite model of Segrest et al. [68]. B: low-resolution model mapped with the corresponding secondary structure prediction. C: possible organization of the apo B-100 on the LDL, after arbitrary motions of the domains. D: representation on a sphere to mimic the LDL arrangement. Image taken from Johs et al. [19].

In parallel, the study of native LDLs remains necessary, especially to decipher the conformation of apo B-100 on the surface of the LDL, which cannot be retrieved from the detergent-solubilized apo B-100, as seen in the SANS study [19] or in cryo-EM [29], where the protein is more elongated. SAXS in various conditions and modelling related to SAXS curves of LDLs are still a topical subject, and were addressed by several groups. We can first mention the study on the kinetics of the core transition by Prassl et al. [20]. They showed by time-resolved SAXS that the transition was taking place quite quickly, in the time resolution limits of the heating and cooling procedures, so respectively 10 milliseconds and 2 seconds. Their study indicated that the core transition can then also occur in peripheral blood capillaries, where the residence time of LDLs can be estimated of the orders of several seconds. Open questions still remain on the consequences of such changes within the body, but it points to the equal importance of investigating LDLs both below and above the phase transition.

Oliveira et al. [89] put in perspective LDLs and HDLs under oxidation, as LDL oxidation is known to be a crucial factor in foam cell formation, which ultimately lead to atheroma plaques in the arteries [14, 52]. Their SAXS measurements outlined changes in LDLs, but not HDLs, under oxidation, and very likely linked to the apo B-100, as the core peak at 1.7 nm^{-1} did not show any differences.

As 17-β-estradiol (E2) hormone was shown to bind to LDLs, and to stabilize apo B-100 conformation, especially under oxidation [90, 91], Papi et al. [21] conducted SAXS experiments on both LDLs and LDLs bound to E2. Here, they used a three-component model to fit the SAXS curve, where the apo B-100 is modelled by three protrusions at the surface, in order to represent the three α-domains of the pentapartite model [68]. Their reconstructed 3D model ended up with the three protrusions being quite distinct, and distributed in an asymmetric way on the surface.

On the contrary, Maric et al. [22] used a core-shell model, which is easier to implement, but already challenging. They modelled the LDL as a super-ellipsoid, and conducted a study through different temperatures to prove its applicability at retrieving known features, like the core transition. Recently, the same model was applied in [92] to compare normolipidemic (N) and triglyceride-rich (TG) LDLs. Here they described changes in the LDL conformation between N- and TG-LDLs, which is of biological relevance as TG-LDLs are closely linked to cardiovascular diseases [10].

Before them, studies on oxidized LDLs, N-LDLs and TG-LDLs, below and above the core transition, but also with high-hydrostatic pressure, were conducted by Golub et al. [36] and Lehofer et al. [23] with SAXS and SANS. Surprisingly, it revealed that LDLs in all forms were resistant to high-pressure, until 3000 bar, and that the observed changes were fully reversible. In addition, they showed that the pressure changes were not isotropic, but more significant on the smallest axis of the LDL,

where a flattening by pressure was observed. These anisotropic changes reflected the distinct compressibilities underlying the LDL surface.

Always with the aim to improve our interpretation of SAXS curves, new mathematical modelling of the LDL is still on-going. We can mention for instance novel approaches from Dr. Wolfang Ring and Dr. Karin Kornmueller at the Medical University of Graz (Austria) [93].

To summarize this section, SAS techniques were crucial to reveal important features of the LDL, like its core organization and transition, and to provide first 3D models for its shape, as well as for the apo B-100 protein. It remains highly complementary to other high-resolution techniques, like electron microscopy. Thanks to its relatively easier and faster implementation, it can help in guiding future direction research, but also in interpreting the 3D maps at higher resolution.

1.3.3 Electron microscopy

Much earlier, negative-staining EM helped to directly visualize lipoproteins, including LDLs, and to estimate their size and shape [94, 95]. It revealed to be a valuable tool to assess the overall quality of sample preparation and LDL isolation, as was reported for example in [18]. However, negative-staining EM also allowed for more detailed studies, using for instance monoclonal antibodies. As each antibody specifically binds to one site of the protein, Chatterton et al. [26] proposed a first mapping of the apo B-100 on the LDL, and introduced the so-called «belt» model, where the protein encircles the LDL. However, as we will see in Chapter 3, negative-staining EM remains limited for higher resolution studies. Moreover, during long times it was hypothesized that the staining was flattening the LDL, and the spherical model remained a consensus until first images from cryo-EM by Antwerpen and Gilkey [96] showed that in fact, their shape is more discoidal. Other images from Spin and Atkinson [27] even revealed «rectangular» views of LDLs. In addition, the application of nanogold clusters, which bind to Cysteine residues, tended to indicate that apo B-100 lies on the LDL surface, which supported the «belt» model. First 3D reconstruction from cryo-EM images of LDLs was proposed by Orlova et al. [28], and is represented in Figure 1.9a. It brought a new visualization of the core layers below the phase transition, which revealed to be lamellar, and not concentric, as it was hypothesized from many studies reviewed previously.

Shortly after, Gantz et al. [29] applied cryo-EM to solubilized apo B-100. The images, shown in Figure 1.10, revealed a high diversity in the conformations of the protein , with more or less elongated molecules. The mean length of 55 different molecules was estimated to be about 342 $\pm$ 120 Å, with lengths going from 160 to 650 Å. In addition, the width of the protein was evaluated to be around 21 $\pm$ 7 Å,

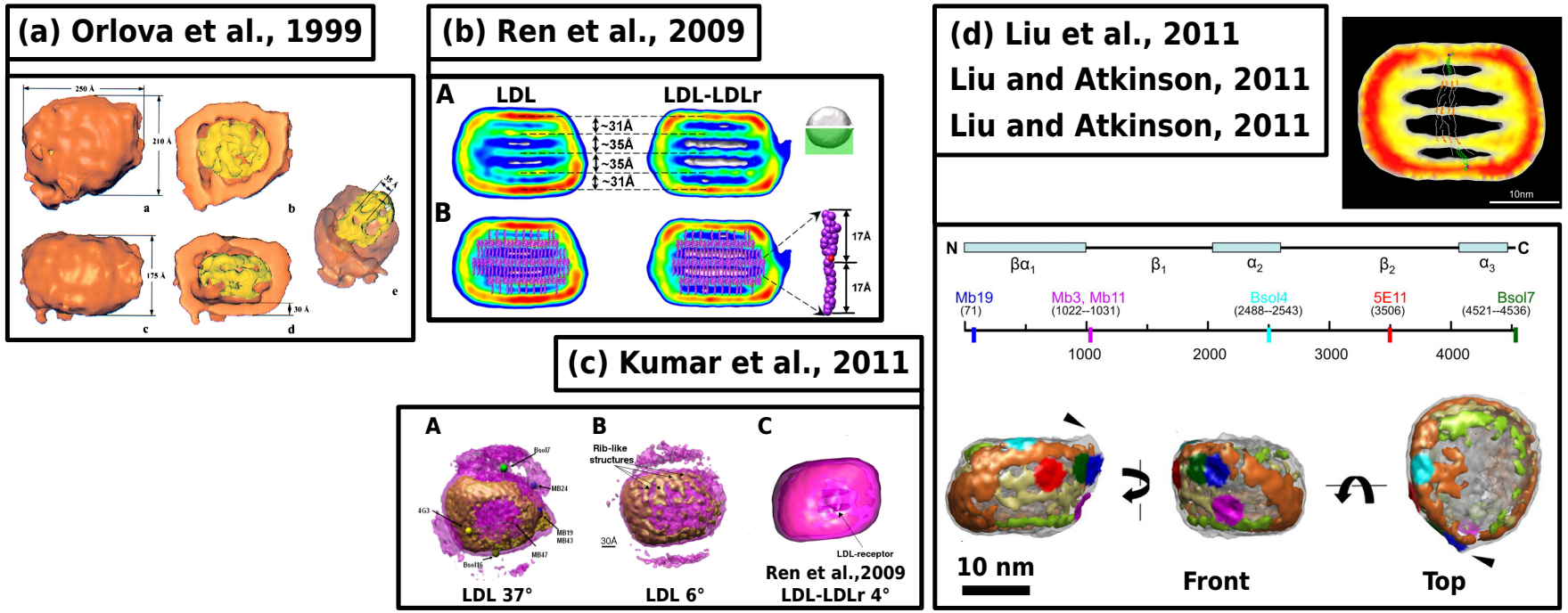


Figure 1.9: Evolution of the 3D reconstructions of the LDL. (a) First 3D map of the LDL from Orlova et al. [28]. (b) Cut views of 3D maps of the LDL and LDL bound to LDL receptor (LDLr) by Ren et al. [30], colored from low (blue) to high threshold values (red). (c) Comparison of 3D reconstructions from Kumar et al. [31] above (A) and below (B) the core transition, with (C) the 3D map of the LDL-LDLr from Ren et al. [30]. (d) Organization of the core (black box), and mapping of monoclonal antibodies proposed by Liu and Atkinson [32, 33], Liu et al. [34]. All images were taken from the cited articles.

ranging between 13 and 45 Å. These measurements fitted well with the circumference of the LDL, and shell thickness estimated in [27, 28].

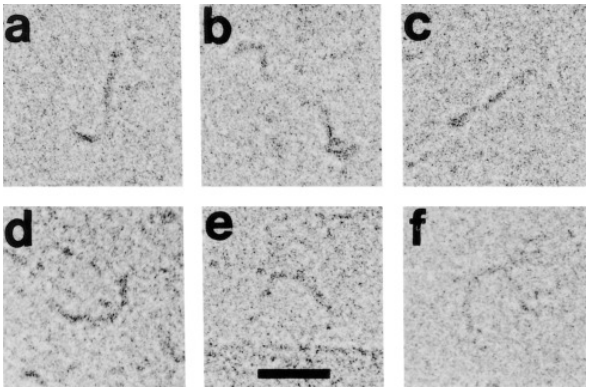


Figure 1.10: Cryo-EM images from apo B-100. Image taken from Gantz et al. [29].

Pursuing characterization studies with cryo-EM, Sherman et al. [97] reported images from triglyceride-rich (TG) LDLs. Compared to normolipidemic (N) LDLs, they did not exhibit any internal layers, in agreement with the fact that it is cholesteryl ester molecules that form layers [86]. While the overall dimensions were the same, the shape of TG-LDLs was more spherical, similar to the N-LDLs above the phase transition. As TG-LDLs present different binding properties [9], it was suggested that the core organization has an influence on the apo B-100 conformation at the surface.

Ten years after the first map from Orlova et al. [28], a new 3D reconstruction of the LDL, as well as the LDL bound to the LDL receptor (LDLr), were proposed by Ren et al. [30], and are depicted in Figure 1.9b. They have retrieved more detailed layers, which definitely endorsed the lamellar model for the core organization. They have assigned the observed protrusion in the LDL-LDLr complex to the $\beta-$propeller domain of the LDL receptor, and proposed that the location is the possible receptor-binding site of the LDL. Finally, they suggested that the flat parts of the LDL on the top and bottom could be assimilated to the amphipathic $\beta-$sheets domains of the pentapartite model.

Soon after, Kumar et al. [31] published maps taken both below and above the core transition, represented in Figure 1.9c. At this point, it is important to mention that all previous LDL samples were vitrified below the phase transition, around 4 or 6°C, and to point out the inherent difficulty to vitrify a sample at 37°C. As measured by Prassl et al. [20] with time-resolved SAXS, the transition can occur in the ms to s time scale, and even probably in smaller times, as the measured values were corresponding to the time resolution. Consequently, the core transition is as fast, or even faster than the vitrification in liquid ethane. Moreover, liquid ethane also presents a small vapor layer, at a temperature low enough to trigger the phase transition just before the vitrification. For all these reasons, sample vitrification of LDLs above the phase

transition requires specific devices and procedures, and the resulting images are likely capturing an intermediate state, as deeply discussed in Liu et al. [34]. Accordingly, the 3D map of Kumar et al. [31] at 37°C is probably an intermediate state, and further discussion should take this into account. Differences remain visible compared to the map at 4°C, as displayed in Figure 1.9c, with a more rounded shape at 37°, and an absence of layers. However, and similarly to Liu et al. [34], this difference can also arise with an intermediate state. In addition, Kumar et al. [31] tried to fit the theoretical location of the commonly used monoclonal antibodies by first docking the homology model of the $NH_2 - \beta\alpha_1$ domain from Richardson et al. [75] into their 3D map, and then by triangulation techniques. From there, and in comparison with the LDL-LDLr map from Ren et al. [30], they agreed with the proposed receptor-binding location. Again, one needs to be careful with this interpretation, given the low resolution, which normally does not allow a precise docking procedure. Interestingly, they applied a bootstrap analysis to compute 3D variance maps, and estimated zones of higher variability in their maps. At low temperature, it was mainly located in specific parts of the shell, while at high temperature, it was more in the core, which they relate to possibly more flexible parts of their maps.

Finally, the last published cryo-EM maps of the LDL were from Liu and Atkinson [32, 33], Liu et al. [34], and are shown in Figure 1.9d. They clarified several points concerning both the core organization and possible localization of apo B-100 on the LDL surface:

- As all previous studies, they retrieved the cholesteryl ester layers in their map [32] (black box in Figure 1.9d). However, contrary to Ren et al. [30], and supported by additional procedures using Undecagold labeling, they assigned the flat top and bottom parts of the LDL to phospholipid monolayers.

- A mapping with 6 different monoclonal antibodies was performed, as detailed in Liu and Atkinson [33]. The results displayed in Figure 1.9d endorsed the «belt» model proposed by Chatterton et al. [26] with similar methods. In addition, they showed that the NH_2 (blue spot) and COOH (green spot) terminals are in close proximity. They hypothesized that interaction between the two termini could modulate the LDL binding, especially if the NH_2 terminal appears to be flexible. Finally, relying on the mapping, the observation of a protrusion following the long axis of the LDL, and similar to the one observed for the LDL-LDLr complex by Ren et al. [30], was attributed to the globular region of the $NH_2 - \beta\alpha_1$ domain (blue spot), and not to the potential binding site (red spot), which is located more on the «side».

Since then, no new cryo-EM maps from LDLs were proposed, and the model presented by Liu and Atkinson [33] remains the most precise and accurate for the localization of apo B-100 on the LDL surface.

Recently, we can mention single-particle tomography experiments performed on VLDLs and IDLs by Yu et al. [98] and Lei et al. [99], respectively. They observed sharp corners in the maps, and presented polyhedral models to account for this specific structure. They could correlate the angle between each flat surface to the size of the particles, and related the corners to possible locations of the apolipoproteins with antibody mapping.

To conclude on this section, cryo-EM allowed for direct visualization of LDLs, and unveiled specific features, like its internal lamellar organization below the phase transition. Localization of apo B-100 on the LDL is still challenging. Monoclonal antibody mapping has led to potential models, always supporting the «belt» view. However, the resolution of all reported maps remains quite low, and no secondary structures could have been identified in lipoprotein maps.

1.4 Investigation of the dynamics

Due to its non-exchangeable nature, apo B-100 remains bound throughout the metabolism of lipoprotein, including the VLDL to LDL transformation in the blood circulation, which involves a huge change of size (from around 50 nm to about 25 nm). This implies apo B-100 to be quite flexible to adapt such conformational changes. Segrest et al. [67] hypothesized that the amphipathic β-sheets' domains are anchored to the LDL, supported by surface pressure experiments [100], while the amphipathic α domains could reversibly bind, and would accommodate size changes, as proposed by Wang et al. [101]. These models call for more dynamical studies of LDLs, as also reviewed in [77].

1.4.1 Molecular dynamics simulations

As reviewed by Pan and Segrest [102], HDL has been extensively studied with molecular dynamics (MD) simulations; its assembly has been modelled [103, 104], as well as precise conformation of apo AI on the HDL, allowing to suggest a model for the interaction of HDL with a relevant protein related to its metabolism [105].

In the case of the LDL, however, its larger size, and much greater number of molecules (around 3000 lipid molecules as summarized in Table 1.2) makes it more complex to be modelled by MD simulations. We can report MD simulations of the TG-rich core of lipoproteins (without the protein) by Hall et al. [106], Chaban and

Khandelia [107, 108], as well as the coarse-grained MD simulations from Murtola et al. [109], whose force-field relies on the MARTINI approach. They performed simulations until tens of μs of the LDL above the core transition, whose snapshots are displayed in Figure 1.11. Here, the apo B-100 protein was modelled in different ways, using the tertiary prediction model from Kriško and Etchebest [78], or including the homology model from Richardson et al. [75].

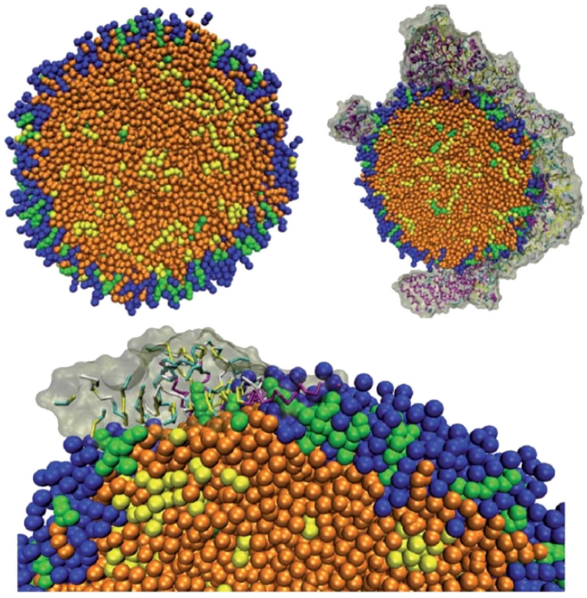


Figure 1.11: Coarse-grained MD simulations of the LDL. POPC and lysoPC are respectively represented in blue and cyan, free cholesterol in green, cholesteryl ester in orange and triglycerides in yellow. The protein is grey. Top: middle cut of a liquid droplet simulation at 18 μs, and the LDL at 10 μs. Bottom: zoom on the interaction between the protein and the core. Image taken from Pan and Segrest [102], Murtola et al. [109].

They retrieved important features of the LDL, with a surface containing the protein, more phospholipids and free cholesterol, while the core is hydrophobic and enriched in cholesteryl esters and triglycerides, with a small amount of free cholesterol (estimated around 13% of the cholesterol, which is less than the estimated ∼30% in [4, 8]). However, they supported the idea of a broad transition in the lipidic distribution from the core to the shell, and not a 2-layers or 3-layers model, as proposed in [4].

An important result from their simulations is the apparent loose packing of the surface, which in fact reveals to be biologically relevant for the functions of the LDL and its interaction with lipid transfer proteins. In addition, it appeared in their models that the apo B-100 protein has hydrophobic parts which directly interact with the core, so cholesteryl ester and free cholesterol. It implies that at these locations the phospholipid monolayer is displaced. However, the protein does not penetrate further in the core.

The diffusion coefficients of lipids were estimated, and all lie around 0.1 to 10 Å^2/ns, with an acceleration of the lipid dynamics on the surface. On the contrary, near the apo B-100 protein, the lipids slow down, especially free cholesterol, which supports a preferential interaction of the protein with free cholesterol, compared to phospholipids. However, compared to the simulation of a liquid droplet of same lipidic distribution, no more effects were visible, and they could not choose between the different protein models they used.

Finally, their MD simulation globally agreed with one hypothesis from Segrest et al. [67] concerning a driving alignment of cholesteryl esters near the β-sheets of the protein, however it was not as strong as Segrest et al. [67] proposed, and is only valid for the surface, and not deep in the core.

Overall, these studies brought new insights in the LDL organization, and complement the aforementioned structural studies. In the future, more precise simulations will be needed, both in the grain of the force-field, or concerning the apo B-100 protein, whose structure was either absent, either fixed in the mentioned simulations.

1.4.2 Incoherent neutron scattering

To probe dynamics of biological samples by experimental methods, incoherent neutron scattering is a technique of choice. The accessible dynamics ranges in the picosecond to nanosecond timescale, which is functionally essential for further long-time processes, as investigated for example in [110] for an enzymatic reaction. In the case of incoherent neutron scattering, further detailed in Chapter 2 and 3, average dynamics of Hydrogen atoms are measured. As a whole, mean-square displacements or resilience can be quantified, and the diffusional processes taking place within the sample estimated [111].

Elastic incoherent neutron scattering (EINS) was first applied to LDLs and VLDLs by Mikl et al. [35], whose mean-square displacements (MSD) are represented in Figure 1.12B. The same samples with the addition of sucrose, which have a stabilizing effect on biomolecules, were also probed (see Figure 1.12C). The study was complemented by spin-label electron-spin resonance (ESR) on the lipids in the core or the shell.

Overall, VLDLs exhibited a higher softness and flexibility, compared to LDLs. However, with the addition of sucrose, dynamics of both lipoproteins were reduced, especially for VLDLs, as seen in Figure 1.12C. The fluidity probed by ESR with and without sucrose revealed to be identical, so the main effect of sucrose should arise from interaction with the surface-exposed parts of apo B-100 protein. The differences in the sucrose effect on VLDLs and LDLs led to the assumption that apo B-100 local mobility is distinct in both lipoproteins. Related to the VLDL to LDL conversion

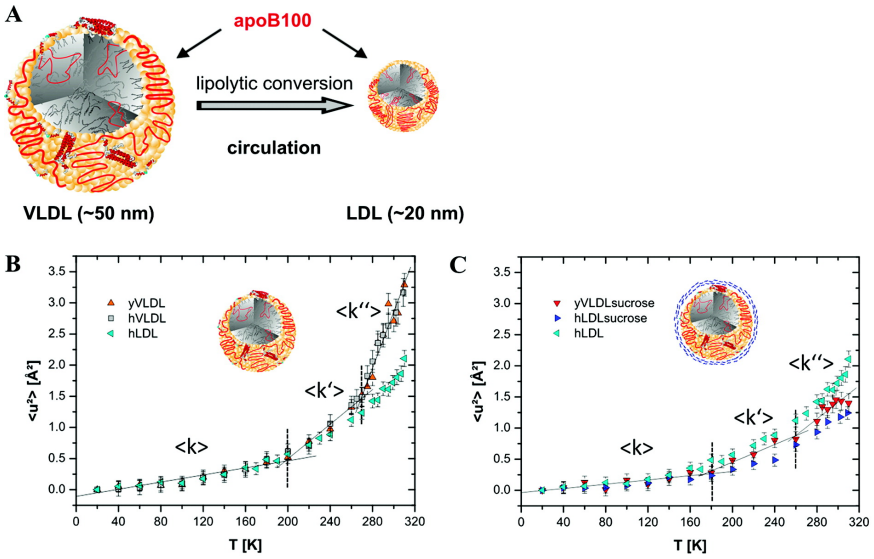


Figure 1.12: Mean-square displacements (MSD) as a function of the temperature for LDLs and VLDLs. A: scheme of the VLDL to LDL conversion. B: MSDs of yolk VLDLs (orange triangles), human VLDLs (grey squares), human LDLs (blue triangles left). C: MSDs of the yolk VLDLs with sucrose (orange down triangles), human LDLs with sucrose (blue triangle right), human LDLs as in B (blue triangles left). Image taken from Mikl et al. [35].

(represented in Figure 1.12A), the higher stiffness of LDLs could be explained by a more rigid apo B-100, as expected by the drastic change in lipoprotein size, and the conformational changes undergone by the protein.

Later on, the effect of hydrostatic high-pressure (HHP) on N-LDLs and TG-LDLs was assessed with elastic and quasi-elastic incoherent scattering by Golub et al. [36] and Peters et al. [37]. In addition to the temperature, the application of HHP is a valuable tool to explore the dynamical landscape of biomolecules, with the advantage of affecting only the volume. Similarly to the studies performed with SANS by the same group described in Lehofer et al. [23], it appeared that the LDL is quite resistant to the pressure, and its molecular dynamics was shown to be identical at all pressure points, until 3000 bar. On the contrary, TG-LDLs revealed to be sensitive to high-pressure, with an estimated slowdown related to a volume reduction at 3000 bar, as represented in Figure 1.13b. Estimation of the MSD in [37] for N-LDLs, TG-LDLs and oxidized LDLs also reported the same observations, with N-LDLs coping with high-pressure, and even a slight increase of its dynamics at 3000 bar, but a huge decrease of the MSDs for TG-LDLs and oxidized LDLs at 3000 bar.

As a whole, these studies indicated a higher compressibility for the modified forms of LDLs, which are hypothesized to be related to conformational changes and a higher flexibility of the apo B-100 protein. Even if the assignment of motions in neutron scattering is more difficult, due to the fact that we measure averages, the

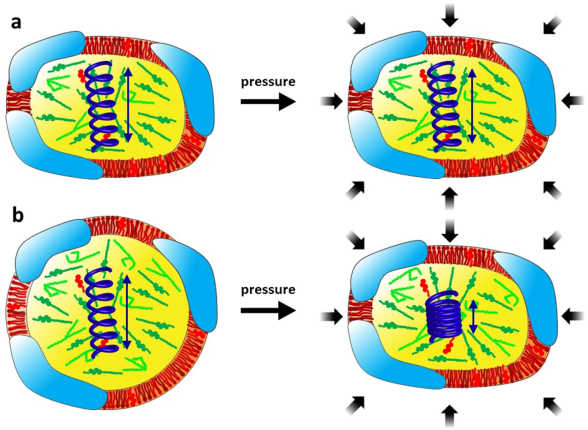


Figure 1.13: Scheme of the effect of hydrostatic high-pressure on (a) N-LDLs and (b) TG-LDLs. The core is represented in yellow, with triglycerides and cholesteryl esters (green), and part of free cholesterol (red). The shell is composed of the phospholipids and free cholesterol (red) and parts of the protein (blue). Image taken from Golub et al. [36].

diffusion coefficients or mean-square displacements that can be extracted are of great significance. As they are estimated experimentally, they complement MD simulations, and could even guide the construction of adapted force-fields.

1.5 Sample preparation and characterization

LDL isolation from human plasma, and above all solubilization of the non-exchangeable apo B-100 protein, remains exceptionally challenging. This section will present an overview of the LDL and apo B-100 preparation, as well as characterization, which relies on the work of Dr. Karin Kornmueller (Medical University of Graz, Austria), who optimized this workflow during the last ten years, and was in charge of it during this thesis. Part of the workflow can also be found in Johs et al. [19], Mikl et al. [35], but is updated in the following section, which is based on the following paper:

> Aline Cisse, Anna-Laurence Schachner-Nedherer, Markus Appel, Christian Beck, Jacques Ollivier, Gerd Leitinger, Ruth Prassl, Karin Kornmueller, and Judith Peters. Dynamics of Apolipoprotein B-100 in Interaction with Detergent Probed by Incoherent Neutron Scattering. The Journal of Physical Chemistry Letters, 2021, 12, 12402–12410. `https://doi.org/10.1021/acs.jpclett.1c03141`.

I learned and participated partly on the preparation and characterization during an exchange in Graz in October 2019 through a Campus France (Amadeus) project. The following workflow corresponds to the preparation of the apo B-100 protein, which

is investigated in Chapter 4, and subfractions of LDLs, from which one fraction was probed with cryo-EM and the results presented in Chapter 5.

1.5.1 LDL isolation

LDLs are directly isolated from CPD-buffered (citrate-phosphate-dextrose) human plasma (250 mL), obtained from the Department of blood group serology and transfusion medicine (Medical University of Graz), after written informed consent, according to a protocol approved by the Institutional Review Board of the Medical University of Graz. The blood plasma is free of pathogens (HBV, HCV, HIV). It is commonly stored in a bag (see Figure 1.14a), at -80 °C, and first needs to be thawed in a 37 °C water bath for about 15 minutes.

1 complete mini EDTA-free protease inhibitor tablet (Roche Diagnostics GmbH, Mannheim, Germany) and 1 mL of a 250 mg/mL EDTA solution are added to the plasma under mild stirring. Plasma is then transferred to centrifuge tubes (Beckman Coulter, Quick Seal), and sealed with a temperature induced hand-held sealing tool.

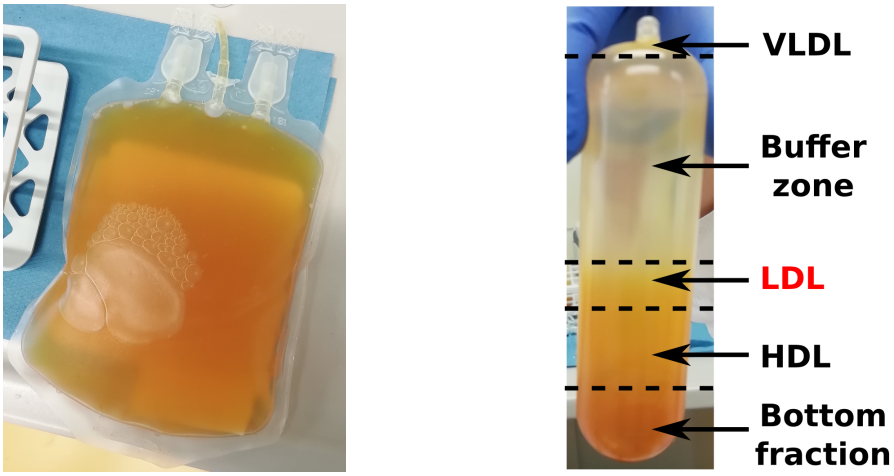
A first centrifugation run is performed with plasma at its own density in an ultracentrifuge (Beckman Coulter with a Beckman TYPE 70 Ti-fixed-angle Titanium rotor, 44000 rpm, 4 °C, 24 h). The appearance of a centrifuge tube after this step is displayed in Figure 1.14b, and the different lipoproteins' fractions are outlined. From there, the LDL fraction is collected, and the density is adjusted with solid potassium bromide (Carl Roth GmbH + Co. KG. Karlsruhe, Germany) to $\rho = 1.063$ g/cm^3 with a digital portable density meter (DMA 35N, Anton Paar Graz, Austria). The sample is transferred to centrifuge tubes, and residual volume is filled up with 10 mM sodium phosphate buffer $\big($$Na_2HPO_4$ x $2H_2O$ (1.44 g/L), KH_2PO_4 (0.26 g/L), 0.1% EDTA, pH 7.4, $\rho = 1.063$ g/cm^3, also adjusted with solid potassium bromide$\big)$.

Two centrifugation runs (44000 rpm, 4 °C, 24 h each) follow, separated by collecting the LDL fraction and adjusting the samples density to $\rho = 1.063$ g/cm^3 between the two runs. To remove potassium bromide after centrifugation, the LDL sample is dialyzed against 10 mM Tris Buffer, pH 7.4 (shown in Figure 1.15a), followed by concentration.

The concentration is performed with centrifugal devices, whose characteristics are the following; Sartorius VIVASPIN 20, PES membrane, 100.000 MWCO, (Sartorius Stedim Lab Ltd, Stonehouse, GL10 3UT, United Kingdom), at 4000 g and 6 °C in a Sigma 3 K 18 tabletop centrifuge (Sigma Laborzentrifugen, Osterode am Harz, Germany).

The characterization of the resulting LDL sample is carried out with:

- SDS-PAGE (3.5% gel, modified after Weber and Osborn);



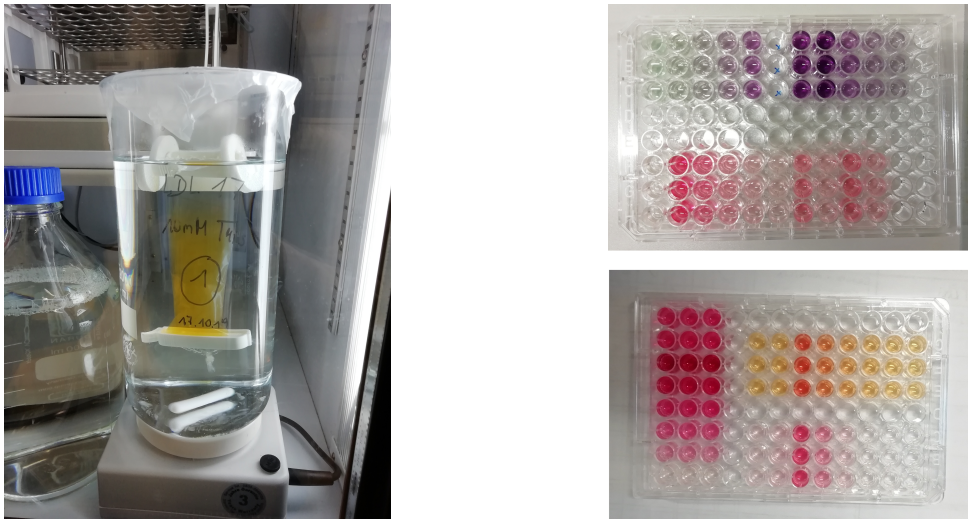
(a) Plasma bag **(b)** Centrifuge tube

Figure 1.14: Photography of (a) a plasma bag and (b) a centrifuge tube after the first ultracentrifugation step.

- Differential scanning calorimetry (DSC, Microcal VP-DSC, Northampton, MA, USA), with a temperature range from 5 to 95 °C and a scan rate of 1 °C/min. An exemplary DSC thermogram of LDLs is presented in Figure 1.16;
- Negative-staining transmission electron microscopy (TEM), whose LDLs pictures are shown in Figure 1.19A.

The chemical composition of LDLs is also determined with different biochemical assays:

- Protein content: BCA™ Protein Assay, (Thermo Scientific, Rockford, USA);
- Phospholipid content: Phospholipids FS (DiaSys Diagnostic Systems GmbH, Holzheim, Germany);
- Free cholesterol content: Free cholesterol FS (DiaSys Diagnostic Systems GmbH, Holzheim, Germany);
- Total cholesterol: Cholesterol FS (DiaSys Diagnostic Systems GmbH, Holzheim, Germany). The measured total cholesterol content is corrected by separating it into its cholesteryl ester contribution and its free cholesterol contribution, where cholesteryl ester = (total cholesterol – free cholesterol) x 1.68. The factor 1.68 represents the ratio of the average molecular weight of cholesteryl ester to free cholesterol;
- Triacylglycerol content: Triglyceride FS (DiaSys Diagnostic Systems GmbH, Holzheim, Germany).



(a) LDL dialysis **(b)** Photometric assays

Figure 1.15: Photography of (a) LDL dialysis against 10 mM Tris Buffer and (b) photometric assays on a CLARIOstar microplate reader (top: cholesterol and protein content, bottom: cholesterol, triacylglycerol, phospholipids content).

All photometric assays are measured in a 96 well plate format on a CLARIOstar microplate reader (BMG LABTECH), as represented in Figure 1.15b. Table 1.3 summarizes the chemical composition of each LDL preparation used in the neutron scattering study of the Chapter 4. In general, the proportions of the components are of the same order of magnitude between each LDL preparation, but variations are noticeable, which reflect the inherent fluctuations between the plasma of each donor.

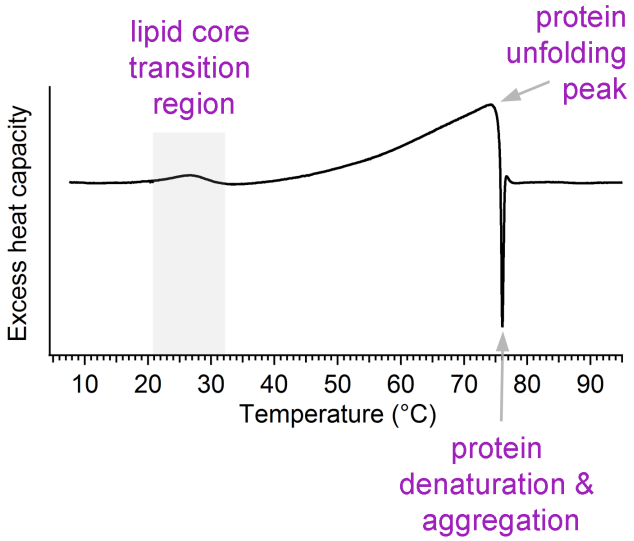


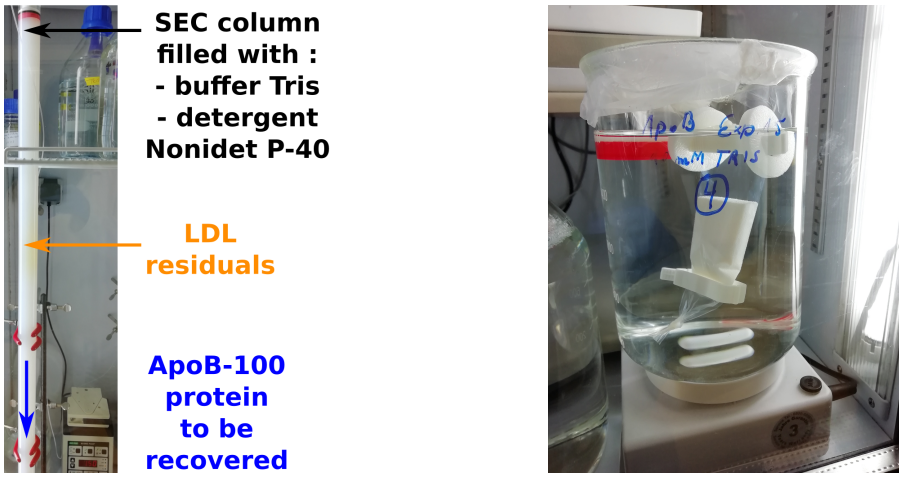
Figure 1.16: Exemplary DSC thermogram of LDLs (preparation 1). (Courtesy of Dr. Karin Kornmueller).

Table 1.3: Chemical composition and phase transition temperatures of each individual LDL preparation

	Protein	Phospholipid	(Total cholesterol corr.)	Cholesteryl ester	Free cholesterol	Triacylglycerol	Core transition temperature (°C)	Protein unfolding peak (°C)
LDL, preparation 1 (mg/mL)	6.0	7.6	20.7	18.0	2.7	2.2	27.6	74.5
LDL, preparation 1 (%)	**16**	**21**		**49**	**7**	**6**		
LDL, preparation 2 (mg/mL)	9.0	12.0	41.4	36.0	5.4	2.0	28.6	72.5
LDL, preparation 2 (%)	**14**	**19**		**56**	**8**	**3**		
LDL, preparation 3 (mg/mL)	12.0	18.5	71.0	62.0	9.0	6.0	20.9	73.8
LDL, preparation 3 (%)	**11**	**17**		**58**	**8**	**6**		
LDL, preparation 4 (mg/mL)	8.6	14.6	42.6	37.1	5.5	3.9	22.7	71.7
LDL, preparation 4 (%)	**12**	**21**		**53**	**8**	**6**		
LDL, preparation 5 (mg/mL)	14.7	17.6	64.9	56.5	8.4	2.9	31.0	72.3
LDL, preparation 5 (%)	**15**	**18**		**56**	**8**	**3**		
LDL, preparation 6 (mg/mL)	14.3	19.6	43.8	38.1	5.7	5.7	20.5	71.7
LDL, preparation 6 (%)	**17**	**24**		**46**	**7**	**7**		

1.5.2 Apo B-100 solubilization with detergent

To solubilize apo B-100, the detergent used is the Nonidet P-40 substitute, also termed NP40, from Roche Diagnostics GmbH (Mannheim, Germany; purchased from Sigma Aldrich). The solubilization and subsequent size exclusion chromatography (SEC) are carried out in aliquots (see Table 1.4 for the number of SEC runs per LDL preparation).



(a) SEC column **(b)** apo B-100 dialysis

Figure 1.17: Photography of (a) LDL / NP40 mixture in a SEC column (b) dialysis of apo B-100 protein against Tris and BioBeads.

For each run, 550 mg NP40 are added to the LDL sample (to an absolute amount of 10 mg protein), and the residual volume is filled up with 10 mM Tris pH 7.4 to a total volume of 2 ml. The mixture is incubated over night on a shaker (1000 rpm) at room temperature. Then, SEC runs are performed in a large column (GE Healthcare, V = 460 mL) packed with Sepharose CL-6B (GE Healthcare Bio-Sciences AB, Uppsala, Sweden), which is shown in Figure 1.17a. The running buffer is a mixture of 10 mM Tris (pH 7.4) and 0.5 mM NP40.

Fractions are collected at a flow rate of 1.5 mL/min with an automatic fraction collection system (4 min. per fraction, 6 mL per fraction) whilst measuring the absorbance at 280 nm. Figure 1.18 shows typical chromatograms of the different SEC runs from one LDL preparation. All fractions within the first peak contain apo B-100 and are collected.

To test for complete delipidation, the protein content, as well as total cholesterol and phospholipid concentration are measured in each fraction as described above. In that case, the protein should be clearly detectable, whereas no cholesterol and phospholipids should be present in the fractions of the first peak.

The pooled fractions are then concentrated using Sartorius centrifugal filter devices, with following characteristics; Sartorius VIVASPIN 20, PES membrane, 100.000

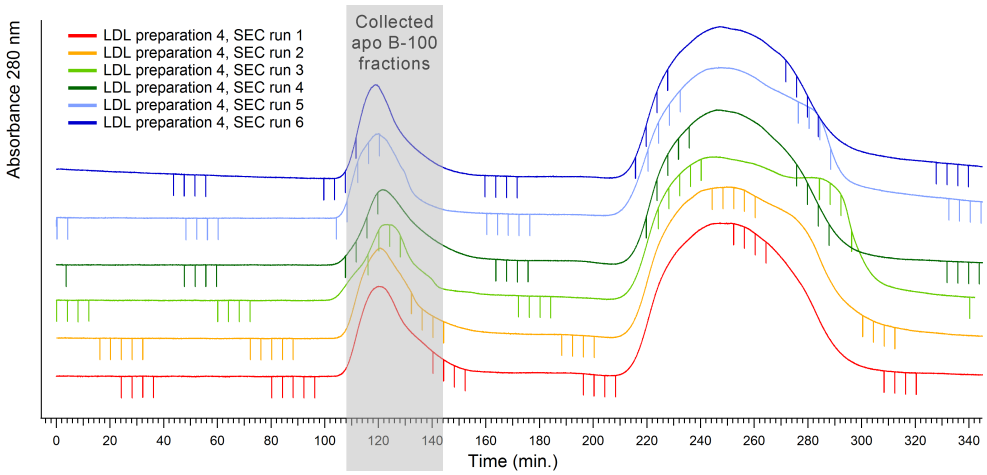


Figure 1.18: Chromatogram of six individual SEC runs of one LDL preparation. The collected apo B-100 fractions are contained in the first peak and highlighted in gray. Curves were shifted in the y-axis for a clearer visibility. (Courtesy of Dr. Karin Kornmueller).

MWCO, Sartorius Stedim Lab Ltd (Stonehouse, GL10 3UT, United Kingdom), at 4000 g and 6 °C, in a Sigma 3 K 18 tabletop centrifuge (Sigma Laborzentrifugen, Osterode am Harz, Germany).

As displayed in Figure 1.17b, the concentrated apo B-100 sample is then dialyzed extensively to reduce the amount of detergent. The sample is charged into a 5 mL dialysis knob (14000 MWCO membrane), and dialyzed against several rounds of 10 mM Tris (2 L, pH 7.4) and Bio-Beads SM-2 resin (Bio-Rad Laboratories, Hercules, CA 94547, USA). The final apo B-100 sample is analyzed regarding protein and detergent content (absorbance measurements at 280 nm and BCA assay, see below), SDS-PAGE, and negative-staining TEM.

1.5.3 Negative-staining TEM

5 µL of sample are dropped onto a glow-discharged (Pelko easiGlow) carbon-coated copper grid (EMResolutions, C400Cu100, 400 mesh) and incubated for 1 min. The following sample concentrations are used:

- LDL sample: optimal $c_{Protein}$ = 0.02 mg/ml
- apo B-100 sample: optimal $c_{Protein}$ = 0.05-0.09 mg/ml
- LDL + NP40 sample: $c_{Protein}$ = 0.02 mg/ml; c_{NP40} = 1.1 mg/ml
- pure NP40: c_{NP40} = 1.1 mg/ml

The sample is blotted with filter paper and immediately 5 µL of 2 % uranyl acetate (UAc) solution is added. After 1 minute, the UAc solution is blotted and replaced with

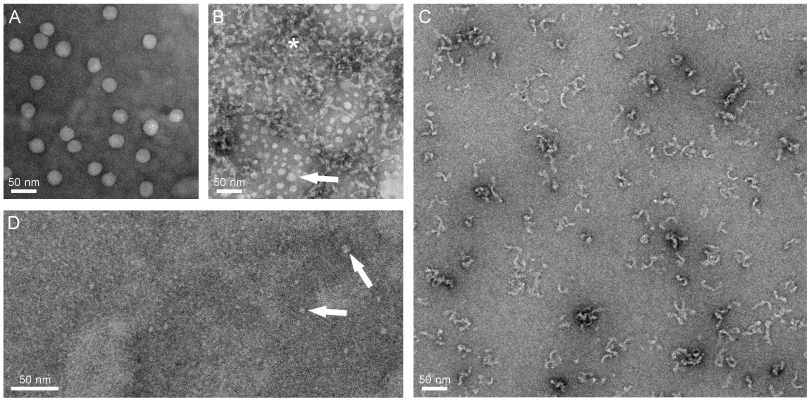


Figure 1.19: Negative-staining transmission electron microscopy (TEM) of all critical sample preparation steps. (A) native LDLs, (B) LDLs mixed with NP-40. Clearly visible are lipid-detergent micelles (arrow) and areas with clustered apo B-100 (asterisk). (C) shows lipid-free detergent-stabilized apo B-100 after dialysis to remove excess detergent. (D) shows pure NP40 as a control to evaluate the size of NP40 micelles. Pure NP40 micelles appear to be smaller than 10 nm (arrows). Scale bars correspond to 50 nm.

another 5 µL of 2 % UAc solution. After 1 minute of incubation, the UAc solution is blotted with filter paper and the grid is air-dried. Visualization of the samples is done using a Fei Tecnai G^2 20 transmission electron microscope (Eindhoven, The Netherlands) operating at an acceleration voltage of 120 kV. Digital images are made using a Gatan US1000 CCD camera at 2K x 2K resolution, and the Digital Micrograph software (Version 1.93.1362, Gatan Inc., Pleasanton, USA). Examples of images are displayed in Figure 1.19.

1.5.4 Determining the protein / detergent ratio

When the absorbance at 280 nm is measured in the apo B-100 / NP40 sample to determine its concentration, both protein and detergent contribute. In contrast, the BCA method is insensitive to NP40 and detects only the protein contribution. By combining both approaches, the protein-detergent ratio can be exactly determined in our samples (Table 1.4).

Every samples' absorbance (A_{total}) is measured on a NanoPhotometer P330 (Implen GmbH, Munich, Germany) at 280 nm, where:

$$A_{\text{total}} = A_{\text{NP40}} + A_{\text{apo B-100}}.$$

According to Beer's law:

$$A_{\text{total}} = \epsilon_{\text{NP40}} \times c_{\text{NP40}} \times d + \epsilon_{\text{apo B-100}} \times c_{\text{apo B-100}} \times d,$$

with ϵ_{NP40} being the molar extinction coefficient of NP40, calculated from the linear regression of a standard dilution series, $\epsilon_{\text{NP40}} = 1973\ \text{M}^{-1}\text{cm}^{-1}$. The molar extinction coefficient of apo B-100, $\epsilon_{\text{apo B-100}}$, is derived from its amino acid sequence by:

$$\epsilon_{\text{apo B-100}} = (n_W \times 5500) + (n_Y \times 1490) + (n_C \times 125),$$

with n_W being the number of tryptophan residues ($n_W = 37$), n_Y being the number of tyrosines ($n_W = 153$), and n_C being the number of cystines. Apo B-100 has 24 cysteine residues, which all form disulfide bridges (cystines), that means $n_C = 12$. We obtain a molar extinction coefficient of $\epsilon_{\text{apo B-100}} = 432970\ \text{M}^{-1}\text{cm}^{-1}$ for apo B-100. The path length d is 0.1 cm and the molar concentration of apo B-100 $c_{\text{apo B-100}}$ is obtained by the BCA method (Pierce BCA™ Protein Assay). All parameters are summarized in Table 1.4.

By solving the expression of Beer's law for the molar concentration of NP40, c_{NP40} can easily be calculated. The amino acid sequence, atomic composition, and molecular weight of apo B-100 were derived from the ExPASy ProtParam database [63, 64]. Table 1.5 summarizes all relevant used constants.

Table 1.4: Determining the detergent contribution.

	No. of pooled SEC runs	Measured absorbance at 280 nm	Dilution factor	$\mathbf{A_{total\ 280nm}}$	$\mathbf{c_{apoB}}$ **(M)** (BCA assay)	$\mathbf{c_{NP40}}$ **(M)** (calculated)	$\mathbf{c_{apoB}}$ **(mg/mL)**	$\mathbf{c_{NP40}}$ **(mg/mL)**
ApoB, preparation 1	7	0.427	8	3.416	6.66E-06	0.013	9.3	8.2
ApoB, preparation 2	6	0.487	8	3.896	7.60E-06	0.016	8.2	10.0
ApoB, preparation 3	7	0.520	10	5.200	1.01E-05	0.022	11.3	13.3
ApoB, preparation 4	6	0.596	10	5.96	1.16E-05	0.026	10.4	15.9
ApoB, preparation 5	7	0.447	10	4.47	8.72E-06	0.020	6.9	12.2
ApoB, preparation 6	8	0.585	10	5.85	1.14E-05	0.025	10.9	15.4

Table 1.5: Relevant constants for separating protein, detergent and buffer contributions.

ϵ_{NP40} ($M^{-1}cm^{-1}$)	1973
ϵ_{apoB} ($M^{-1}cm^{-1}$)	432970
d (cm)	0.1
Apo B-100 molecular weight (g/mol)	512858
Apo B-100 formula	$C_{23073}H_{36367}N_{6077}O_{6918}S_{102}$
Apo B-100 no. of Hydrogens/molecule	36367
NP40 molecular weight (for n=9, g/mol)	617
NP40 formula (for n=9)	$C_{33}H_{60}O_{10}$
NP40 no. of Hydrogens/molecule (for n=9)	60
Tris molecular weight (g/mol)	121.14

1.5.5 Lyophilization

Finally, all apo B-100 samples of every two LDL preparations are pooled, frozen at -80 °C, and lyophilized using a VirTis BenchTop freeze dryer 3L (Virtis SP Scientific, US) for 24 h, as shown in Figure 1.20.

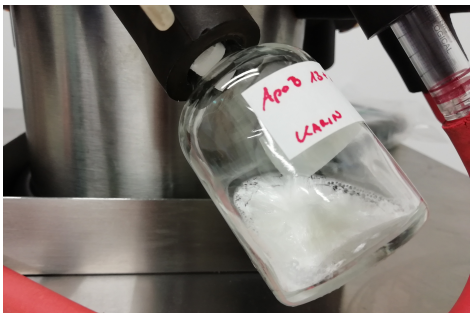


Figure 1.20: Lyophilization of apo B-100.

As a reference, also NP40 in 10 mM Tris pH 7.4 was frozen and lyophilized. Due to the gel-like appearance of this sample the lyophilization process was extended to 48 h. Table 1.6 summarizes the total amounts of all components. Before hydration under D_2O atmosphere for EINS and QENS experiments, the lyophilized powder of all preparations is pooled.

Table 1.6: Apo B-100 lyophilization

	c_{apoB} (mg/mL)	c_{NP40} (mg/mL)	Sample volume before lyophilization (mL)	Total amount of apo B-100 (mg)	Total amount of NP40 (mg)	Total amount of Tris, theoretical (mg)	Weighted amount of powder after lyophilization (mg)
ApoB, preparation 1	9.3	8.2	3.847	35.8	31.6	4.7	137.9 (preparations 1 + 2)
ApoB, preparation 2	8.2	10.0	3.291	27.0	33.0	4.0	
ApoB, preparation 3	11.3	13.3	2.614	29.5	34.7	3.2	141.7 (preparations 3 + 4)
ApoB, preparation 4	10.4	15.9	2.418	25.1	38.4	2.9	
ApoB, preparation 5	6.9	12.2	3.897	26.9	47.4	4.7	169.5 (preparations 5 + 6)
ApoB, preparation 6	10.9	15.4	3.422	37.3	52.7	4.1	

1.5.6 Determining the number of Hydrogens in the sample

As detailed in Chapter 2 and 3, neutron scattering experiments exploit the large incoherent cross-section of Hydrogen atoms compared to other atoms or the D_2O hydration layer. Thus, the contribution of Hydrogen atoms of the sample dominates the spectra. It is then of high relevance to separate the number of Hydrogens that derive from the protein contribution as well as from the detergent contribution. We therefore calculate, based on the determined concentrations of both components and based on the constants shown in Table 1.5 the number of Hydrogen atoms in the total lyophilized powder (see Table 1.7):

$$\text{Total amount of apo B-100 (mol)} = \frac{\text{Total amount of apo B-100 (mg)}}{\text{MW}_{\text{apoB}} \times 1000}$$

$$\text{No. of apo B-100 molecules in the preparation} = \text{Total amount of apo B-100 (mol)} \times N_A,$$

where N_A, the Avogadro constant $= 6.022 \times 10^{23}$ mol^{-1}.

$$\text{Apo B-100, no. of Hydrogens total} = \text{No. of apo B-100 molecules in the preparation} \times 36367,$$

with 36367 being the number of Hydrogens in one apo B-100 molecule, derived from its chemical formula [63, 64] (see chemical formula in Table 1.5). Similar calculations were performed for NP40 with the corresponding constants shown in Table 1.5.

$$\text{Apo B-100, Hydrogens (\%)} = \frac{\text{Apo B-100, no. of Hydrogens total} \times 100}{\text{Apo B-100, no. of Hydrogens total} + \text{NP40, no. of Hydrogens total}}.$$

We find that 36 % of all Hydrogens in the sample can be attributed to apo B-100, whereas 64 % of all Hydrogens might be attributed to NP40. The parameter z defines

the ratio of apo B-100 Hydrogens to the number of total Hydrogens in the sample, thus $z = 0.36$.

Indeed, this small proportion is not ideal for neutron scattering experiments, where we would prefer having the majority of the signal coming from the protein, and not the detergent. Nevertheless, multiple attempts to reduce the proportion of detergent have been conducted by Dr. Karin Kornmueller. For example, further size-exclusion chromatography on the apo B-100 / NP40 sample was applied, but the tests have all ended with the total loss of the protein, revealing a strong interaction between apo B-100 and the detergent NP40. To take into account this interaction, and still extract the protein contribution from incoherent neutron scattering data, we developed a novel methodology for the analysis, as will be presented in Chapter 4.

Table 1.7: Determination of the total number of Hydrogens in the sample and separation into protein and detergent fraction.

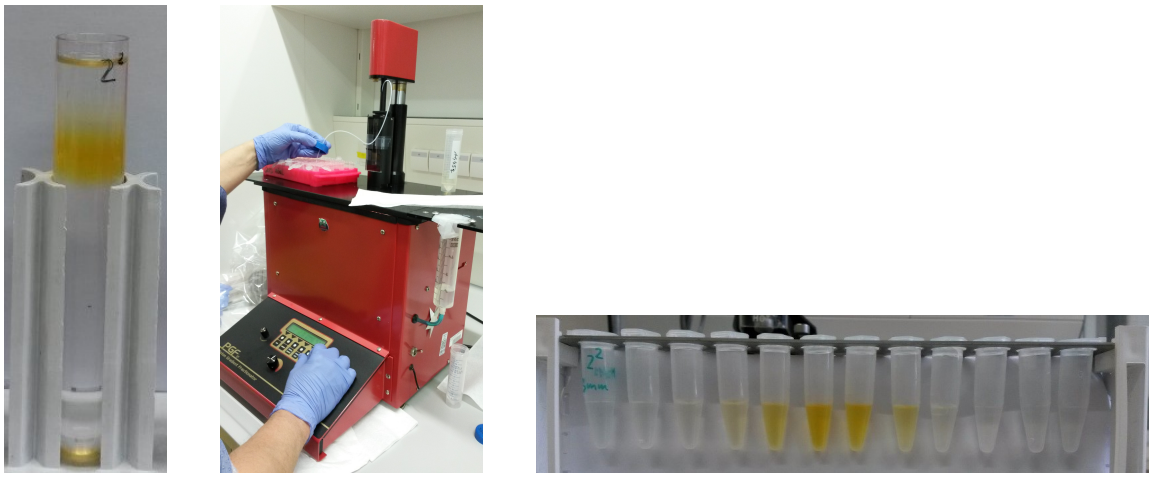
	Total amount of apo B-100 (mg)	Total amount of NP40 (mg)	Total amount of apo B-100 (mol)	Total amount of NP40 (mol)	Number of apo B-100 molecules in the preparation	Number of NP40 molecules in the preparation	Apo B-100, no. of Hydrogens total	NP40, no. of Hydrogens total	**Apo B-100, Hydrogens (%)**	**NP40, Hydrogens (%)**
ApoB, preparation 1	35.8	31.6	7.0E-08	5.1E-05	4.2E+16	3.1E+19	1.53E+21	1.85E+21	45	55
ApoB, preparation 2	27.0	33.0	5.3E-08	5.3E-05	3.2E+16	3.2E+19	1.15E+21	1.93E+21	37	63
ApoB, preparation 3	29.5	34.7	5.8E-08	5.6E-05	3.5E+16	3.4E+19	1.26E+21	2.03E+21	38	62
ApoB, preparation 4	25.1	38.4	4.9E-08	6.2E-05	3.0E+16	3.8E+19	1.07E+21	2.25E+21	32	68
ApoB, preparation 5	26.9	47.4	5.2E-08	7.7E-05	3.2E+16	4.6E+19	1.15E+21	2.77E+21	29	71
ApoB, preparation 6	37.3	52.7	7.3E-08	8.5E-05	4.4E+16	5.1E+19	1.59E+21	3.09E+21	34	66
Sum of preparation 1-6	**181.6**	**237.9**	**3.5E-07**	**3.9E-04**	**2.1E+17**	**2.3E+20**	**7.76E+21**	**1.39E+22**	**36**	**64**

1.5.7 LDL subfractionation

In order to work with homogeneous LDL samples in cryo-EM, a LDL sample needs to be further separated in several subfractions, whose density and size will differ. This step is done by Iodixanol density gradient ultracentrifugation, as described for plasma in Hörl et al. [112], directly after the LDL isolation step described in Section 1.5.1.

Subfractionation of a LDL sample is done in several runs. In each run, an open-top polyclear tube (Seton; Sciences Services GmbH; Munich, Germany) is filled with a mixture of 3 mL LDL sample ($c_{prot} = 3.55$ mg/mL, $\rho = 1.051$ g/cm^3), 5.8 mL phosphate buffered saline (PBS; pH 7.4), and 3.2 mL Iodixanol 60% (OptiPrep™ density gradient medium), and then centrifuged in an ultracentrifuge (Beckman Coulter GmbH with a swinging bucket rotor SW 41 Ti) at 36000 rpm and 16 °C, for 67 h. A tube after this centrifugation step is shown in Figure 1.21a.

Afterwards, a total of 25 subfractions (the bottom fraction was excluded) is extracted by steps of 3 mm (starting from top to bottom) using a piston gradient fractionator (Biocomp; Sciences Services GmbH; Munich, Germany), which is displayed in Figure 1.21b. Eppendorf tubes containing each subfraction are exposed in Figure 1.21c.



(a) Tube **(b)** Fractionator **(c)** Eppendorf tubes after subfractionation

Figure 1.21: Photography of (a) tube containing the mixture of LDL and OptiPrep after ultracentrigation, (b) the fractionator of Dr. Gerd Hörls laboratory (Medical University of Graz), (c) Eppendorf tubes after 3mm fractionation with the fractionator. (Courtesy of Dr. Karin Kornmueller).

All subfractions, from every tube, are analyzed for their cholesterol composition (as described previously), whose results are displayed for example in Figure 1.22. As seen visually, the main fractions with an intense yellow color are containing the most cholesterol. The main fractions of interest are then pooled together from different

runs, e.g. Fraction 7 from tube 1, 2 and 3 (respectively blue, orange and green in Figure 1.22).

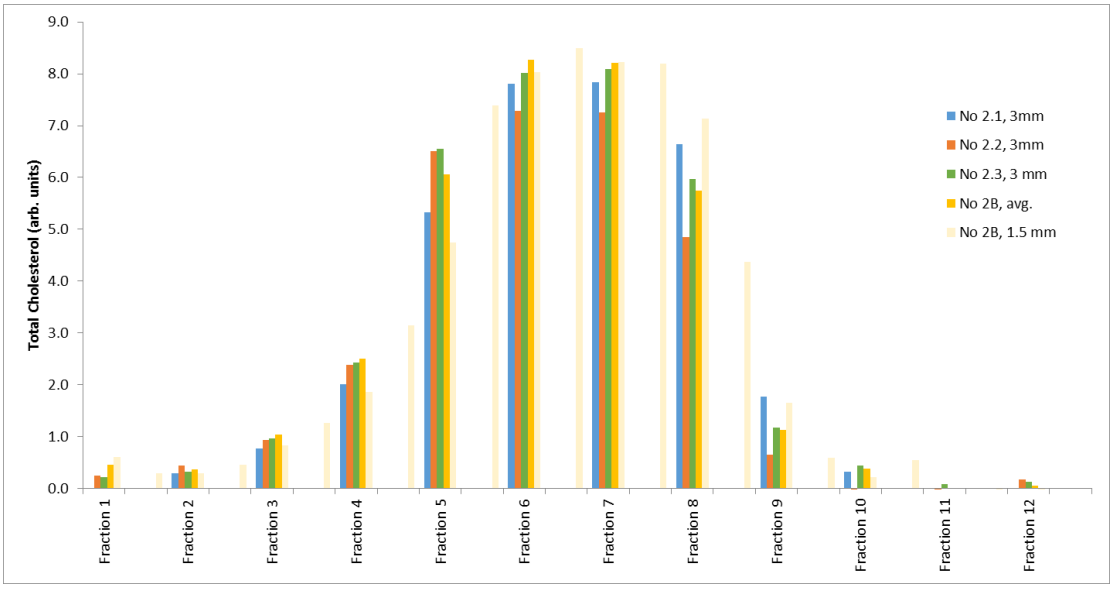


Figure 1.22: Total cholesterol measured for each subfraction. The different colors refer to different tubes. (Courtesy of Dr. Karin Kornmueller).

The resulting subfractions are dialyzed against 10 mM HEPES (2 L, pH 7.4) to remove the Iodixanol. Several buffer changes are needed to completely get rid of it. After dialysis, various sample characterization are performed, as described above; SDS-PAGE (see for example in Figure 1.23), DSC, negative-staining EM, protein and lipid content.

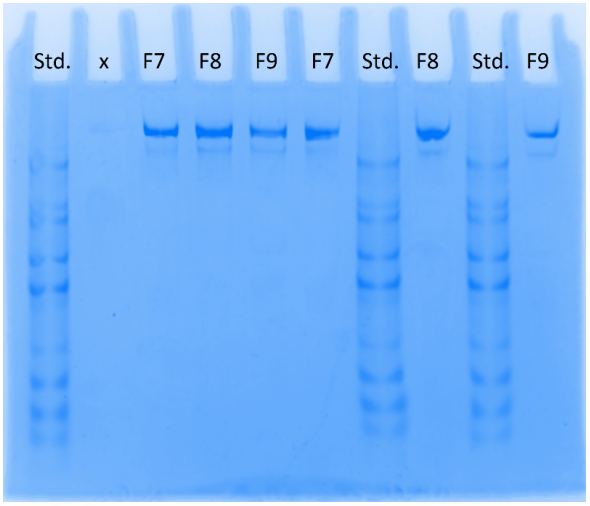


Figure 1.23: SDS-PAGE of different subfractions. (Courtesy of Dr. Karin Kornmueller).

To correlate a density to each fraction, absorbance measurements are performed at 350 nm with a NanoDrop. Standard series were performed on Iodixanol in PBS

for different concentrations. From there, the following relation between the measured absorbance at 350 nm, A_{350}, and density ρ (g/cm^3) could have been derived:

$$\text{In PBS:} \quad A_{350} = 1.9215\rho - 1.9308 \tag{1.1}$$

Fractions of pure Iodixanol, ultracentrifugated and fractionated similar to the mixture of LDL and Ioduxanol, were measured for their absorbance, and their densities estimated with Equation 1.1. Table 1.8 summarizes the measured absorbance and estimated densities for each pure Iodixanol fraction (second and third columns).

Table 1.8: Estimation of the density for each fraction. In blue are highlighted the subfractions commonly related to LDL [112], and where the density was linearly extrapolated (sixth column) from the estimated one (fifth column, not highlighted rows). (Courtesy of Dr. Karin Kornmueller).

Fraction	A_{350} pure Iodixanol	Eq. 1.1 Density ρ	A_{350} LDL + Iodixanol	Eq. 1.1 Density ρ	Extrapolated Density ρ	Reference [112] Density ρ
	blank corrected	g/cm^3	blank corrected	g/cm^3	g/cm^3	g/cm^3
1	0.007	1.0085	0.012	1.0111	1.0058	1.0039
2	0.010	1.0099	0.010	1.0100	1.0096	1.0049
3	0.010	1.0100	0.014	1.0121	1.0134	1.0066
4	0.013	1.0118	0.022	1.0163	1.0172	1.0089
5	0.018	1.0140	0.034	1.0225	1.0210	1.0121
6	0.025	1.0177	0.105	1.0595	1.0248	1.0163
7	**0.033**	**1.0218**	**0.171**	**1.0938**	**1.0286**	**1.0215**
8	0.036	1.0236	0.196	1.1068	1.0324	1.0278
9	0.047	1.0295	0.743	1.3915	1.0362	1.0356
10	0.054	1.0329	0.127	1.0709	1.0400	1.0440
11	0.067	1.0399	0.085	1.0491	1.0438	1.0522
12	0.076	1.0444	0.087	1.0501	1.0476	1.0591
13	0.091	1.0522	0.092	1.0527	1.0514	1.0648
14	0.101	1.0572	0.101	1.0574	1.0552	1.0686
15	0.099	1.0564	0.105	1.0595	1.0590	1.0704
16	0.103	1.0584	0.107	1.0605	1.0628	1.0721
17	0.102	1.0581	0.126	1.0704		1.0723
18	0.103	1.0583	0.115	1.0647		1.0723
19	0.100	1.0569	0.107	1.0605		1.0732
20	0.103	1.0583	0.112	1.0631		1.0732
21	0.104	1.0590	0.111	1.0626		1.0734
22	0.106	1.0600	0.111	1.0626		1.0751
23	0.103	1.0586	0.121	1.0678		1.1017
24	0.117	1.0659	0.130	1.0725		1.2038
25*	0.430	1.2286	0.285	1.1532		1.3239

For the LDL + Iodixanol sample, from which the LDL density needs to be extracted, the absorbance was also measured. However, as seen in Figure 1.21c, at the highest densities, LDLs have an inherent yellow color, so their values cannot be directly derived from absorbance measurements as done for the pure Iodixanol. At the

same time, not all LDL subfractions are yellow, and outside the peak, it is possible to estimate the density from absorbance using the Equation 1.1, which is shown in the fourth and fifth columns of Table 1.8. The values in column 4 can be linearly interpolated as a function of the fraction. Thus, at the regions where the yellow color interferes with absorbance measurements, mainly from fractions 5 to 12 (highlighted in blue in the table), the density values are extrapolated from this interpolation. We can remark slight differences between the estimated values, and those from the reference values (last column) [112]. It should be noted that the reference values were calculated from a direct subfractionation of the whole plasma, and not isolated LDL as performed here, so that the initial conditions differ. However, the values in the two last columns generally agree with two digits, which support the estimated densities calculated for each fraction.

In the case of the cryo-EM experiments described in Chapter 5, all measurements were done on the LDL fraction 7, so an estimated density $\rho \simeq 1.0286$ g/cm^3, which is close to the LDL densities investigated by cryo-EM in all previous studies [30, 32, 34].

1.6 Objectives of the thesis work

Following the characterization of the molecular properties of the LDL and apo B-100, this thesis focuses on two major aspects:

- The molecular dynamics of the detergent-solubilized apo B-100, which is probed with elastic and quasi-elastic incoherent neutron scattering (EINS, QENS). A detailed introduction of the techniques is given in Chapter 2 (Section 2.4) for the theoretical framework, while the Chapter 3 (Section 3.1) includes the instrumentation and technical parts. The results are presented in Chapter 4. They address the separation of protein and detergent contributions in the context of both hydrogenated samples, with the application of a new methodology for QENS data.

- The structure of the whole LDL, with the underlying question of the localization of apo B-100 on the LDL surface. To this end, cryo-EM experiments are performed. An overview of cryo-EM is exposed in Chapter 2 (Section 2.5) with some elements of theory, and the Chapter 3 (Section 3.2) covers the constitution of a microscope, as well as the image processing pipeline. Proposed 3D maps of the whole LDL are reported in Chapter 5, and examined in detail to draw an actualized model of the LDL organization.

Chapter 2

Theory

In 1897, Thomson's discovery of the electron changed the way atoms were seen. His first model of the atom was again rectified in 1911, where an interpretation by Rutherford of scattering experiments, using α particles on gold foil, showed that all the mass is concentrated in a positively charged nucleus, surrounded by a negatively charged cloud. The Rutherford model was quickly followed by the discovery of the proton in 1919, of the neutron by Chadwick in 1932, and other elementary particles continue today to be uncovered, like the Higgs boson in 2012. Thus, one of the points history shows us is the importance of scattering experiments in revealing the inner properties of matter, and how particles can be used as a probe.

This chapter focuses on neutron scattering and image formation in electron microscopy, which constitute the core techniques of this thesis.

2.1 Fundamental forces and constituents of atoms

Interactions are separated in physics in four fundamental forces: gravity, electromagnetism, strong and weak nuclear forces.

Gravity acts with an infinite range, decreasing with the distance, following a $1/r^2$ law. It is a force of attraction between massive objects, and was first described by Newton's law of universal gravitation, before Einstein's theory of general relativity overcame it. At molecular scale however, gravity is considered as negligible, and governed by the other forces.

Electromagnetism defines how charged particles interact between each other, and within an electric or magnetic field. Its range is also supposed to be infinite, dropping with distance. Electromagnetism is described by Maxwell's equations. Yet, at the nuclear scale, it is not sufficient to explain how nuclei are stabilized, because positively charged protons should repel each other.

It is the strong nuclear force, an attractive force on the 10^{-15} m (femtometer) range, which explains the stability of the nuclei. On the contrary, the spontaneous decay of a neutron into an electron, called β-radiation, a proton p^+ and an antineutrino $\bar{\nu}$, is explained by the weak nuclear force. Its range is also on the femtometer scale, but is much less powerful.

The Standard Model of particle physics, reviewed for example in Cottingham and Greenwood [113], aims at unifying all the forces under one unique theory. Within this framework, the forces are driven by bosons, whose spin is integer. For example, electromagnetism is conducted through photons, whereas strong nuclear force is mediated by gluons. Until now, gravity is the only fundamental force which does not enter in this unified description, and its mediator particle is an object of speculation.

The particles with a spin multiple of 1/2 are called fermions, and are the constituents of matter. The properties of protons and neutrons, which compose the nucleus of an atom, and of electrons, are displayed in Table 2.1.

Particle	Charge (10^{-19} C)	Mass (10^{-27} kg)	Spin
Proton	+1.602177	1.672622	1/2
Neutron	0	1.674927	1/2
Electron	-1.602177	0.000911	1/2

Table 2.1: Properties of the main particles constituting atoms.

Normally buried inside atoms, these particles can be used as a probe, as well as photons with different wavelengths (visible light, X-rays...). Due to their distinct properties, they interact differently with matter. The charged particles (protons and electrons) or photons interact through electromagnetic forces, and preferentially with the electron cloud of atoms. Whereas neutrons, having no charge, interact through the strong nuclear force with the nucleus.

In the thesis, we will focus on the use of neutrons and electrons, for probing dynamics of biomolecules with neutron scattering and their structure with electron microscopy, respectively.

2.2 Wave-particle duality

It is important to remember that all particles share the same wave-particle duality. This quantum property explains why neutrons and electrons can be diffracted, scattered or reflected, similar to light, and is at the basis of the techniques used here.

The wave-particle duality is expressed first in Planck's law:

$$E = h\nu = \hbar\omega, \tag{2.1}$$

with E the energy of the considered object, $h = 6.62607015 \times 10^{-34}$ kg·m^2/s the Planck constant, ν the wave frequency, $\hbar = h/(2\pi)$ the reduced Planck constant, and ω the angular frequency. Similarly, the momentum $\vec{p}$ of a particle can be described by De Broglie's equation:

$$\vec{p} = \hbar\vec{k}, \text{ with } |\vec{k}| = \frac{2\pi}{\lambda}. \tag{2.2}$$

$\vec{k}$ is the wave vector, indicating the direction of the wave, and λ is the wavelength of the radiation.

Finally, we should also mention Heisenberg's uncertainty principle, which is a direct consequence of the wave-particle duality. It states that we can not determine with precision and in a simultaneous way the position x of a particle and its momentum in x direction p_x (so its velocity):

$$\Delta x \Delta p_x \geq \frac{\hbar}{2}. \tag{2.3}$$

Its equivalent in the time t / energy E domain expresses as:

$$\Delta t \Delta E \geq \frac{\hbar}{2}. \tag{2.4}$$

2.3 Schrödinger's equation

As the quantum mechanics description holds for any particle, the propagation of either neutrons or electrons can be described by Schrödinger's equation, which expresses as:

$$i\hbar\frac{\partial\psi(\vec{r},t)}{\partial t} = \mathcal{H}\psi(\vec{r},t) = \Big[-\frac{\hbar^2}{2m}\Delta + V(\vec{r})\Big]\psi(\vec{r},t). \tag{2.5}$$

i is the imaginary number, $\psi(\vec{r},t)$ is the wave function, depending on position $\vec{r}$ and time t, and $\mathcal{H}$ is the Hamiltonian of the system. This Hamiltonian can be separated in two terms. $-\frac{\hbar^2}{2m}\Delta$ represents the kinetic potential, with m the mass of the object, and Δ the Laplacian operator, that writes in Cartesian coordinates (x, y, z) as: $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. $V(\vec{r})$ stands for the potential experienced by the particle.

Considering a static potential, and separating time and space dependencies, Schrödinger's equation can be written in a time-independent form:

$$\psi(\vec{r},t) = u(\vec{r})\mathrm{e}^{-\frac{iEt}{\hbar}}, \tag{2.6}$$

$$\Big[-\frac{\hbar^2}{2m}\Delta + V(\vec{r})\Big]u(\vec{r}) = E \cdot u(\vec{r}), \tag{2.7}$$

with $u(\vec{r})$ being the static part of the wave function, and E the total energy of the system. $u(\vec{r})$ et E can be seen as the eigenfunctions and eigenvalues of the Hamiltonian $\big[-\frac{\hbar^2}{2m}\Delta + V(\vec{r})\big]$, respectively. The last equation can be interpreted in a classical way as the conservation of the energy, the first term representing the kinetic energy, and the second one the potential energy.

Equation 2.7 remains the initial point of all calculations presented below, both for neutrons or electrons. The fundamental differences lie first in the expression of the potential, which depends on the type of interaction the particle will have with its environment; electromagnetic for electrons, strong nuclear force for neutrons. Then, the Schrödinger's equation for electrons will be relativistic, including the rest mass $E_0 = mc^2$ of the electron in the formal expression of E, contrary to the neutrons where a non-relativistic analysis remains valid due to their small velocity (~ 1000 m/s) compared to the speed of light c.

2.4 Neutron scattering

Neutrons represent an ideal probe for biomolecules for several reasons. Their neutral charge, a weak interaction with the sample (and consequently negligible radiation damage) are first examples. In addition, their wavelength is about $\lambda \sim 1$ Å; in the order of magnitude of the amplitude of atomic displacements, or of atomic distances. Finally, neutrons have particular interaction with Hydrogen, what we will see in more details in the following.

The theoretical elements of this part mainly rely on the following books and articles from Sivia [114], Bée [111], Schober [115], Squires [116] and Lovesey [117].

2.4.1 Neutron cross-section

The interaction of a neutron with a nucleus can be divided in two different processes:

- The neutron can be absorbed by the nucleus, which is always associated to a nuclear reaction. The decay of the excited nucleus can lead to γ emission, charged particles emission, or fission. Neutron absorption is of particular interest for neutron detectors: for example ^{3}He or ^{10}B are elements which are used to

«convert» neutrons into charged particles, that can then be detected. Whereas in a reactor core, neutron absorption by Uranium initiates fission.

- The neutron can be scattered by the nucleus; its momentum, and so direction, can change, as well as its energy. When there is no exchange of energy between the neutron and the nucleus, we talk about *elastic scattering*, whereas *inelastic scattering* happens when energy exchange occurs. *Quasi-elastic scattering* is an intermediate case where the exchange remains small. We will precise later on each of these cases.

The two processes described above can be represented by a quantity called the *cross-section*.

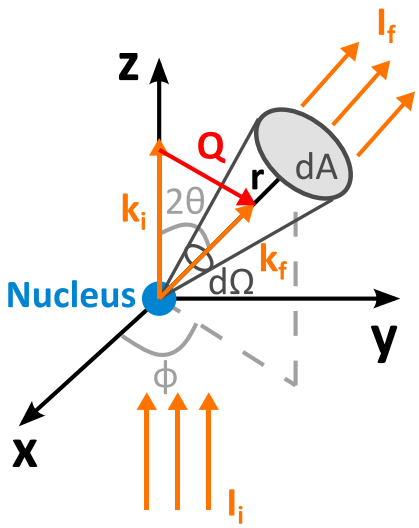


Figure 2.1: Scheme of neutron scattering by a single nucleus. 2θ is the scattering angle, while ϕ is the azimuthal angle. The solid angle is written as $\mathrm{d}\Omega = \mathrm{d}A/r^2$.

Fig. 2.1 shows a schematic neutron scattering process; an incident flux I_i is deviated by a nucleus, resulting in a final flux I_f. When neutrons interact with nuclei, they exchange the momentum $\Delta\vec{p}$, and in the case of inelastic scattering, the energy ΔE. The differences after and before the interaction, with respective subscripts f and i, are written as:

$$\Delta\vec{p} = \hbar(\vec{k_f} - \vec{k_i}) = \hbar\vec{Q}, \tag{2.8}$$

with $\vec{Q}$ being the *momentum transfer*.

$$\Delta E = \hbar(\omega_f - \omega_i) = \hbar\omega, \tag{2.9}$$

with $\hbar\omega$ being the *energy transfer*.

(2.10)

What is of interest in a scattering experiment is the ratio between final and initial flux:

$$\mathrm{d}\sigma = \frac{\text{flux scattered into } \mathrm{d}\Omega}{\text{incident flux}} = r^2 \mathrm{d}\Omega \frac{I_f(r,\theta,\phi,t)}{I_i(t)}. \tag{2.11}$$

The quantity $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ is called the *differential cross-section*. Because neutrons can also exchange energy, another quantity can be introduced to extend the previous formalism; the *double-differential cross section* $\frac{\partial^2\sigma}{\partial\Omega\partial E_f}$. By integrating these quantities over all the possible angles, as well as over all possible final energies, we find the total cross-section, generally given in barns (10^{-24} cm^2):

$$\sigma_{total} = \int\int \frac{\partial^2\sigma}{\partial\Omega\partial E_f}\mathrm{d}E_f\mathrm{d}\Omega = \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\mathrm{d}\Omega = \int \mathrm{d}\sigma. \tag{2.12}$$

In the classical approximation, where the nucleus is seen as a hard sphere of radius r_0, this value represents the area where the neutrons are scattered by the nucleus, so πr_0^2.

In general, neutron cross-section gives an indication about the probability of interaction with the nucleus ; either by scattering (noted σ_s), either by absorption (noted σ_a), so that in total:

$$\sigma_{total} = \sigma_s + \sigma_a \tag{2.13}$$

The neutron cross-sections can not be theoretically predicted, and experimental values are then tabulated, as compiled for example in the Neutron Data Booklet [118]. Some examples of values are represented in Table 2.2.

If we focus on the absorption and scattering cross-sections' columns, we can retrieve the elements with high neutron absorption ; ^{3}He, ^{6}Li or ^{10}B found in neutron detectors, or ^{235}U in the core reactors. By comparison, the incoherent neutron scattering cross-section is very high for Hydrogen (around 80 b), but much smaller for its isotope, Deuterium (around 7 b). This observation has two consequences:

- Neutrons have a particular interaction with Hydrogen, which makes them adapted for biomolecules (containing around 50% of H atoms), but also unique compared to other techniques, like X-rays. The interaction of X-rays with atoms depends on the atomic number Z, so X-rays interact preferentially with bigger atoms. Consequently, Hydrogens are almost invisible to them. In contrary, neutrons reveal to be complementary to get precious information on the Hydrogen atoms. A comparison between neutron and X-ray cross-sections can be found in Figure 2.2.

- The striking difference between Hydrogen and Deuterium cross-sections allows specific labeling of parts of the sample by a H-D exchange, and thus detailed studies of specific parts of the molecules. In addition, isotopic exchange on the buffer enables to limit its contribution. For example, heavy water D_2O is commonly used for investigation of proteins in solution.

Element	Cross-sections (in barns)			
	Absorption	Scattering	Coherent	Incoherent
^{1}H	0.3326(7)	82.03(6)	1.7583(10)	80.27(6)
^{2}H / D	0.000519(7)	7.64(3)	5.592(7)	2.05(3)
^{3}He	5333.0(7.0)	6.0(4)	4.42(10)	1.532(20)
^{6}Li	940.0(4.0)	0.97(7)	0.51(5)	0.46(5)
^{10}B	3835.0(9.0)	3.1(4)	0.144(6)	3.0(4)
C	0.00350(7)	5.551(3)	5.551(2)	0.001(4)
N	1.90(3)	11.51(11)	11.01(5)	0.50(12)
O	0.00019(2)	4.232(6)	4.232(6)	0.000(8)
Al	0.231(3)	1.503(4)	1.495(4)	0.0082(6)
Si	0.171(3)	2.167(8)	2.1633(10)	0.004(8)
V	5.08(4)	5.10(6)	0.01838(12)	5.08(6)
Fe	2.56(3)	11.62(10)	11.22(5)	0.40(11)
Ni	4.49(16)	18.5(3)	13.3(3)	5.2(4)
^{235}U	680.9(1.1)	14.0(2)	13.78(11)	0.2(2)

Table 2.2: Neutron cross-sections for some elements. (Values reproduced from [118]). Error on the last digits is shown between brackets.

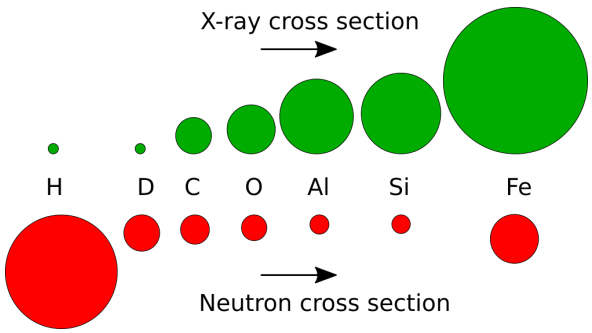


Figure 2.2: Comparison between X-ray and neutron cross sections. Reproduced from ref. [119].

2.4.2 Asymptotic solution of the Schrödinger's equation

As said before, the initial point to derive all the formalism relies in Schrödinger's equation, in our case the stationary case presented in Equation 2.7. In the region near the potential $V(\vec{r})$, knowledge of $V(\vec{r})$ is necessary to solve the equation, but not in the far-field region. This hypothesis remains consistent with neutron instruments, where the distance between the scattering event and the detector is more than one meter, to be compared to the size of the sample holder ($\sim$cm^2), or the size of a protein ($\sim$nm^2). Moreover, the strong nuclear force occurring between the neutron and nuclei ranges in the 10^{-15} m, as outlined previously, meaning that the potential $V(\vec{r})$ decreases quickly, which is again consistent with the far-field regime. With these considerations in mind, we can focus on the asymptotic solution of the Schrödinger's equation ($r \to \infty$). Formal derivation of this asymptotic solution can be found in Section 3 of the paper of Schober [115], which makes use of the partial waves formalism. We will consider only the most general form, and focus on the physical significance of each term. The eigenfunctions of Equation 2.7 write then in the far-field limit as [115]:

$$u_{\vec{k}}(\vec{r}) \to \frac{1}{\sqrt{V}}\left(\mathrm{e}^{i\vec{k}\cdot\vec{r}} + f_{\vec{k}}(\theta,\phi)\frac{\mathrm{e}^{ikr}}{r}\right),\ r \to \infty, \tag{2.14}$$

with $V = (2\pi)^3$ the normalization volume as introduced in quantum mechanics, $\vec{k}$ the wave vector of the neutron, of modulus $k = \sqrt{2mE}/\hbar$, and $f_{\vec{k}}(\theta,\phi)$ with dimension of a length is the *scattering amplitude*.

Before describing in more details Equation 2.14, it is necessary to introduce the formalism of wave packages. Indeed, scattering has by definition a certain chronology: at $t \to -\infty$ the neutron is created, then propagates freely until the target nucleus, interacts with its potential $V(\vec{r})$ at $t = 0$, where it is scattered, before propagating again freely, until being counted on a detector at $t \to +\infty$. This time dependence can not be expressed by one stationary solution, but instead by a superimposition of stationary states in wave packages:

$$\psi(\vec{r},t) = \frac{1}{\sqrt{V}}\int \mathrm{d}^3k\ w(\vec{k})u_{\vec{k}}(\vec{r})\mathrm{e}^{-i\omega_k t}. \tag{2.15}$$

The wave packages can be seen as traveling particles, that is why we can infer from them a time dependence. The major information is contained in the function $w(\vec{k})$, which represents the distribution of stationary waves in the wave package. This distribution remains constant, even after the scattering, but the weight of each of these stationary waves is modulated by the corresponding scattering channel $u_{\vec{k}}$.

A convenient interpretation of the wave packages is to introduce a probability current. First, we can write a probability density for the particle as:

$$\rho(\vec{r},t) = \psi(\vec{r},t)\psi(\vec{r},t)^* = |\psi(\vec{r},t)|^2 \tag{2.16}$$

The local variations of $\rho(\vec{r},t)$ can lead to a probability current $\vec{J}(\vec{r},t)$, that obeys to the continuity equation:

$$\frac{\partial}{\partial t}\rho(\vec{r},t) + \vec{\nabla}\cdot\vec{J}(\vec{r},t) = 0, \tag{2.17}$$

and writes as:

$$\vec{J}(\vec{r},t) = \frac{\hbar}{2im}(\psi^*\vec{\nabla}\psi - \psi\vec{\nabla}\psi^*). \tag{2.18}$$

Because we can always write wave functions as a function of an amplitude A and a phase Φ as follows:

$$\psi(\vec{r},t) = A(\vec{r},t)\mathrm{e}^{i\Phi(\vec{r},t)}, \tag{2.19}$$

the probability density is then expressed by the square of the amplitude:

$$\rho(\vec{r},t) = A^2(\vec{r},t), \tag{2.20}$$

and the probability current by:

$$\vec{J}(\vec{r},t) = \rho(\vec{r},t)\vec{\nabla}\Big(\frac{\hbar\Phi}{m}\Big). \tag{2.21}$$

By comparison with fluid dynamics, $\vec{\nabla}(\hbar\Phi/m)$ represents the velocity of the probability flow.

To come back to the general form of the asymptotic solution Equation 2.14, we can decompose it in two terms:

- A part containing the free unscattered particles, also seen as a plane wave:

$$\frac{1}{\sqrt{V}}\mathrm{e}^{i\vec{k}\cdot r}. \tag{2.22}$$

  In that case, the probability density is constant over all space: $\rho(\vec{r},t) = 1/V$, and the probability current can be expressed as a function of the momentum $\vec{p}$:

$$\vec{J}(\vec{r},t) = \frac{1}{V}\frac{\hbar\vec{k}}{m} = \frac{1}{V}\frac{\vec{p}}{m}. \tag{2.23}$$

The modulus of $\vec{J}(\vec{r},t)$ represents the flux of initial particles:

$$J_{ini} = \frac{1}{V}\frac{\hbar k}{m} \tag{2.24}$$

- The second term contains a spherical wave propagating in space:

$$\frac{1}{\sqrt{V}} f_{\vec{k}}(\theta,\phi)\frac{\mathrm{e}^{ikr}}{r}. \tag{2.25}$$

This part represents the scattered particles. Because we considered the time-independent Schrödinger's equation and a time-independent potential $V(\vec{r})$, we are considering elastic scattering, and thus the momentum modulus remains constant after scattering. As a consequence, the wave vector modulus of scattered particles is still $k = 2\pi/\lambda$. The probability density is equal to:

$$\rho(\vec{r},t) = \frac{1}{V}\frac{1}{r^2}|f_{\vec{k}}(\theta,\phi)|^2. \tag{2.26}$$

The probability density retrieves a $1/r^2$ dependency, corresponding well to the conservation of the particle over its spherical propagation. Next, the probability current of the scattered wave writes as:

$$J_{scatt} = \frac{1}{V}\frac{1}{r^2}|f_{\vec{k}}(\theta,\phi)|^2\frac{\hbar k}{m} = J_{ini}\frac{1}{r^2}|f_{\vec{k}}(\theta,\phi)|^2. \tag{2.27}$$

We notice that the information on the interaction between the neutron and the potential $V(\vec{r})$ is only contained in the scattering amplitude $f_{\vec{k}}(\theta,\phi)$. Its angular dependencies θ and ϕ express how the particle's direction is changed upon elastic scattering. By retrieving the definition Equation 2.11, we can write the differential cross section as:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = r^2\frac{J_{scatt}}{J_{ini}} = |f_{\vec{k}}(\theta,\phi)|^2. \tag{2.28}$$

In the far-field limit, the differential cross-section then depends only on the scattering amplitude.

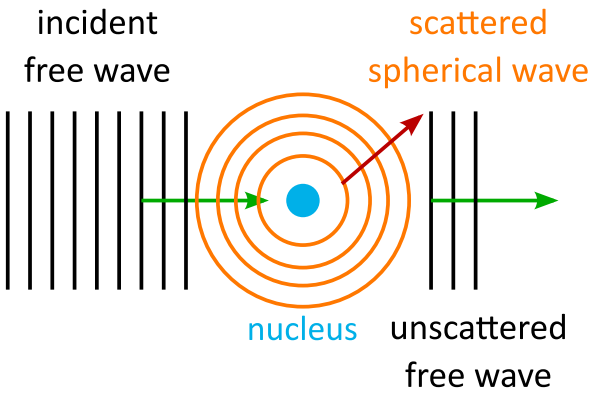


Figure 2.3: Scattering by one nucleus as seen in the far-field limit.

Figure 2.3 displays a representation of the scattering process in the current far-field regime, with an incident wave interacting with a nucleus and ending up in the two terms previously described; an unscattered part and a scattered part seen as a spherical wave. From this representation, and the fact that the final wave only contains the sum of the two terms, we see that this regime supposes that the unscattered and scattered waves do not interact with each other, in other words, the scattered part is much weaker than the free wave. We will focus in more details on this hypothesis in the next section.

The partial wave formalism introduced in [115] gives more insights into the scattering amplitude. In particular, for short-range potential, like the strong nuclear force we are dealing with, and in the limit of long wavelength compared to the dimensions of the target, in other words by considering a point-like nucleus, it can be shown that the scattering is then isotropic. In that case, we can introduce a *scattering length* b as:

$$b = -\lim_{k \to 0} f(\theta) \tag{2.29}$$

We can visualize b as the extension of the interaction between the neutron and the target, which is ranging in the order of the femtometer (10^{-15} m). If we take the example of the hard-sphere case, b is simply the radius r_0 of the target, and the differential cross-section is written as:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = b^2 \Rightarrow \sigma_{total} = 4\pi b^2 = 4\pi r_0^2. \tag{2.30}$$

Equation (2.30) shows that the differential cross-section is different from the classical case (which is simply the area πr_0^2). This simple example demonstrates the necessity to use a quantum formalism, which underlies the partial waves treatment, and thus the current form of the asymptotic solution.

In general, b is enough to describe the interaction occurring in neutron scattering.

As shown in Figure 2.4, there is no clear trend of b with the atomic number Z, and it needs to be tabulated for each element, and even isotopes [118]. That is why the consequent cross-sections, as shown in Table 2.2, are not regular neither.

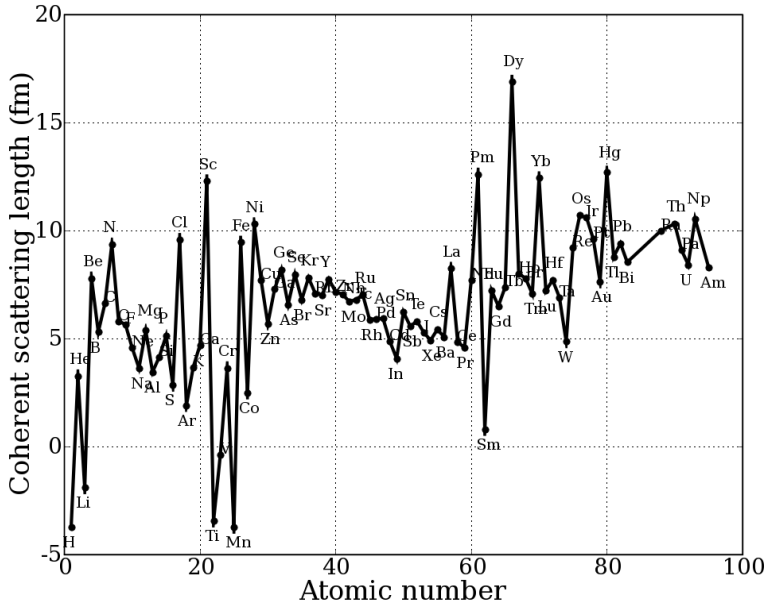


Figure 2.4: Neutron scattering length as a function of the atomic number. (Reproduced from `http://gisaxs.com`).

The results introduced here remain however the limit of a more general case. In particular, it only considers one kind of targets. For an ensemble of nuclei, and always aiming at a generic solution of the Schrödinger's equation, as no analytical solution is possible, a perturbation approach can be applied, and constitutes what is called the Born approximation.

2.4.3 The Born approximation

To apply a perturbation approach, we can rewrite the Schrödinger's equation 2.7 to make appearing a source term, similar as in electrostatics where the electric field $\vec{E}$ writes as:

$$\vec{\nabla} \cdot \vec{E}(\vec{r}) = \frac{\rho(\vec{r})}{\epsilon_0}. \tag{2.31}$$

In Equation 2.31, the source term is constituted by the charge density $\rho(\vec{r})$, at the origin of the electric field $\vec{E}$ we want to determine.

In the scattering process, we would want to apply a similar formalism, where the source of perturbation is assumed to be the potential $V(\vec{r})$. For that we rewrite the Schrödinger's equation 2.7 by placing all the terms containing the potential $V(\vec{r})$ in

the right-hand side:

$$-\frac{\hbar^2}{2m}\Big(\Delta + k^2\Big)u_{\vec{k}}(\vec{r}) = -V(\vec{r})u_{\vec{k}}(\vec{r}). \tag{2.32}$$

By explicitly writing the right-hand side of the previous equation as a source term:

$$j_{\vec{k}}(\vec{r}) = -V(\vec{r})u_{\vec{k}}(\vec{r}), \tag{2.33}$$

and introducing an operator $\hat{D}$ for the left-hand side:

$$\hat{D} = -[\hat{H}_0 - E] = -\frac{\hbar^2}{2m}\Big(\Delta + k^2\Big), \tag{2.34}$$

we end up with the equation:

$$\hat{D}u_{\vec{k}}(\vec{r}) = j_{\vec{k}}(\vec{r}). \tag{2.35}$$

Because Schrödinger's equation is linear, and in the context of point-like nuclei, we can express the source $j_{\vec{k}}(\vec{r})$ as a superimposition of single-point sources (represented mathematically by Dirac distributions):

$$j_{\vec{k}}(\vec{r}) = \int \mathrm{d}^3 r' j_{\vec{k}}(\vec{r}')\delta(\vec{r} - \vec{r}'). \tag{2.36}$$

Practically speaking, it means that the scattering length b is considered to be much smaller than the distance between nuclei. Within this framework, the solution of Equation 2.35 for a single source $\delta(\vec{r} - \vec{r}')$ is a Green function $G_0(\vec{r}, \vec{r}')$:

$$\hat{D}G_0(\vec{r}, \vec{r}') = -\frac{\hbar^2}{2m}\Big(\Delta + k^2\Big)G_0(\vec{r}, \vec{r}') = \delta(\vec{r} - \vec{r}'). \tag{2.37}$$

From there, using the superimposition property, among other steps that we will not detail here, it is possible to derive the general solution of Equation 2.35 [115]:

$$u_{\vec{k}}(\vec{r}) = u^0_{\vec{k}}(\vec{r}) + \int \mathrm{d}^3 r' G_0(\vec{r}, \vec{r}')j_{\vec{k}}(\vec{r}'), \tag{2.38}$$

with $u^0_{\vec{k}}(\vec{r})$ the free unscattered particles, as introduced in Equation 2.22, and the Green function expressed as:

$$G_0(\vec{r}, \vec{r}') = -\frac{1}{4\pi} \cdot \frac{2m}{\hbar^2} \cdot \frac{\mathrm{e}^{ik\cdot|\vec{r}-\vec{r}'|}}{|\vec{r} - \vec{r}'|}, \tag{2.39}$$

which corresponds to a spherical wave centered on the location $\vec{r}'$ of the point source. By making explicit the source term with the potential $V(\vec{r})$, we end up with the

integral form of Schrödinger's equation:

$$u_{\vec{k}}(\vec{r}) = u^0_{\vec{k}}(\vec{r}) + \int \mathrm{d}^3 r' G_0(\vec{r}, \vec{r}') V(\vec{r}') u_{\vec{k}}(\vec{r}'). \tag{2.40}$$

We remark that the wave function $u_{\vec{k}}(\vec{r})$ we are interested in can be found on both sides of Equation 2.40. Physically, we retrieve the same picture as in the asymptotic solution Equation 2.14, with an unscattered part represented by $u^0_{\vec{k}}(\vec{r})$, and a scattered part corresponding to the integral. The latter expresses how a source potential gives rise to spherical waves $G_0(\vec{r}, \vec{r}')$, whose strength will be directly modulated by the potential, but also by the wave function itself. Indeed, if we consider the probability density picture of the wave function, a higher value will mean a higher perturbation. At the end, the equation displays the two necessary conditions to have scattering at a certain point $\vec{r}'$; a non-zero potential $V(\vec{r}')$, and a non-zero probability of presence of the particle.
Equation 2.40 can be formulated in a compact form using operators, which constitutes the Lippman-Schwinger equation:

$$|u_{\vec{k}}> = |u^0_{\vec{k}}> + \hat{G}_0 \hat{V} |u_{\vec{k}}> . \tag{2.41}$$

Equation 2.41 can be iterated, by replacing the $|u_{\vec{k}}>$ term on the right-hand side, we have:

$$|u_{\vec{k}}> = |u^0_{\vec{k}}> + \hat{G}_0 \hat{V} |u^0_{\vec{k}}> + \hat{G}_0 \hat{V} \hat{G}_0 \hat{V} |u_{\vec{k}}> . \tag{2.42}$$

We can still iterate the previous equation. At the end, Equation 2.41 can be written as a series, denoted as the Born series:

$$|u_{\vec{k}}> = \sum_{n=0}^{\infty} (\hat{G}_0 \hat{V})^n |u^0_{\vec{k}}> = \hat{\Omega} |u^0_{\vec{k}}>, \tag{2.43}$$

where $\hat{\Omega} = \sum_{n=0}^{\infty} (\hat{G}_0 \hat{V})^n$ is called the wave operator, which acts on the incident free wave $|u^0_{\vec{k}}>$ to give the scattered wave $|u_{\vec{k}}>$. The wave operator can also be written as a function of the scattering operator $\hat{T}$ as:

$$\hat{\Omega} = 1 + \hat{G}_0 \hat{T}, \text{ with } \hat{T} = \hat{V} + \hat{G}_0 \hat{V} + ... + \hat{G}_0^{(n-1)} \hat{V}^n. \tag{2.44}$$

Truncation of the Born series Equation 2.43 is directly related to the interaction between the neutron and the target nuclei. $n = 0$ corresponds to the case where there is no scattering process, and the wave function is simply the unscattered $|u^0_{\vec{k}}>$. By assuming that the perturbation caused by the scattering process is small compared to the free unscattered wave, we can truncate the Born series to the first order $n = 1$.

This case is the so-called Born approximation, then the wave function, the wave and scattering operators write as:

$$|u_{\vec{k}} > = |u^0_{\vec{k}} > + \hat{G}_0 \hat{V} |u^0_{\vec{k}} >, \tag{2.45}$$

$$\hat{\Omega} = 1 + \hat{G}_0 \hat{V} \tag{2.46}$$

$$\hat{T} = \hat{V} \tag{2.47}$$

As a consequence, the previous coupling between the scattered wave and the potential disappears. The Born approximation has several consequences:

- It only considers cases where the scattered wave is weak compared to the incoming wave. Even near the target, the incoming wave remains unperturbed, so no attenuation effect is present;
- Multiple scattering is completely neglected. We see it through the wave operator Equation 2.46 where we only consider the action of the potential (through operator $\hat{V}$), then propagation (through operator $\hat{G}_0$), and no further interaction with the potential or additional propagation. Furthermore, the scattering operator $\hat{T}$ in Equation 2.47 simplifies to the sole action of the potential operator $\hat{V}$, representing a single scattering process.

This approximation holds for slow and thermal neutrons commonly used in neutron scattering experiments, where the incoming energy is in the order of meV, the neutron flux quite low (few to thousands neutrons per second), and the thickness of the sample holder $\lesssim 1$ cm remains reasonable compared to the probability of interaction with the sample. Figure 2.5 gives a representation of the scattering for an array of point-like nuclei in the Born approximation.

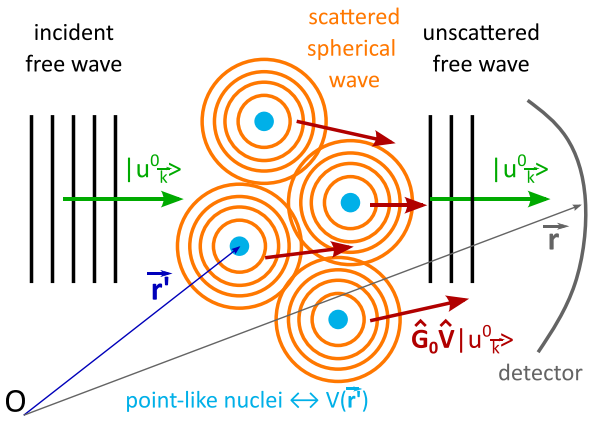


Figure 2.5: Scattering by point-like nuclei in the Born approximation.

If we come back to an integral form, and by replacing the Green function $G_0(\vec{r},\vec{r}')$ by its expression presented in Equation 2.39, Equation 2.45 can be written as:

$$u_{\vec{k}}(\vec{r}) = u^0_{\vec{k}}(\vec{r}) - \frac{1}{4\pi}\frac{2m}{\hbar^2}\int \mathrm{d}^3r' \frac{\mathrm{e}^{ik\cdot|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}V(\vec{r}')u^0_{\vec{k}}(\vec{r}'). \tag{2.48}$$

From there, if we consider the asymptotic case of the far-field limit, so that the position $\vec{r}$ is much greater that the range of the potential $\vec{r}'$, we can simplify the following terms:

$$|\vec{r}-\vec{r}'| \simeq r\Big(1-\frac{\vec{r}\cdot\vec{r}'}{r^2}\Big), \tag{2.49}$$

$$\frac{\mathrm{e}^{ik\cdot|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|} \simeq \frac{\mathrm{e}^{ik\cdot r}}{r}\mathrm{e}^{-ik\cdot(\hat{r}\cdot\vec{r}')}, \text{ with } \hat{r}=\frac{\vec{r}}{r}. \tag{2.50}$$

Following that, the wave function can be simplified as:

$$u_{\vec{k}}(\vec{r}) = u^0_{\vec{k}}(\vec{r}) - \frac{1}{4\pi}\frac{2m}{\hbar^2}\frac{\mathrm{e}^{ik\cdot r}}{r}\int \mathrm{d}^3r' \mathrm{e}^{-ik\cdot(\hat{r}\cdot\vec{r}')}V(\vec{r}')u^0_{\vec{k}}(\vec{r}'). \tag{2.51}$$

By directly comparing Equation 2.51 with the asymptotic solution presented before in Equation 2.14, we can retrieve a theoretical expression for the scattering amplitude in the context of the Born approximation:

$$f_{\vec{k}}(\theta,\phi) = -\frac{\sqrt{2\pi}m}{\hbar^2}\int \mathrm{d}^3r' \mathrm{e}^{-ik\cdot(\hat{r}\cdot\vec{r}')}V(\vec{r}')u^0_{\vec{k}}(\vec{r}') \tag{2.52}$$

$$\Rightarrow f_{\vec{k}}(\theta,\phi) = -\frac{m}{2\pi\hbar^2}\int \mathrm{d}^3r' \mathrm{e}^{-ik\cdot(\hat{r}\cdot\vec{r}')}\mathrm{e}^{i\vec{k}\cdot\vec{r}'}V(\vec{r}'), \text{ with } u^0_{\vec{k}}(\vec{r}') = \frac{1}{(2\pi)^3}\mathrm{e}^{i\vec{k}\cdot\vec{r}'}. \tag{2.53}$$

The following notations are introduced:

$$\vec{k}_i = \vec{k}, \tag{2.54}$$

$$\vec{k}_f = k\hat{r}, \tag{2.55}$$

$$\vec{Q} = \vec{k}_i - \vec{k}_f. \tag{2.56}$$

$\vec{Q}$ corresponds to the momentum transfer, and from there the scattering amplitude finally expresses as:

$$f_{\vec{Q}}(\theta,\phi) = -\frac{m}{2\pi\hbar^2}\int \mathrm{d}^3r' \mathrm{e}^{i\vec{Q}\cdot\vec{r}'}V(\vec{r}') \propto \mathcal{FT}[V(\vec{r'})], \tag{2.57}$$

with $\mathcal{FT}$ the spatial Fourier Transform. It means that within the framework of the Born approximation, the scattering amplitude, containing all the information about the interaction between the neutron and the potential, is directly the Fourier transform of the potential.

2.4.4 Fermi's golden rule

To generalize the current formalism in the case of inelastic scattering (so possible energy exchanges between the neutron and the sample), and to relate it to the neutron cross-section, it is necessary to introduce conditional probabilities and Fermi's golden rule (a more detailed derivation is shown in [115] Section 5). Fermi's golden rule gives the probability $w_{\vec{k_i}\to\vec{k_f},\lambda_i\to\lambda_f}$ to go from a state $|\lambda_i>$ of wave vector $\vec{k_i}$ to a state $|\lambda_f>$ with $\vec{k_f}$:

$$w_{\vec{k_i}\to\vec{k_f},\lambda_i\to\lambda_f} = \frac{2\pi}{\hbar}\Big| < \lambda_f, \vec{k}_f|\hat{T}|\lambda_i, \vec{k}_i > \Big|^2 \rho(E_f), \tag{2.58}$$

with $\hat{T}$ the scattering operator introduced previously, and $\rho(E_f)$ the density of final states. What this relation expresses is how the scattering operator $\hat{T}$ acts on the incoming neutrons $|\lambda_i, \vec{k}_i >$, the result being then projected on $< \lambda_f, \vec{k}_f|$ during the measurement.

In the case of the Born approximation, the scattering operator is simply the action of the potential, as shown in Equation 2.47, so $\hat{T} = \hat{V}$, and Fermi's golden rule simplifies to:

$$w_{\vec{k_i}\to\vec{k_f},\lambda_i\to\lambda_f} = \frac{2\pi}{\hbar}\Big| < \lambda_f, \vec{k}_f|\hat{V}|\lambda_i, \vec{k}_i > \Big|^2 \rho(E_f), \tag{2.59}$$

This transition probability enables to access conditional cross-sections through the master equation of scattering. Before introducing it in more details, it is necessary to present Fermi's pseudo-potential, which is equivalent to the hypothesis of an isotropic scattering, in the case of point-like targets:

$$V_{eff}(r)\psi(r) = \frac{2\pi\hbar^2}{m} b\delta(\vec{r}) \Rightarrow V(\vec{Q}) = \frac{2\pi\hbar^2}{m} b, \tag{2.60}$$

with $\delta(\vec{r})$ the Dirac function, b the scattering length introduced in the previous sections, and $V(\vec{Q})$ the form factor. In that case, the form factor reveals to be constant as a function of Q, meaning that the scattering is homogeneous in space.
Fermi's pseudo-potential is a specificity of neutrons, because the strong nuclear forces allow to consider the target nuclei as point-like. In the case of incoming X-rays or electrons, the form factor would not be constant, but would depend on the extension of the electron cloud of the target atoms.

2.4.5 The master equation of scattering

Considering Fermi's golden rule in the case of the Born approximation, the double-differential cross-section can thus be derived, constituting the master equation of scattering [115]:

$$\frac{\mathrm{d}^2\sigma_{\vec{k_i}\to\vec{k_f}}}{\mathrm{d}\Omega\mathrm{d}E_f} = \frac{k_f}{k_i}\mathrm{S}(\vec{Q},\omega), \tag{2.61}$$

with $\mathrm{S}(\vec{Q},\omega)$ the neutron scattering function.

$$\mathrm{S}(\vec{Q},\omega) = \frac{1}{2\pi\hbar}\sum_{\lambda_i} p(\lambda_i) \sum_{j,j'=1}^{N} b_j b_{j'}^* \int_{-\infty}^{\infty} dt < \lambda_i|\mathrm{e}^{-i\vec{Q}\hat{R}_{j'}^{\circ}}\mathrm{e}^{i\vec{Q}\hat{R}_j(t)}|\lambda_i > \mathrm{e}^{-i\omega t}, \tag{2.62}$$

with $p(\lambda_i)$ the probability of each initial state λ_i (Boltzmann equilibrium for example), and b_j the scattering length of the jth nucleus. $\hat{R}_j(t)$ is the time-dependent Heisenberg operator, defined as [116]:

$$\hat{R}_j(t) = \mathrm{e}^{\frac{i\hat{H}t}{\hbar}}\vec{r}_j\mathrm{e}^{-\frac{i\hat{H}t}{\hbar}}, \tag{2.63}$$

with $\hat{H}$ the Hamiltonian of the scattering system, and $\vec{r}_j$ the position of the jth nucleus. In this formalism, we have the following expressions:

$$\mathrm{e}^{i\vec{Q}\cdot\hat{R}_j(t)} = \mathrm{e}^{\frac{i\hat{H}t}{\hbar}}\mathrm{e}^{i\vec{Q}\cdot\vec{r}_j}\mathrm{e}^{-\frac{i\hat{H}t}{\hbar}}, \tag{2.64}$$

$$\hat{R}_j(0) = \vec{r}_j. \tag{2.65}$$

$\hat{R}_{j'}^{\circ} = \hat{R}_{j'}(0)$ is the Heisenberg operator representing the initial positions of the nuclei. As a consequence of these notations, $\hat{R}_{j'}^{\circ}$ and $\hat{R}_j(t)$ do not commute, and the product of the two exponential terms in Equation 2.62 does not lead to the sum of their exponents. The two operators commute only in these specific cases:

- At the initial time $t_0 = 0$, as shown in Equation 2.65, where the Heisenberg operator simplifies to the vector position of the jth scatterer;
- In the classical limit, where the operators are treated as position vectors, so $\hat{R}_{j'}^{\circ} = \vec{r}_{j'}(0)$ and $\hat{R}_j(t) = \vec{r}_j(t)$. This approximation implies that the discrete nature of energy transfers is neglected, as they are assumed to be much smaller than the thermal energy $\frac{1}{2}k_BT$. In that case, we pass from a quantum formalism to a statistical description of the system.

The neutron scattering function displayed in Equation 2.62 is a pure function of momentum and energy transfer, and more importantly, does not depend on the detail of the experimental setup. In practice, it is $\mathrm{S}(\vec{Q},\omega)$ which is measured during a neutron

scattering experiment, through counting the scattered neutrons as a function of the momentum transfer, and energy transfer. The way it is measured and the strategies to access momentum and energy transfers will be discussed in Chapter 3.

We can remark that Equation (2.62) displays a Fourier Transform in time. This way, an intermediate scattering function $\mathrm{I}(\vec{Q},t)$ can be introduced:

$$\mathrm{S}(\vec{Q},\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{I}(\vec{Q},t)\mathrm{e}^{-i\omega t}\mathrm{d}t. \tag{2.66}$$

In the classical limit, the intermediate scattering function $\mathrm{I}(\vec{Q},t)$ writes as [116]:

$$\mathrm{I}(\vec{Q},t) = \frac{1}{N}\sum_{j,j'}^{N} < \mathrm{e}^{-i\vec{Q}\vec{r_{j'}}(0)}\mathrm{e}^{i\vec{Q}\vec{r_j}(t)} >, \tag{2.67}$$

with N the total number of scatterers in the sample. Both scattering functions can be interpreted in terms of correlation functions. By introducing the classical Van Hove pair correlation function $\mathrm{g}(\vec{r},t)$ [120]:

$$\mathrm{g}(\vec{r},t) = \frac{1}{N}\sum_{j,j'}^{N} < \delta\Big(\vec{r}+\vec{r_{j'}}(0)-\vec{r_j}(t)\Big) >, \tag{2.68}$$

we can write the intermediate scattering function as a Fourier transform of the Van Hove pair correlation function:

$$\mathrm{I}(\vec{Q},t) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} \mathrm{g}(\vec{r},t)\mathrm{e}^{i\vec{Q}\cdot\vec{r}}\mathrm{d}\vec{r}. \tag{2.69}$$

In theory, the Van Hove pair correlation function $\mathrm{g}(\vec{r},t)$ contains all the information on the sample we would like to extract, and is directly related to what we are measuring, $\mathrm{S}(\vec{Q},\omega)$, through a double Fourier transform in time and space. This result is fundamental in the neutron scattering field, and all data analyses rely on it.

It is however important to remember that this result is always derived in the context of the Born approximation and its corresponding hypotheses. Furthermore, we are dealing now with the classical limit, which generally remains consistent with the properties of neutron scattering [115]. First, due to the nature of the neutron beam, whose source presents huge thermal fluctuations (Chapter 3 will detail how neutrons are produced), so the incident beam is not a pure quantum state, but is better described by a Maxwellian distribution around 300 K. Then, interaction of incident neutrons with the system will lead to decoherence, justifying a statistical description over the previous quantum formalism. Nevertheless, we need to keep this classical limit in mind: in some extreme cases, like at temperatures close to 0 K, it does not apply anymore and can even be visible in the sample dynamics, through the so-called zero-point fluctuations.

2.4.6 Coherent and incoherent contributions

Within the framework of the Born approximation at the classical limit, the Van Hove correlation function can be interpreted by separating the sum on j and j' in Equation 2.68:

- For $j = j'$, it is written as:

$$\mathrm{g}_i(\vec{r}, t) = \frac{1}{N} \sum_{j}^{N} < \delta\Big(\vec{r} + \vec{r_j}(0) - \vec{r_j}(t)\Big) >, \tag{2.70}$$

  and reads as the probability to find a scatterer at time t on the position $\vec{r_j}(t)$ knowing that at the initial time $t_0 = 0$ it was at $\vec{r_j}(0)$. This part is the so-called *incoherent* part, and is directly related to the dynamics of the sample.

- For $j \neq j'$, we are dealing with the *coherent* part, which represents how distinct scatterers are related to each other. If we take the case at $t = 0$, we recall the static pair-correlation function:

$$\mathrm{g}_s(\vec{r}, t = 0) = \frac{1}{N} \sum_{j,j'}^{N} < \delta\Big(\vec{r} + \vec{r_{j'}}(0) - \vec{r_j}(0)\Big) >, \tag{2.71}$$

  which is directly related to the static structure of the sample.

As a consequence, Equation (2.61) can be split into two terms:

$$\frac{\mathrm{d}^2\sigma_{\vec{k_i}\to\vec{k_f}}}{\mathrm{d}\Omega \mathrm{d}E_f} = \frac{k_f}{k_i}\frac{N}{4\pi}\Big[\sigma_{coh}\mathrm{S}_{coh}(\vec{Q}, \omega) + \sigma_{inc}\mathrm{S}_{inc}(\vec{Q}, \omega)\Big], \tag{2.72}$$

$$\text{with} \ : \quad \sigma_{coh} = 4\pi(\bar{b})^2 \quad ; \quad \sigma_{inc} = 4\pi\overline{(b^2)} - 4\pi(\bar{b})^2, \ \text{for a homogeneous sample.} \tag{2.73}$$

Depending on the mean value of the scattering length $\bar{b}$ or the mean-square value $\overline{b^2}$, each atom will contribute more or less to the coherent or incoherent part. In the case of Hydrogen, $\bar{b}$ is close to 0, in contrary to $\overline{b^2} \simeq 6.5$ barns [115], which results in a big incoherent part. We can also refer to the coherent and incoherent cross-sections displayed in Table 2.2.

In the following of the thesis, we will focus on the incoherent neutron scattering techniques, which give access to dynamical information on a given sample. Different approaches can be used, depending on the energy transfer $\Delta E = \hbar\omega$ between the incoming neutrons and the target sample:

- *Elastic incoherent neutron scattering (EINS)*, where there is no energy exchange. It gives information about the geometry of atomic movements.

- *Quasi-elastic incoherent neutron scattering (QENS)*, where small energy exchanges arise due to diffusional motions inside the sample. It gives access to the geometry of the motions, through the elastic part (Elastic Incoherent Structure Factor, abbreviated EISF), along with the nature of atomic movements: rotation or translation for example, with their timescales and diffusion coefficients.

- *Inelastic incoherent neutron scattering (INS)*, where specific energy transfers occur between the neutron and the target, which can be caused for example by phonons.

Figure 2.6 shows the different cases as they appear in the incoherent neutron scattering function. In the following, we will only focus on the elastic and quasi-elastic incoherent neutron scattering techniques, so small energy transfers. Usually in the case of thermal neutrons, they are in the order of magnitude of the μeV to meV.

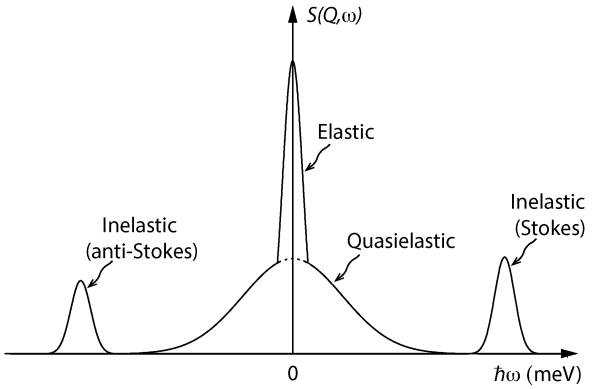


Figure 2.6: Representation of the incoherent neutron scattering function as a function of the energy transfer $\hbar\omega$ for one momentum transfer Q-value.

2.4.7 Elastic Incoherent Neutron Scattering (EINS)

For deriving the neutron scattering function in the case of EINS, the following hypotheses are assumed:

- There is no energy exchange between the neutron and the target sample, however we need to consider an instrumental energy resolution ΔE, with $\hbar\omega = 0 \pm \Delta E$;

- The positions of the nuclei after a long time $\vec{r}_j(\infty)$ and at the initial time $\vec{r}_{j'}(0)$ are uncorrelated, and both tend towards $\vec{r}_j$.

Consequently, the intermediate scattering function Equation (2.67) can be expressed in a time-independent way [116]:

$$\mathrm{I}_{inc}(\vec{Q}, \infty) = \frac{1}{N} \sum_j | < \mathrm{e}^{i\vec{Q}\cdot\vec{r_j}} > |^2. \tag{2.74}$$

By assuming that the motions of the atoms around their equilibrium positions follow a Gaussian distribution, hypothesis widely used in EINS known as the Gaussian approximation, the Fourier Transform of Equation (2.74) gives [121]:

$$S_{inc}(\vec{Q}, \omega = 0) \approx e^{-\frac{1}{3}<u^2>Q^2}, \tag{2.75}$$

which is valid for $< u^2 > Q^2_{max} \lesssim 1$, with $< u^2 >$ being the *Mean-Square Displacements (MSD)*. MSD represent the average dynamics of the individual atoms, especially Hydrogen atoms, as their incoherent cross-section is much higher than for other atoms [118]. Thus, to access the dynamics' information of a sample, the MSD are retrieved from the $S_{inc}(\vec{Q}, \Delta E)$ through:

$$< u^2 > \approx -3 \frac{d \ln[S_{inc}(\vec{Q}, \Delta E)]}{dQ^2}\Big|_{Q \leq Q_{max}}. \tag{2.76}$$

In elastic scattering, the measured intensity depends only on the scattering angle 2θ. If the wavelength λ is fixed by the monochromator of the instrument, it is possible to express the intensity as a function of the momentum transfer Q, through Bragg's law (see Figure 2.7) at first order of diffraction (assumed to be the sole order measured by the detector):

$$2d \sin(\theta) = \lambda. \tag{2.77}$$

On the other hand:

$$Q = \frac{2\pi}{d}. \tag{2.78}$$

By using Eq. (2.77) and (2.78), Q can be expressed as:

$$Q = \frac{4\pi}{\lambda} \sin(\theta). \tag{2.79}$$

Therefore, we can rewrite the measured intensity $S_{inc}(\theta, \Delta E)$ as $S_{inc}(\vec{Q}, \Delta E)$, described by the Gaussian approximation Equation (2.75), and the MSD can be extracted as explained through Equation (2.76).

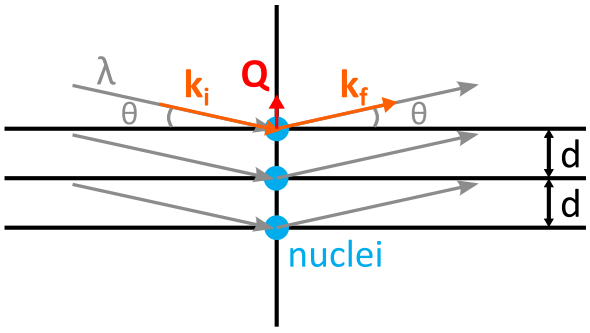


Figure 2.7: Scheme linked to Bragg's law. θ is the half-deviation angle. d is the inter-planar spacing. Bold characters correspond to vectors.

2.4.8 Quasi-elastic Incoherent Neutron Scattering (QENS)

To go further, we remind here the incoherent part of the intermediate scattering function:

$$I_{inc}(\vec{Q},t) = \frac{1}{N}\sum_j \left\langle \mathrm{e}^{i\vec{Q}\cdot\vec{r}_j(t)}\mathrm{e}^{-i\vec{Q}\cdot\vec{r}_j(0)}\right\rangle. \tag{2.80}$$

We consider that within the energy transfers corresponding to quasi-elastic scattering, the following dynamical motions are considered for a molecule:

- The internal molecular vibrations of the nuclei around their equilibrium position;
- The motions of the molecule as a whole: external translational and rotational vibrations and motions.

Moreover, another important point assumed here is the independence of the dynamics. In this context, each motion can be examined separately, which supposes no coupling. A subsequent condition is then that the different motions occur at distinct and defined time scales. We can then write the position vector of the target nuclei $\vec{r}(t)$ as:

$$\vec{r}(t) = \vec{r}_e(t) + \vec{u}(t), \tag{2.81}$$

where $\vec{r}_e(t)$ is the location of the equilibrium position at time t, so corresponding to the molecular motions as a whole, and $\vec{u}(t)$ the displacement of nuclei around their equilibrium position, due to internal vibrations. $\vec{r}(t)$ for each system is unknown, and is precisely containing the dynamical information we would want to extract from a system. In molecular dynamics (MD) simulations, $\vec{r}(t)$ is derived from a force field we need to characterize and build, which constitute a starting point and it requires to formulate strong hypotheses. In QENS, we can infer $\vec{r}(t)$, so the molecular motions occurring in our sample, from the measurements of the neutron scattering function $\mathrm{S}(\vec{Q},\omega)$. In this section, we will see how we deal with neutron scattering functions, and the forms we can derive from general and common diffusional models, like Brownian diffusion.

Separation of different contributions

In the following derivation, we will consider the case of a non-crystalline sample, in which the neutrons are mainly interacting with Hydrogen atoms. In addition, all H atoms are assumed to be equivalent, implying that QENS analysis infers an average dynamics, and not individual motions. In this context, the dynamic independence and separation of the time scales allow the intermediate scattering function to be written

as [111]:

$$I_{inc}(\vec{Q},t) = I^{T}_{inc}(\vec{Q},t) \cdot I^{R}_{inc}(\vec{Q},t) \cdot I^{V}_{inc}(\vec{Q},t), \tag{2.82}$$

with $I^{T}_{inc}(\vec{Q},t)$ the translational term, $I^{R}_{inc}(\vec{Q},t)$ the rotational part and $I^{V}_{inc}(\vec{Q},t)$ corresponding to the vibrational motions. Applying the convolution theorem, and because the scattering function $S_{inc}(\vec{Q},\omega)$ is the Fourier transform of $I_{inc}(\vec{Q},t)$ in time, we then write:

$$S_{inc}(\vec{Q},\omega) = S^{T}_{inc}(\vec{Q},\omega) \circledast S^{R}_{inc}(\vec{Q},\omega) \circledast S^{V}_{inc}(\vec{Q},\omega), \tag{2.83}$$

with the convolution being defined as:

$$S^{A}(\vec{Q},\omega) \circledast S^{B}(\vec{Q},\omega) = \int \mathrm{d}\omega' S^{A}(\vec{Q},\omega') S^{B}(\vec{Q},\omega-\omega'). \tag{2.84}$$

The vibrational part $S^{V}_{inc}(\vec{Q},\omega)$ constitutes the Debye-Waller factor, as derived in Bée [111]:

$$S^{V}_{inc}(\vec{Q},\omega) = \mathrm{e}^{-<u^2_V>Q^2}\Big[\delta(\omega) + S^{V}_{inel}(\vec{Q},\omega)\Big]. \tag{2.85}$$

$S^{V}_{inel}(\vec{Q},\omega)$ can be related to the different vibrational levels of a molecule, which can be written as a sum of delta-functions centered at different energy transfers ω_μ. In practice, however, these levels are far from the quasi-elastic regions, and can then be neglected in QENS, so that the Debye-Waller factor is only:

$$S^{V}_{inc}(\vec{Q},\omega) \simeq \mathrm{e}^{-<u^2_V>Q^2} \cdot \delta(\omega). \tag{2.86}$$

This factor damps the total neutron scattering intensity at high-Q values. Finally, coming back to Equation 2.83, we end up with the general expression for the quasi-elastic contribution:

$$S^{q}_{inc}(\vec{Q},\omega) = \mathrm{e}^{-<u^2_V>Q^2}\Big[S^{T}_{inc}(\vec{Q},\omega) \circledast S^{R}_{inc}(\vec{Q},\omega)\Big]. \tag{2.87}$$

Expressing the translational and rotational terms is the subject of numerous studies in QENS, and we will give later a survey of the main models used in this thesis.

The Elastic Incoherent Structure Factor (EISF)

We saw in Equation 2.74 that the intermediate scattering function can be written at infinite time $t \to \infty$ only as a function of $\vec{Q}$. The scattering function in QENS can then be expressed in a general way as a sum of the elastic part (at infinite time) and

the quasi-elastic part depicted previously:

$$I_{inc}(\vec{Q},t) = I_{inc}(\vec{Q},\infty) + I^{q}_{inc}(\vec{Q},t) \tag{2.88}$$

$$\Rightarrow S_{inc}(\vec{Q},\omega) = I_{inc}(\vec{Q},\infty)\delta(\omega) + S^{q}_{inc}(\vec{Q},\omega). \tag{2.89}$$

The $I_{inc}(\vec{Q},\infty)\delta(\omega)$ will constitute the pure elastic component of the scattering function. Due to the infinite time considered, it corresponds to the final, and average, distribution of the target nuclei. $I_{inc}(\vec{Q},\infty)$ accounts for the so-called *Elastic Incoherent Structure Factor (EISF)*, and is the fraction of elastic scattering contained in the total quasi-elastic intensity. A non-zero EISF indicates the presence of an elastic peak $\delta(\omega)$, and so the presence of a scatterer which is localized in space. As a consequence, the EISF accesses the geometry of the motions within a sample. In practice, it is determined as the ratio between the integrated intensities of the elastic part $\mathcal{A}^{el}(Q)$ and the total ones:

$$A_0(Q) = \frac{\mathcal{A}^{el}(Q)}{\mathcal{A}^{el}(Q) + \mathcal{A}^{q}(Q)}, \tag{2.90}$$

with $\mathcal{A}^{q}(Q)$ the integrated intensity of the quasi-elastic broadening.
If we look at the two extreme cases, a purely elastic line will lead to an EISF equal to 1, which corresponds to nuclei remaining immobile. Whereas a purely translational disordered sample, like water, for which $G_{inc}(\vec{r},\infty) = 0 \Rightarrow I_{inc}(\vec{Q},\infty) = 0$, will give an EISF equal to 0, so no elastic peak.
Nevertheless, it is necessary to recall that instrumental resolution is finite (what we call ΔE in the calculations of EINS). As a consequence, we never measure a purely elastic line; the latter is always broadened, as can be displayed in pink in Figure 2.8. This broadening ΔE gives an indication about the time scale accessible by the instrument, and has several consequences. First, if motions are too slow compared to the accessible time scale (towards magenta curves in Figure 2.8), the resultant quasi-elastic broadening will be considered as negligible compared to the broadening of the elastic peak, and so these slow motions will be neglected. Secondly, the fast motions will turn to big broadening, that can be seen in its extreme case as a flat background in the quasi-elastic spectra (towards cyan in Figure 2.8). These two instrumental limits result in the impossibility to determine an absolute EISF, which will change with the instrument resolution.

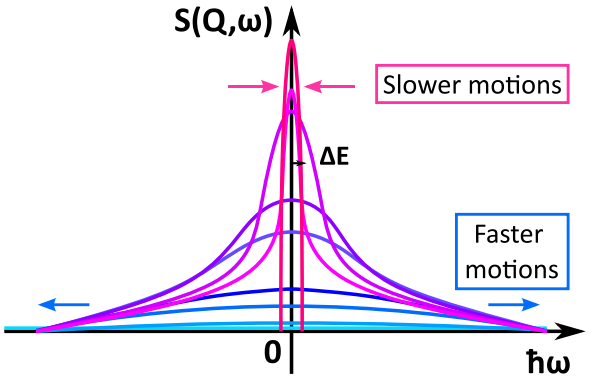


Figure 2.8: Broadening of the QENS spectrum $S(\vec{Q},\omega)$ depending on the motions time scale. Elastic peak in shown in pink with a finite width ΔE corresponding to the instrumental resolution.

2.4.9 Common diffusion models in QENS

Fickian diffusion

If we consider a liquid sample, and long time scales, notably compared to the mean time between the collisions of the molecules, the Van Hove pair correlation function $g(\vec{r},t)$ will obey the equation of diffusion, or Fick's law;

$$\frac{\partial n(\vec{r},t)}{\partial t} = D\nabla^2 n(\vec{r},t), \tag{2.91}$$

with $n(\vec{r},t)$ the number of atoms per unit volume at position $\vec{r}$ and time t, and D is the diffusion coefficient. For isotropic diffusion, the solution is of the form [116]:

$$g(\vec{r},t) = \left[\frac{1}{2\pi\sigma^2(t)}\right]^{\frac{3}{2}} \mathrm{e}^{-\frac{r^2}{2\sigma^2(t)}} \tag{2.92}$$

$$\Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}\sigma^2(t) = 2D \tag{2.93}$$

$$\Rightarrow \sigma^2(t) = 2D|t| + c, \tag{2.94}$$

with c a constant that can be considered as small compared to $2Dt$ in the case of long time scales, and will be neglected in the following. From Equation 2.92, we can determine the intermediate neutron scattering function by taking the Fourier transform in space:

$$\mathrm{I}(\vec{Q},t) = \mathrm{e}^{-Q^2 D|t|}. \tag{2.95}$$

From there, a Fourier transform in time gives the measured neutron scattering function $S(\vec{Q}, \omega)$:

$$S(\vec{Q}, \omega) = \frac{1}{\pi} \cdot \frac{DQ^2}{(DQ^2)^2 + \omega^2}, \tag{2.96}$$

which has the form of a Lorentzian function centered in $\omega = 0$ (schematized in Figure 2.9), whose half-width at half-maximum (HWHM) writes then as:

$$\Gamma(Q) = DQ^2. \tag{2.97}$$

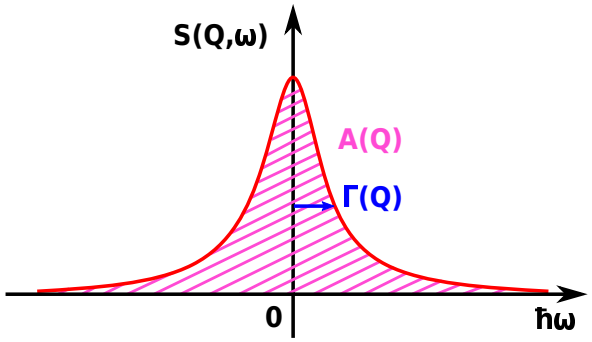


Figure 2.9: Representation of a Lorentzian function, with its HWHM $\Gamma(Q)$ in blue, and its area $A(Q)$ (also termed amplitude) in pink.

From there, we remark that the quasi-elastic broadening measured in the spectrum $S(\vec{Q}, \omega)$ can be modeled by a Lorentzian function. If we fit the measurements with the expression presented in Equation 2.96, extracting from it the HWHM for each Q-value, and fitting the HWHM as a function of Q^2 with Equation 2.97, we can retrieve information on the dynamics of the system through the diffusion coefficient D. This method is at the basis of QENS spectra analysis.

In general, the measured quasi-elastic broadening will be modeled by several Lorentzian functions of different HWHM, depending on the time scale of the different motions considered. The trend of the HWHM as a function of Q^2 gives the type of motion. A linear trend corresponds to Fickian diffusion, as depicted in Figure 2.10 by the orange line. Deviation from this linear case can be given by other models, that will be exposed in the next sections.

Jump-diffusion

In the jump-diffusion model, the molecules do not diffuse in a continuous way like in the previous case. We need to consider sites around which the molecule oscillates during a time τ, before «jumping» to another site with diffusion coefficient D, repeating again the oscillation during τ, jump again etc. In that case, Singwi and Sjölander [122] show

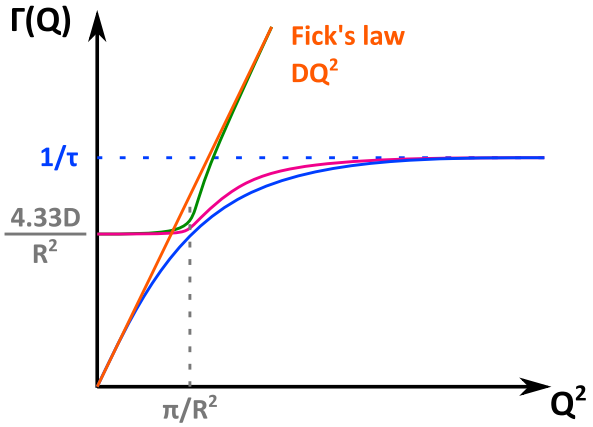


Figure 2.10: HWHM as a function of Q^2 represented for different type of motions. In orange is shown the Fickian diffusion, in blue the jump-diffusion. In green and pink the respective confined motions.

that the corresponding neutron scattering function $S(\vec{Q}, \omega)$ can also be written as a Lorentzian function, but with a different HWHM:

$$\Gamma(Q) = \frac{DQ^2}{1 + DQ^2\tau}. \tag{2.98}$$

At low-Q values, so high r values, we can verify that $\Gamma(Q) \to DQ^2$, indicating that at big space scale the molecule follows a Fickian diffusion behavior. Reversely, at high Q-values, the HWHM tends towards a plateau equal to $1/\tau$, as shown in Figure 2.10 (blue curve).
It can be shown that the corresponding EISF for a two-site jump-diffusion writes as [123]:

$$A_0(Q) = 1 - 2\phi(1 - \phi)[1 - j_0(Qd)], \tag{2.99}$$

with d the distance between the two sites, and $\phi \in [0; 0.5]$ the proportion of particles performing jump-diffusion.
The jump-diffusion is a common model to describe the motions of Hydrogen atoms in methyl groups.

Confined diffusion

Finally, a common class of diffusion models includes confined diffusion. Hall and Ross [124] calculated that a corresponding EISF can be simply written as:

$$A_0(Q) = \left[\frac{3j_1(QR)}{QR}\right]^2, \tag{2.100}$$

with j_1 the spherical Bessel function of first order, and R the radius of confinement. The corresponding neutron scattering spectra $S(\vec{Q},\omega)$ can still be written as a superposition of Lorentzian functions, but their HWHM have more complicated expressions. Hall and Ross [124] show, however, that the confinement is visible at low Q-values through a plateau, whose value depends on the diffusion coefficient D and the radius of confinement R. Volino and Dianoux [125] calculated numerically the values of this plateau, and in a good approximation, it can be estimated to:

$$\Gamma(Q \to 0) \simeq \frac{4.33D}{R^2}. \tag{2.101}$$

At high Q-values, so small r values, the confinement is not visible, and the HWHM tends toward either a DQ^2 law in the case of confined Fickian diffusion (green curve in Figure 2.10), either a plateau of around $1/\tau$ for confined jump-diffusion (pink line in Figure 2.10).

To model directly the HWHM, we can use a heuristic model incorporating the two asymptotic cases, and sigmoid functions to represent the intermediate trend (a factor 10 is arbitrarily applied to have a steeper variation), as done for example in [126].

Confined Fickian diffusion:

$$\Gamma(Q) = \frac{1}{1+\mathrm{e}^{+10\left[Q^2-(\pi/R)^2\right]}} \cdot \frac{4.33D}{R^2} + \frac{1}{1+\mathrm{e}^{-10\left[Q^2-(\pi/R)^2\right]}} \cdot DQ^2, \tag{2.102}$$

Confined jump-diffusion:

$$\Gamma(Q) = \frac{1}{1+\mathrm{e}^{+10\left[Q^2-(\pi/R)^2\right]}} \cdot \frac{4.33D}{R^2} + \frac{1}{1+\mathrm{e}^{-10\left[Q^2-(\pi/R)^2\right]}} \cdot \frac{DQ^2}{1+DQ^2\tau}. \tag{2.103}$$

Obviously, there are many other types of short-time diffusion that can occur in a system, like rotational diffusion [127], and they are described by numerous models, as reviewed for example in Gabel et al. [128], Bée [129], Khodadadi and Sokolov [130], Smith et al. [131], Grimaldo et al. [132]. New models continue to be developed today, tending towards more complex and complete pictures of the dynamics of biomolecules, as for example in the following studies on lipids [133, 134]. Alternatively, to overcome the limitations of some approximations, novel theoretical frameworks are developed for EINS and QENS. We can cite the approach introduced by Peters and Kneller [135] to analyze EINS data beyond the Gaussian approximation, or the work of Kneller [136] to consider a full quantum framework, over the classical limit of Van Hove [120].

2.5 Electron microscopy

Compared to visible light, electrons accelerated in vacuum can reach much smaller wavelengths. As a consequence, electron microscopes can achieve better resolutions, and access the nanometer and sub-nanometer scale. Nowadays, they are able to reveal the details of various systems, from whole cells, bacteria and viruses, to individual proteins or fibers. The resolution goes from the nm to the Å lengthscale, which even allows to build atomic models.

In this section, the basics of image formation in electron microscopy will be presented, then emphasis will be put on single-particle reconstruction. This part will rely on the following books and articles from Frank [137], Reimer [138], Vulović et al. [139], Koeck and Karshikoff [140], and online courses from Jensen [141]. Introduction of the Fourier slice theorem draws on an online presentation from Dynerman [142].

2.5.1 Electron-sample interaction

In a sample, free electrons will interact with the Coulomb potential of the nuclei, as well as the bound electrons, through electromagnetic forces. There is a wide variety of possible reactions, that are summarized in Figure 2.11:

- Electron-sample interaction can trigger light emission; through deexcitation of excited bound electrons, giving rise to catholuminescence or X-rays, or through Bremsstrahlung losses.
- Excitation of the bound electrons can also generate Auger or secondary electrons.
- Absorption of the electrons by the sample or creation of electron-hole pairs can also occur, leading to a measurable electrical current within the sample.
- Different scattering processes can happen. First, the electrons can be backscattered. Their detection is at the basis of Scanning-Electron Microscopy (SEM). The electrons can also be transmitted, which constitute the ground of Transmission-Electron Microscopy (TEM). The transmitted beam is composed of unscattered electrons, elastically scattered electrons (no energy exchange within the sample), and inelastically scattered electrons in the case of energy exchange. An image is formed by the interferences of scattered electrons with the unscattered ones. However, the image formed by inelastically scattered electrons is not in the focal plane, and can cause chromatic aberrations. In practice, the inelastically scattered electrons can be removed by using energy-filters, so only elastic scattering is measured. Nevertheless, inelastic scattering remains the

major cause of radiation damage within the sample, which still contributes to add noise in the images.

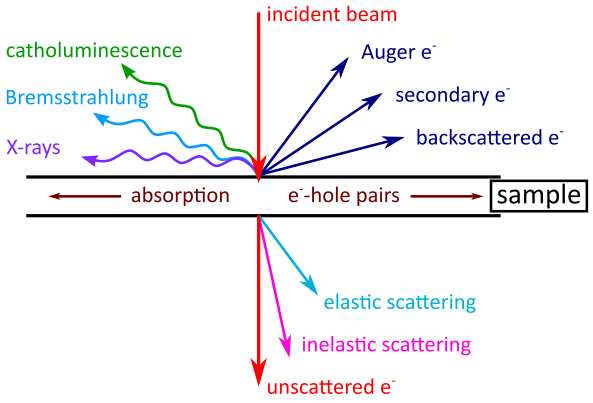


Figure 2.11: Reactions taking place between an incident beam of electrons and a sample. Adapted from Williams and Carter [143].

2.5.2 The two modes of transmission electron microscopy

Before going into the details of electron scattering, it is necessary to introduce the two main configurations of transmission electron microscopy:

- Single-particle: the grid where the sample is deposited (the different types of sample preparation will be outlined in Chapter 3) remains perpendicular to the electron beam, as represented in Figure 2.12a. In that mode, the sample is composed of numerous copies of the same protein, but all measured in different orientations. By assuming a homogeneous sample, the identical projections are averaged to increase the signal-over-noise ratio (SNR), and the different orientations are back-projected to give a 3D reconstruction of the protein.

- For bigger samples, like viruses or entire cells, where homogeneity can not be assumed and direct averaging not used, the sample can be rotated according to the electron beam, as schematized in Figure 2.12b. In that case, we are talking about electron tomography (ET). The resulting tilt series is used to directly reconstruct the full object (we talk then about a 3D tomogram). In some cases like viruses, where several copies of the spike proteins can be assumed to be similar enough, sub-tomogram averaging (STA) can be used to increase the SNR in some regions, and obtain better reconstructions.

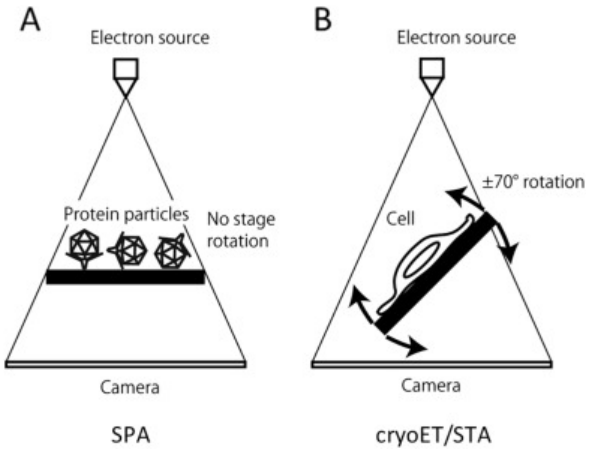


Figure 2.12: Differences between the configurations of single-particle (A) and tomography (B). Figure taken from [144].

While the methods of reconstruction are differing, they share common basics that we will present in the following, before focusing on single-particle analysis (SPA).

2.5.3 Electron scattering cross-section

In the case of electrons, the potential of the atoms within the sample includes both the Coulomb potential of the nucleus, but also the screening of the bound electrons by their own potential. The picture here is much more complicated than for neutron scattering, where all the atoms were seen as point-like nuclei. Here for example, the extension of the electronic cloud needs to be taken into account.

Nevertheless, we can still define the scattering process in the far-field regime in a similar way as for neutrons. The exit wave can then be described by an unscattered plane wave and a scattered spherical wave [138]:

$$\psi_s = \psi_0 \Big[\mathrm{e}^{2\pi i k z} + i f(\theta) \frac{\mathrm{e}^{2i\pi k r}}{r} \Big], \tag{2.104}$$

with $f(\theta)$ the scattering amplitude. We notice there that compared to neutrons, the scattered part presents a phase shift $i = \mathrm{e}^{i\pi/2}$. Moreover, $f(\theta)$ is complex:

$$f(\theta) = |f(\theta)| \mathrm{e}^{i\eta(\theta)}, \tag{2.105}$$

with $\eta(\theta)$ an additional phase shift, depending on the scattering angle. However, in the far-field regime, the scattering cross-section and amplitude are still related like in Equation 2.28 where:

$$\frac{\mathrm{d}\sigma_e}{\mathrm{d}\Omega} = |f(\theta)|^2. \tag{2.106}$$

Calculations of the scattering amplitude, and then the subsequent elastic cross-sections, can be done using several methods, outlined in Reimer [138]. As for neutrons, some of them rely on a partial wave analysis and finding an exact solution for Schrödinger's equation.

An approximate formula for the elastic cross-section of electrons, considering Hartree-Fock-Slater or Dirac-Slater atoms, can be written as [145]:

$$\sigma_e = \frac{1.4 \cdot 10^{-6}}{\beta^2} Z^{3/2} \left[1 - 0.26 \frac{Z}{137\beta}\right] \text{ nm}^2, \text{ for } \frac{Z}{137\beta} < 1.2, \tag{2.107}$$

where Z is the atomic number, β the ratio between the velocity of electrons and the speed of light c, $\beta^2 = 1 - \left[\frac{mc^2}{V_0 + mc^2}\right]^2$, V_0 the acceleration voltage, and mc^2 the rest energy of electrons. Equation 2.107 displays a clear relation between the cross-section of electrons, and the extension of the electronic cloud, with the atomic number Z. Indirectly through β, it also depicts how the cross-section depends on the acceleration voltage of the microscope.

Determination of inelastic cross-sections is harder, and relies on much more parameters, like formation of chemical bonds within the sample. In general it can be shown that the inelastic cross-section is higher than the elastic one, which justifies the importance of removal of inelastically scattered electrons from the final signal. More recently, Peet et al. [146] studied in more details the elastic and inelastic cross-sections of electrons for Carbon atoms, whose results are displayed in Figure 2.13.

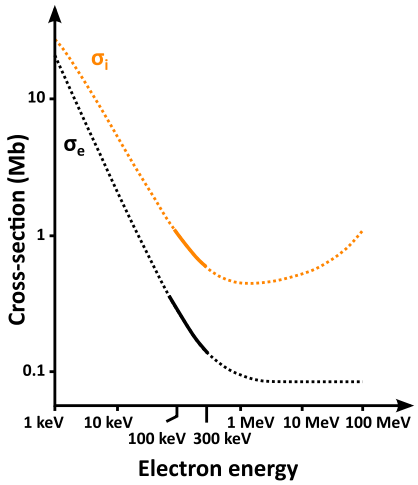


Figure 2.13: Theoretical elastic (black line) and inelastic (yellow line) cross-sections of the electron for Carbon atoms as a function of the beam energy. Solid line represent experimentally verified regions, whereas extrapolated ones are depicted with dotted lines. Adapted from Peet et al. [146].

What we remark first is the decrease of electron scattering cross-sections with the acceleration voltage, then the high difference (more than the double) between

the inelastic and elastic parts, especially for high voltage. The order of magnitude of electron cross-sections, around 10^6 barns, can be put in parallel with neutron cross-sections displayed in Table 2.2, that remain much smaller, in the barn to 10 barns range, in accordance with the point-like nuclei hypothesis.

2.5.4 Common approximations for image formation

To understand what the measured signal corresponds to, several approximations are assumed, as reviewed in Vulović et al. [139].

The small angle approximation

This approximation is similar to the Gauss conditions in optics [147]. The electron beam travels mainly following the axis z of the microscope, so the corresponding wave function can be separated in two parts;

$$\psi(\vec{r}) = \Psi(x, y, z)\mathrm{e}^{ikz}, \tag{2.108}$$

where e^{ikz} is the plane wave traveling in z direction, and $\Psi(x, y, z)$ the part varying slowly with z. Rearranging the Schrödinger's equation 2.7, and considering the small angle approximation $|\frac{\partial^2\Psi}{\partial z^2}| << |k\frac{\partial\Psi}{\partial z}|$, i.e. $k^2 >> k_x^2 + k_y^2$, we get:

$$\frac{\partial\Psi(x, y, z)}{\partial z} = \left[\frac{i\lambda}{4\pi}\nabla_{x,y}^2 + i\sigma V(\vec{r})\right]\Psi(x, y, z), \tag{2.109}$$

where λ is the incident wavelength of the electron, and $\sigma = \frac{2\pi m|e|\lambda}{h^2}$ the interaction constant, which includes the mass m of the electron, and its charge $|e|$. Equation 2.109 displays two parts; a propagation term $(i\lambda/4\pi)\nabla_{x,y}\Psi$, and an interaction term $i\sigma V(\vec{r})\Psi$.

The projection assumption

If the sample is thin enough, the propagation term can be neglected compared to the interaction term. In that case, we can solve Equation 2.109 and end up with the following expression of $\Psi(x, y, z)$:

$$\Psi(x, y, z) = \mathrm{e}^{i\sigma\int_{-\infty}^{z} V(\vec{r})dz'}. \tag{2.110}$$

The integral corresponds to the projected potential $V_z(\vec{r})$, and so the resulting wave function is written as:

$$\Psi(x, y, z) = \mathrm{e}^{i\sigma V_z(\vec{r})}. \tag{2.111}$$

The weak-phase object approximation

In the case of weak scattering, so $\sigma V_z(\vec{r}) << 1$, and neglecting multiple scattering, we can simply write Equation 2.111 as:

$$\Psi(x, y, z) \simeq 1 + i\sigma V_z(\vec{r}). \tag{2.112}$$

As a consequence of these approximations, the resulting wave function, i.e. the detected signal, can be directly seen as the linear projection of the sample potential.

The interesting part of this result is that the sample potential is actually what we want to measure when performing microscopic measurements. When aiming at the structure of a protein for example, we probe the Coulombic potential of the protein, that we will note $\rho(x, y, z)$ (to not mix with the general potential of the sample $V(x, y, z)$ that can refer to any atom in the sample, protein or background). In this context, the detected image $I(x, y)$ of one protein corresponds then directly to the linear projection of $\rho(x, y, z)$, as schematized in Figure 2.14:

$$I(x, y) = \int \rho(x, y, z)\mathrm{d}z. \tag{2.113}$$

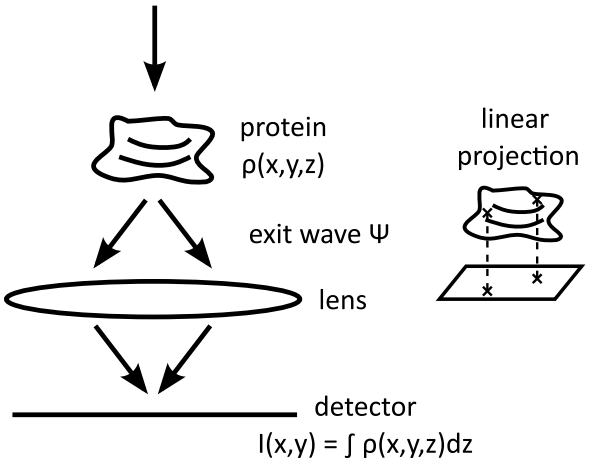


Figure 2.14: Representation of the imaging in the weak-phase object approximation.

If the protein, so $\rho(x, y, z)$, takes different orientations, the projection will be along these new orientations. The consequences are crucial in image reconstruction, and can be introduced with the Fourier slice theorem [148].

2.5.5 Fourier slice theorem

We can compute the 3D Fourier transform of the potential $\rho(x, y, z)$ as:

$$\mathcal{FT}[\rho(x,y,z)] = \hat{\rho}(s_x, s_y, s_z) = \int\int\int \rho(x,y,z) \mathrm{e}^{-2i\pi[s_x x + s_y y + s_z z]} \mathrm{d}x\mathrm{d}y\mathrm{d}z. \tag{2.114}$$

If we consider the central section of that Fourier transform, i.e. when $s_z = 0$, then we can write:

$$\begin{aligned}
\hat{\rho}(s_x, s_y, s_z = 0) &= \int\int\int \rho(x,y,z) \mathrm{e}^{-2i\pi[s_x x + s_y y]} \mathrm{d}x\mathrm{d}y\mathrm{d}z \\
&= \int\int \left[\int \rho(x,y,z)\mathrm{d}z \right] \mathrm{e}^{-2i\pi[s_x x + s_y y]} \mathrm{d}x\mathrm{d}y \\
&= \int\int I(x,y) \mathrm{e}^{-2i\pi[s_x x + s_y y]} \mathrm{d}x\mathrm{d}y \\
\Rightarrow \hat{\rho}(s_x, s_y, s_z = 0) &= \mathcal{FT}[I(x,y)].
\end{aligned} \tag{2.115}$$

The last equation is fundamental, it means that the 3D Fourier transform of the central section is directly related to the 2D transform of the projection.

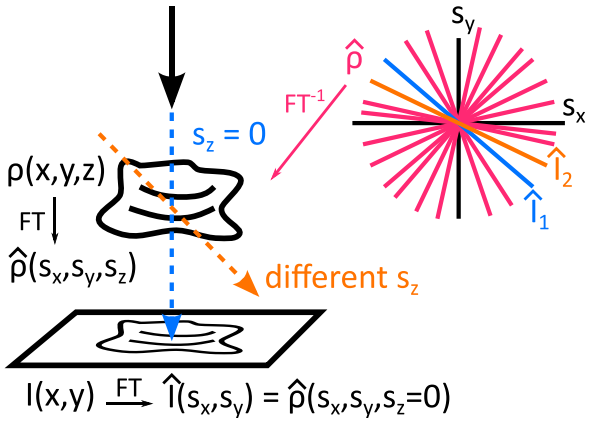


Figure 2.15: Representation of the Fourier slice theorem.

As represented in Figure 2.15, if we get images from different orientations, so different central sections, we can reconstitute the full Fourier transform of the potential $\hat{\rho}(s_x, s_y, s_z)$, and by Fourier inversion, retrieving the potential of the object of interest $\rho(x, y, z)$. This method is at the basis of 3D reconstruction from 2D projections.

2.5.6 Direct Fourier Inversion

Based on the direct Fourier inversion, a basic algorithm for 3D reconstruction, taking as input 2D images $I(x, y)$ from different orientations of the protein $\rho(x, y, z)$, and giving as output the 3D map of $\rho(x, y, z)$, can be written in the following way [142, 149]:

1. Initialization of an empty 3D voxel grid for $\hat{\rho}(s_x, s_y, s_z)$.

2. For each image $I(x, y)$:

   (a) Compute the 2D Fourier transform $\hat{I}(s_x, s_y)$.

   (b) Place $\hat{I}(s_x, s_y)$ along the corresponding central section in the 3D voxel grid. This step can be visualized in Figure 2.15 with a cut at $s_z = 0$ of the voxel grid in the top right. An individual $\hat{I}(s_x, s_y)$ is represented by a line.

   (c) Repeat.

3. Invert $\hat{\rho}(s_x, s_y, s_z)$, output $\rho(x, y, z)$.

There are critical steps in this basic algorithm. The first one, which is central in electron microscopy, is obviously the step 2(b); how to place $\hat{I}(s_x, s_y)$ in the 3D voxel grid, or in other words, how to determine with confidence to which orientation a 2D image corresponds compared to the unknown 3D structure. On top of that, the previous algorithm requires to fill in all the Fourier space. In practice, it is not possible, and one is limited by the finite number of projections that can be measured. Even worse, preferential orientations of the sample can limit further the full covering of Fourier space, and lead to artifacts in the final reconstruction.

Another pivotal step is purely numerical. As the 3D voxel grid is discretized, the computed variables need to be interpolated. Linear interpolation is often used, with a weighted average of the function according to the lattice. However, the discretization also requires to use discrete Fourier transforms, where aliasing effects can occur and need to be handled. It leads to additional constraints on the choice of the grid and sampling, or of appropriate filters in the image analysis.

Finally, one should consider that the image recorded on the detector is not directly the pure projection $I(x, y)$. On the one hand, and especially in the case of cryo-EM, noise needs to be considered. In cryo-EM, it mainly comes from the scattering by the vitrified ice, whose cross-section is of the same order of magnitude as the protein or lipids constituents. In image analysis, the noise is considered to be additive. On the other hand, the image is convoluted to the point-spread function of the electron microscope, the so-called Contrast Transfer Function (CTF). The CTF (presented in more details in Chapter 3) includes the lens aberration, but also the defocus, which is applied during the data collection in order to optimize the phase-contrast. If we call $M(x, y)$, and its corresponding Fourier transform $\hat{M}(s_x, s_y)$, the measured signal, we write:

$$\hat{M}(s_x, s_y) = \hat{C}(s_x, s_y) \cdot \hat{I}(s_x, s_y) + \hat{N}(s_x, s_y), \tag{2.116}$$

with $\hat{C}(s_x, s_y)$ the CTF, and $\hat{N}(s_x, s_y)$ the noise function in Fourier space.

To retrieve the original image signal $\hat{I}$, it would be tempting to directly divide the measured signal $\hat{M}$ by the CTF $\hat{C}$. However, the CTF exhibits some zero values, rending impossible a direct deconvolution by dividing. Different strategies for the CTF correction can be applied, like phase-flipping, but the most optimal way is through a *Wiener filter*, of the form [150]:

$$\hat{I} = \frac{\hat{C}\hat{M}}{\hat{C}^2 + k}, \tag{2.117}$$

$$\text{with } k \text{ the Wiener constant: } k = \frac{1}{SNR} = \frac{|\hat{N}|^2}{|\hat{I}|^2}. \tag{2.118}$$

In theory, the Wiener filter has also low-pass filtering effects, which prevent overfitting. But its implementation in the first 3D reconstruction algorithms (like EMAN [151] or FREALIGN [152]) did not necessarily retrieve these effects. When Bayesian approaches were introduced, it became possible to have a direct implementation of the Wiener filter in 3D reconstruction, which is the case for instance in the RELION (REgularized LIkelihood OptimizatioN) package [150, 153].

2.5.7 Euler angles

The assignment of a precise orientation to each 2D projection is one of the central points in cryo-EM image processing. Formally, a projection can be described in terms of the so-called *Euler angles*. Among the twelve existing conventions, Heymann et al. [154] proposed the ZYZ convention as a common convention for all cryo-EM software. In this context, a rotation can be divided in three subsequent rotations, whose angles are represented in the Cartesian system in Figure 2.16:

- A rotation around the z-axis of an angle ϕ, referred as the *azimuthal angle*;
- A rotation around the y-axis of an angle θ, called the *elevation angle*;
- A final in-plane rotation of angle ψ around the new z-axis.

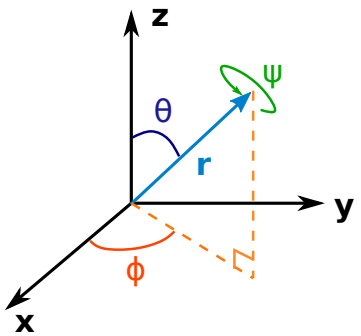


Figure 2.16: Representation of the Euler angles [154].

The general rotation matrix $R(\psi, \theta, \phi)$ can be written as:

$$R(\psi, \theta, \phi) = R_z(\psi) \cdot R_y(\theta) \cdot R_z(\phi), \tag{2.119}$$

with the following matrices:

$$R_z(\phi) = \begin{bmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.120}$$

$$R_y(\theta) = \begin{bmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \tag{2.121}$$

$$R_z(\psi) = \begin{bmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.122}$$

We remark that the in-plane angle ψ does not play a role in the projection, so that the assignment of an orientation to each projection with regard to the 3D initial volume reduces to the determination of the azimuthal and elevation angles, respectively ϕ and θ.

2.5.8 Bayesian methods for 3D reconstruction

The problem of single-particle reconstruction in cryo-EM is quite challenging; it is a so-called incomplete problem in mathematics, because we miss the Euler angles of each projection and need to determine them. In addition, the cryo-EM images are noisy, which make the 3D reconstruction *ill-posed*, and requires regularization, through notably filtering methods. To limit user intervention, and avoid as much as possible overfitting, especially building on noise, Bayesian methods are commonly used in cryo-EM. Introduction about the subject and implementation of such algorithm in the software RELION can be found in Scheres [150, 153], Sigworth et al. [155], Scheres [156].

In a Bayesian framework, we deal with probabilities, and the purpose is to find the model with parameter set Θ which has the highest probability to be correct, knowing the images X and the prior information Y. This probability is called the *posterior distribution* and can be written with Bayes' law as [150]:

$$P(\Theta|X, Y) \propto P(X|\Theta, Y)P(\Theta, Y), \tag{2.123}$$

where $P(X|\Theta, Y)$ is called the *likelihood*, and represents the probability of having the dataset X, knowing the model, and $P(\Theta, Y)$ is the *prior*, corresponding to how probable is the model compared to the prior information. The prior contains the essential condition of « smoothness », meaning that the potential $\rho(x, y, z)$ of a sample slowly varies in space, which remains physically consistent. Θ represents the model, containing for example the Euler angles to optimize by iterative maximization of the posterior probability $P(\Theta|X, Y)$.

Derivation of such a procedure leads to the so-called Expectation-Maximization algorithm [150], which is iterative, as schematized in Figure 2.17:

- In the Expectation step, *projection matching* is performed, by comparing the experimental images to the computed projections of the last reconstruction. This comparison enables to assign new Euler angles to the images. In statistical methods however, it is not one unique set of Euler angles which is assigned, but probabilities for each set of Euler angles.

- Then, in the Maximization step, the Bayes' law as introduced in Equation 2.123 is computed, taking into account the new likelihood, as well as the prior. Because new set of Euler angles were assigned in the previous step, it leads to a new 3D reconstruction, and a new Expectation-Maximization iteration. Iterations are stopped either when the changes in the parameters and the reconstruction are small, either after a fixed number that we define.

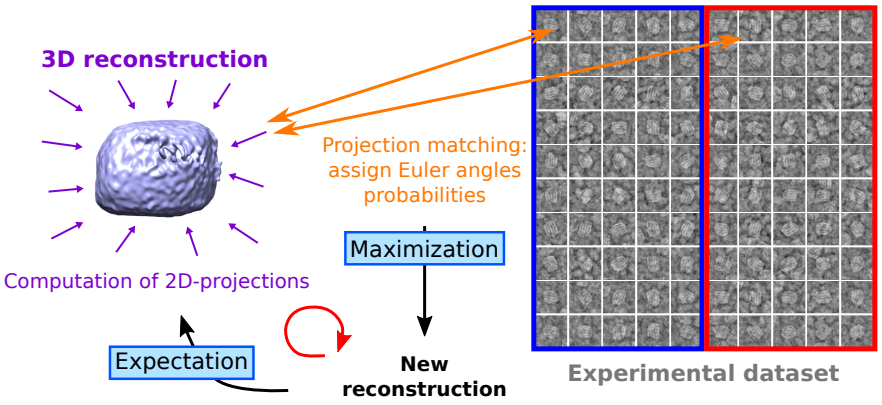


Figure 2.17: Representation of the Expectation-Maximization algorithm. The 3D map was produced with UCSF ChimeraX [157, 158].

Scheres [150] showed that this Bayesian method also allows for the CTF correction, by a direct implementation of the Wiener filter, as introduced in Equation 2.117. In general, Bayesian strategies enable to limit arbitrary settings from the user, except for some precise parameters. One of the main limiting factors, however, remain in the convergence of the 3D reconstruction, which can be stuck in local minima. It is also sensitive to the initial volume we set to start the iterations.

Implementation of Bayesian approaches in single-particle reconstruction was initiated with software packages XMIPP [159, 160], then RELION (REgularized LIkelihood OptimizatioN) [150, 153], and continues to be the rule, as for instance with the more recent software cryoSPARC (Cryo-EM Single Particle Ab-initio Reconstruction and Classification) [161].

2.6 Conclusion

Scattering techniques remain essential to access the atomic scale, whether it be structural or dynamical information. Each probe, X-rays, electrons or neutrons, presents its own advantages and drawbacks, as was already reviewed in 1995 by Henderson [162] for microscopy, which make them all complementary. At the end, the main point to consider is the amount of information we can retrieve compared to the radiation damage. In terms of dynamical studies, the ability of neutrons to capture the motions of Hydrogen atoms, with nearly no radiation damage, is definitely an asset for the investigation of short-time diffusion. On the contrary, due to the low flux, coherence volume and brilliance of neutrons, atomic-resolution microscopy is not achievable, while electrons offer valuable phase-contrast images, and overcome the necessity to produce crystals, as for X-ray diffraction. However, electron microscopy necessitates averaging methods for image processing and 3D reconstruction, which pose other challenges.

In the next chapter, we will shortly walk through the production of either neutrons or electrons, until their detection, by passing through some details on the sample preparation and the instruments, both spectrometers and microscopes, we used during the thesis. An overall picture of image processing in cryo-EM, with some associated software packages, will be exposed as well.

Chapter 3

Instrumentation, Software

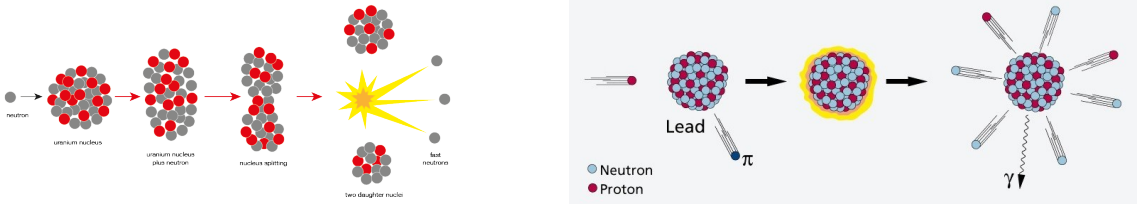
3.1 Neutron scattering

3.1.1 Neutron production and moderation

Neutrons are produced through nuclear reactions. There are two main ways to extract neutrons from a nucleus; through fission or spallation.

Fission

In nuclear fission, the absorption of an incoming neutron by a nucleus of Uranium 235 leads to the splitting of the nucleus in two fragments of smaller atomic number, and the release of 3 free neutrons in average, as represented in Figure 3.1a.



(a) Nuclear fission **(b)** Spallation

Figure 3.1: Comparison of the two types of neutron production. Images taken from `https://mlz-garching.de` and `https://www.psi.ch`.

The free neutrons have a kinetic energy of around 2 MeV and a half-life of about ten minutes [163]. Neutron sources based on nuclear reactors can deliver a high-flux of neutrons, which is continuous. For example, the Institut Laue-Langevin (ILL), which has the brightest flux of thermal neutrons, delivers around $1.5{\cdot}10^{15}$ neutrons per cm^2 and per s. Its nuclear reactor has a thermal power of around 58 MW, and works by cycles of around 50 days, separated by changes of the fuel element [164]. Other reactor

sources are for instance the High Flux Isotope Reactor at the Oak Ridge National Laboratory (USA) or the FRM II at the Heinz Maier-Leibnitz Zentrum (Germany).

Spallation

Spallation depicted in Figure 3.1b is based on evaporation of a nucleus after being hit by a high-energy particle (for example a proton). It necessitates to produce first high-energy particles through an accelerator, before sending them on heavy-target nuclei, like liquid mercury. Spallation gives around 20 neutrons per incident proton, whose energy is also around MeV. Compared to a reactor, the flux is higher, but is pulsed, with a repetition rate of around 50 Hz. This characteristic requires then to adapt the instruments. Among the different existing sources, we can cite the Spallation Neutron Source at the Oak Ridge National Laboratory (USA), the ISIS Neutron and Muon Source at the STFC Rutherford Appleton Laboratory (UK) or the Japan Proton Accelerator Research Complex (Japan).

Moderation

Typical processes we are interested in, like diffusion, are in the order of the meV. As a consequence, the produced MeV-neutrons need to be moderated. Heavy water D_2O can be used as a moderator; collisions with D_2O will enable to go from high-energy neutrons ($\sim$MeV) to thermal neutrons ($\sim$25 meV). Cold and hot sources can then be used to decrease or increase the energy, respectively, depending on the systems that are investigated. Table 3.1 shows the different ranges of energy covered after moderation.

Neutrons	Energy range (meV)	Boltzmann temperature (K)
Cold neutrons	0.1 - 10	$\sim$25
Thermal neutrons	10 - 100	$\sim$300
Hot neutrons	100 - 500	$\sim$2400
Epithermal neutrons	> 500	> 2500

Table 3.1: Characteristic of the different neutron energy ranges after moderation.

3.1.2 Neutron optics and detection

Beam-tubes and neutron guides

From the reactor to the instruments, the neutrons need to be delivered with minimal loss. Neutron guides are based on total internal reflection. This condition can be derived from the Snell-Descartes' law, and necessitates the angle between incoming neutrons and the guide to be lower than a critical value θ_c. The θ_c value depends

strongly on the type of materials, and the condition for total reflection is fulfilled in an optimal way for Nickel, or for an alloy of Ni/Ti [165]. The main difference between the beam-tubes, located directly in the reactor core, and the neutron guides, lies in the external material; Aluminum for beam-tubes, which guarantees a strong mechanical resistance, and glass for the guides, which also enables to absorb the escaping neutrons.

Neutron detectors

Neutron detection relies on the big absorption cross-sections of some elements, as seen in Table 2.2, to convert neutral neutrons into charged particles [166]. Two absorption reactions involving ^{3}He and ^{6}Ni can be written as:

$$^3\mathrm{He} + n \rightarrow ^3\mathrm{H} + p^+ + 0.764MeV \tag{3.1}$$

$$^6\mathrm{Li} + n \rightarrow ^4\mathrm{He} + ^3\mathrm{H} + 4.79MeV \tag{3.2}$$

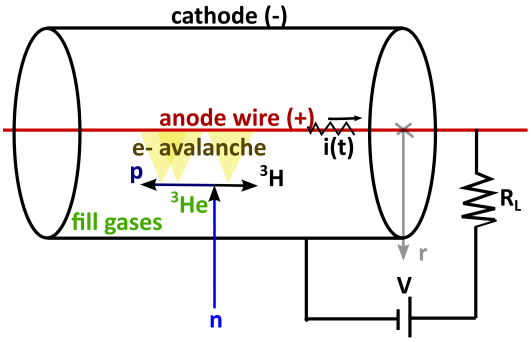


Figure 3.2: Schematic representation of a proportional counter.

The first reaction in Equation 3.1 is at the basis of proportional counters, where tubes are filled with ^{3}He gas, as schematized in Figure 3.2. Because the tubes are submitted to a high voltage V, the exiting protons will ionize the gas, create an electron avalanche, that will give rise to a measurable electrical current on the anode. From there, we see that a neutron detector is basically a neutron counter. The information on the neutron momentum or energy is lost during the detection process, and needs to be measured through other strategies. The momentum transfer can be derived from the deviation angle of the scattered neutrons. To determine this angle, individual detectors can be aligned side by side, and each angle will be related to one specific detector. 1-D position-sensitive detectors (PSD) are also a solution, by detecting the location of a neutron absorption along the anode, as well as multiwire gas detectors. Due to the high increase of ^{3}He prices, other types of detectors are considered, like ^{6}Li scintillators, based on Equation 3.2. In that case, the reaction gives rise to light, whose intensity can be increased with photomultipliers. The light detection technology

enables to accommodate high count rates, and the scintillators are in general more precise in the determination of the neutron time-of-flight.

3.1.3 Instruments

In practice, the neutron scattering function $S(\vec{Q}, \omega)$ is measured through the neutron count rate as a function of three measurable quantities;

- The change in the direction of the neutron, so the angle of deviation θ as represented in Figure 3.3. To access this angle, multi-arrays of detectors whose angles are known are used, as outlined in the previous section. From the angle θ, we can then access the momentum transfer Q with Bragg's law, as derived in Equation 2.78.
- The change in neutron's energy. As explained previously, during the neutron detection all information about its energy is lost. To get the energy transfer $\hbar\omega$, two main strategies are used;
  - To measure the neutron time-of-flight, deriving then its velocity and thus its energy. This approach is used in time-of-flight spectrometers.
  - To measure only at fixed energy transfers, which are then known, and to scan the system for different energy transfers. It is the method underlying backscattering spectrometers.
- The change in the neutron spin orientation. It is of particular interest for magnetic samples, but we will not cover these methods.

We see that in a basic neutron experiment, as we only count the neutrons, we do not separate between the incoherent or coherent part. Polarization analysis can specifically discriminate between the two parts (see for examples papers of Moon et al. [167] and Gaspar et al. [168]), but in general both contributions will be measured in a neutron scattering experiment. The instruments dedicated to structural (coherent) or dynamical (incoherent) studies will have different ranges in momentum and energy transfers. Diffraction or small-angle neutron scattering instruments, focusing more on the coherent part, will measure at lower Q-ranges (so higher deviation angles), there the incoherent contribution will be considered as a background. In addition, elastic processes will be favored in the measurements, and energy transfers are not considered. Reversely, for dynamical studies, higher Q-ranges are measured, and energy transfers are of interest. If the sample has internal structures, the coherent part of the signal can translate into peaks at specific Q-values, which are more difficult to correct. However, at the end, the measurement of the coherent or incoherent part strongly depends on the sample, if it is in a crystalline or amorphous state, or in solution. For instance, a

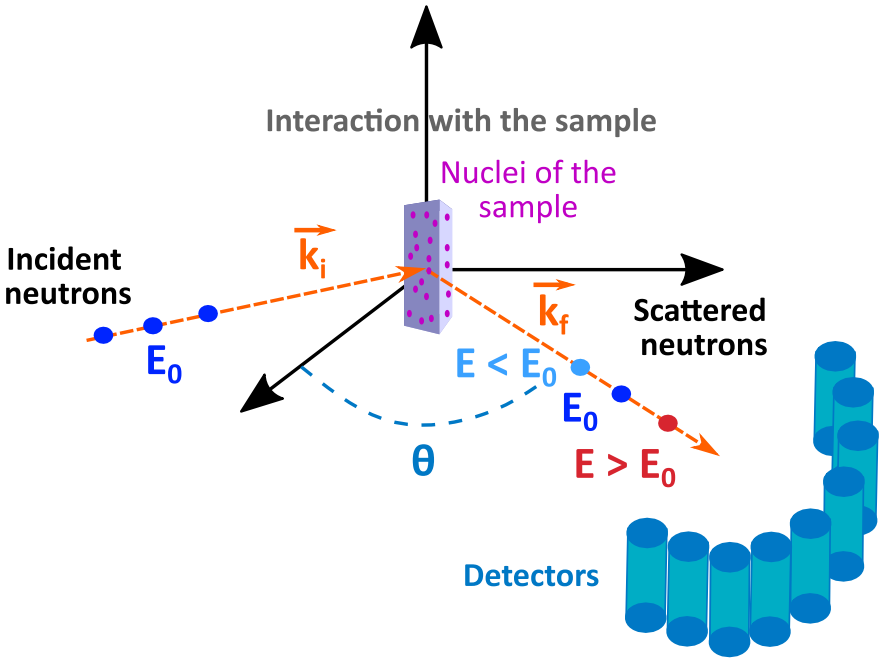


Figure 3.3: Schematic representation of the neutron scattering process.

solution measured in a diffractometer will give a major incoherent part, and no Bragg's peaks. Reversely a crystal in a spectrometer will present a much smaller incoherent contribution compared to the Bragg's peaks. At the very end, what we aim to measure depends on both the sample state and preparation, and the type of instrument used for the experiments.

Characteristics of the instruments used within this thesis are presented in Table 3.2, which summarizes some important parameters to be taken into account for choosing a specific instrument, like the Q-range, the dynamical range (energy transfer), or the energy resolution.

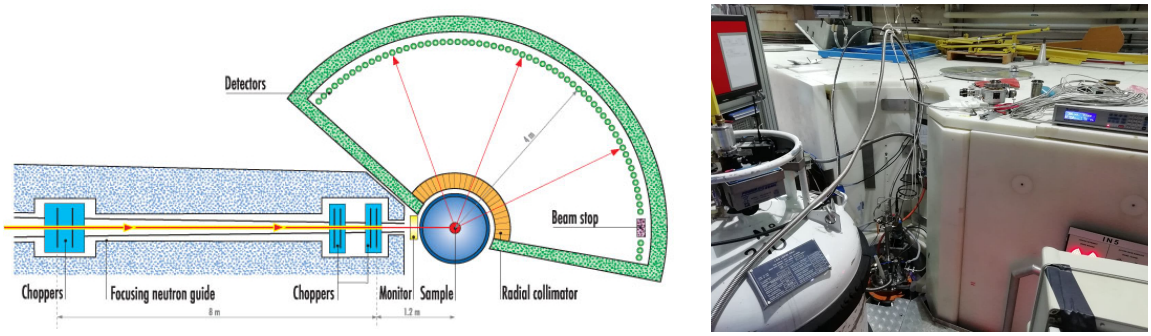
Instruments	**IN13**	**IN5**	**IN16b BATS**	**IN16b Si(111)**
Wavelength (Å)	2.23	10	$\Delta\lambda/\lambda \sim 12$ %	6.271
Accessible Q-range ($Å^{-1}$)	0.2 - 4.9	0.02 - 1.16[a]	~0.2 - 1.8	0.1 - 1.8
Energy transfer (μeV)	/	$\Delta E/E \sim 1$ %	± 150	± 31
Energy resolution (μeV)	8	10	4	0.75
Accessible timescale (ps)	~100	~100	~200	~1000

[a] $\sim 0.2/\lambda < Q_{elastic} < 11.6/\lambda$ (angles $\sim 5°$ to $135°$), i.e., $\sim$ 0.02 - 1.16 with 10 Å

Table 3.2: Characteristics of the spectrometers from Institut Laue-Langevin (ILL) used in this thesis. λ_0 is the initial wavelength, and E the energy transfer between the neutron and the sample.

3.1.4 Time-of-flight spectrometers

The neutron flux needs to be «time-structured» for time-of-flight measurements. It is necessary to know precisely the time t_0 from which we start to measure the time-of-flight τ. To perform that, the neutron flux, which is polychromatic and continuous at the exit of the reactor, is converted into a pulsed monochromatic beam. This conversion is performed using choppers, which are, for example, rotating disks parallel to the neutron beam, coated with Gadolinium. Choppers have also slits transparent to neutrons. Then, by adjusting the angular speed of the chopper, we are able to «cut» the neutron beam, and to select the wavelengths of interest. In the cold time-of-flight spectrometer IN5 [169], represented in Figure 3.4, a set of 6 choppers is used to successively create the neutron pulse, avoid frame overlap (e.g. fast neutrons of a pulse coinciding with slow neutrons of the previous pulse), and to select the initial wavelength λ_0. The choppers constitute the so-called primary spectrometer, while the sample and consecutive detectors corresponds to the secondary spectrometer.



(a) Schematic representation (from `www.ill.eu`) **(b)** Photo of the instrument

Figure 3.4: IN5 time-of-flight spectrometer. (a) Scheme of the instrument and its different components. (b) Photography on the top of IN5, where we mount the sample.

After being scattered, the neutrons will be counted as a function of their deviation angle (multi-array detectors), and their time-of-flight τ will be measured at this moment. As the distance L between the sample and the detectors is known, we can calculate the velocity v of the neutrons, and by that going back to their wavelength, following:

$$\lambda = \frac{h}{mv} = \frac{h}{m} \cdot \frac{\tau}{L}. \tag{3.3}$$

From Planck's law we can then retrieve the energy of the neutron, and so the energy transfer $\hbar\omega$. In time-of-flight spectrometers, we thus access in once the full spectrum $S(\vec{Q}, \omega)$. Also, compared to other spectrometers, a high dynamical range is accessible, over the order of the meV. As to be noted, there is a dependence between the accessible

Q-range and the dynamical range, the latter being shifted as a function of Q. As a consequence, in time-of-flight instruments, the maximum energy transfer is dependent on Q, so that the quasi-elastic peak can be precisely analyzed only in a certain Q-range.

3.1.5 Backscattering spectrometers

In backscattering spectrometers, the final energy to detect is fixed, and the other neutrons are not counted. As a consequence, we need to scan for several energy transfers to get the full spectra $S(\vec{Q}, \omega)$, contrary to the time-of-flight measurements where it is recorded in once. The advantage of backscattering spectrometers lies in their high-energy transfer resolution. Indeed, if we differentiate Bragg's law $\lambda = 2d\sin(\theta)$, we can access to the energy transfer resolution as:

$$\frac{\Delta\Delta E}{\Delta E} = \frac{2\Delta\lambda}{\lambda} = 2\cot\theta\Delta\theta + \frac{2\Delta d}{d}. \tag{3.4}$$

What we notice is that this resolution is optimized for the backscattering geometry, at $\theta = 90°$. The incident energy E_0 of the neutrons will then be changed by using a 90° Bragg reflection. There are two methods:

- Doppler drivers: the extra energy ΔE will be given by the Doppler motion of the monochromator. It is used in IN16b spectrometer, where high quality silicon crystals are mounted on a Doppler driver.
- Monochromator lattice: modification of the lattice parameter of the monochromator, for example by heating or cooling, will select the neutrons with a specific energy $E_0 + \Delta E$. This option is exploited on IN13 instrument, with monochromators and analyzers composed of fluorine crystals.

We see that choppers are also necessary here to have a monochromatic beam of fixed energy E_0 before the Bragg reflection. After the reflection, the neutrons are deflected until the sample, where scattering occurs. The scattered neutrons travel towards analyzers whose lattice and composition are the same as the monochromator. When arriving to the analyzers, only neutrons of energy E_0 (so the neutrons having lost the energy ΔE after scattering) are reflected until the detectors.

In the case of backscattering spectrometers, the so-called primary spectrometer will include the choppers and the monochromator, while the secondary spectrometer consists of the sample, analyzers and detectors.

IN13 thermal neutron backscattering spectrometer

The IN13 instrument [170], shown in Figure 3.5, allows accessing a high-Q range, until 5.5 $Å^{-1}$, thanks to a low incoming wavelength of 2.23 Å and a covering of small

deviation angles. A CaF_2 monochromator is used to backscatter the incident neutrons, and the control of the energy transfer ΔE is tuned with temperature. This control can be modeled by the following Bragg equation depending on temperature T [111]:

$$\lambda = 2d_{422}\Big(1 + \beta_1 T + \frac{\beta_2 T^2}{2}\Big)\sin(\theta), \tag{3.5}$$

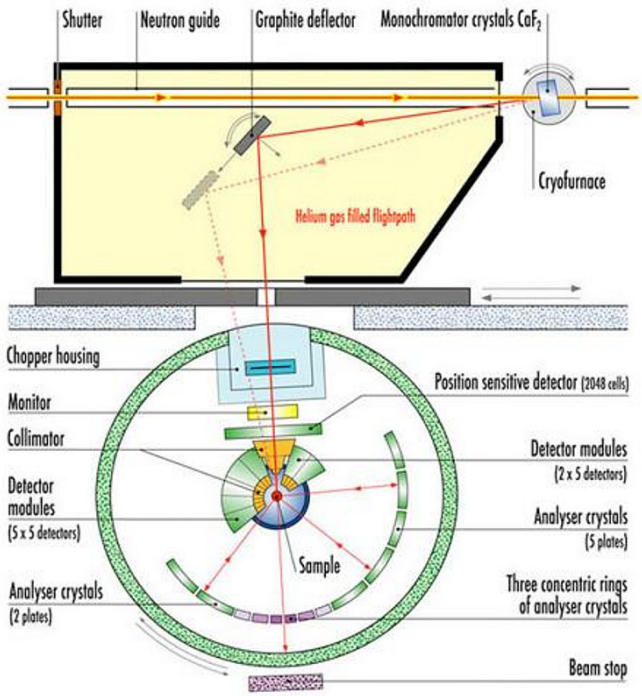
with d_{422} the spacing of the lattice at equilibrium, and $\beta_1 = 1.9406 \times 10^{-5}$ K^{-1} and $\beta_2 = 1.8744 \times 10^{-8}$ K^{-2} the expansion coefficients of the fluorine crystals CaF_2.

In practice, however, due to the low neutron flux on IN13, together with the necessity to measure for long times to have good statistics, the monochromator is often used at its equilibrium, so without energy exchange. That is why EINS experiments are commonly carried out on IN13.

IN16b high flux backscattering spectrometer

An improved neutron guide, and a specific Phase Space Transformation (PST) chopper enables the backscattering spectrometer IN16b, depicted in Figure 3.6, to reach one of the highest fluxes of neutrons at ILL. Combined to a sub-μeV resolution, it enables to access the ns time scales for dynamics, which makes it well-suited for investigating slow diffusional motions [171]. The monochromator is constituted by high quality Si(111), Si(311) or GaAs(200) crystals, mounted on a Doppler driver, which can take different velocity profiles depending on the measurements.
The new time-of-flight option BATS [172] extends the dynamical range of ± 31 μeV until ± 150 μeV, for an energy resolution which remains quite high, around 1.5 μeV. In the latter case, the energy window is centered, but it can be shifted to access even higher energy exchanges. The BATS option relies on a set of four high-speed choppers, containing multiple slits, which will structure the neutron beam similarly as in time-of-flight instruments, like IN5. In that case, monochromator is not used, as displayed in the scheme Figure 3.6c. The flexibility of the four choppers allows to tune the energy resolution, like on IN5, and allows to choose it from 1.5 to 8 μeV.

Overall IN16b instrument covers a high range of possible measurements, and is now adapted for several time scales, with a larger dynamical range. It complements the time-of-flight spectrometers well, which have a huge dynamical range, but lower energy resolution, and neutron spin-echo instruments (not covered here) whose accessible time scales range from the ns to the μs.



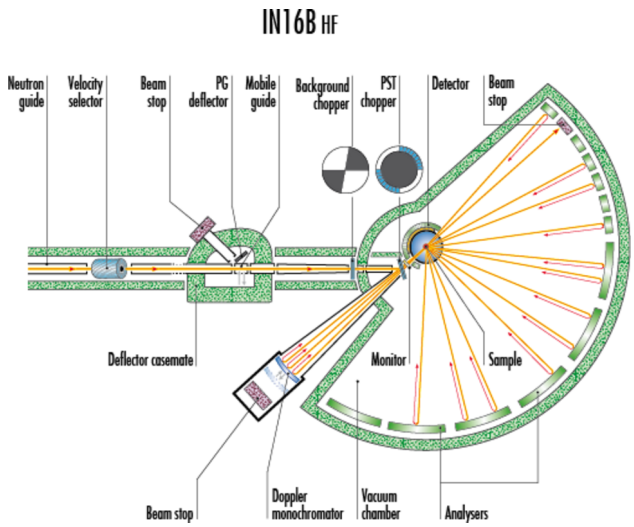
(a) Scheme

(b) Photo

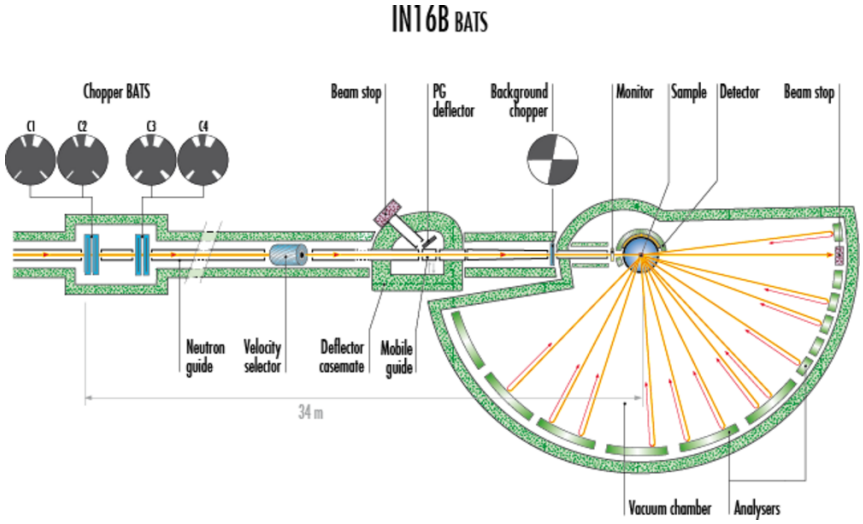
(c) Photo of the inside

Figure 3.5: IN13 backscattering spectrometer. Images taken from `www.ill.eu`. (a) Scheme of the instrument and its different components. (b) Photography of the whole spectrometer. (c) Photography of the inside; the detectors are visible on the left, and parts of the CaF_2 analyzers array on the right.

(a) Photo of the instrument



(b) Scheme of IN16b Si(111)

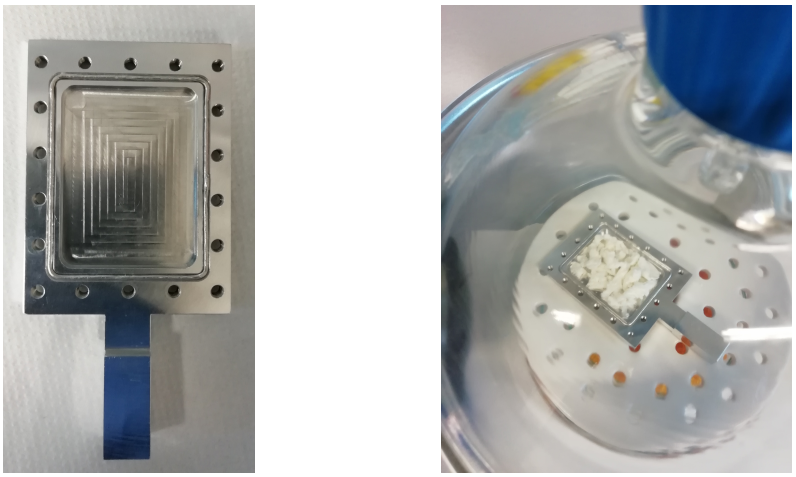


(c) Scheme of BATS configuration

Figure 3.6: IN16b backscattering spectrometer. (a) Photography of the instrument, where part of the neutron guide is visible on the left, the Doppler driver is in the middle, and in white the secondary spectrometer is shielded. Schemes of (b) the classical Si(111) configuration of IN16b and (c) the time-of-flight configuration IN16b BATS. (Taken from [164]).

3.1.6 Sample preparation

Because the neutron beam has commonly a size of around the cm^2, and to optimize the statistics with a huge quantity of sample (> 100 mg), sample holders have usually the dimensions of the cm^2. They can be found both flat for solid or powder samples, or cylindrical for liquids. In incoherent neutron scattering, to limit the contribution of the cell, Aluminum sample holders are used, as the neutron incoherent cross-section for Al is quite small (see in Table 2.2).
During the thesis, flat Al sample holders, as shown in Figure 3.7a, were used as the sample was in a powder form. Its thickness can be adjusted with several types of lids, from 0.1 mm to 3 mm.



(a) Sample holder

(b) In a desiccator

Figure 3.7: Photography of (a) an Aluminum sample holder commonly used in EINS ans QENS, and (b) a sample holder with a protein powder in a desiccator, to control its hydration.

When dealing with powder sample, hydration is necessary to study its dynamics. Usually samples are hydrated until a value of 0.4 g D_2O / g protein, corresponding to one layer of D_2O [173]. Here, heavy water is employed to limit the signal from the buffer, as the cross-section for Deuterium is much smaller, as depicted in Table 2.2. Hydration is usually done in a desiccator, which is displayed in Figure 3.7b.

3.1.7 Data correction and normalization

Before analysis and comparison of the experimental data to the models presented in Chapter 2, it is necessary to correct the data for the contribution of the empty cell or the buffer, and to normalize it to the contribution of the instrument through measurements of Vanadium, whose scattering signal is mainly incoherent and elastic in the Q-range investigated.

Bée [111] introduced an expression for the correction of the raw data. The limit of the experiments is that we never access the intensity of the pure sample, which we call I^{s}, but we always measure both the sample and the cell, whose intensity is noted I^{s+c}_{s+c}. However, we can not simply say that this intensity is the sum of the pure sample I^{s} and the pure empty cell $I^{\mathrm{c}}_{\mathrm{c}}$, because the intensity of the sample in presence of the empty cell, or reversely the empty cell in presence of the sample, is different from the pure case. Introducing attenuation coefficients for each case, Bée [111] derived, however, an expression to retrieve the signal of interest I^{s} as:

$$I^{\mathrm{s}} = \frac{1}{A^{\mathrm{s}}_{\mathrm{s+c}}}\left[I^{\mathrm{s+c}}_{\mathrm{s+c}} - \frac{A^{\mathrm{c}}_{\mathrm{s+c}}}{A^{\mathrm{c}}_{\mathrm{c}}} I^{\mathrm{c}}_{\mathrm{c}}\right], \tag{3.6}$$

where $A^{\mathrm{s}}_{\mathrm{s+c}}$ is the attenuation coefficient of the sample in presence of the cell, $A^{\mathrm{c}}_{\mathrm{s+c}}$ the attenuation coefficient of the cell in presence of the sample, and $A^{\mathrm{c}}_{\mathrm{c}}$ the attenuation coefficient of the pure empty cell. Attenuation coefficients are accessible through transmission T measurements ($A = 1/T$). Among all the instruments presented earlier, only IN13 can directly measure the transmission of a given sample. On other instruments, approached transmission coefficients can be computed from the tabulated cross-sections of each element constituting the sample.

In practice, within the terms introduced in Equation 3.6, we can access $A^{\mathrm{c}}_{\mathrm{c}}$ and $I^{\mathrm{c}}_{\mathrm{c}}$ by measurements of the pure empty cell. $I^{\mathrm{s+c}}_{\mathrm{s+c}}$ constitute the measurements of the sample in the cell. Distinction of the attenuation coefficients $A^{\mathrm{s}}_{\mathrm{s+c}}$ and $A^{\mathrm{c}}_{\mathrm{s+c}}$ from the measured value $A^{\mathrm{s+c}}_{\mathrm{s+c}}$ is not feasible, and we will commonly write $A^{\mathrm{s}}_{\mathrm{s+c}} = A^{\mathrm{c}}_{\mathrm{s+c}} = A^{\mathrm{s+c}}_{\mathrm{s+c}}$, so that at the end we have:

$$I^{\mathrm{s}} = T^{\mathrm{s+c}}_{\mathrm{s+c}} \cdot I^{\mathrm{s+c}}_{\mathrm{s+c}} - T^{\mathrm{c}}_{\mathrm{c}} \cdot I^{\mathrm{c}}_{\mathrm{c}}, \tag{3.7}$$

with $T^{\mathrm{s+c}}_{\mathrm{s+c}}$ and $T^{\mathrm{c}}_{\mathrm{c}}$ the respective transmission coefficients of the sample in the cell, and the empty cell. Then normalization by the Vanadium is done by dividing the corrected intensity by the Vanadium signal and its corresponding transmission coefficient. In addition, the effect of the geometry of the empty cell, especially for flat sample holders, and its orientation compared to the neutron beam can also be corrected using Paalman-Pings equations [174].

Usually the data correction and normalization is done using one of the following software packages:

- The Large Array Manipulation Program from ILL (LAMP) [175], which is based on IDL language. IN13 raw data were handled using LAMP.
- The Manipulation and Analysis Toolkit for Instrumentation Data (Mantid) [176], which will soon replace LAMP for all data reduction. Its core is based on C++

language, and can be used through a dedicated Graphical User Interface (Mantid Workbench) and Python scripts [177]. All quasi-elastic data from IN16b, IB16b BATS and IN5 were corrected through Mantid.

3.1.8 Future developments in incoherent neutron scattering

Thorough the years, the increase of neutrons flux, associated to the constant upgrade of the instruments, have expanded the understanding of biological systems, for both their structure and dynamics, and are now allowing more complex and realistic samples [178].

The sample preparation remains obviously a challenge in itself, where milligrams of sample are required for dynamical investigation, especially for native samples, or for the production of partially or fully deuterated components, as reviewed in [179].

Then, dynamical studies naturally call for time-resolved experiments. If it is sometimes applied to SANS studies, as mentioned for example in [180, 181], it is still scarce in incoherent neutron scattering. The main reason is the long measurement time required for EINS and QENS, compared to SANS. First time-resolved QENS experiments are reported for light-sensitive systems [182–184], which make use of pump-probe approach, with a laser exciting the sample, precisely followed by neutron measurements. In this configuration, kinetics in the range of μs to ms is accessible, and requires a high repetition of cycles to acquire enough statistics. Other strategies make use of slower processes, occurring in minutes to hours, like protein crystallization [185]. Improvements in the neutron flux, instrumentation, or modelling, associated to the rise of time-resolved studies in many complementary techniques, like X-rays / XFEL scattering or electron microscopy, will certainly put forward time-resolved experiments in EINS and QENS.

Finally, the analysis of incoherent neutron scattering data continuously requires the development of new models. First, more accurate descriptions of the measurements are needed, which is addressed as mentioned earlier by new theoretical frameworks, and the consideration of quantum effects [136]. Then, novel approaches become essential to investigate more complex samples, as well as native specimen, which are composed of many components, all containing Hydrogens, that must be distinguished.

3.2 Electron microscopy

3.2.1 Sample preparation

Electron microscopy is one of the techniques requiring the less amount of sample, around 4 µL, for an average protein concentration around 1 - 10 mg/mL. The main challenge of sample preparation is to find a strategy to image a liquid solution, while remaining in the high-vacuum imposed by working with electrons.

To perform that, grids of ∼3 mm diameter are used, on which the sample is deposited. Excess water is then removed by « blotting » with paper filter, and the sample remains on the grid by capillarity. The thickness of a sample layer, estimated recently around 70-100 nm [186], allows the electron beam to interact with the sample, while limiting multiple scattering, absorption effects or inelastic events.

The type of grid depends on the sample preparation method (negative-staining or cryo-EM, detailed below). Negative-staining EM makes use of copper grids, associated to a continuous carbon film, whereas in cryo-EM, it is a holey carbon film, as shown in Figure 3.8b. Cryo-EM employs a great variety of grids to optimize the sample preparation and acquisition. We can mention the use of carbon-coated grids, alternatives to the copper like gold or rhodium, or Lacey grids, where the distribution of hole sizes in the film is not periodic anymore, but irregular.

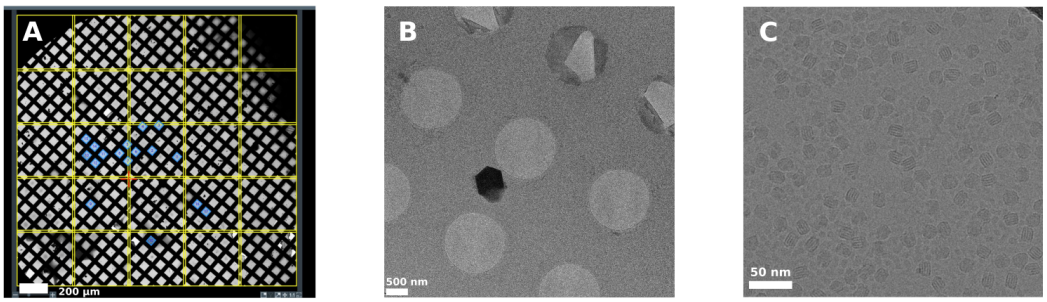


Figure 3.8: Different levels of a cryo-EM grid. (A) Atlas of the whole grid, where we see the different squares. (B) Zoom inside a square, where empty holes are visible. (C) Data collection inside a hole, where the particles are standing.

There are two distinct methods for preparing the sample in single-particle EM, which require different type of microscopes or additional devices.

Staining

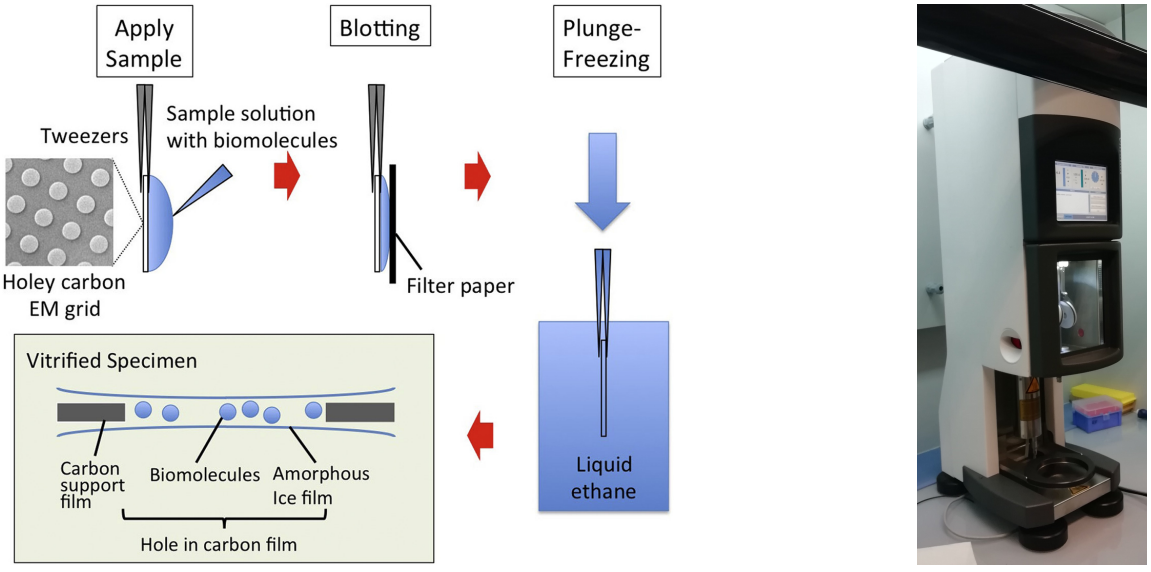
In negative-staining EM, a staining agent, like uranyl acetate (UAc), is added to fix the sample, which then enhances its contrast. This method allows for very good signal-over-noise ratio, but is limited by several factors. First, the resolution will be

limited by the finite size of the staining agent, to around 15 Å [187]. Then, the agent can deform the sample. Proteins were shown to be flattened in the process, as reviewed in several examples in Frank [137].

However, negative-staining EM remains a method of choice for retrieving first images of new samples, to test the conditions of preparation, and already gives numerous information about the sample. It is the first step of the pipeline before going to more advanced techniques.

Vitrification

The works of Dubochet et al. [188], McDowall et al. [189] isolated a specific state of water; the vitrified ice, whose amorphous properties preserves the sample in a near-native environment. Contrary to crystalline ice, the absorption of electrons is limited, so that the phase-contrast image is visible. To reach this state and avoid ice formation, the vitrification is usually done in an automatic way by plunge-freezing in liquid ethane, as reviewed in [144] and schematized in Figure 3.9a. Figure 3.9b displays the photography of a Vitrobot system, which allows to control multiple parameters for the vitrification, like the hydration and temperature prior to the plunging, or the blot time and force.



(a) Vitrification scheme **(b)** Vitrobot

Figure 3.9: Vitrification process represented in (a) by a scheme (reproduced from Murata and Wolf [144]) and (b) by a photography of the Vitrobot from the Institut Europeen de Chimie et Biologie (IECB, Bordeaux, France).

In addition, to overcome the previous limits of staining, the low temperature imposed by this preparation limits the thermal motions, thus improving the quality of the image, but also restricts radiation damage [190]. The main counterpart is a low

signal-over-noise ratio, as the electron cross-section for vitrified ice is on the same order of magnitude of Carbon or other elements of the sample, around 1 Mb, as computed by Angert et al. [191].

Figure 3.10 displays a comparison between images of LDLs taken from negative-staining EM and cryo-EM. With staining-EM, the internal layers of the LDL are not accessible, contrary to cryo-EM where they are clearly visible. However as precised earlier, the signal-over-noise ratio appears rather weak when the sample is embedded in vitreous ice, which will require averaging methods to process the images.

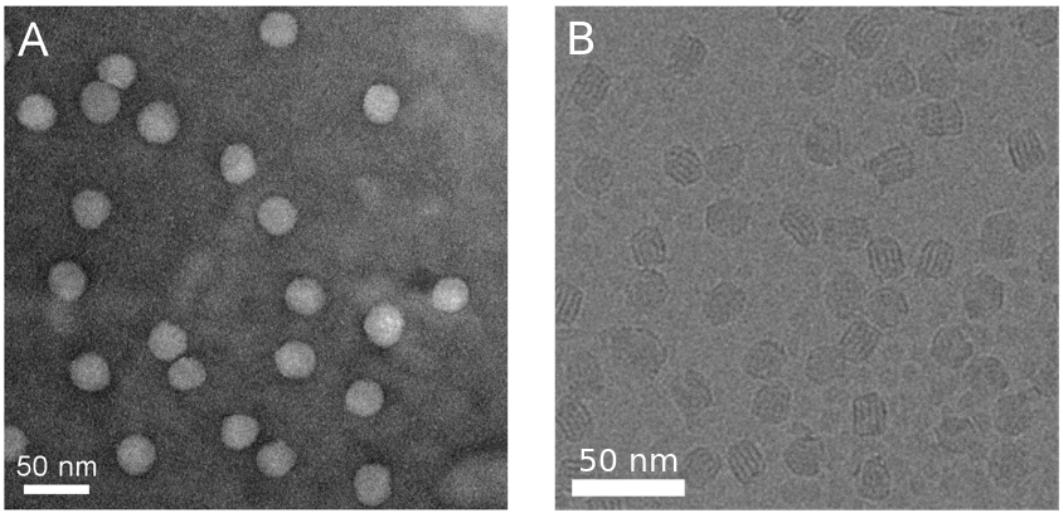


Figure 3.10: Comparison between LDLs screened in (A) negative-staining EM and (B) cryo-EM.

3.2.2 Main components of an electron microscope

Since the construction of the first electron microscope by Max Knoll and Ernst Ruska in 1931 [192], its primary frame, as illustrated in Figure 3.11, remains the same. However, across the last years, exceptional improvements of the EM maps were reported, which we often refer by the «Resolution Revolution». Behind this revolution is the introduction of direct electron detectors, which enhanced both the quality of recorded images and their quantity, as well as concomitant developments in image processing algorithms.

Electron source

Electrons are produced through *electron guns* submitted to a high-voltage (100 - 300 keV) [138]. The first electron guns relied on thermionic emission. The cathode, a

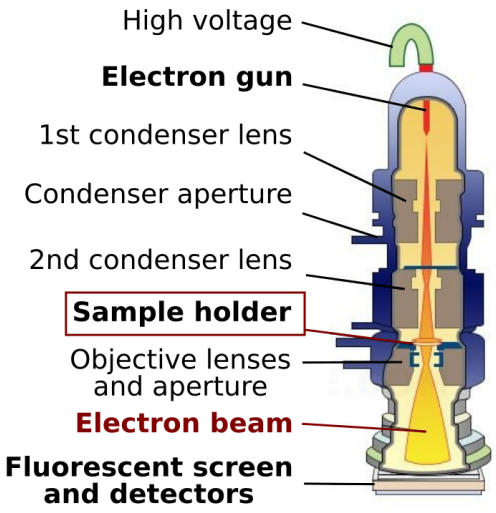


Figure 3.11: Scheme of an electron microscope.

Tungsten filament, or a LaB_6 crystal tip, is heated, leading to the release of high-energy electrons, which are then focused by a Wehnelt anode. The form of this anode determines the brightness and other properties of the electron beam.

Recently, the brightness was increased by several order of magnitude by field-emission guns (FEG). They also consist in a pointed tip, as seen in Figure 3.12. Here the curvature radius is so small (less than 0.1 µm), that the generated electric field is much higher, allowing electron emission at cold temperatures, through quantum tunnelling effect.

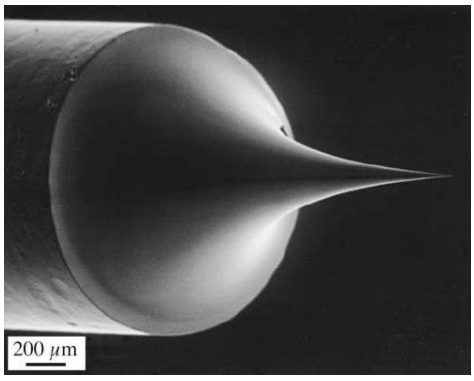


Figure 3.12: Image of a FEG tip (reproduced from [143]).

Electromagnetic lenses

The idea behind the manipulation of an electron beam is the same as for light, with the use of lenses. In the case of electrons however, the deflection is directly induced

by electromagnetic fields through the Lorentz force [138]:

$$\vec{F} = -e(\vec{E} + \vec{v} \times \vec{B}), \tag{3.8}$$

with e the charge of an electron, $\vec{E}$ the electric field, $\vec{v}$ the velocity of the electron and $\vec{B}$ the magnetic field. The purpose of the electromagnetic lenses is to produce and finely tune $\vec{E}$ and $\vec{B}$. Properties of the lenses, for example the magnification, are controlled by changing their current.

There are two main systems for lenses in an electron microscope, similar to a light microscope [138, 193]. The condenser lenses, or illumination system, focus the electron beam to the sample, and control the spot size and intensity. Then, generally placed after the sample holder, the objective lenses (or imaging system) will contribute to the image formation on the detector, and to the magnification.

As for light, electromagnetic lenses are also subjected to aberrations [138, 193]:

- Astigmatism can be observed if the lenses are not perfectly symmetrical.
- Spherical aberrations: electrons passing at the center of the lenses, or near the coil are deflected in different ways, which will end up with different focal planes. By placing an aperture before the focal point, the number of different focal planes can be limited, and so effects on the image are reduced. Spherical aberrations are quantified through the spherical aberration coefficient C_s, that lies around 2.7 mm for the electron microscopes used during the thesis.
- Chromatic aberrations: when there is inelastic scattering, the exit electron beam can contain electrons of different wavelengths, which will be deflected on different ways and can deform the image. Analogy with light is when objects in a photography are contoured with small rainbow-like motifs. In an electron microscope, a way to correct for these aberrations is to remove the inelastically scattered electrons with energy-filters, keeping only a pure elastic signal.

Electron detectors

It is important to note that the signal measured at the exit of an electron microscope is real, contrary to a light microscope whose image is virtual. The second remark is that the image contains all the information (amplitude and phase), in contrast to X-ray crystallography where the phase information is lost.

Historically, images from electron microscopy were recorded on photographic films, before being scanned for further processing. Introduction of Charged Coupled Devices (CCD) cameras allowed to collect a much higher quantity of data. However, CCD require to first convert the electron signal into light, through a set of scintillators

and fiber optic coupling, which can lead to loss of the signal. In addition, all these intermediate steps limit the measurement rates.

In the last decade, the establishment of Direct Electron Detectors (DED) was crucial in the passage to the Å resolution [194], and is really the basis of the «Resolution Revolution». Here the electrons can be directly counted, without any intermediate. In addition, short «movies» instead of images can be measured thanks to the high-frame rate. These movies have highlighted the importance of beam-induced motions, and have led to the development of algorithms for motion correction. They also allow to limit the effects of radiation damage in the images, by weighting differently the first and last frames.

3.2.3 Characteristics of the electron microscopes used

During the thesis, the data collection for single-particle analysis on LDLs were performed on two electron microscopes, which are both instruments from the company Thermo Fisher Scientific [195]. The tests for grid preparation and first screening were performed on the Talos Arctica of the Institut Europeen de Chimie et Biologie (IECB, Bordeaux, France), shown in Figure 3.13a. Subsequent data collection was carried out on the Titan Krios from the CM01 beamline in the European Synchroton Radiation Facility (ESRF, Grenoble, France), reviewed in Kandiah et al. [196] and displayed in Figure 3.13b. Characteristics of the different instruments are summarized in Table 3.3.

(a) Talos Arctica **(b)** Titan Krios

Figure 3.13: Photography of (a) the Talos Arctica from the IECB (Bordeaux, France) and (b) the Titan Krios from the CM01 beamline in the ESRF (Grenoble, France) [196].

Instrument	Talos Arctica	Titan Krios	Titan Krios
Date data collection	28.02.2019	31.01.2020	12.04.2021
Voltage (kV)	200	300	300
Magnification	120k	130k	130k
Pixel size (Å/pix)	1.24	1.053	1.053
Dose per frame (e^-/Å.s)	2.66	0.99	1.01
Number of frames	20	30	30
Total dose (e^-/Å.s)	53.2	29.76	30.32
Spherical aberration (mm)	2.7	2.7	2.7
Detector	Falcon 3	Gatan K2 Summit	

Table 3.3: Characteristics of the electron microscopes on which LDL data were collected.

3.2.4 Image processing pipeline

Nowadays, all the mentioned improvements in the instruments lead to a high throughput in data collection, which commonly includes thousands of micrograph movies. The image processing is performed in numerous software packages. The first of their type include: EMAN [151], FREALIGN [152], IMAGIC [197], SPIDER [198] or XMIPP [199].

In the following, we will present the pipeline we follow for image analysis, with a focus on the two software packages we employed; RELION [150, 153] and cryoSPARC [161]. This pipeline is also represented in Figure 3.14, with a distinction on the procedures running on the whole movies or micrographs (in orange), and those on the individual particles in blue and green.

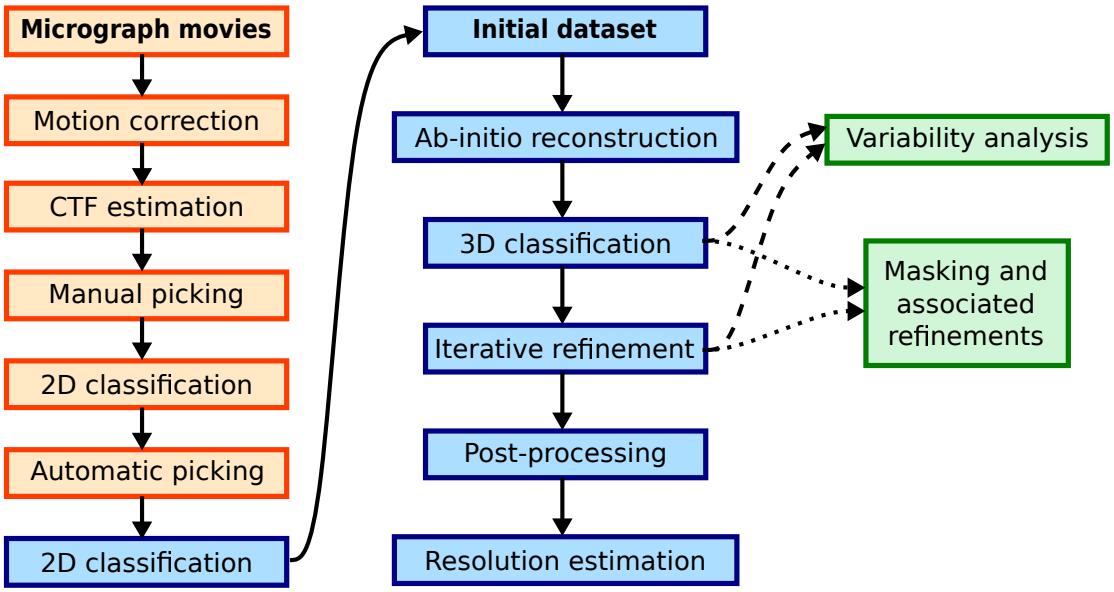


Figure 3.14: Representation of the image processing pipeline. In orange, the whole movies or micrographs are treated, whereas in blue and green it is the individual particles. (Adapted from [200]).

3.2.5 Motion correction

The recording of short movies instead of images enables to correct for beam-induced motions, as can be reviewed in [201, 202]. Algorithms like MOTIONCORR, and the more recent MotionCor2 [202, 203], compute cross-correlation between all the frames to derive motion paths. Then, the frames are all aligned, which improves the quality of the image, as can be shown in Figure 3.15.

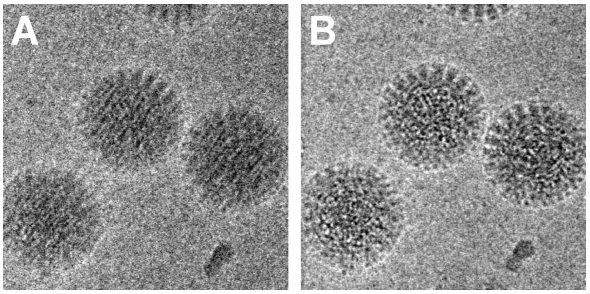


Figure 3.15: Examples of images (A) before and (B) after motion correction. Image reproduced from [201].

3.2.6 CTF estimation

The CTF (Contrast Transfer Function) depends on both the general characteristic of the microscope, like the acceleration voltage or the coefficient of spherical aberrations, but is also related to the defocus applied to each image. Indeed, to access the phase-contrast of an image, the electron beam is never set in the focal plane, but at underfocus, to values corresponding to information at the Å-nm scale. However, the

interferences leading to the formation of an image present also destructive interferences at some frequencies of interest. It is then necessary to measure at several values of defocus to access a broad range of frequencies, so space scales. When collecting data, a first estimation of the defocus is given by the instrument, but a more precise estimation is required for each individual micrograph.

The CTF estimation is performed on the Fourier space, by fitting the so-called Thon rings of the image by the following model [138]:

$$\hat{C}(q,z) = 2\sin\Big(\chi(q) - \pi\lambda z q^2\Big), \tag{3.9}$$

$$\text{with } \chi(q) = \frac{2\pi}{\lambda}\Big(\frac{1}{4}C_s\lambda^4 q^4 + \frac{1}{2}\Delta f\lambda^2 q^2\Big), \tag{3.10}$$

where λ is the electron wavelength, z is the position on the column axis, C_s is the spherical aberration coefficient (set to 2.7 mm for Talos Arctica or Titan Krios microscopes), Δf is the defocus at the sample position. Because we are at underfocus, $\Delta f < 0$. Finally $q = |\vec{q}|$ is the norm of the wave vector. Figure 3.16 shows an example of the fit of the Thon rings, using the Ctffind algorithm implemented in RELION [204].

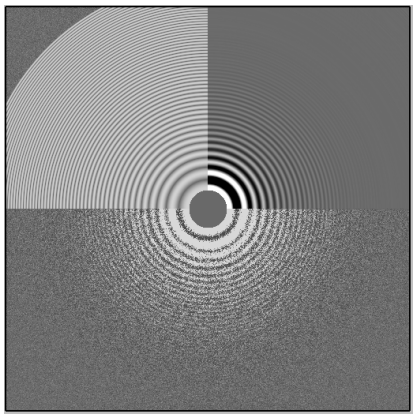


Figure 3.16: Fit of the Thon rings on a micrograph of our dataset (visualization in RELION).

The CTF estimation relies on all the approximations introduced in Chapter 2. Among all, the projection assumption and the subsequent weak-phase object approximation assume that the specimen is thin enough. In practice, however, the ice layer measures around 100 nm [186], and Koeck and Karshikoff [140] showed that already for proteins larger than 50 nm, small artifacts appear at high resolution. Zhang and Zhou [205] derived an expression for the resolution limit, and calculated that for a 300 kV microscope, and aiming at 3 Å resolution, the above approximations are valid only for particles smaller than 640 Å. For large viruses, single-particle analysis requires to take into account the defocus gradient within the sample, otherwise it remains limited to a certain resolution. In the case of LDLs, the diameter lies around 250 Å, so that the above approximations still remain valid. In the case of tomography,

where samples are much thicker, and the sample grid is tilted, the CTF estimation is supported by different frameworks, as presented in [206, 207].

As mentioned in Chapter 2, the CTF is not directly deconvoluted because of its zero values. Ultimately, its correction is done through Wiener filtering during the 3D reconstruction, as expressed in Equation 2.117.

3.2.7 Particle picking

Once the micrographs are corrected overall, the particles of interest, in our case the LDL, need to be extracted in boxes to be further processed. This procedure, denoted as the particle picking, is generally performed in several steps.

Commonly, the first step is to carry out a manual picking of thousands of particles, in different micrographs, with various orientations and defocus. The goal is to create templates for a subsequent automatic picking. The templates are computed by 2D classification, where close orientations are averaged, in order to increase the signal-over-noise ratio (detailed in the next section). Automatic picking remains a challenging task as it should detect a signal within a noisy background, while avoiding false positives and artifacts. For instance, ice crystals in a micrograph appear in black in the micrographs (as seen for example in Figure 3.17), and have a huge signal, compared to the particles we want to select. Numerous algorithms were proposed through the years. For instance, Roseman [208] used local correlation functions to find peaks and localize particles, with a Fast Fourier transforms (FFT) implementation to speed up calculations. Scheres [209] pursued the use of FFT, but with a probability-based approach. Recently, the rise of deep-learning methods using neural networks also applied to cryo-EM, as reviewed in [210], especially for the particle picking, which remains a classical problem of these techniques. Among all the algorithms proposed, we can cite the recent Topaz from Bepler et al. [211], that also propose micrograph denoising, or the new DeepPicker job from cryoSPARC [212].

In Figure 3.17, we notice that false positive (black ice crystals) are picked. That is why inspection of the micrographs is essential, and constitutes the next step of a particle picking job. Because there are thousands of micrographs, it is done in a semi-automatic way, by manually tuning parameters like the power of the signal or noise we want to include, then checking the resulting picking on some micrographs, and repeating the process if needed.

At the end, each particle is extracted from the micrographs in a square box, as shown in Figure 3.18, and they will be processed in that form all along the next steps. Once extracted, we will not refer to the term image anymore (which corresponds to the whole micrograph), but to particles or projections.

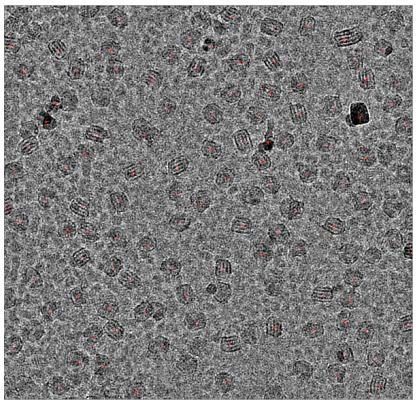


Figure 3.17: Examples of particle picking in a micrograph using the classical template picker of cryoSPARC. Red points correspond to local peaks, where particles are localized.

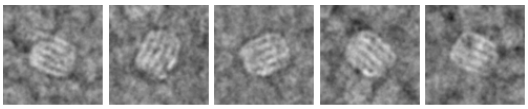


Figure 3.18: Examples of LDLs extracted in 400x400 pixels boxes through cryoSPARC. The images are lowpass filtered for better visualization.

3.2.8 2D classification

As introduced earlier, 2D classification is a step whose purpose is to classify the different particles, based on their orientation, in order to improve the signal-over-noise ratio. It relies on the previous assumption, introduced in Equation 2.116, of an additive Gaussian noise, whose average tends then towards zero.

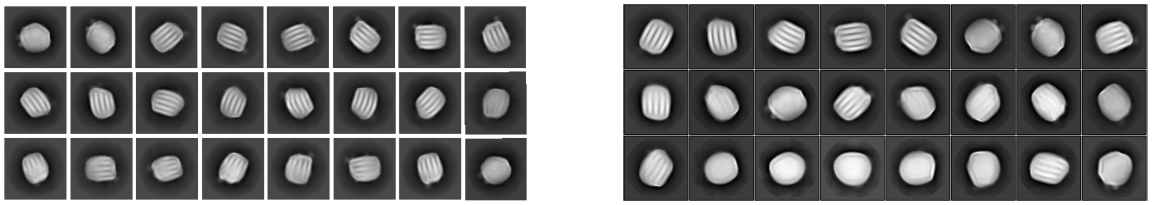
The 2D classification is an iterative process, and first requires to align the particles:

- They are centered by translation in the x and y direction;
- They are rotated so that projections of same Euler angles ϕ and θ (see scheme in Figure 2.16) can be gathered. The rotation consists in looking for a common in-plane angle ψ.

First methods for the classification of each particle in a 2D class used maximum cross-correlation approaches, as mentioned in Sigworth et al. [155]. In that approach, one particle is assigned to one unique class, the one for which the cross-correlation is the highest. In the Bayesian framework, as introduced in XMIPP package [159, 160], and then in RELION software [150, 153], one particle is not assigned to one unique 2D class, but probabilities for each particle to belong to each 2D class are computed. A 2D class is then the result of a probability-weighted average of all particles. As reviewed in Sigworth et al. [155], Bayesian approaches have great advantages when the data are noisy, as it is the case in cryo-EM. Using probabilities, and not a unique

assignment, improves the convergence, by limiting the contribution of false positives or wrong assignment.

The enhancement of the signal-over-noise ratio after 2D classification can be observed by comparing a selection of 2D classes of LDLs in Figure 3.19 and previous picked particles in Figure 3.18. In the 2D classes, the background noise was highly reduced, and the signal from the LDL particle is much better. In addition, 2D classes computed from RELION or cryoSPARC were nearly identical in our case, as can be observed in Figure 3.19. In general, the reproducibility of the 2D classification across the software is rather good.



(a) cryoSPARC 2D classes **(b)** RELION 2D classes

Figure 3.19: Examples of 2D classes of LDL particles in (a) cryoSPARC (400×400 pixels box) and (b) RELION (352×352 pixels box). Within the circles, the particles are normalized to zero mean and unity variance [200].

Several rounds of 2D classification are commonly performed to clean the dataset, but also to extract interesting orientations. From this initial dataset, 3D reconstruction is processed using different algorithms, and various strategies can be followed depending on the software used, the sample we are treated, and the questions we are trying to answer.

3.2.9 Ab-initio reconstruction

The determination of a starting 3D model from the 2D projections remains a critical step in image processing, because it will be at the basis of all the iterative refinements. The initial structure can be either obtained from other measurements (tilt-pairs, homology modeling...), either reconstructed *de novo*.

In the case of LDL reconstruction, previous EM maps were available, but several reasons pushed us to favor the *de novo* reconstruction; first the differences in the LDL batches, then in the instruments and algorithms used for reconstruction, so direct comparison was more difficult. For *de novo* reconstruction, both RELION and cryoSPARC use a Stochastic Gradient Descent (SGD) as introduced by Punjani et al. [161], through cryoSPARC. The advantage of this algorithm is its convergence for any random initialization, in addition to be less sensitive to local maxima. Once an ab-initio model is obtained, it needs to be further refined to improve its resolution.

3.2.10 Iterative refinement and resolution estimation

As introduced in Chapter 2, the iterative refinement from an initial structure is based on Bayesian approaches and can be summarized in the Expectation-Maximization algorithm. In the ideal case, refinement needs few or no user intervention, and its convergence should depend scarcely on the initial model or parameters used for refinement.

In addition, due to the high amount of noise in cryo-EM dataset, refinement should avoid overfitting, that is to say, to build on noise. A way to prevent it is contained in the so-called Gold-standard Fourier shell correlation (FSC) procedure [213, 214]. In that framework, the dataset is randomly split in two subsets, which are independently refined. Then, a FSC is computed as a function of the frequency, as described in Rosenthal and Henderson [215], and represented in Figure 3.20. At low frequencies (so low resolution), the half-maps are similar, so that the FSC is high, and close to 1. Then by going to higher frequencies (so higher resolutions), details of each map begin to differ, and the FSC decreases. Ultimately, at the very high resolutions, there is only random noise, so the FSC tends towards zero. Cryo-EM reconstructions are then considered « reliable » and interpretable until a certain limit, which is determined by the 0.143 FSC criterion [215]. This limit represents the overall resolution of the map, and should coincide with the structural details seen on the reconstruction.

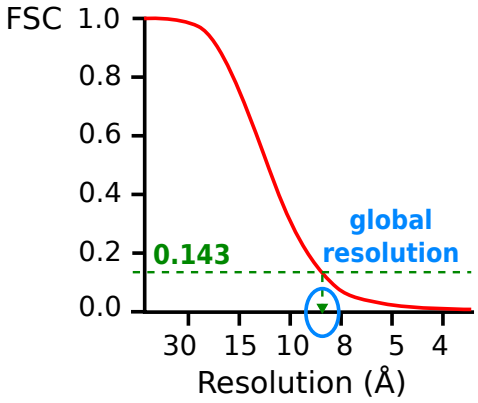


Figure 3.20: Representation of the FSC as a function of the resolution.

Implementation of the iterative refinement and gold-standard procedure is different from one software to another. RELION uses an adaptive Expectation-Maximization algorithm [153], where the Euler angles are first searched with a coarse sampling. Then, the search domain reduces, and the sampling becomes finer and finer over the different iterations. It implies that at each iteration, the new determined Euler angles are close to the true ones. As a consequence, the main drawback is to get stuck in a local maximum.

In contrast, to optimize the Expectation step, cryoSPARC uses a branch-and-bound algorithm [161], which turns out to be a global search on the Euler angles, and not

a local one as for RELION. It limits then the probability of being stuck in a local maximum. In addition, this approach speeds up the calculations, allowing refinements in 1 or 2 hours, compared to around 20 hours for RELION, on the same machine [161]. However, compared to RELION, cryoSPARC tends to over-estimate the resolution, notably through a harsh removing of the solvent contribution. Commonly, post-processing jobs of RELION are used on the resulting reconstructions of cryoSPARC. It enables to end up with a good estimate of the resolution, as well as consistent structural features.

During the thesis, we began by processing the LDL data with RELION, before switching to cryoSPARC for the following reasons:

- New algorithms were implemented in cryoSPARC during the time of the thesis, like variability analysis, which we will describe later on.
- Time considerations were also decisive. As our dataset became more populated ($\sim$700k particles), the tests and calculations were much faster with cryoSPARC.
- Possible local maxima could be avoided, especially regarding the strong signal of the lipid layers.

Previous considerations on the possible over-estimation of the resolution were, however, always in mind.

The question of the resolution of a 3D reconstruction is central in cryo-EM, and still subject to strong debates, from the 0.143 threshold in the FSC proposed by Rosenthal and Henderson [215], to new ways of calculating the FSC, as suggested by Penczek [216]. In between, Cardone et al. [217], and then Kucukelbir et al. [218] with the ResMap algorithm, introduced local resolution estimation for a better interpretation of the maps. For instance, it is often encountered that regions with lower resolution present enhanced flexibility.

3.2.11 Handling heterogeneity : from 3D classification to variability analysis

If the sample is not homogeneous, in the sense that it can take several conformations for example, the 3D reconstruction can be blurred by averaging all the different conformations. A first approach to handle heterogeneity is to perform 3D classification, as reviewed in Scheres [200]. In that case, it is not just one map that is iteratively refined, but K volumes, with K an user-defined parameter. As an example, application of extensive 3D classification on a spliceosomal complex separated open and close conformations [219]. Recent extensions of the 3D classification aim at reconstructing continuous variability, so not distinct conformations, which is biologically more relevant. Independent strategies were followed:

- Multi-body refinement as proposed by Nakane et al. [220] models molecular motions by splitting the particle in several regions, considered as moving from each other in an independent way.

- 3D variability analysis (3DVA) implemented in cryoSPARC [221] uses a Principal Component Analysis (PCA) approach, where all the different possible conformations are represented by a linear subspace.

- In cryoDRGN from Zhong et al. [222], deep neural networks are exploited, which will learn a 3D model from the dataset. Each particle will be embedded in a latent space, and a corresponding structure can be generated. In that way continuous heterogeneity, as well as discrete states, can be modelled.

3.2.12 Post-processing

In general, the loss of high-resolution contrast is a common rule in cryo-EM, for all the reasons mentioned previously, such as beam-induced motions, radiation damage or inelastic scattering. In addition, the transfer function of the detector needs to be taken into account. All these contributions to the loss of high-resolution can be compiled in a Gaussian blurring on the form $e^{B/(4d^2)}$, with d the resolution and B the so-called B-factor [215]. This form is analog to the Debye-Waller factor, as introduced in X-rays or neutron scattering (see Equation 2.86 for example), where the damping factor represents the thermal motions of the atoms. In the case of cryo-EM, the B-factor does not only describe the effect of temperature, but by analogy B is often referred as the «temperature» factor. To improve the quality of the final map, it can be interesting to correct for the B-factor, in a process called *sharpening*. This B-factor can either be determined from a fit of a Guinier plot (logarithm of the structure factor as a function of the frequency), either be manually tuned.

In the meantime, during the iterative refinement and calculation of the FSC, the contribution of the solvent background can also lower the resolution of the final map. On the contrary, noise can have contributed to overfitting, which would appear as false artifacts in the map. With the purposes of reducing the solvent contribution, while avoiding overfitting artifacts, and end up with a reliable value of resolution, post-processing of the final maps is performed. This step consists of several stages;

- Masking of the protein (or the LDL in our case) to remove the solvent background. One should take care that the mask is large enough, and contains soft edges, to avoid artifacts in the Fourier space, and so wrong estimation of the resolution. Masking can be performed in external software like UCSF Chimera [223], or within RELION or cryoSPARC jobs.

- High-resolution noise substitution, by phase randomization, and calculation of the true FSC, as presented in Chen et al. [224]. Here, all Fourier components beyond a certain resolution, determined by the FSC computed during the iterative refinement, will be phase-randomized, so that they do not contribute to the half-maps anymore. The FSC of these phase-randomized maps evaluates the effect of the solvent masking. Theoretically, it is supposed to tend quickly towards 0 as the signal was substituted by random noise. If it is not the case, it means that the mask introduced « fake » features, that will be then corrected to give the true FSC, and more reliable resolution estimation.

All these computations are gathered under the term *post-processing*. Corresponding algorithms can be found in both RELION and cryoSPARC software, but the post-processing job of RELION was favored to compute good estimates of the resolution.

3.2.13 Masking and associated refinements

To focus on some parts of a protein, that would be for instance flexible in relation to each other, masking can be an option. Several strategies can be pursued:

- Particle subtraction and subsequent refinement. This strategy was successfully used in the studies of Nguyen et al. [219] or Nakane et al. [220], and enabled to improve the resolution of some flexible parts, while recovering different conformations of the protein of study, which can be linked to their functionality.

- Local refinement, with a finer sampling and local search of the Euler angles on the masked part, which can be useful in the case of flexible parts.

Obviously, one needs to be careful with the interpretation of the resulting maps, and to avoid possible artifacts, especially on the borders of the mask. As a consequence, a mask needs to be large enough and to be softened, i.e. the edges are not sharp, but are passing from 1 to 0 in a gradual manner. As a rule of thumbs, the number of voxels for such gradient should be at least equal to three times the estimated resolution.

RELION has several options for controlling the search extent: coarse, local, and alignment can even be skipped for 3D classification. It also implements particle subtraction, and now multi-body refinement jobs, as described in [220]. CryoSPARC includes particle subtraction and local refinement since the version 3, but some features are still in beta version.

3.2.14 Future developments in single particle cryo-EM

Since the «Resolution Revolution» and the 2017 Nobel Prize of Chemistry, cryo-EM is continuously growing and evolving, achieving 3D maps comparable to X-ray

crystallography, which enables now atomic building [225, 226]. The rise of deep neural networks in the algorithms for many steps of the image processing, from the picking, to model building, and obviously the 3D reconstruction, constitutes also a big step in image analysis, as reviewed in Si et al. [210]. Still, many challenges remain to be tackled, among them we can mention the sample preparation and handling sample heterogeneity.

First, the vitrification is a powerful technique, which really pushed the cryo-EM towards the Å resolution, however, the grid preparation suffers from several drawbacks, as can be reviewed in Carragher et al. [227], Glaeser [228]. For instance, we can cite agglomeration, denaturation at the air-water interface, or migration of the proteins to the carbon grid. These difficulties can be dealt with various methods, from changing the biochemical conditions (adding of detergents or salts), the vitrification procedure, the type of grids etc. Unfortunately, all these tests can take time, and the resources (like a Vitrobot and screening) are not available to everyone, or their access remains restricted. It limits both the resolution of some maps, and investigation of challenging proteins, that can remain stuck to the sample preparation step. Solutions to overcome these problems are continuously proposed, like in [229], or with new systems like Spotiton [230–232], which replaces the mechanical pipetting and blotting by an inkjet dispensing droplets of the sample onto the grid.

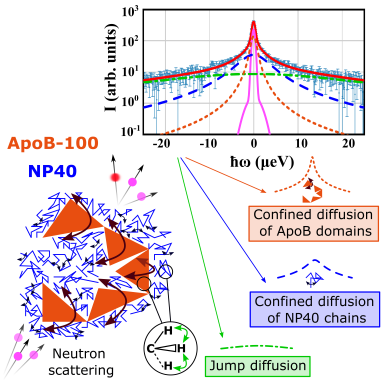
Then, the treatment of sample heterogeneity and flexibility is crucial, as mentioned among the current challenges in [233, 234], for example. In single-particle analysis, it can only be handled by computational means, as already started by various packages [221, 222]. However, for the moment these packages require the starting 3D reconstruction to be already at an excellent resolution, with visible secondary structures, and a good estimate of the orientations. This can be the case if the sample flexibility is limited to distances in the Å range, which still enables to get good averages, even if some regions are slightly blurred, while remaining distinguishable. If the flexibility is too high and continuous, the regions can be blurred so that no secondary structures can be retrieved and the whole 3D map is probably not fully optimized. We can hope that future developments of the already proposed packages, or new algorithms, will make it possible to go beyond that limit.

Chapter 4

Dynamics of apolipoprotein B-100 in interaction with detergent probed by incoherent neutron scattering

This chapter is based on the following paper :

> Aline Cisse, Anna-Laurence Schachner-Nedherer, Markus Appel, Christian Beck, Jacques Ollivier, Gerd Leitinger, Ruth Prassl, Karin Kornmueller, and Judith Peters. Dynamics of Apolipoprotein B-100 in Interaction with Detergent Probed by Incoherent Neutron Scattering. The Journal of Physical Chemistry Letters, 2021, 12, 12402–12410. `https://doi.org/10.1021/acs.jpclett.1c03141`.



Apolipoprotein B-100 (apo B-100) is the protein moiety of both low- and very-low-density lipoproteins, whose role is crucial in the cholesterol and triglyceride transport. Aiming at the molecular dynamics' details of apo B-100, scarcely studied, we performed elastic and quasi-elastic incoherent neutron scattering (EINS, QENS) experiments combining different instruments and timescales. Similar to classical membrane proteins, the solubilization results in remaining detergent, here Nonidet P-40 (NP40). Therefore, we propose a framework for QENS studies of protein – detergent complexes, with the introduction of a combined model, including the experimental apo B-100 / NP40 ratio. Relying on the simultaneous analysis of all QENS amplitudes, this approach is sensitive enough to separate both contributions. Its application identified two points : (i) apo B-100 slow dynamics, and (ii) acceleration of NP40 in presence of apo B-100. Direct translation of the exposed methodology makes now achievable the investigation of more membrane proteins by neutron spectroscopy.

4.1 Introduction

With a molecular mass of about 550 kDa, apolipoprotein B-100 (apo B-100) is the sole protein constituent of low-density lipoproteins (LDLs) and part of very-low density lipoproteins (VLDLs). LDLs and VLDLs have major roles in the triglyceride and cholesterol metabolism, acting as their transporter. Lipoproteins are natural nanoparticles with an amphiphilic shell, ensuring a clear separation with their environment. For LDLs, it is composed of a monolayer of phospholipids and apo B-100 moiety [41], whose parts are either exposed to the surface, or partially embedded in the monolayer [67]. Due to its amphiphilic nature, and from estimations of its hydrophobicity [65, 66], apo B-100 is considered a membrane-like protein, with very similar characteristics as classical membrane proteins have.

Apo B-100 mediates the conversion from the VLDL to the LDL, which involves a drastic change of particle size, from around 500 Å for a VLDL to about 250 Å for a LDL, indicating that the protein is going through significant conformational changes. Until now, apo B-100 was scarcely studied in its detergent-solubilized form. Investigations of its structure are reported using cryo-electron microscopy [29], small angle-neutron scattering [19], or tension measurements [101], as reviewed in Chapter 1, but no studies on its molecular dynamics were conducted so far. Neutron spectroscopy is a method of choice to access picosecond - nanosecond dynamics. It was applied previously on the whole LDL, in the normolipidemic or triglyceride-rich (pathological-like) forms [35, 36]. In this context, we are aiming at performing measurements on apo B-100 using neutron spectroscopy, which will complete the data published on whole LDL.

One of the main limiting factors for conducting such a study is the remaining presence of Nonidet P40 substitute (NP40) detergent used for apo B-100 solubilization, which we described previously in Chapter 1. In general, membrane(-like) proteins are stabilized by detergents in order to prevent their aggregation in an aqueous environment [235]. For that reason, their characterization often requires to take the detergent contribution into account [236]. Usually, protein - detergent complexes are examined using neutron crystallography or small angle neutron scattering. They give the possibility to focus either on the protein or detergent structure by contrast-matching techniques, like reported in the following studies [237–239], which can be enhanced by the use of deuterated detergents [240, 241].

In dynamical studies, as Deuterium has a much smaller incoherent cross-section compared to Hydrogen [118], isotopic exchange can be performed on the buffer to reduce its signal ; for example heavy water (D_2O) is commonly used for studies of proteins in solution. Deuterated detergents were therefore proposed for quasi-elastic neutron scattering (QENS) measurements on bacterial photoreaction center [242].

However, isotopic exchange also occurred in the surface exposed parts of the protein, rendering their dynamics unreachable. As a consequence, the signal came mainly from the buried parts, estimated around 85%. Moreover, deuterated detergents remain an expensive solution, and the proposed protocols are not applicable to all membrane(-like) proteins.

Recently, a QENS study on G-protein-coupled receptor was conducted by Shrestha et al. [243], whose sample contained 25% of remaining CHAPS[1] detergent. Both protein / detergent and pure detergent samples were measured. It appeared that the detergent signal was nearly identical in both cases, meaning that the CHAPS and protein dynamics were decoupled. This decoupling approach made the detergent correction straightforward, and the extraction of the protein contribution successful.

On the contrary, slaving effects are generally observed for protein - solvent systems. One of the first points concerns hydration, that was proved to play a crucial role in protein dynamics, as studied in detail for myoglobin [244, 245]. Frauenfelder et al. [246] showed that proteins can be conditioned by solvent fluctuations [246, 247]. In parallel, some solvents, like glycerol or trehalose, were shown to slow down lysozyme dynamics [248–251]. Due to the strong interactions that occur between detergents and solubilized proteins [235], coupling dynamics is then expected to be a general behavior, and should be tackled for QENS measurements.

The present study aims to provide a complete QENS study on a complex system involving protein - detergent interaction. We propose a new model to separate the different contributions, that overcomes the issues encountered when working with membrane(-like) proteins : (i) the presence of detergent in a non-negligible proportion, (ii) the possibility of coupling dynamics between the solubilized protein and its detergent. Applied to apo B-100 / NP40, this theoretical framework allows a first characterization of the lipid-free apo B-100 dynamical properties, and opens the way how membrane-like proteins can be treated in presence of detergent.

4.2 Sample preparation

Apo B-100 was prepared as described in Chapter 1. In brief, LDL was isolated from human plasma via ultracentrifugation. Then LDL was incubated with NP40, and apo B-100 was purified by size exclusion chromatography and subsequent dialysis to remove excess detergent. An overview of each step is shown in Figure 1.19 with negative-staining TEM pictures. To achieve the high quality and quantity required for neutron scattering studies, the sample solutions of each individual preparation were characterized (see in Chapter 1, section 1.5 in Tables 1.3-1.7), frozen at -80 °C and

[1] zwitter-ionic detergent 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate

lyophilized. For the neutron scattering experiments, the resulting powder was hydrated under D_2O atmosphere to a hydration level of 0.4 g D_2O / g sample. In addition, pure NP40 detergent was also lyophilized, during 48 h, but displayed a gel-like appearance.

Both samples were scanned using elastic and quasi-elastic incoherent neutron scattering (EINS and QENS), in order to access their picosecond - nanosecond dynamics. As the neutron incoherent cross-section for Hydrogen is much higher than for other atoms [118], the measured signal mainly comes from Hydrogen atoms.

4.3 QENS measurements

In QENS experiments, diffusional processes are accessible, like rotation [127], or translational diffusion [122]. For hydrated powders, only internal motions are considered, and no center of mass self-diffusion is observable.

4.3.1 Experiments and instruments

QENS measurements were carried out at 300 K at Institut Laue-Langevin (ILL), Grenoble, France, on the backscattering spectrometer IN16b in Si(111) configuration [171, 252], in BATS time-of-flight configuration [172, 253], and on the cold time-of-flight spectrometer IN5 [169, 254], whose characteristics are detailed in Table 4.1.

Parameters	**IN13**	**IN5**	**IN16b BATS**	**IN16b Si(111)**
Wavelength (Å)	2.23	10	$\Delta\lambda/\lambda \sim 12$ %	6.271
Accessible Q-range ($Å^{-1}$)	0.2 - 4.9	0.02 - 1.16[a]	∼0.2 - 1.8	0.1 - 1.8
Energy transfer (μeV)	/	$\Delta E/E \sim 1$ %	± 150	± 31
Energy resolution (μeV)	8	10	4	0.75
Accessible timescale (ps)	∼100	∼100	∼200	∼1000

[a] $\sim 0.2/\lambda < Q_{elastic} < 11.6/\lambda$ (angles $\sim 5°$ to $135°$), i.e., $\sim$ 0.02 - 1.16 with 10 Å

Table 4.1: Characteristics of the spectrometers from Institut Laue-Langevin (ILL) used in this study.

Additionally, empty cell and Vanadium measurements were carried out for correction and normalization purposes. All data reduction and correction, including consideration of absorption and orientation of the empty cell according to Paalman-Pings expressions [174], were completed with the Manipulation and Analysis Toolkit for Instrument Data (Mantid) [176].

4.3.2 QENS model

Corrected data were then fitted in Python, using the least-squares minimization algorithm implemented in the package *lmfit* from Python [255]. The model used to fit

the spectra is the following one [111]:

$$S(Q,\omega) = \mathcal{B}(Q,\omega) + \mathcal{R}(Q,\omega) \circledast \beta(Q)\Big[A_0(Q)\delta(\omega) + \sum_{i=1}^{n} A_i(Q)\mathcal{L}\Big(Q,\omega;\Gamma_i(Q)\Big)\Big], \quad (4.1)$$

with $S(Q,\omega)$ the corrected spectra, $\mathcal{B}(Q,\omega)$ a constant (IN16b) or linear (IN5) background, accounting for the instrumental background and unresolved solvent contribution. $\mathcal{R}(Q,\omega)$ represents the instrumental resolution, characterized through the fit of Vanadium measurements by one (IN5) or several Gaussian curves (IN16b). $\beta(Q)$ is an intensity scaling factor. The symbol $\circledast$ stands for the convolution operation, defined as $(f \circledast g)(x) = \int_{-\infty}^{\infty} f(x-u)g(u)\mathrm{d}u$. $A_0(Q)\delta(\omega)$ includes the elastic contribution, with $A_0(Q)$ the elastic incoherent structure factor (EISF), accounting for the geometry of localized motions, and $\delta(\omega)$ the Dirac distribution. The quasi-elastic contribution, including diffusional processes, is expressed by a sum of n Lorentzian functions $\mathcal{L}(Q,\omega)$ of areas $A_i(Q)$ (which we will call amplitudes in the following) and half-width at half-maximum (HWHM) $\Gamma_i(Q)$ (also termed as linewidths). For a schematic representation of the amplitude and linewidth of a Lorentzian function, we can recall the Figure 2.9 of Chapter 2.

In the QENS model presented in Equation 4.1, each Lorentzian function accounts for one type of motion. In theory their number can be infinite, but in practice n is often less than 4, because of the limited signal-over-noise ratio. Here, n was empirically determined by optimizing the data fitting, especially the reduced chi-square value, which should lie around 1. Attention was also taken on the Lorentzian parameters, like A_i or Γ_i, so that their shape as function of Q remains regular (no 0 or infinite values, clear extrema and trends). Bayesian analyses as presented in Sivia et al. [256] were also conducted, but were not decisive to separate between $n = 2$ or $n = 3$, as shown in Figure 4.1.

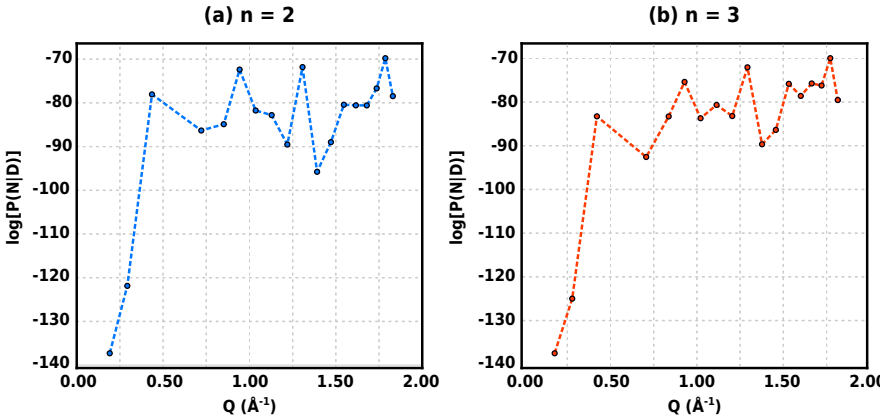


Figure 4.1: Bayesian analysis performed as presented in Sivia et al. [256], for apoB-100 / NP40 sample measured on IN16b, for (a) n = 2 and (b) n = 3.

The final analyses lead to n equal to 2 in the case of the pure detergent NP40 spectra. For apo B-100 / NP40 sample, n is equal to 2 for the lowest energy resolutions (IN5 and IN16b in BATS mode), and equal to 3 for IN16b spectra in classical mode. Examples of spectra fitting are shown in Figure 4.2.

4.3.3 Analysis of the amplitudes : combined model

To assign a motion and a specific contribution to each Lorentzian function, a model for describing both apo B-100 / NP40 and pure NP40 samples was built. Commonly the sole EISF (amplitude A_0) is investigated, but our model takes advantage of all amplitudes A_i. As all parameters are shared between the A_i, such a combined model allows a more robust data fitting. This approach is directly derived from a model for lipid motions called the dynamical Matryoshka model [134] (see also in Appendix A).

In the case of apo B-100 / NP40, our model considers that apo B-100 domains and NP40 chains undergo distinct confined diffusion [125], while at the same time their methyl groups perform jump-diffusion [122, 123]. This combined model is expressed in the amplitudes as follows :

$$A_0(Q) = p_{\text{imm}} + \epsilon_s(1-p_{\text{imm}})[zB_{\text{apoB}}(Q) + (1-z)B_{\text{NP40}}(Q)]B_{\text{jump}}(Q), \tag{4.2}$$

$$A_{\text{apoB}}(Q) = C + \epsilon_s(1-p_{\text{imm}})z[1-B_{\text{apoB}}(Q)]B_{\text{jump}}(Q), \tag{4.3}$$

$$A_{\text{NP40}}(Q) = C + \epsilon_s(1-p_{\text{imm}})(1-z)[1-B_{\text{NP40}}(Q)]B_{\text{jump}}(Q), \tag{4.4}$$

$$A_{\text{jump}}(Q) = C + \epsilon_s(1-p_{\text{imm}})[1-B_{\text{jump}}(Q)], \tag{4.5}$$

with z and $(1-z)$ the ratio of apo B-100 Hydrogens, respectively NP40 Hydrogens, to the total number of Hydrogens in the sample. When apo B-100 protein contribution is visible (classical IN16b), z is the determined experimental ratio $z = 0.36$, as determined from the sample preparation (see Chapter 1, Table 1.7). When it is not within the time window (IN16b BATS, IN5), or for pure NP40, this ratio will be equal to $z = 0$; in that case, the model reduces to the three amplitudes (respective elastic peak and $n = 2$ Lorentzian functions) expressed in equations (4.2), (4.4) and (4.5). The constant

$$C = \frac{(1-p_{\text{imm}})(1-\epsilon_s)}{n} \tag{4.6}$$

is a normalization constant accounting for the immobile fraction of Hydrogen atoms p_{imm}, and a multiple scattering factor ϵ_s acting mainly at low Q-values. $B_{\text{apoB}}(Q)$ represents the confined diffusion of subunits from apo B-100 protein (H atoms or

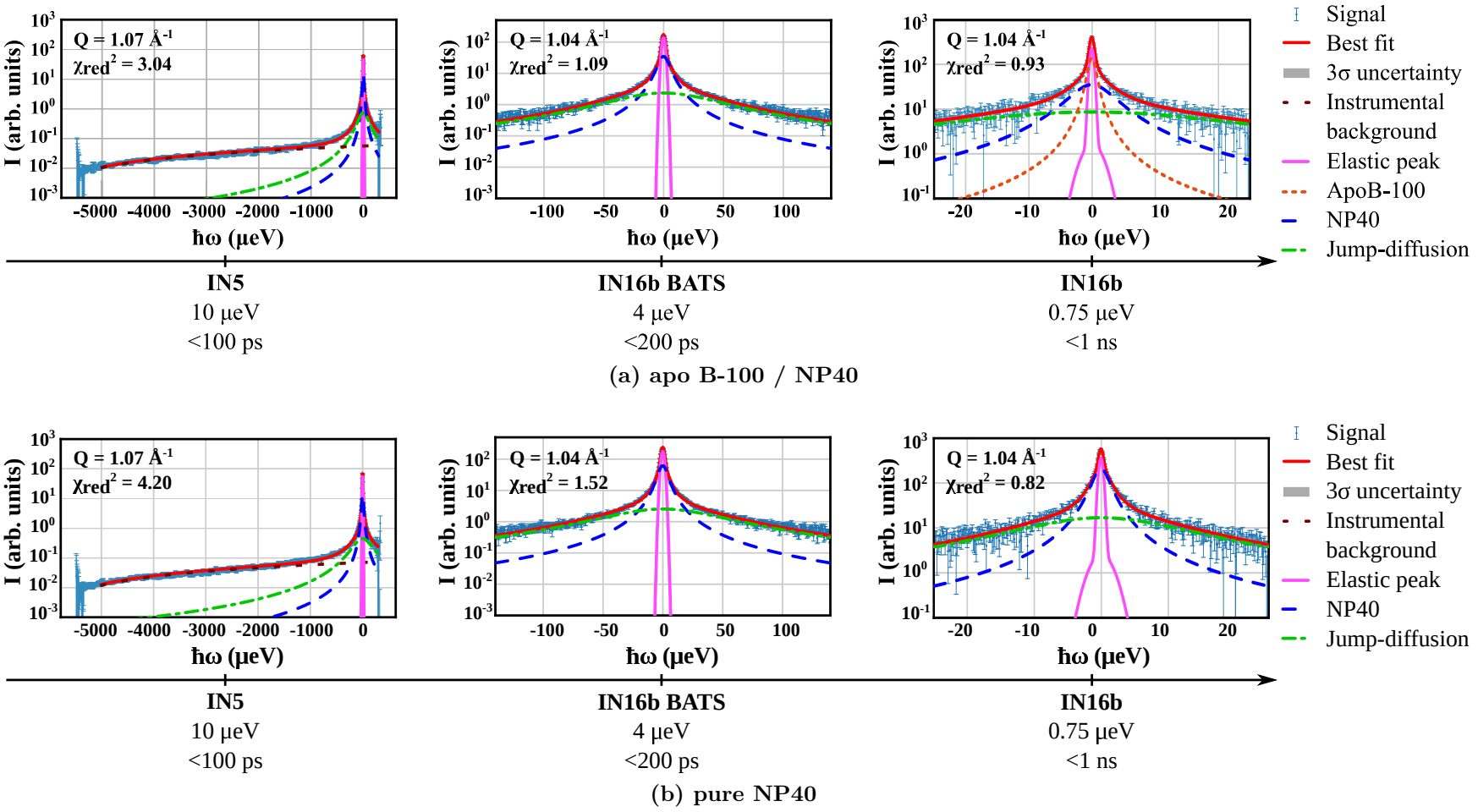


Figure 4.2: Example of data fitting on each instrument around Q = 1.05 Å^{-1} for (a) apo B-100 / NP40 and (b) pure NP40 samples. $n = 2$ for the pure NP40 sample on all instruments, and for apo B-100 / NP40 on IN5 and IN16b in BATS mode. $n = 3$ for apo B-100 / NP40 on IN16b in classical mode, indicating an extra contribution from the apo B-100 compared to pure NP40. 3σ uncertainty is the fit uncertainty as computed by *lmfit*, but is too small to be visible at this Q-value. Y-axis is on log scale.

domains), and writes as [125]:

$$B_{\mathrm{apoB}}(Q) = \left[\frac{3\mathrm{j}_1(Qr_{\mathrm{apoB}})}{Qr_{\mathrm{apoB}}}\right]^2, \tag{4.7}$$

with r_{apoB} the radius of confinement. Correspondingly, $B_{\mathrm{NP40}}(Q)$ describes the confined diffusion of NP40 H atoms and chains:

$$B_{\mathrm{NP40}}(Q) = \left[\frac{3\mathrm{j}_1(Qr_{\mathrm{NP40}})}{Qr_{\mathrm{NP40}}}\right]^2, \tag{4.8}$$

with r_{NP40} the radius of confinement for NP40. Finally, the jump-diffusion motions of methylene and methyl groups of both apo B-100 and NP40 chains embedded in:

$$B_{\mathrm{jump}}(Q) = 1 - 2\phi(1-\phi)[1 - \mathrm{j}_0(Qd)], \tag{4.9}$$

with $\phi \in [0; 0.5]$ the proportion of Hydrogen atoms performing jump-diffusion motions, and d the distance between two sites [123].

The amplitudes A_i are simultaneously fitted by the combined model presented previously, using the least-squares minimization algorithm implemented in the package *lmfit* from Python [255]. Figure 4.3 represents the global fit of amplitudes A_i with the combined model, for $n = 3$, and Figure 4.4 depicts the case when $n = 2$ for apo B-100 / NP40 on IN16b. Table 4.2 summarizes the extracted parameters.

Sample	Parameters	IN5 (100 ps)	BATS (200 ps)	IN16b (1ns)
apo B-100 / NP40	p_{imm}	0.49 ± 0.02	0.24 ± 0.01	0.10 (fixed)[a]
	r_{apoB} (Å)	-	-	3.2 ± 0.7
	r_{NP40} (Å)	3.1 ± 0.3	2.7 ± 0.1	4.6 ± 0.4
	d_{jump} (Å)	3.2 ± 1.4	2.2 ± 0.1	3.1 ± 0.1
	ϕ_{jump}	0.10 ± 0.03	0.50 (max. value)	0.50 (max. value)
pure NP40	p_{imm}	0.51 ± 0.01	0.23 ± 0.01	0.17 ± 0.01
	r_{NP40} (Å)	3.4 ± 0.1	2.7 ± 0.1	4.2 ± 0.2
	d_{jump} (Å)	3.4 ± 0.4	2.2 ± 0.1	2.7 ± 0.1
	ϕ_{jump}	0.11 ± 0.01	0.28 ± 0.02	0.5 (max. value)

[a] Fixed for numerical reasons to the plateau of A_0 at high Q-value.

Table 4.2: Parameters retrieved from the amplitudes analysis.

As seen in Figure 4.3, we achieved a good superimposition between the fitting curves and data points, good reduced chi-square values, and an overall quality among all samples, demonstrating the robustness of the combined model. In addition, it proves to be enough sensitive to choose the value of n, which is the main critical point in QENS analyses. Tests on IN16b data with $n = 2$ for apo B-100 / NP40 revealed that the combined model is not applicable in that case (see Figure 4.4f), whereas it

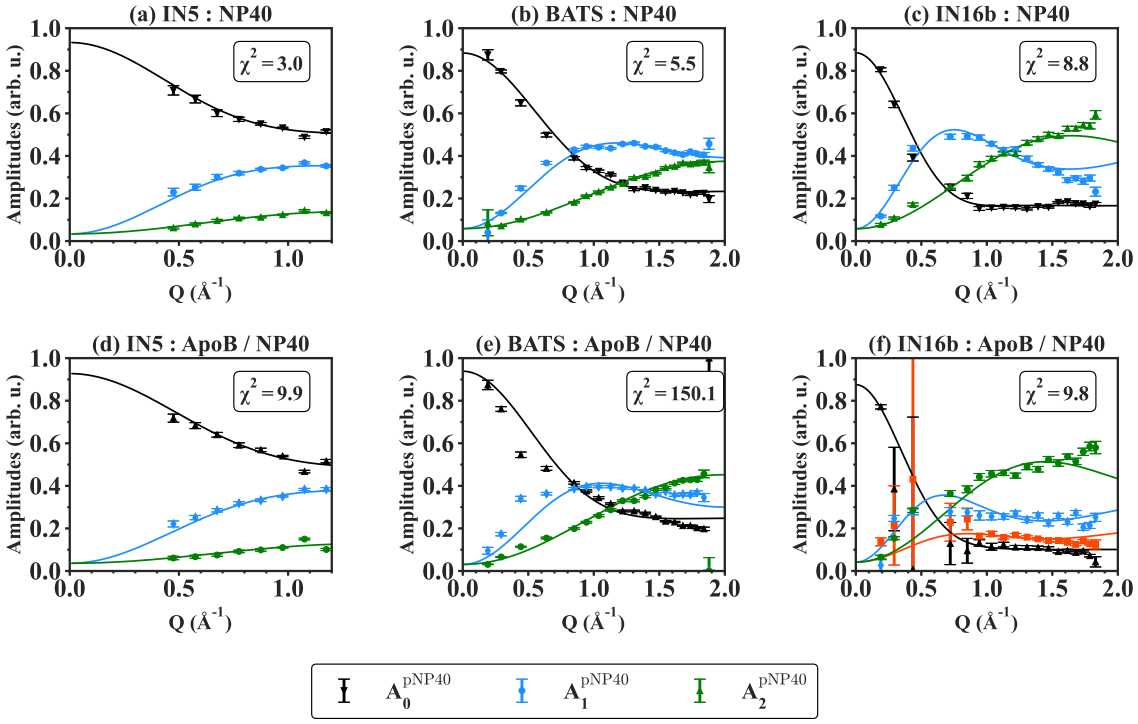


Figure 4.3: Fitting of the amplitudes A_i by the combined model presented in Equations 4.2-4.5 with $n = 3$. The first row corresponds to the Lorentzian amplitudes of pure NP40 measurements, extracted from the spectra measured on IN5, IN16b BATS, and classical IN16b, respectively. The second row includes the amplitudes from apo B-100 / NP40. Data points are superimposed by their corresponding fit (plain line, same color). Fit of the IN5 spectra at low Q-values exhibited high uncertainties, so corresponding parameters A_i and Γ_i were neglected.

applies well for pure NP40 with $n = 2$, as shown in Figure 4.4c. Reversely, by looking in more details at apo B-100 / NP40 measured on IN16b BATS, the reduced chi-square displayed in Figure 4.3e is much higher (∼150), and the low-Q value range is less well fitted, which could indicate a missing apo B-100 contribution. However tests with $n = 3$ Lorentzian functions on this dataset lead to overfitting. The time window of IN16b BATS likely includes only a part of apo B-100 motions.

The extracted parameters from Table 4.2 display good order of magnitude; d_{jump}, ranging between 2 and 3 Å, is consistent with H sites distances in methyl groups [257], while the confinement radii and immobile fractions have consistent values with other proteins [132]. The confinement radii from IN16b results indicate that apo B-100 domains are more restricted than for NP40. This could be attributed to a stabilizing effect on the protein due to NP40 presence.

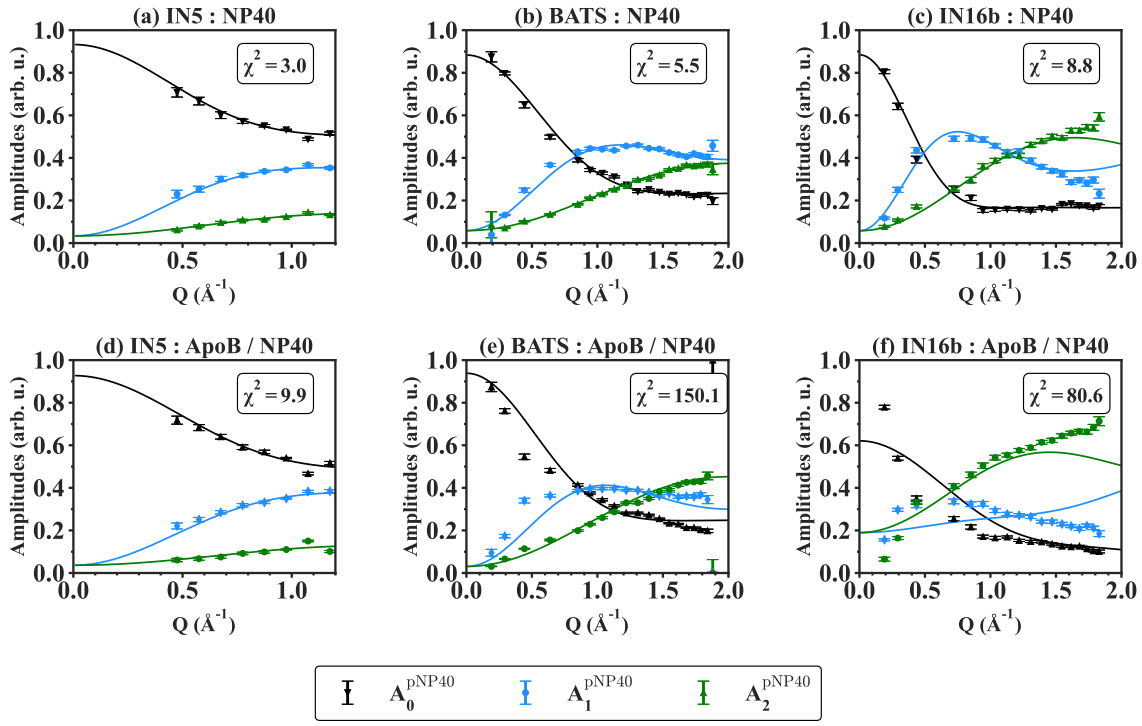


Figure 4.4: Fitting of the amplitudes A_i by the combined model presented in Equations 4.2,4.4,4.5 with $n = 2$ for apo B-100 / NP40 measured on IN16b.

4.3.4 Analysis of the linewidths : jump-diffusion

To go further, and extract dynamical parameters for each contribution, the linewidths Γ_i were investigated. This time, individual fitting was applied, using jump-diffusion models, either the original model from Singwi and Sjölander [122] for IN5 data (see in Equation 2.98), or a confined version [126] for IN16b BATS and IN16b parameters as introduced in Equation 2.103. Data fitting is shown for IN16b in Figure 4.5, and for all cases in Figure 4.6. Extracted parameters are displayed in Table 4.3.

All numerical values from Table 4.3 are within the same order of magnitude as compared to other proteins [132]. For IN5 measurements, some diffusion coefficients are above the values from bulk water (around 230 Å^2/ns at 300 K [258]). Similarly as in Trapp et al. [259], these outliers come from working with apparent diffusion coefficients. All the relaxation processes are not taken into account, especially for fast processes, which can lead to apparent nonphysical values. Residence times are less affected by these limits and stay within an adequate range, in the order of picosecond.

Striking differences in the values of dynamical parameters are observed between pure NP40 and NP40 in presence of apo B-100 (Table 4.3), which strongly suggest coupling dynamics. To our surprise, they indicate an acceleration of NP40 dynamics in presence of the protein. As an example, IN16b results display a diffusion coefficient

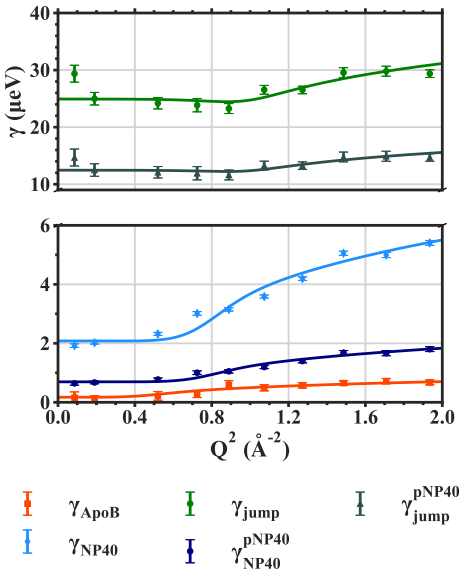


Figure 4.5: Fitting of the linewidths as a function of Q^2 for IN16b data. Data points are superimposed by their corresponding fit (plain line, same color). The exponent pNP40 refers to the pure NP40 sample.

Sample	Parameters	IN5 (100 ps)	BATS (200 ps)	IN16b (1 ns)
apo B-100 / NP40	D_{apoB} (Å^2/ns)	-	-	1.3 ± 0.4
	τ_{apoB} (ps)	-	-	550 ± 120
	D_{NP40} (Å^2/ns)	99 ± 47	12.8 ± 0.9	9.5 ± 0.6
	τ_{NP40} (ps)	42 ± 6	68 ± 4	67 ± 5
	D_{jump} (Å^2/ns)	1162 ± 730	160 ± 30	89 ± 17
	τ_{jump} (ps)	4.8 ± 0.7	8.4 ± 0.6	15.5 ± 1.1
pure NP40	D_{NP40} (Å^2/ns)	28 ± 4	9.5 ± 0.8	3.2 ± 0.2
	τ_{NP40} (ps)	15 ± 5	73 ± 6	200 ± 14
	D_{jump} (Å^2/ns)	317 ± 52	219 ± 46	44 ± 9
	τ_{jump} (ps)	0.9 ± 0.6	8.2 ± 0.5	31.0 ± 2.2

Table 4.3: Parameters retrieved from the linewidths analysis (confinement radii a in the SI).

D_{NP40} of 3.2 Å^2/ns for pure NP40, whereas in presence of apo B-100, the same value increases threefold to reach 9.5 Å^2/ns. Equivalent observations stand remarkably well for IN5 and BATS, with similar increase in magnitude; 1.5 to 3 fold. This accelerated dynamics is in line with simulation studies from Grimaldo et al. [260]. In their multi-component system, larger molecules are slowed down compared to the same monodisperse system, whereas smaller molecules are accelerated, as noted for NP40. Additionally, a QENS study from Wolanin et al. [261] showed that water confined in a hydrophobic matrix is accelerated under pressure. Nevertheless, all these studies focused on the translational diffusion coefficient, whereas in the apo B-100 / NP40 case, we are dealing with lyophilized samples. As a consequence, the present results express acceleration of the detergent internal dynamics, namely its chains, in presence of a large solubilized protein. To our knowledge, this kind of observation is

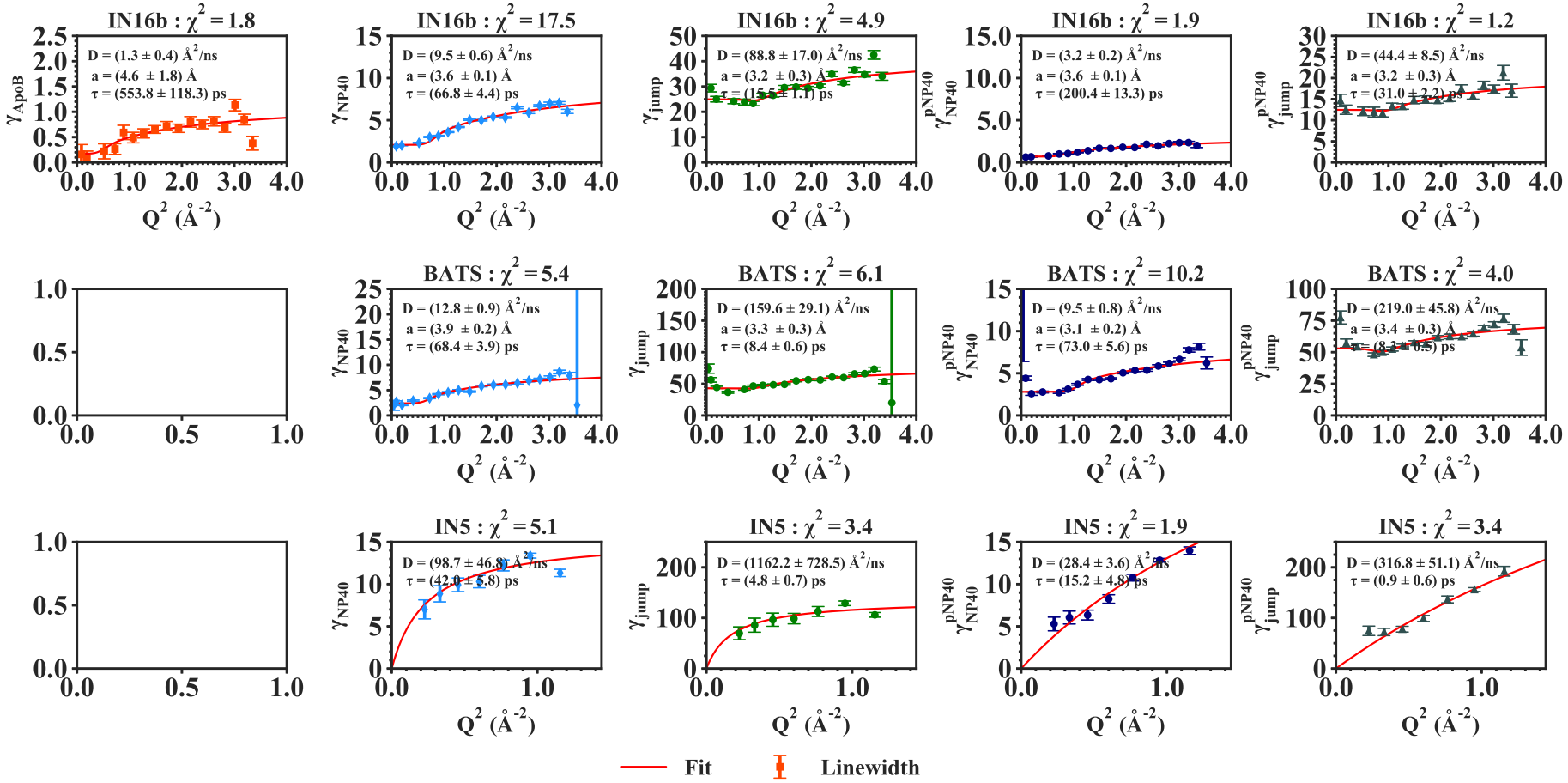


Figure 4.6: Fitting of the linewidths Γ_i with a confined jump-diffusion model (IN16b, first and second rows) and classical jump-diffusion model (IN5, third row).

unprecedented. One possible explanation could lie in the differences between micellar structures in the pure NP40, and individual detergent monomers in apo B-100 / NP40.

From the numerical values in Table 4.3, the internal dynamics of apo B-100 protein reveals to be very slow, which can be partly due to the previous considerations. The corresponding diffusion coefficient, $D = 1.3$ Å^2/ns, is around ten times slower than common values for proteins [132]. Furthermore, the residence time τ is as high as 550 ps, whereas it usually lies around one or two orders of magnitude less. Similar high values were measured for water in a halophilic organism [262], or for human acetylcholinesterase [263], at the same instrumental resolution. Since the residence time τ is representative of the protein interaction with the neighboring molecules, a high value indicates a strong interaction. Associated to the small confinement radius r_{apoB} lying around 3 Å (see Table 4.2), this slow dynamics could be explained by high confinement of apo B-100 domains, notably by the surrounding NP40 detergent, but could also be an intrinsic feature of apo B-100.

4.4 EINS measurements

To have a comparison with previous measurements on LDLs and VLDLs published in Mikl et al. [35], and to probe the effect of temperature, elastic scans were performed from 20 K to 310 K on the backscattering spectrometer IN13 (ILL), whose characteristics are presented in Table 4.1. In EINS experiments, averaged vibrational and very localized motions are probed, up to 100 ps in the case of IN13, similar to IN5. Additionally, empty cell and Vanadium measurements were carried out for correction and normalization purposes, which were treated using the Large Array Manipulation Program (LAMP) [175].

Mean-Square Displacements (MSDs) $< u^2 >$ were extracted at each temperature from the measured spectra assuming the Gaussian approximation [121], in the range of $Q = [0.5\ ;\ 2.0]$ Å^{-1}:

$$< u^2 > \simeq -3 \frac{\partial \ln \mathrm{S}(Q, 0 \pm \Delta E)}{\partial Q^2}, \tag{4.10}$$

with Q the momentum transfer, ΔE the energy resolution of the instrument, and $\mathrm{S}(Q, 0 \pm \Delta E)$ the measured spectrum.

Subsequently, a force constant analysis as introduced by Zaccai [264] was performed on the MSDs as a function of the temperature, from T = 240 K to T = 310 K. From there, an average force constant of the sample, $< k' >$, characterizing its resilience,

can be retrieved:

$$< k' >= \frac{0.00138}{\mathrm{d} < u^2 > /\mathrm{d}T} \tag{4.11}$$

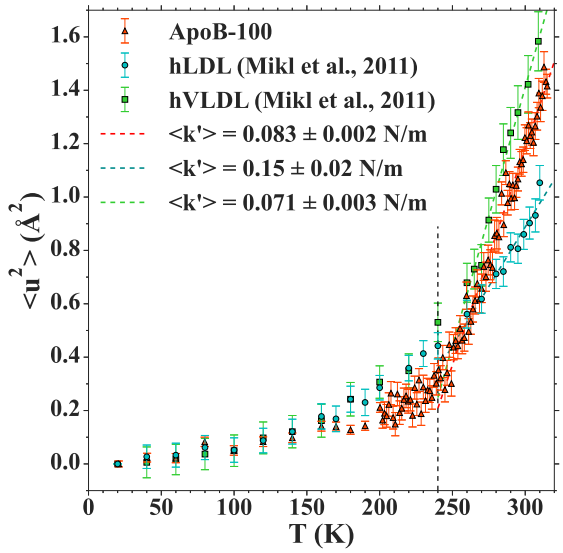


Figure 4.7: Mean-square displacements $< u^2 >$ as a function of temperature T of apo B-100 corrected signal (orange triangles), compared to LDLs (blue circles) and VLDLs (green squares) results published in Mikl et al. [35] (https://pubs.acs.org/doi/10.1021/ja203679g, further permissions related to the material excerpted should be directed to the ACS).

The MSDs in Figure 4.7 increase with temperature, as gain of thermal energy by the sample leads to enhanced motions. A steep increase is observed around T = 220 K, and corresponds to the so-called dynamical transition [244]. It is a common behavior of proteins in hydrated powder or in solution, characterizing the transition from harmonic motions in the low temperature range, to anharmonic motions which enable protein functionality [264].

It is important to note that the accessible motions on IN13 correspond to the jump-diffusion of methyl groups, both present in apo B-100 or NP40, as derived previously by our combined model. To isolate apo B-100 contribution, the pure NP40 signal was subtracted, weighted by its measured transmission coefficient. Comparison between MSDs from apo B-100 corrected signal and pure NP40 measurements is shown in Figure 4.8. As for the radii of confinement displayed in Table 4.2, the NP40 detergent displays slightly larger dynamics than for apo B-100 protein.

Absolute values of the apo B-100 MSDs reveal to be quite high, as compared with other proteins at the same temperature and hydration level in Table 4.4. Meanwhile, its average force constant value is much lower, which indicates a smaller resilience, so

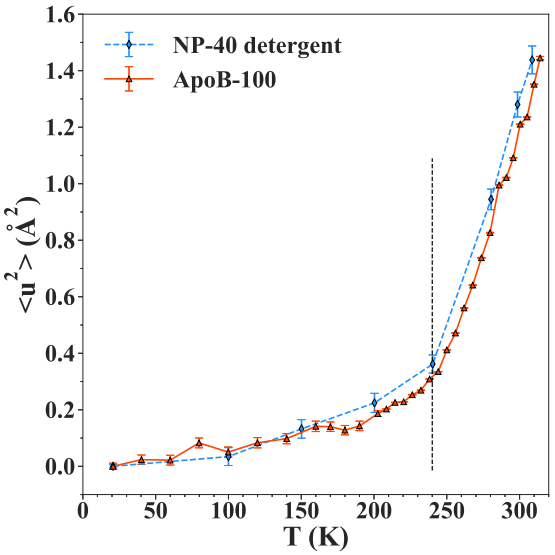


Figure 4.8: Mean-square displacements $< u^2 >$ as a function of temperature T of apoB-100 corrected and rebinned signal (orange triangles), compared to NP40 measurements (blue diamonds). Lines are guide for the eyes.

higher softness of apo B-100 protein. In parallel, comparison with published MSDs of LDLs and VLDLs [35] (Figure 4.7) reveals that apo B-100 MSDs are higher than for LDLs and close to VLDLs. In the same manner, the force constant is again intermediate between the LDLs and VLDLs values.

Table 4.4: Comparison of apo B-100 MSD and $< k' >$ values with other systems. Error on the last digit is shown in brackets.

System	MSD (300 K) (Å^2)	$< k' >$ (N/m)	Reference
human apo B-100	**1.24 (4)**	**0.083 (2)**	
myoglobin	0.5	0.3	Doster et al. [244] Zaccai [264]
photosystem membrane II	0.7 (2)	0.12 (5)	Nagy et al. [265]
bacteriorhodopsin (purple membrane)	1.1 (1)	0.12	Lehnert et al. [266]
human acetylcholinesterase	1.0 (1)		Trapp et al. [267]
human LDL	0.90 (6)	0.15 (2)	Mikl et al. [35]
human VLDL	1.4 (1)	0.071 (3)	Mikl et al. [35]

Overall, the elastic measurements indicate that apo B-100 is a highly flexible protein at the picosecond range, both in terms of MSD and force constant. One might speculate that it could enable the protein to quickly accommodate changes in the lipidic distribution during the VLDL to LDL transformation.

4.5 Conclusion

To summarize, the present study shows that it is possible to investigate dynamics of a solubilized protein in interaction with its detergent, using neutron scattering. The quasi-elastic measurements give rise to a combined model, including the experimental protein / detergent ratio. Such a model is directly tested by taking advantage of all Lorentzian amplitudes, and not just the EISF. Their simultaneous analysis makes the tests more sensitive, and allows selecting the most accurate model. Within this framework, apo B-100 dynamics could have been precisely delineated and revealed to be quite confined and slow. In addition, the analyses shed light on the dynamical effects of protein - detergent interactions, with an observed acceleration of the NP40 detergent compared to the pure NP40 sample. Alternatively, at the picosecond time scale, elastic scans reveal highly flexible properties of apo B-100 methyl groups, and an intermediate dynamics between LDL and VLDL, which could be put in perspective with the VLDL to LDL transformation.

To conclude, we strongly believe that the methodology exposed here for apo B-100 can be translated to classical membrane proteins. Accounting for at least one third of the proteome, their extraction often comes with detergent that remains bound, mostly due to the hydrophobic nature of the transmembrane domains. However, detergent presence should not prevent studies with neutron scattering. With the present investigation, and the introduction of combined models, we hope to open the field of neutron spectroscopy to more studies on membrane protein dynamics.

Chapter 5

LDL organization examined by cryo-electron microscopy

5.1 Introduction

Since the last 3D maps from Ren et al. [30], Kumar et al. [31] and Liu and Atkinson [33], that allowed accessing crucial information about the organization of the LDL, the cryo-EM technique has gone through considerable improvements, especially on the detectors and algorithms for image processing. It naturally called for a revised cryo-EM model of the LDL, with the same underlying questions as in the cited studies, namely the localization of the apo B-100 protein on the LDL surface, and the arrangement of the various lipids within the LDL.

The interpretation of cryo-EM maps of LDL remains challenging, as the lipids, as well as the protein, have similar electron cross-sections, resulting in similar densities, especially on the shell [15]. The main differences reside in the denser rings of cholesterol in its free or esterified forms (around 0.42 e^-/nm^3), or the phospholipids' head with stronger densities (about 0.40-0.45 e^-/nm^3 as seen for instance in [268]) compared to the tails ($\sim$0.29 e^-/nm^3) [15]. These variations allowed to propose in early cryo-EM maps of LDL from Orlova et al. [28] a new model for the core organization, with the lamellar arrangement of the cholesteryl ester molecules below the core transition, as presented in Chapter 1. However, the similarities in the electron densities from the protein and phospholipids' heads have made it more difficult to separate the contributions. Meanwhile, many membrane proteins have been investigated by cryo-EM, as reviewed in [269, 270], and recently a reconstruction of a membrane protein in a liposome, so in presence of a lipid bilayer, was reported [271]. In that context, new explorations of the LDL organization with cryo-EM seemed promising, and are the subject of the present chapter.

5.2 Screening and data collection

5.2.1 Single-particle

The sample preparation and screening of the grids were conducted on the cryo-EM platform at the Institut Européen de Chimie et Biologie (IECB, Bordeaux, France) by the platform responsible, Armel Bézault, on a Talos Arctica microscope (presented in Chapter 3, Figure 3.13a, Table 3.3). We conducted several tests to determine the optimal conditions for cryo-EM, especially the concentration, type of grid, or the temperature of equilibrium before the vitrification.

Prior to the sample preparation, the grids were glow-discharged with an intensity of 25 mA and a negative polarity, to make them hydrophilic. Two types of grids were tested:

- « Normal » Quantifoil grid R2/2 Cu 200, which were glow-discharged during 45 seconds. On these grids, the optimal concentration was c = 3 mg/mL, and an example of image is shown in Figure 5.1A;
- Carbon-coated grid R2/2 300 mesh with a carbon layer of 2 nm. Glow-discharge was applied during 30 seconds. As the samples attach better to this type of grid, due to the carbon layer, the optimal concentration was lower: c = 1 mg/mL. Images are displayed in Figures 5.1B-C.

As outlined in Chapter 1, the LDL core undergoes a phase transition around 20°C, so the choice of the temperature before vitrification is crucial. We tested temperatures below (4 °C) and above (37 °C) the phase transition. At 4 °C, the crystalline phase is well recovered, with characteristic striations on some orientations of the LDL (Figures 5.1A,B). However, the cholesteryl ester layers also appear in the images at 37°C in Figure 5.1C. As emphasized in Chapter 1, the core transition seems to be faster than the vitrification process, and at least the brief passage in the vapor layer of liquid ethane is sufficient enough to trigger the transition, which the images corroborate. For that reason, we did not collect any data from the grid vitrified at 37 °C. Moreover, even if we would have an intermediate state, we would not be sure that it is homogeneous in all the grids and the 3D reconstruction would be even more complicated. Finally, between the normal and carbon-coated grids, we chose to collect data preferentially on carbon-coated grids, because of the better distribution of LDLs, as well as ice quality.

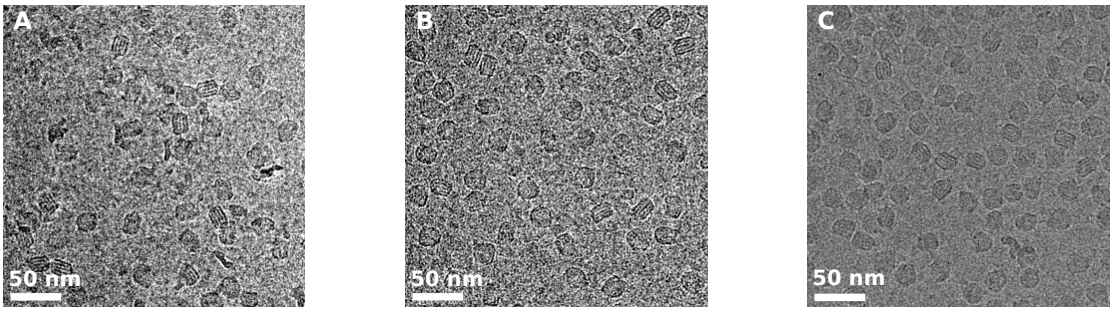


Figure 5.1: Screening tests on LDLs. A: One LDL preparation (4°C) on a normal grid (data collected). B: Another LDL preparation (4°C) on a carbon-coated grid (data collected). C: LDLs (37°C) on a carbon-coated grid. Scale bars are all corresponding to 50 nm.

In addition, grids containing the detergent-solubilized apo B-100 protein, and formerly vitrified by Dr. Guy Schoehn (Institut de Biologie Structurale, IBS, Grenoble, France), were also screened at the IECB. Images are shown in Figure 5.2.

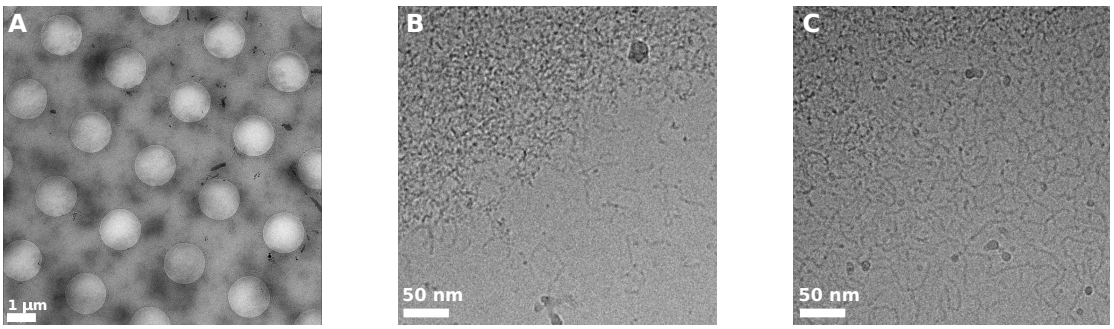


Figure 5.2: Screening tests on apo B-100 grids. A: overview of grid holes (scale bar: 1 µm). B, C: examples of images (scale bar: 50 nm).

However, aggregation of the protein was reported, and an overview of the grid (see Figure 5.2A) revealed that the protein was mainly accumulating on the carbon parts of the grid, and not within the holes. in Figures 5.2B and C, outside the aggregates, some individual proteins are visible, and are similar to what Gantz et al. [29] visualized with cryo-EM on apo B-100 (see also Figure 1.10). Different conformations are also visible, with a more or less extended protein. Nevertheless, the aggregation was too critical for a data collection. As the negative-staining images were of good quality, as seen in Figure 1.19, this are mainly the conditions for vitrification that would need to be optimized. Moreover, as a membrane-like protein, apo B-100 probably also needs to be stabilized with other methods, like nanodiscs, or reconstituted liposomes. These tests are planned by Assoc. Pr. Ruth Prassl's team at the Medical University of Graz as future research directions.

During this thesis, we concentrated on the reconstruction of the whole LDL, and performed two data collections on the Titan Krios from the European Synchroton

Radiation Facility (ESRF, Grenoble, France) with Dr. Eaazhisai Kandiah and Dr. Daouda Traoré. Characteristics of the data collection are reported in the Table 3.3 from the Chapter 3. In both cases, motion correction was performed through the MOTIONCORR module from RELION in the ESRF cluster. Then the resulting micrographs were processed either in RELION, either in cryoSPARC. First, the CTF was estimated, as introduced in Chapter 3, in RELION with the CTFFIND module from Rohou and Grigorieff [204], in cryoSPARC with a patch-based CTF estimator. In RELION, the particle picking was performed in several steps, as advised by Scheres [200], with a first manual picking, followed by the creation of a template, and then automatic picking based on the template. Additional sorting was performed with the calculation of a Z score [153], and discarding of all particles with a Z score bigger than 1. In cryoSPARC, the template for automatic picking was directly the reference 3D reconstruction computed from RELION. In both software packages, the picking parameters were optimized to reduce the number of false positives (like ice crystals), while recovering enough particles for statistics. Subsequent sorting of the particles was conducted with different rounds of 2D classification, which will be detailed in the next sections, as well as the final 3D reconstructions.

5.2.2 Tomography

During the project, tomography was also explored as an idea, notably to deal with the huge size of the LDL for a single-particle analysis, and possible heterogeneity of the apo B-100 protein on the LDL surface. Grid preparation, screening and a tilt series acquisitions were conducted at the Institut de Biologie Structurale (IBS, Grenoble, France) by Dr. Irina Gutsche's team, including Dr. Ambroise Desfosses and Dr. Maria Bacia-Verloop, on the Glacios of IBS.

Images of LDLs on a Lacey grid are displayed in Figure 5.3. In that case, the carbon film is not constituted by a periodic array of holes (as seen for example in Figure 5.2a), but are more randomly distributed, which is depicted in Figures 5.3B-F. In general, the images were of very good quality, especially within smaller holes from the Lacey grids (see Figures 5.3G-I). However, it quickly appeared that tomography approaches will be limited by several factors. First, the LDL remains a quite small particle for tomography, and the apo B-100 protein is not clearly visible in the images, like it is for instance for spike proteins from viruses. In addition, individual particle tomography gives results at very low resolution. The main reason is a lower dose used for each frame, as it is necessary to collect at the same location for several tilt angles, while limiting the radiation damages. The only solution to overcome the low-resolution is to perform sub-tomogram averaging, which requires enough homogeneous particles, as in single-particle, whereas the throughput in tomography is much lower. Finally,

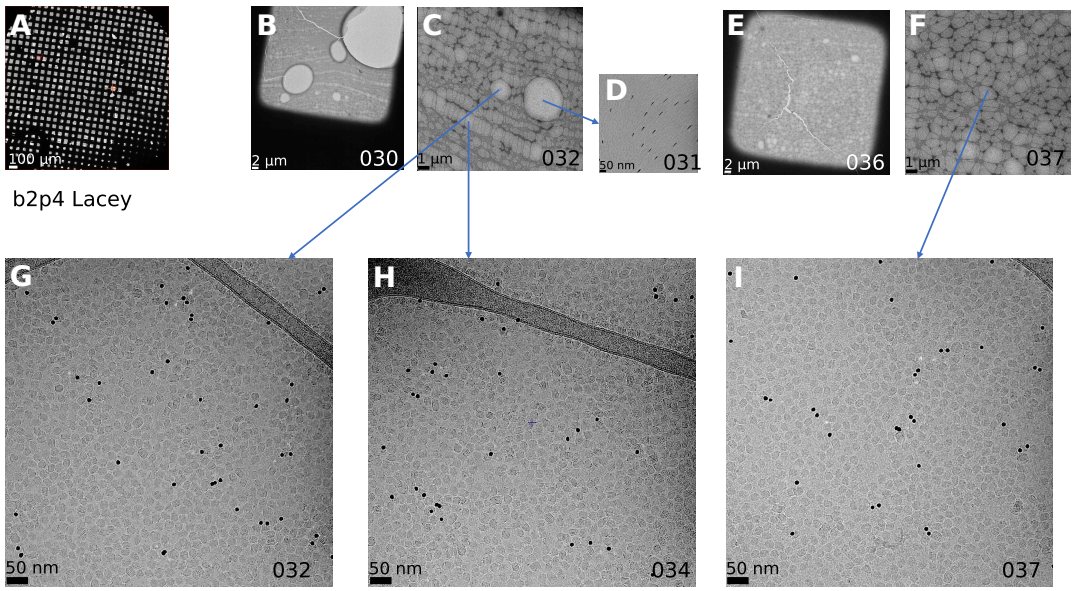


Figure 5.3: Screening tests for tomography on LDLs grids. A: overview of a Lacey grid. B: Zoom on a square, C: higher magnification. D: images from a large hole, and G,H: images from a smaller hole. E: Zoom on another square, F: higher magnification, and I: image on a small hole. Black dots on the images G-I correspond to gold beads used as fiducial markers. (Courtesy of Dr. Irina Gutsche).

tomography suffers from missing-wedge artifacts affecting sub-tomogram averaging procedures, whereas our single-particle dataset showed a uniform distribution of views. For all these reasons, we then focused more on the single-particle dataset, with deeper image analyses, with advice from Dr. Ambroise Desfosses (IBS, Grenoble, France).

5.3 Implementation of the workstation

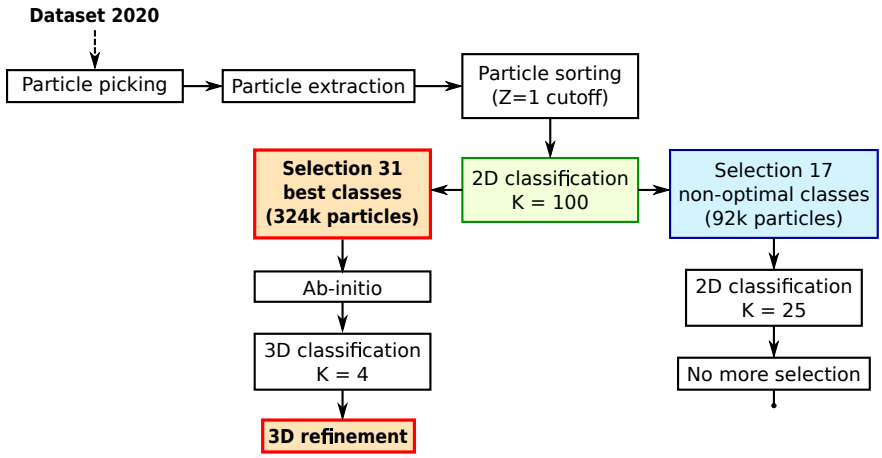
To process the data, and because the cluster from ILL did not include any graphics card, it was necessary for our group to buy and install a workstation, with specific characteristics, on which cryo-EM software packages could run, in our case RELION [153] and cryoSPARC [161]. The characteristics are summarized in the Table 5.1.

Table 5.1: Characteristics of the workstation Dell Precision 5820 XCTO (basic).

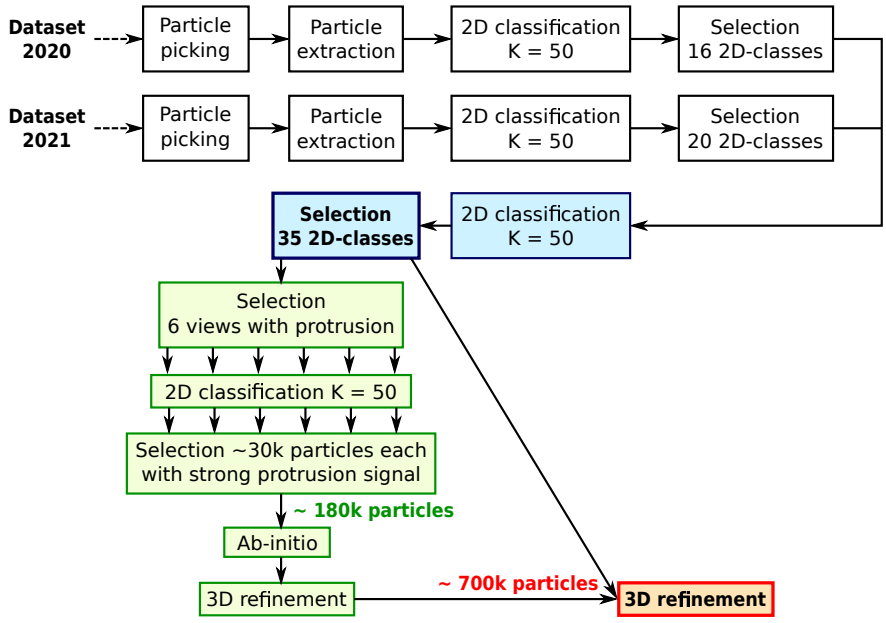
Components	Details
Core	Intel® Core™ i7-9800X 3.8 GHz, 4.5 GHz Turbo, 8 cores, 16.5 Mo Cache, HT, (165W, DDR4-2666 Non-ECC)
Memory	64 Go (4 x 16 Go) UDIMM DDR4 non ECC 2666 MHz
Hard drive	512 Go SSD SATA Class 20 2.5"
Chipset controller	Integrated Intel AHCI SATA (8 x 6.0 Gb/s), SW RAID 0.1,5,10
Graphics card (x2)	Nvidia Quadro RTX5000, 16 Go, 4DP, VirtualLink (7920T)
Internal hard drive (x2)	Seagate Barracuda Pro ST10000 DM0004, 10 To internal 3.5", SATA 6 Gb/s, 7200 rpm, 256 Mo cache memory

5.4 2D classification

After motion correction, CTF estimation and particle picking, as shortly described earlier, 2D classification was performed with two objectives; increasing the signal-over-noise ratio with an average on different particles sharing a common orientation, and cleaning the dataset by removing the outliers, that should go in lower quality classes. Different procedures were followed in RELION and cryoSPARC, and are summarized in Figure 5.4. In RELION, a harsh sorting was performed after the classification: only the best classes, with a strong signal from the LDL and low noise in the background, were selected, and examples are shown in Figure 5.5a. Additional 2D classification was carried out on the other 2D classes, however, they led to the same non-optimal classes, and the corresponding particles were then all discarded from the dataset. In cryoSPARC, the two datasets from 2020 and 2021 were first processed separately, and gathered only after a selection of 2D classes. Additional sorting of the particles was done with a 2D classification from all particles. From there, six 2D classes were selected, which present a good signal of the shell and «protrusions» of the LDL (see the densities designated by arrows in Figure 5.5), while ensuring a correct distribution of Euler angles. This step aimed at isolating a representative subset of particles exhibiting these extra-densities of interest, while balancing the number of particles in the different views of the LDL (to avoid missing views and missing-wedge artifacts). Subsequent 2D classification of each of the six 2D classes was conducted, followed by a selection of classes with strong protrusion signal to achieve around 30k particles for each view. Examples are displayed in Figure 5.5b. Then 3D ab-initio reconstruction and refinement were performed to give a first 3D map, serving as a reference to process all 700k particles.



(a) RELION procedure

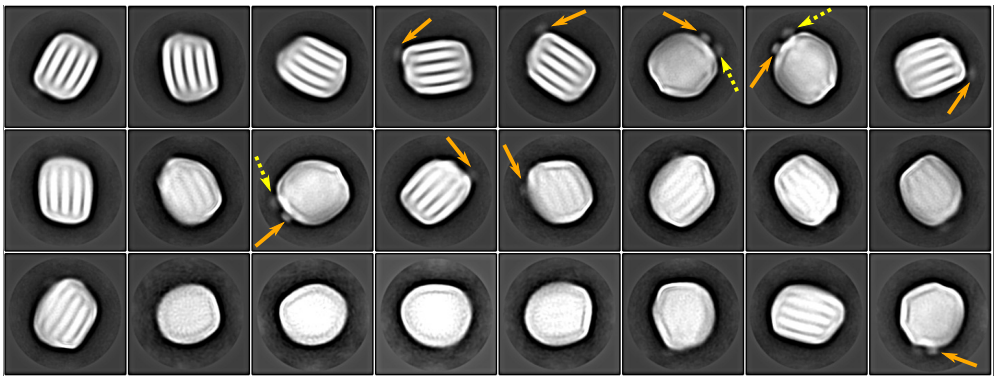


(b) cryoSPARC procedure

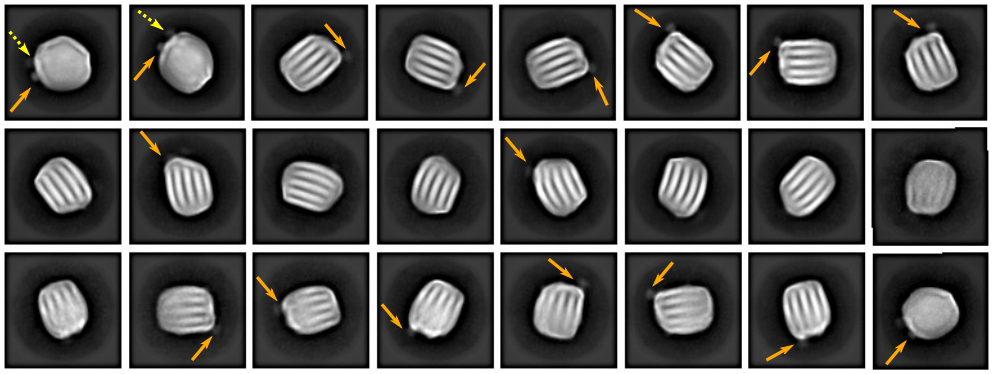
Figure 5.4: 2D classification procedures for (a) RELION and (b) cryoSPARC.

We remark a high reproducibility of the 2D classification in RELION and cryoSPARC, as evidenced by the comparison between Figures 5.5a and b. Common features are retrieved in both cases:

- Three characteristic layers within the LDL core, that we can relate to the crystalline phase of the core, and the arrangement of the cholesteryl esters, as described in Chapter 1, and attested by various techniques, including SAXS and cryo-EM (see [22, 23, 33] for the most recent studies we presented earlier);
- The presence of two protrusions in some classes, designated in Figure 5.5 by plain orange and dotted yellow arrows. The one indicated by the orange arrow is retrieved in other views, along the longest axis of the LDL, and was already assessed by most of the previous cryo-EM studies, from Ren et al. [30], Kumar



(a) RELION 2D classes



(b) cryoSPARC 2D classes

Figure 5.5: Examples of 2D classes of LDL particles in (a) RELION (352×352 pixels box) and (b) cryoSPARC (400×400 pixels box). Within the circles, the particles are normalized to zero mean and unity variance [200]. Orange arrow designates the so-called « front » protrusion, and dashed yellow arrow the « side » protrusion (the names are arbitrary). The contrast of the figures were enhanced with an external software for better visualization, even if it remains better on a screen than printed.

et al. [31], Liu and Atkinson [32] and Liu and Atkinson [33]. This protrusion is commonly associated to the globular $NH_2 - \beta\alpha_1$ domain of the apo B-100 protein. On the contrary, the second protrusion designated by the dotted yellow arrow was never observed before. To check if it could be another conformation of the first protrusion, we performed additional 2D classification on the classes with the two protrusions. However, even with only few hundreds of particles within a class, all 2D classes were presenting the two protrusions, indicating that this extra density should be accounted for another possible part of the apo B-100 protein.

5.5 3D reconstruction

Ab-initio reconstruction, followed by iterative refinement were processed in both RELION [153] and cryoSPARC [161] with the selected particles from 2D classification.

In RELION, the dataset only included the particles from the first data collection on the Titan Krios, which comprises around 320000 (320k) particles, whereas in cryoSPARC it contained both datasets, with around 700000 (700k) particles.

The resolution was estimated for all following maps by a post-processing job in RELION, whose parameters are displayed in the Table 5.2. They were chosen following

Table 5.2: Parameters for post-processing jobs run in RELION.

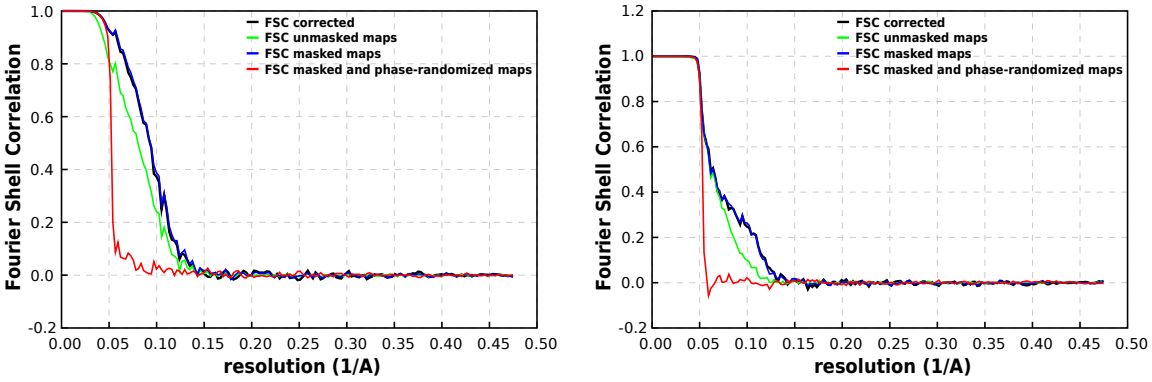
3D maps	Low-pass filter (Å)	B-factor (Å^2)	High- and low-resolution limits B-sharpening (Å)	Mask dilatation radius (pixels)	Mask soft-padding width (pixels)	**Final resolution (Å)**
RELION reference	15	-500	5 - 20	10	15	**8.8**
cryoSPARC reference	10	-500	5 - 20	10	20	**9.0**
cryoSPARC map S	10	-500	5 - 20	10	20	**11.4**
cryoSPARC map F	10	-500	5 - 20	10	20	**9.8**
cryoSPARC map SF	10	-500	5 - 20	10	20	**9.8**

first estimations given by the 3D refinement, and a close verification of the FSC curves. The 3D reconstructions displayed in the following correspond to the post-processed maps. As the post-processing parameters were identical in all cases, it allows for direct comparison. Fourier Shell Correlation (FSC) curves, which we introduced in Chapter 3 (section 3.2), are computed by the post-processing from RELION and are displayed in Figure 5.6 for the RELION and cryoSPARC reference maps, and in Figure 5.13 for the 3D classes. The FSC curves all present a regular trend, with a decrease at increasing resolution. In addition, the masks were checked for potential effects in the post-processed maps, which often occur if they are too tight, through the calculation of the FSC for phase-randomized maps (in red). As it quickly tends towards 0, it confirms that the parameters for the masks were suitable.

Finally, the 3D maps were displayed with UCSF ChimeraX [157, 158], and the threshold value for all maps was chosen to enclose a volume of V = $3.325 \cdot 10^6$ Å^3, which corresponds to the theoretical volume of a LDL particle, as introduced in Orlova et al. [28].

5.5.1 RELION map

An overview of the RELION map is presented in Figure 5.7 for different views, and measured distances are displayed in Figure 5.8.



(a) RELION reference **(b)** cryoSPARC reference

Figure 5.6: FSC curves computed in RELION of the final reference maps from (a) RELION and (b) cryoSPARC. The green and blue curves correspond respectively to the FSC for unmasked and masked maps. To check for mask effect, the FSC for masked and phase-randomized maps was computed and displayed in red. Corrected FSC is then shown in black.

Overall, the LDL exhibits dimensions around 200 Å, with two opposite flat surfaces, that we arbitrarily designated as the «top» and «bottom» of the LDL, as shown in Figure 5.7. When viewed from the front or the left (arbitrary views), it presents a rounded rectangular shape, similar to the early cryo-EM images from Spin and Atkinson [27]. The more visible features remain the characteristic cholesteryl ester layers, that we measured to be separated by about 37 Å (see Figure 5.8), which is consistent with the 36 Å periodic spacing measured and estimated with SAXS [17, 22, 23] or cryo-EM [30, 32, 33].

The higher resolution achieved with our current map allowed us to perform more precise measurements, as summarized in Figure 5.8. As a result, we estimated the layers' thickness to be about 21 Å. In previous reconstructions [30, 32, 33], each layer was associated to a monolayer formed by the ring of cholesteryl esters (which is estimated to have a length of about 10 Å), whereas our map suggests instead a bilayer. This is also supported by a closer look at the cut view in Figure 5.8a, where the layers are not homogeneous, but seem to be split in two superimposed layers. On the contrary, in between the layers, the low-density regions are estimated to have a length of around 17 Å, which correlate well with the length of a cholesteryl ester «tail».

Our map also reveals an extra layer between the core and the shell (indicated by a dotted pink line in Figure 5.8a and b), which measures on the top and bottom around 10 to 12 Å, and is thinner on the side (about 7 Å). Hevonoja et al. [4] already suggested that the free (unesterified) cholesterol concentrates on the shell, which was supported by MD simulations from Pan and Segrest [102], but in a more gradual way. The extra shell in our map supports the presence of regions enriched in free cholesterol, whose ring length is about 10 Å. These cholesterol-rich regions are largely found on the top

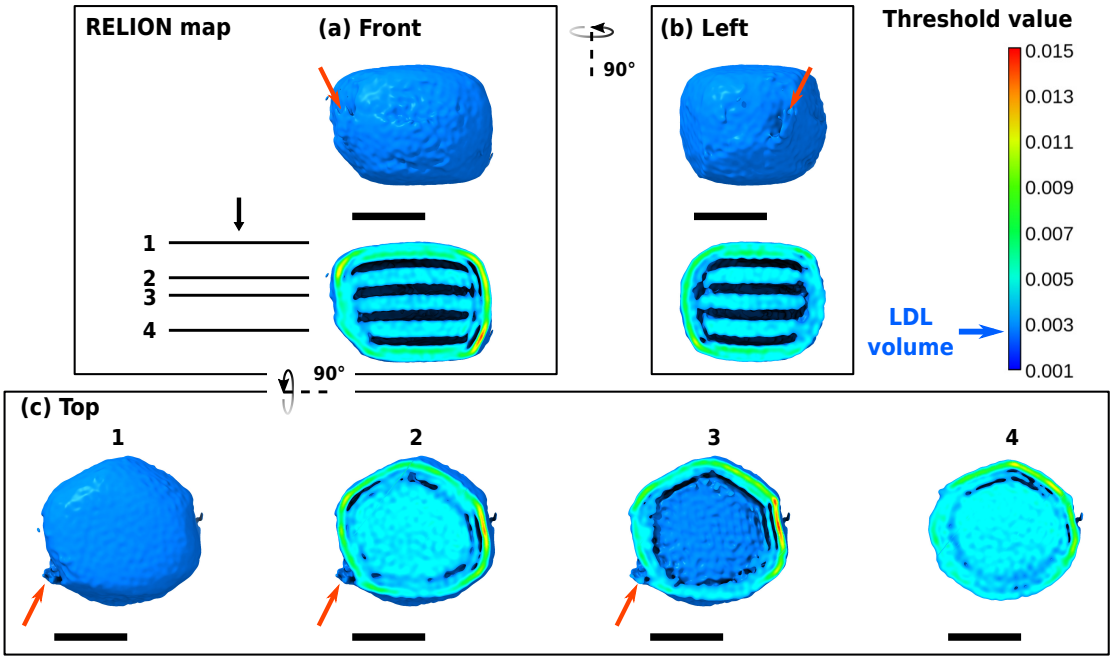


Figure 5.7: Overview of the 3D map from RELION. Views are arbitrary. (a) Front view and a corresponding cut in the middle. (b) Left view, and the corresponding middle cut. (c) Top view, with different cuts through the LDL. Maps are colored according to the threshold value, whose legend is shown in the right. The arrow indicates the protrusion. Scale bar: 100 Å. 3D map images produced with UCSF ChimeraX [157, 158].

and bottom, where the density is much higher, as represented by greener parts in Figure 5.7 and 5.8 (so a threshold density about the double of the one from internal layers). The 7 Å gap sometimes observed between the shell and this intermediate layer is too short to correspond to the full length of the phospholipids' tails (more about 10-20 Å as seen for example in [268, 272]), which would indicate a gradual incorporation of the free cholesterol in the phospholipid monolayer.

Concerning the shell, its thickness stands from around 10 Å to 20 Å in thicker regions, but apart from the protrusion, there is no additional clue of the apo B-100 protein parts. As already suggested by Laggner et al. [18], the protein does not extend far from the phospholipid monolayer, between 5 and 10 Å maximum. Furthermore, due to the number of phospholipid molecules, the protein is also expected to be directly in contact with the core at some locations [84], and parts can even be buried, as was suggested by the trypsin-digestion experiments [65]. The current map supports the idea of a tight association of the protein with the phospholipid monolayer, or even the core, as could be speculated for the thickest regions of the shell (left side in Figure 5.8a and b, next to the protrusion).

Contrary to the 2D classes in Figure 5.5a, the 3D map from RELION in Figure 5.7 presents only one protrusion, at the threshold corresponding to the LDL volume. We remark that the highest thresholds (so strongest signal) come from the shell, or

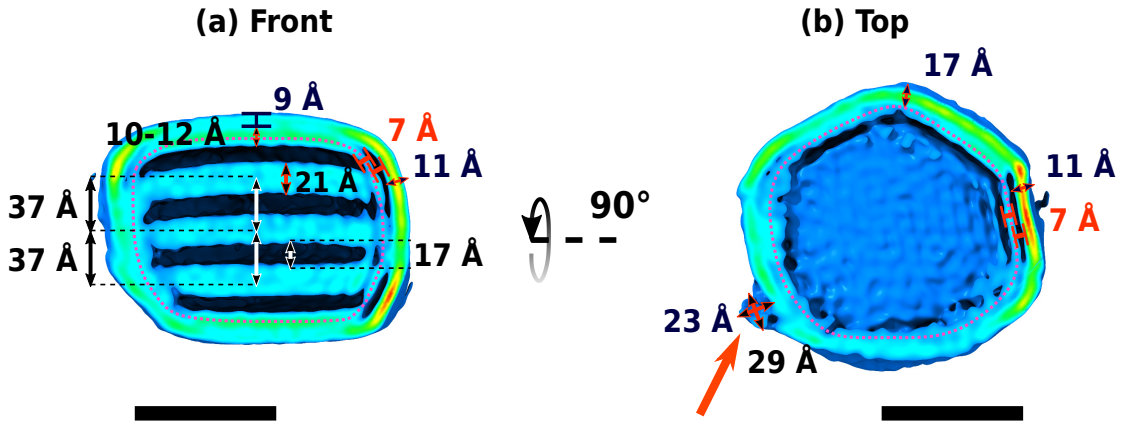


Figure 5.8: Distances within the 3D map from RELION. (a) Front view. (b) Top view. The arrow designates the protrusion. The dotted pink line indicates the intermediate layer. Scale bar: 100 Å. All measured distances are approximate values. 3D map images produced with UCSF ChimeraX [157, 158].

the internal layers, but the signal from the protrusion is slightly lower. At even lower thresholds (not shown) more density is visible in the protrusion part, but is mixed with noise. Already in the 2D classes, the signal from the protrusions is weaker compared to the shell or layers. One possible explanation could reside in the flexibility of the apo B-100 protein outside the LDL, which would result in fuzzy densities, and so lower threshold densities in the 3D reconstruction.

This possible flexibility was supported by 3D classification (not shown) with K = 10 classes, which all exhibited slightly different shapes, especially in the protrusion part. We have run numerous tests with a focus on the shell, either with masks around the region of interest, either by subtraction of the core, with local alignments, or by skipping it, but the results were all hardly interpretable. In addition, most of the tests we ran suggest that RELION is aligning preferentially to the lipids internal layers, or lipidic shell, which can blur even more the protein parts. In this context, and for several other reasons including a higher computation speed and new algorithms to treat heterogeneity, we then ran most of our image analysis within the cryoSPARC package.

5.5.2 CryoSPARC map

The final «reference» map from cryoSPARC is represented in Figure 5.9, and the measured distances are displayed in Figure 5.10.

Concerning the core organization, the cryoSPARC map retrieves the same characteristics as in RELION, as measured in Figure 5.10: the 37 Å periodic layers, each displaying a 20 Å thickness indicating a bilayer arrangement, as well as an

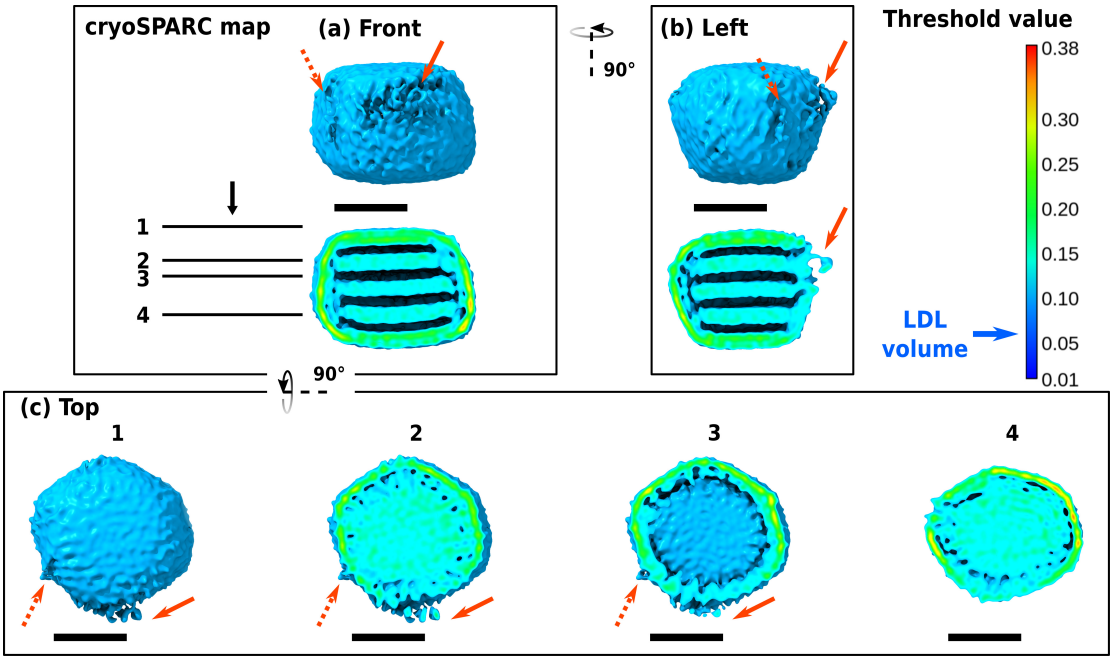


Figure 5.9: Overview of the 3D map from cryoSPARC. Views are arbitrary. (a) Front view and a corresponding cut in the middle. (b) Left view, and the corresponding middle cut. (c) Top view, with different cuts through the LDL. Maps are colored according to the threshold value, whose legend is shown in the right. The plain arrow indicates the «front» protrusion, and the dashed arrow the «side» protrusion. Scale bar: 100 Å. 3D map images produced with UCSF ChimeraX [157, 158].

additional 7 to 10 Å intermediate layer between the core and the shell (indicated by a dotted pink line), suggesting free cholesterol-enriched regions.

However, contrary to RELION, the shell presents higher density (as seen through greener parts in the maps). Similar tests as in RELION were performed on cryoSPARC, with a focus on the shell, and led to an identical map compared to the 3D map of reference presented in Figure 5.9, indicating that cryoSPARC is less driven by the lipidic layers for the refinement. More importantly, cryoSPARC retrieved the two visible protrusions of the 2D classes (Figure 5.5b) in the 3D reconstruction. We termed them arbitrarily «front» and «side» protrusions, respectively, indicated by the plain and dashed arrows in Figure 5.9 and 5.10.

The «front» protrusion presents three characteristic lines, as seen in the inset of Figure 5.10a, measured between 20 Å and 35 Å in UCSF ChimeraX, whereas from the left view in b, it has a triangular shape, whose base is estimated about 45 Å, and height around 24 Å (see Figure 5.10c). This protrusion is not plain, but encloses an empty cavity. From its location, along one of the longest axis of the LDL, as visualized in the left view (Figure 5.9b and 5.10c), non-negligible density threshold, and globular shape, this protrusion can be related to the $NH_2 - \beta\alpha_1$ domain of apo B-100, similar to the previous cryo-EM 3D maps of Ren et al. [30], Kumar et al. [31], Liu and Atkinson [32].

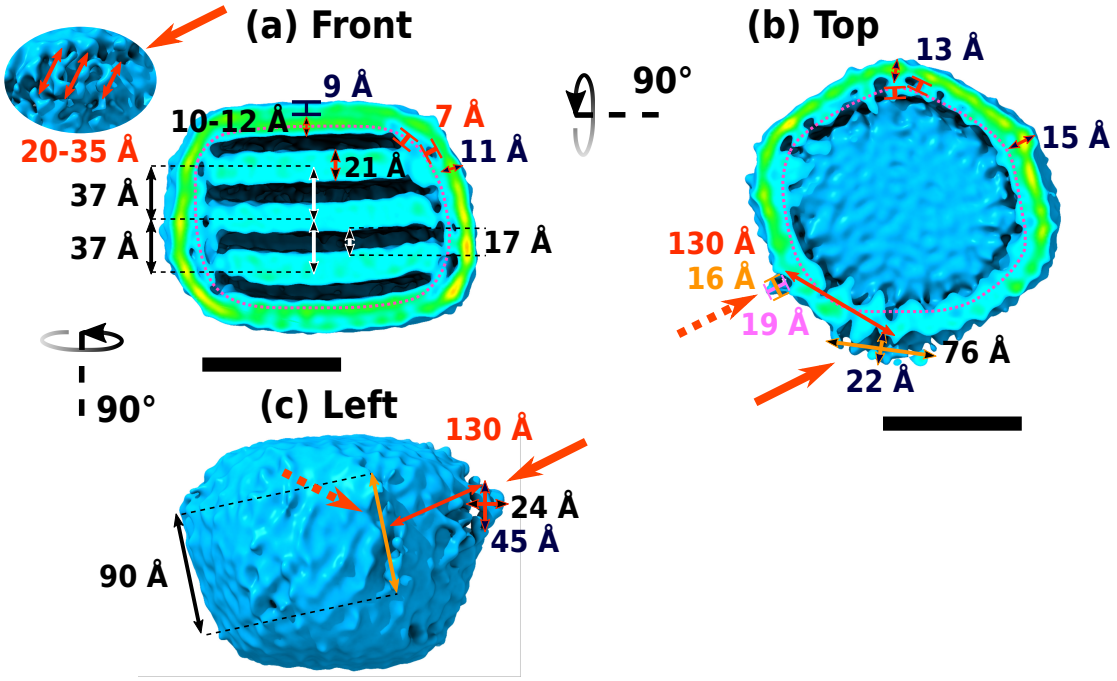


Figure 5.10: Distances within the 3D map from cryoSPARC. (a) Front view. An inset on the front protrusion is shown on the top left. (b) Top view. (c) Left view. The plain arrow indicates the «front» protrusion, and the dashed arrow the «side» protrusion. The dotted pink line indicates the intermediate layer. Scale bar: 100 Å. All measured distances are approximate values. 3D map images produced with UCSF ChimeraX [157, 158].

As introduced in Chapter 1, several groups proposed an homology model for this domain, based on the structure of the lamprey lipovitellin, like in [73, 75]. Dimensions of the N-βC region of this homology domain (the β-barrel domain of Figure 1.6) were estimated with UCSF ChimeraX to be about 40 Å in height, and 25-30 Å in width, which correlate with the dimensions of the protrusion measured in the left view in Figure 5.10c. Due to the low resolution of our map, we did not try to dock the homology model in our map, but could approximately place the β-barrel domain on the «front» protrusion region, and the others parts containing α-helices, including the lipid pocket domain, within the LDL shell. In the top view in Figure 5.10b, the extension of the protrusion was estimated around 76 Å, meaning that other parts outside the β-barrel domain could potentially stand at the LDL surface.

In contrast, the «side» protrusion is less globular in its shape: as outlined in Figure 5.10c, the extra densities are aligned along a line estimated around 90 Å. From the top, this protrusion is much thinner, extending within a 16×19 Å^2 rectangle, and remains at about 130 Å from the «front» protrusion, as estimated in Figure 5.10b. In the absence of any secondary structures, the assignment of this protrusion to a specific part of the apo B-100 protein is hardly feasible. We can only propose hypotheses based on its location in relation to the «front» protrusion, and with a comparison

to the monoclonal antibody mapping performed by Liu and Atkinson [33], shown in Figure 5.12e. In the map from Liu and Atkinson [33], the globular $NH_2 - \beta\alpha_1$ domain was mapped with the Mb19 antibody, and is highlighted in dark blue. As mentioned earlier, this domain is found along the longest axis of the LDL, and we highlighted in a similar way the «front» protrusion in Figure 5.12a. From the location of the «side» protrusion, and its close proximity with the globular domain, we can correlate it with the 5E11 binding in the map from Liu and Atkinson [33], colored in red. This location also corresponds to the potential binding site of the LDL, which is enriched in N-glycosylation sites containing carbohydrates [65, 66]. Because they are localized at the surface, the visible «side» protrusion could in fact turn out to be the glycans, and not directly parts of the protein. Several observations support this hypothesis: the total absence of secondary structures or recognizable domains, or the difficulty to retrieve it on the 2D classes, and even more on the 3D reconstruction.

Nevertheless, despite a better resolved shell, we did not get more information on the apo B-100 protein location, through notably secondary structures, like α-helices or β-sheets. Generally in cryo-EM, highly flexible parts result in blurred densities, as seen for instance in [219, 220, 273–275], or in extreme cases in their total disappearance from the final map, like the central part of the nuclear pore complex [276, 277]. Wang et al. [101] proposed that the LDL size can be modulated by reversible association of the protein with the lipids through its α-enriched domains, in the form of more-or-less extensible spring-like domains. From our map, the fuzzy densities corresponding to the protrusions in the 2D classes, and the absence of secondary structures in the 3D map, are then likely explained by a high flexibility of the apo B-100 protein on the LDL surface, including both the surface-exposed parts (like the «front» protrusion), but also the lipid-associated parts.

A comparison of our 3D map with an ab-initio SANS model from Lehofer et al. [23] is shown in Figure 5.11. The sample preparation was carried out by the same group, but without subfractionation, so that the sample for SANS is more heterogeneous than in cryo-EM. Moreover, the comparison remains qualitative as the alignment is arbitrary, notably with two low-resolution maps. However, some information can be retrieved. First, the shapes are similar, with a long axis and two flat surfaces at the top and bottom (see in Figure 5.11a). Nevertheless, the cryo-EM map is smaller than the SANS model, until 30 Å of discrepancy, and more along preferential directions (see in the front and top views). The differences mainly come from the sample preparation (more heterogeneity in SANS), but it is interesting to note that they are not isotropic, and concentrated in the «belt» of the LDL. As noted by Laggner et al. [18] from early SANS experiments on LDLs, the protein does not extend far beyond 10 Å from the phospholipids' surface. However, it is likely that our cryo-EM map «misses» some parts because of the protein flexibility, which we directly see from the comparison

in Figure 5.11, especially in the «belt» part, often associated to apo B-100 regions on the LDL surface. Following these considerations, we tried to explore the LDL heterogeneity within our cryo-EM dataset with several image processing strategies.

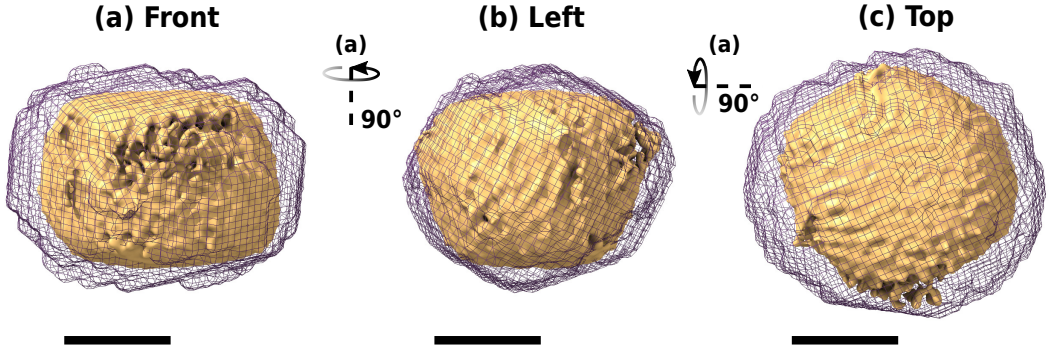


Figure 5.11: Comparison between the 3D map from cryoSPARC (in plain ochre) and a SANS model from Lehofer et al. [23] (purple mesh). (a) Front view. (b) Left view. (c) Top view. Scale bar: 100 Å. 3D map images produced with UCSF ChimeraX [157, 158].

5.6 3D classification

Different strategies for the 3D classification were followed, but we will focus on the last classification, which reveals the clearest differences. To summarize the main trials:

- On RELION, 3D classification before and after the refinement from the first dataset (∼320k particles) was performed with different K values: K = 4 or K = 10. Differences in the global shape of the LDL were observed, or even the layers, with slightly different curvatures, and the presence or not of the protrusion was reported. In addition, different T values were tested (4 or the maximum 25), enhancing in different ways the signal in the reconstruction algorithms. But the final 3D reconstruction remains unchanged, with a smooth map and no structural details from the protein outside the protrusion.

- Still on RELION, various local 3D classifications focusing on the protrusion were performed, by limiting the alignment, or even skipping it, with different K or T values. Strong mask effects were reported, in addition to a preferential alignment to the shell as mentioned earlier, and still the absence of recognizable secondary structures.

- In cryoSPARC, many trials were performed on the first dataset as well, and then on both datasets (∼700k particles), through the so-called heterogeneous refinements, with various K values, and successive classifications to try reducing the dataset into homogeneous subsets. Some tests were hardly interpretable

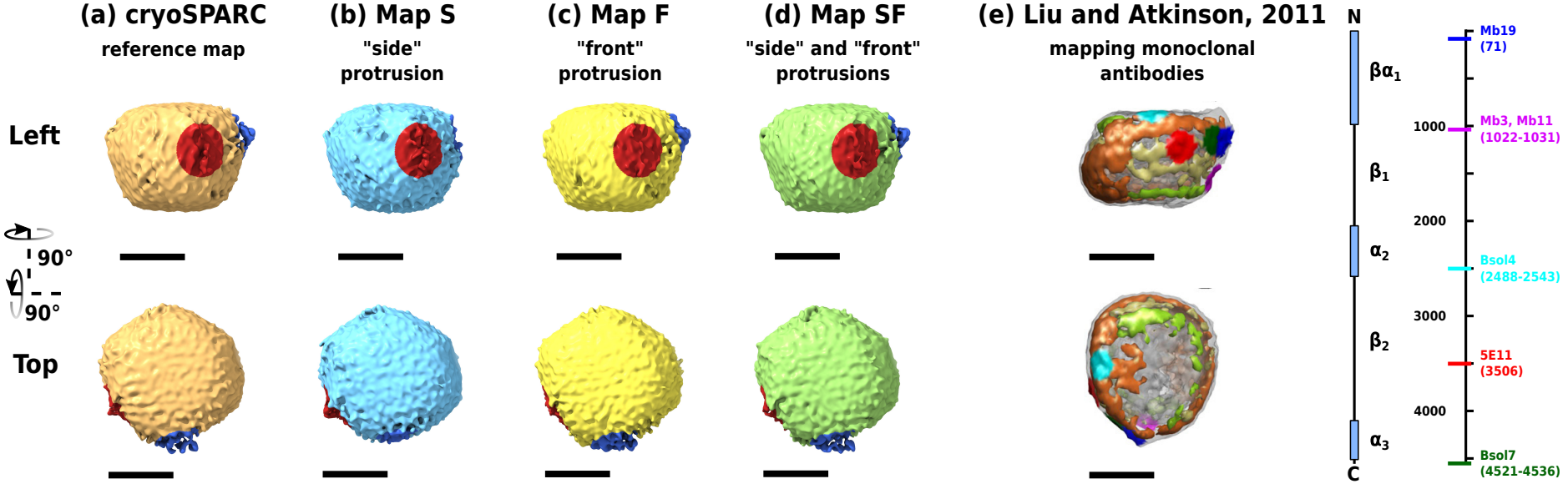


Figure 5.12: Highlights on the protrusions in cryoSPARC maps. Red areas correspond to the «side» protrusion, while blue zones refer to the «front» protrusion. Scale bar: 100 Å. 3D map images produced with UCSF ChimeraX [157, 158].

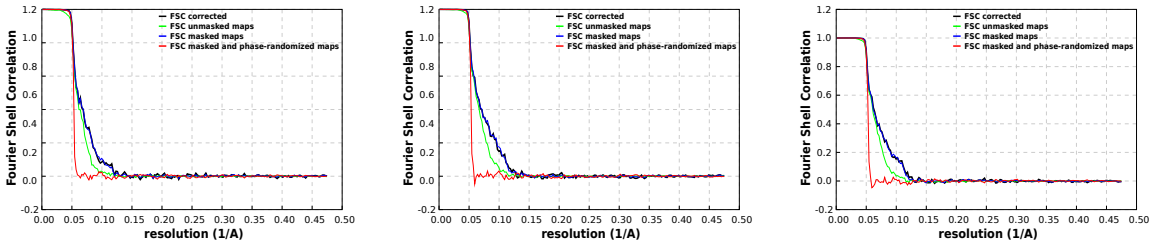
as cryoSPARC was separating the particles by «views» (Euler angles), which resulted in 3D classes with preferential orientations. This problem could be overcome by doing successive classifications with only K = 3. However, by reducing the number of particles in each class, we did not end up with more structural details, as could be expected if we would separate distinct conformations, meaning that apo B-100 flexibility would be more continuous. Still, to get valuable information, we isolated small subsets, and then gathered similar classes. From there, we could already observe that the protrusions were not present in all 3D classes.

- Local refinement algorithms from cryoSPARC were also tested, especially with a focus on the regions around the protrusions. As in RELION, strong mask effects could have been reported, and no further enhancement of the resolution with interpretable structural details.

- On the last classification strategy, we have randomly split the whole dataset in 100k batches, reconstructed a 3D map from there, and then performed 3D classification with K = 3 on each reconstruction. We have ended up with around 25 classes, with less than 50k particles in each of them. Close inspection of all classes have led to a rough separation in three categories. Similar classes were then gathered and refined to lead to the three maps presented in Figure 5.14, with their corresponding Fourier Shell Correlation (FSC) in Figure 5.13:

  - Map termed as «S», for «side» protrusion, as no «front» protrusion was visible, and bigger densities on the «side» protrusion region were reported (see Figure 5.14a). This class was less populated, with 150k particles.
  - Alternatively, a map termed as «F», built from about 272k particles, where the «front» protrusion was mainly present (see Figure 5.14b), especially on the top view).
  - A map containing both protrusions, from 258k particles, that we called map «SF» (Figure 5.14c).

In all three classes, the global shape remains similar, but strong differences are concentrated on the regions of protrusions. On the one hand, the presence or not of the «side» protrusion brings changes only in surfaces, as seen in the different cuts from the top view in Figure 5.14. This observation remains consistent with the hypothesis we advanced in the previous section, where the side protrusion corresponds in fact to glycans. On the other hand, the variations on the «front» protrusion also translate in the LDL shell, as seen in the cut from all views. It looks like the «front» protrusion is desorbing from the LDL surface from the map S (Figure 5.14a) to the map



(a) FSC Map S **(b)** FSC Map F **(c)** FSC Map SF

Figure 5.13: FSC curves from the final 3D classes, as computed by RELION. (a) Map S. (b) Map F. (c) Map SF. Legend colors same as in Figure 5.6.

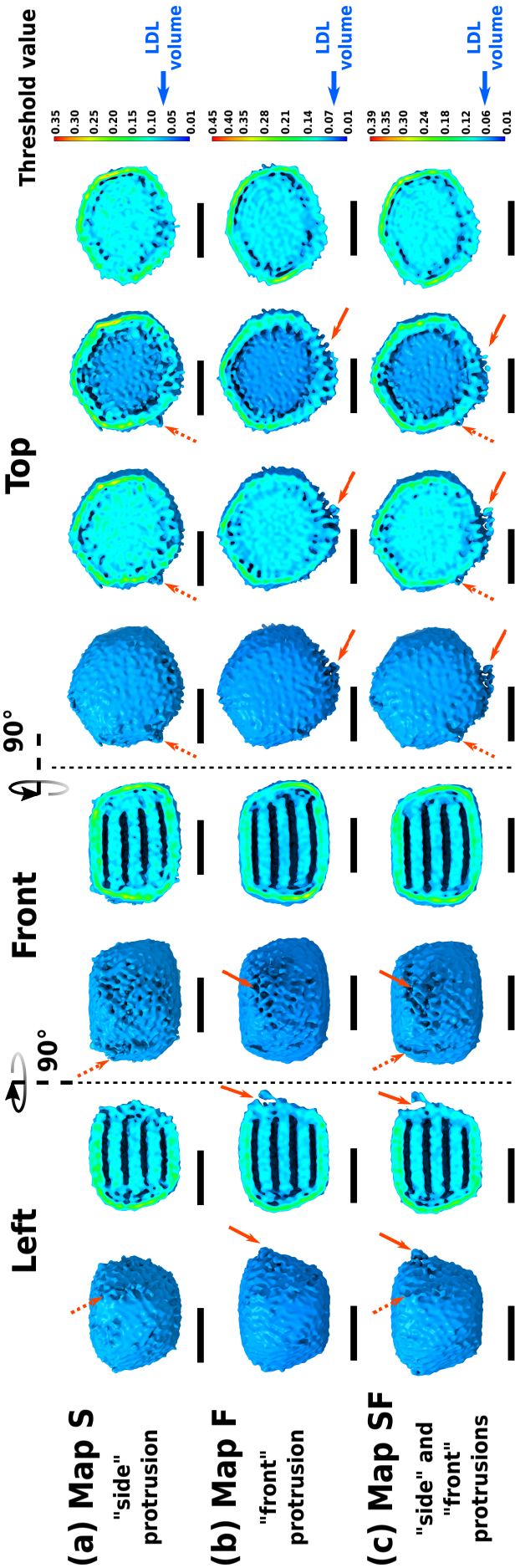


Figure 5.14: Final 3D classes in cryoSPARC. Scale bar: 100 Å. 3D map images produced with UCSF ChimeraX [157, 158].

SF (Figure 5.14c). This motion remains speculative, but is compatible with a non-negligible flexibility of apo B-100, as well as hypotheses from Wang et al. [101] about the absorption/re-absorption of α-domains of the protein to accommodate various LDL sizes.

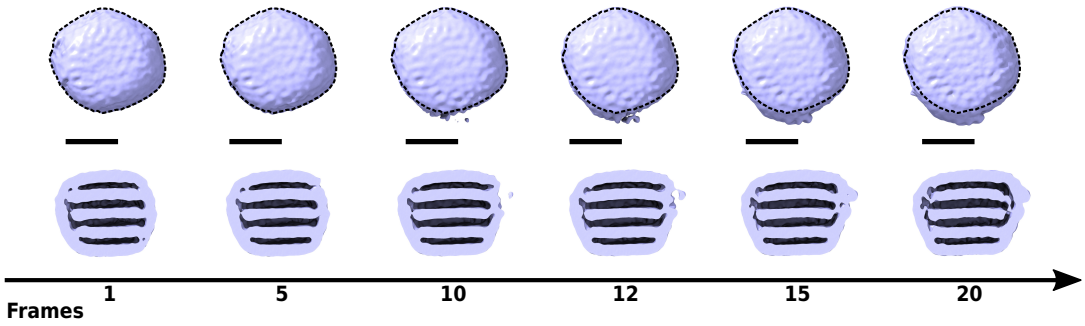
5.7 Trials on variability analysis

Similar to the bootstrap analysis and the calculation of a 3D map variance on the LDL reconstruction by Kumar et al. [31], we aimed at exploring the LDL flexibility in more details. In that purpose, we ran some trials with the two main current algorithms for the study of continuous heterogeneity: the 3D Variability Analysis (3DVA) from cryoSPARC [221], and the package cryoDRGN (Deep Reconstructing Generative Networks), whose algorithm is based on neural networks [222, 278].

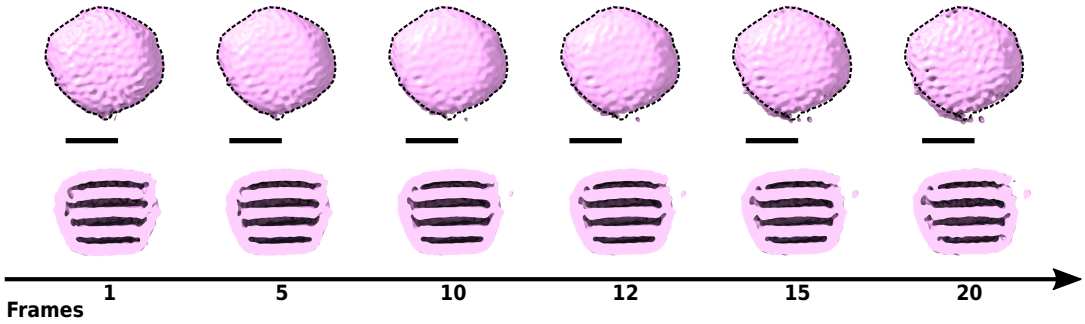
5.7.1 3D Variability analysis (cryoSPARC)

In cryoSPARC, the 3DVA option considers a family of volumes to reconstruct, and not one unique 3D volume. This family is created in a linear way, using a Principal Component Analysis (PCA), from a 3D refinement V_0 assumed to be the reference map. Different volumes will be reconstructed by adding positive or negative values from this map of reference V_0. In our case, we used the reference map from cryoSPARC presented in the subsection 5.5.2 and the Figure 5.9, and computed three main components. For each component, a series of volumes (also termed as frames) is reconstructed along the so-called latent coordinate, which indicates where is the volume along the variability component. Such a series is displayed in Figure 5.15 for the three components, in two views of the LDL: the top view, and a cut from the left view (same notations as before).

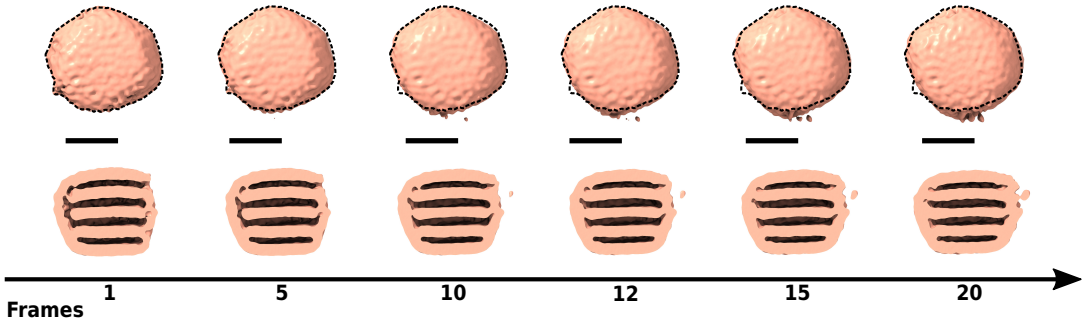
In general, the shape of the LDL remains the same, as well as the layers, thorough the three components of the Figure 5.15. The variations are localized on the surface, as noticeable in the top views, and around the regions of the protrusions. In the first and second components (Figure 5.15a and b), the «front» protrusion extends from the LDL surface to different distances. In the first frames of Figure 5.15a, it is not visible, while in the last ones it extends until the $\sim$22 Å, as measured previously in Figure 5.10. Whereas in Figure 5.15b, it is more the protrusion shape which is slightly changing. In the third component (Figure 5.15c), the «side» protrusion is also noticeable, but only in the first frames, reversely to the «front» protrusion, that appears only in the last frames. It correlates with the 3D classes, where the map S or map F only presented the «side» or «front» protrusion.



(a) Component 1. Top: top view. Bottom: cut from the left view.



(b) Component 2. Top: top view. Bottom: cut from the left view.



(c) Component 3. Top: top view. Bottom: cut from the left view.

Figure 5.15: Alignment of different frames from cryoSPARC 3DVA for the three principal components in (a), (b) and (c). On the top view, the first volume is encircled by a black dashed line, which is repeated along all frames to have a comparison. Scale bar: 100 Å.

Overall, the variability analysis corroborates with a variance localized on the shell of the LDL, especially on the two visible protrusions, and remains in line with the previous 3D classification. However, both types of calculations are run inside cryoSPARC, and so rely on similar hypotheses and algorithms. To check how the variability is distributed on the LDL, we also conducted tests within the software package cryoDRGN.

5.7.2 Continuous reconstructions (cryoDRGN)

Contrary to cryoSPARC, cryoDRGN draws on neural networks to parametrize the structures. In addition, the conformational space is not handled by a linear description,

as done in 3DVA, which potentially accesses more complex variations. All particle images are encoded in a latent space (as represented in the center of Figure 5.16), in which the heterogeneity is described by a latent variable z. At each point of this conformational space, a corresponding structure is reconstructed through a decoder. Depending on the type of heterogeneity, this latent space can take various shapes. For instance, clusters can be retrieved, similar to 3D classes, except that in this case, we do not need to give a fixed number of classes (the K value). Whereas if the «motions» are more continuous, the latent space will not exhibit a precise clustering.

For the LDL, we took the reference map from cryoSPARC, and trained first the dataset for 50 epochs, at dimension 8 for the latent variable z, as advised by Zhong et al. [222]. From there, the latent space exhibited a huge cluster of particles, and small subsets that we visualized as junk particles. As a consequence, we removed the latter, and kept only the particles of the cluster for an additional training, until 75 epochs, always at dimension 8. The resulting latent space representation and examples of LDL reconstructions are shown in Figure 5.16.

As for 3DVA or previous 3D classification, the global shape of the LDL remains unchanged, and all variations occur at the surface. The variability seems more continuous as there are few clusters, and the different conformations do not exhibit clear trends in each location of the latent space. The only difference with cryoSPARC 3DVA is the presence of some structures with a disrupted shell at the location of the front protrusion. These cases could potentially relate to LDLs with parts of the core being exposed. However, the encoder of cryoDRGN examines the whole image, and all sources of heterogeneity, from the particle itself, protein or lipids, or even background noise, which can also reflect in the reconstruction, as seen in Figure 5.16 with some volumes surrounded by noise.

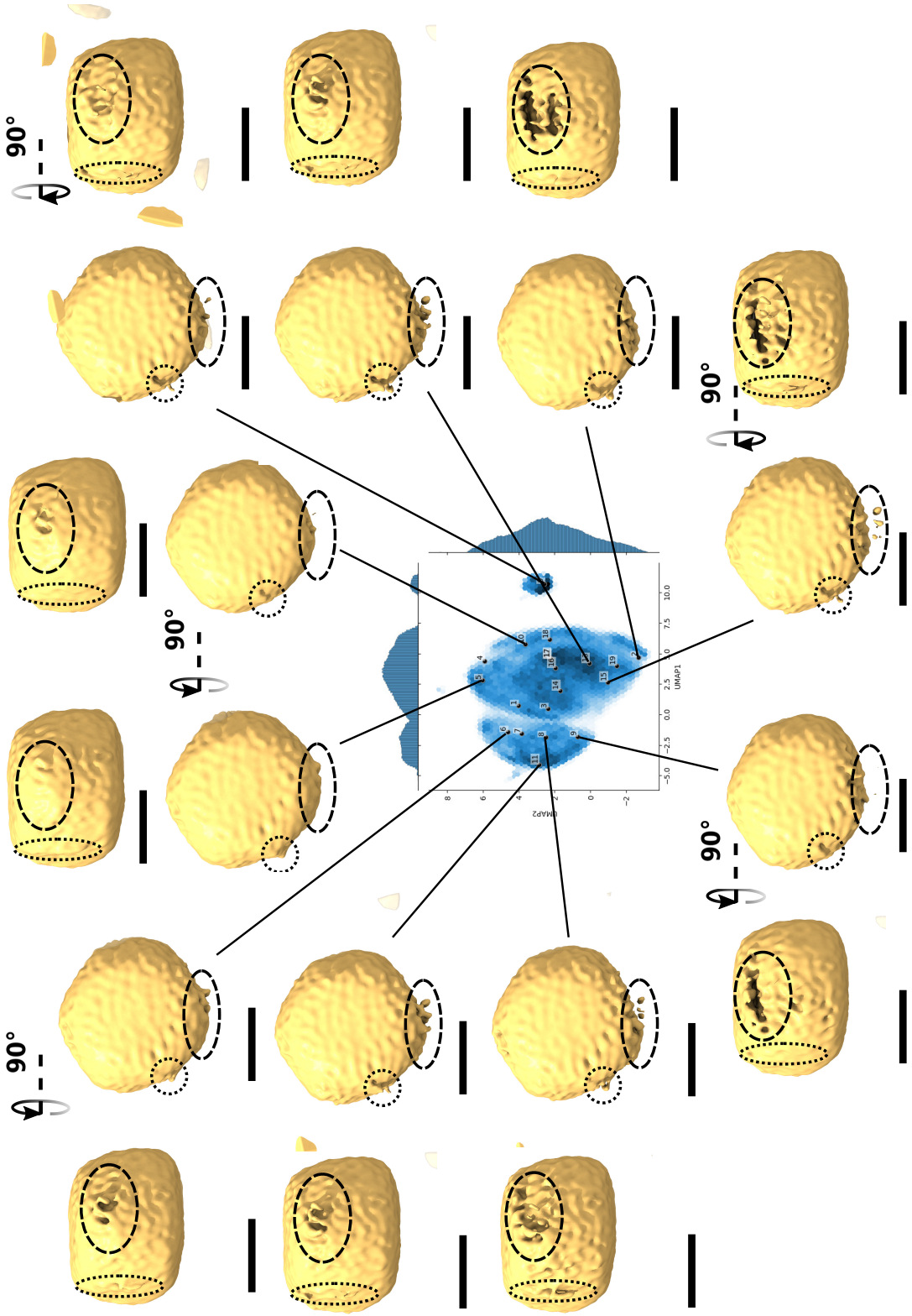


Figure 5.16: Examples of 3D maps from cryoDRGN. In the center, UMAP representation of the latent space, where each point corresponds to one particle image. Top and front views are shown for all volumes. The «front» and «side» protrusions are respectively encircled by a dashed and dotted line. Scale bar: 100 Å. 3D map images produced with UCSF ChimeraX [157, 158].

As a whole, both variability analyses have pointed out the important variations of the LDL shell, compared to the quite constant internal organization. Differences were mainly observed in the protrusions: presence or not, extension from the LDL surface, or even disrupted shell. Nevertheless, the main limitations of both cryoSPARC 3DVA and cryoDRGN is the need for a good estimation of the Euler angles as starting point, and even a 3D reference map already presenting secondary structures, which is not our case. The interpretation of the results must be taken with great care, but overall they corroborate with the 3D classification, and the observed variations concentrated in the shell support the hypothesis of a great flexibility of the apo B-100 protein on the LDL.

5.8 Conclusion

We report in this chapter new 3D reconstructions of LDLs, using the last state-of-the-art techniques offered by cryo-EM. LDL reconstruction remains obviously challenging due to the nature of this complex, composed in majority by various types of lipids, while the protein moiety apo B-100 lies both within and at the surface of the LDL. We propose here 3D maps at 9 Å resolution, which allow us to derive a revised model of the LDL organization.

First, the lamellar ∼37 Å-arrangement of the cholesteryl esters, known since long time from other techniques like SAXS [82] and initially visualized by cryo-EM [28], reveals to be constituted of bilayers, and not monolayers. Given the high proportion of cholesteryl esters, this arrangement involves a greater packing of the molecules and better stabilization. Second, the observation of an intermediate shell agrees with the general hypotheses concerning regions enriched in free cholesterol towards the shell. In general, the distances we measured and proposed in the present chapter can help in the development of more precise LDL models, as done currently in SAXS, or to build initial models in MD simulations.

The localization of the apo B-100 protein remains an open question, as our maps did not reveal structural details of the protein, like secondary structures. However, we could visualize two protrusions in the maps, that we arbitrarily termed «front» and «side» protrusions. The largest one, displaying a globular shape, and sitting along the longest axis of the LDL, was associated with previous observations in cryo-EM, thus assigned to the $NH_2 - \beta\alpha_1$ domain, notably the β-barrel, as compared to existing homologous models. The second protrusion, which was never assessed before, is located at the potential receptor binding site of the LDL, as compared with the early monoclonal antibody mapping of Liu and Atkinson [33]. Because this region is enriched in glycosylated sites, we hypothesize that the visible extra-densities could in fact be the glycans. Still, the absence of any structural hint from the protein was

intriguing, especially due to its huge size, and the high statistics we accumulated for the image analysis. Nevertheless, as remarked in many cryo-EM reconstructions, flexible parts commonly translate in blurred zones, or even total disappearance in the final map.

In that context, we postulate that the apo B-100 protein is highly flexible on the LDL, both its surface-exposed parts, or the β domains in tight association with the lipids. This hypothesis is supported by further examination of the LDL heterogeneity, through 3D classification or variability analyses, where the general LDL shape is conserved, as well as the core organization, but changes in the shell are observed, notably in the regions of the protrusions. A huge flexibility of apo B-100 on the LDL surface remains in line with its functionality, especially in the context of the VLDL to LDL conversion which implies huge conformational changes. In addition, we have shown in the previous chapter that apo B-100 in its detergent-solubilized form exhibits a non-negligible internal diffusion. At the nanosecond time scale, its order of magnitude resembles those of center-of-mass translational diffusion, and at the picosecond scale a large dynamics and great flexibility comparable to lipids was reported. The dynamics of apo B-100 on the LDL could be expected to be similar: highly constrained by its environment, but with significant internal dynamics, allowing adaptation to the various lipoproteins' sizes.

Finally, we plan to release the rawdata associated to the current 3D maps online, in particular through the EMPIAR database [279], to foster new discussions, and help the cryo-EM community to build new methods to handle such complexes.

Conclusion and perspectives

English version (EN)

The molecular details of LDLs and apo B-100 are far from being totally resolved, and still contain challenging and fascinating questions. The high proportion of lipids constituting the LDLs, and the non-exchangeable nature of apo B-100, are among the many reasons of such complexity. In the thesis work we tried, however, to deal with them, and to bring new insights about the LDL organization and apo B-100 dynamics, using the latest cutting-edge techniques such as incoherent neutron scattering and single-particle cryo-EM.

In Chapter 4, the remaining presence of NP40 detergent in association with apo B-100 was addressed by applying a novel approach to the QENS data. Directly inspired by a model for lipid motions, called the dynamical Matryoshka model (presented in Appendix A), our combined model takes advantage of all Lorentzian amplitudes to incorporate the experimental ratios of apo B-100 and NP40 in the sample, and then to separate their contributions. Combined to traditional analyses of the linewidths, outlined in Chapter 2, we derived several dynamical aspects of the apo B-100 / NP40 sample. We first highlighted the acceleration of the NP40 internal dynamics in presence of apo B-100, compared to measurements on pure NP40. Similar observations could have been reported in MD simulations [260] or other systems in QENS [261] for center of mass translational diffusion, but never for internal dynamics, nor directly on a protein / detergent sample. Furthermore, we were able to quantify the molecular dynamics of apo B-100 at the ps-ns timescale.

With the methodology introduced and employed here for QENS analyses, we hope to open incoherent neutron scattering to more dynamical studies on systems such as membrane proteins, which constitute a large proportion of the proteome, with various functionalities (recognition, channels...), but where the presence of detergent must often be taken into account.

In parallel, the LDL structure was investigated using single-particle cryo-EM, what is described in Chapter 5. The proposed 3D reconstructions, at about 9 Å resolution, made it possible to derive an actualized model of the LDL organization, where, for

instance, the core lamellae reveal to be bilayers, as opposed to the classical monolayer picture. The present 3D maps may be considered to help in the interpretation of SAXS curves and the development of associated models [93], or in the creation of starting models for MD simulations. However, despite a high statistics (around 700k particles), and extensive image analyses, no secondary structures from the apo B-100 protein could have been revealed. Only two extra-densities were identified, that we arbitrarily termed as «front» and «side» protrusions. The «front» protrusion was associated to the $NH_2 - \beta\alpha_1$ globular domain of the protein, while the «side» protrusion, which was never assessed before, was hypothesized to represent the glycans attached to the potential receptor binding domain. The absence of any secondary structure, even after extensive image analyses, was postulated to come from an intrinsic flexibility of the apo B-100 protein on the LDL surface.

At the present time, the large proportion of lipids, and the inherent flexibility of apo B-100, make the 3D reconstruction of LDL at high-resolution unachievable with the current tools. By presenting the state of the current work and the multiple attempts we performed, while releasing the raw data online, notably on the EMPIAR database [279], we wish to initiate new discussions in the cryo-EM community, and the emergence of new algorithms to reconstruct such complexes.

As a whole, the studies conducted in the thesis unraveled a more dynamical picture of the apo B-100 protein. In its detergent-solubilized form, it appears at the slowest timescale (ns) to be quite constrained by its environment (namely NP40 detergent), while displaying large dynamics at the ps scale, even comparable to lipids (see in Chapter 4). In association with the LDL, apo B-100 seems to present an intrinsic flexibility as well, which is indirectly evidenced by the absence of any secondary structure in the cryo-EM maps, and an overall blurring for some surface-exposed regions, like the $NH_2 - \beta\alpha_1$ globular domain (Chapter 5).

An obvious outlook of these studies is the elucidation of apo B-100 structure, which could help to decipher the whole LDL map, and to get more insights into several functionalities of the LDL, like its binding properties. To this end, the sample conditions of the detergent-solubilized apo B-100 still need to be optimized for cryo-EM, in order to avoid the aggregation observed in the screening tests (see in Chapter 5). First possible approaches are the tuning of buffer conditions, or adding of surfactants, which are common practices in cryo-EM, as reviewed in [227]. More elaborate strategies can also be considered to reconstitute a minimal lipidic environment for apo B-100, such as the use of nanodiscs [280] or liposomes [271]. However, the sample preparation will also need to be optimized at the level of vitrification, with trials on different types of grids (for example, carbon-coated grids gave better results with LDL), or blotting parameters. They will be further discussed in the context of new

theses starting in the Medical University of Graz (Austria), and the venue of Dr. Karin Kornmueller in the group of Pr. Judith Peters (Grenoble, France) for several months in 2022.

Meanwhile, with the recent release of AlphaFold for predicting protein structure [281], a theoretical model for apo B-100 could be derived from its primary sequence, and will be under consideration as well. It would complete the secondary structure prediction from Segrest et al. [68], whose «pentapartite» model remains a reference today, and the latest $NH_2 - \beta\alpha_1$ homologous model [75]. Furthermore, because it would be directly derived from the primary sequence of apo B-100, the computed tertiary structure would be more accurate than the proposed model from Kriško and Etchebest [78]. Always in view of cryo-EM experiments, an initial 3D model would help in the image processing and interpretation of the apo B-100 reconstructions. In addition, if the final 3D maps have sufficient resolution, it could allow building an atomic model of the protein. Finally, a tertiary model is required for developing MD simulations, which constitute another potential research direction.

In the case of the whole LDL, the previous MD simulations [102, 109] need to be extended. First, they should be further applied at temperatures below the core transition, to investigate the core organization, but also to incorporate the information from the 3D maps we are proposing here. In addition, the intrinsic dynamics of the protein on the LDL surface should be taken into account. Moreover, contrary to the MD simulations proposed in [109] where apo B-100 sits as a whole on the LDL surface, buried parts should be considered, or β-sheets at the same level of the phospholipids monolayer. Indeed, even in the raw cryo-EM micrographs, except for the protrusions, we did not observe very large densities outside the LDL surface. However, such MD simulations are obviously computationally expensive, especially with the number and variety of lipid molecules that are composing a LDL. Coarse-grained simulations are adapted to begin with, and to reach long time scales. Still, to have a more precise picture of apo B-100 dynamics, it could be of interest to run MD simulations on the apo B-100 alone. They would additionally be a great complement to potential cryo-EM studies of the detergent-solubilized apo B-100, as mentioned earlier. In particular, if the protein presents several conformations, and it was already observed for Gantz et al. [29] or in our negative-staining EM images, they must be distinguished in the single-particle analysis workflow. For that, MD simulations are a method of choice, and in extreme cases they even allowed to distinguish the various conformations of a spike protein directly in raw tomograms [282].

In conjunction to MD simulations, dynamical studies as we carried out with incoherent neutron scattering remain essential, as they allow to quantify directly and experimentally diffusional processes at the ps-ns scale. The investigation we performed on the detergent-solubilized apo B-100 could help in building the associated MD

simulations. Continuation of these experiments can be envisaged to probe even more precisely apo B-100 dynamics. For instance, neutron spin-echo measurements could be considered to reach longer time scales. It would require however some adjustments as the sample is lyophilized, and not in solution. Then, to attain apo B-100 dynamics on the whole LDL, and with the improvements of the last few years in deuteration, parts of LDL could be deuterated to focus on the protein, such as the phospholipids (as done in early SANS studies [18]), or the cholesterol, whose fully perdeuterated counterpart can now be produced [283].

Finally, molecular characterization of the modified forms of LDL is crucial as well. Following the previous SAXS, SANS, and incoherent neutron scattering measurements [23, 36], cryo-EM examination could be carried out, notably on triglyceride-rich (TG)-LDL, which was shown in early images to have a more rounded shape and to lack the cholesteryl ester arrangement [97]. An interesting question could concern the protein apo B-100, and how its structure, but also dynamics, is affected on modified LDL forms. What is more, if ever the dynamics of apo B-100 on the LDL surface would decrease on TG-LDL, the determination of part of its structure could potentially become more feasible. Obviously, the contrary could also be true, with a larger dynamics of apo B-100 on TG-LDL, but the image alignment would be at least less driven by the core organization, as it is expected to be disordered similarly to above the phase transition.

To conclude, the results presented in this thesis are part of a bigger picture. The dynamical information we retrieved of apo B-100, the actualized 3D model of the LDL, completed later on by future structural details of the protein, better interpretation of SAXS curves, or more elaborate MD simulations, among many other research directions, constitute all together essential information to better understand LDL functionality, and its link with cardiovascular diseases.

Version française (FR)

Les détails moléculaires des LDLs et de la protéine apo B-100 sont loin d'être totalement résolus et soulèvent encore des questions fascinantes. La forte proportion de lipides constituant les LDLs, et la nature non échangeable d'apo B-100, sont autant de raisons à l'origine de cette complexité. Dans ce travail de thèse, nous avons cependant essayé de les surmonter et d'apporter de nouveux éclairages sur l'organisation des LDL et la dynamique d'apo B-100, en utilisant les dernières techniques de pointe telles que la diffusion incohérente des neutrons et la cryo-EM.

Dans le chapitre 4, nous avons traité la présence résiduelle du détergent NP40 en association avec apo B-100 en appliquant une nouvelle approche aux données QENS.

Directement inspiré d'un modèle pour les mouvements lipidiques, appelé modèle Matryoshka (présenté en annexe A), notre modèle combiné tire parti de toutes les amplitudes des Lorentziennes pour incorporer les ratios expérimentaux d'apo B-100 et de NP40, puis pour séparer leurs contributions. En combinaison avec des analyses traditionnelles des largeurs de Lorentziennes, décrites dans le chapitre 2, nous avons pu déterminer plusieurs aspects dynamiques de l'échantillon d'apo B-100 / NP40. Nous avons d'abord mis en évidence l'accélération de la dynamique interne de NP40 en présence d'apo B-100, par rapport aux mesures de NP40 pur. Des observations similaires ont pu être rapportées dans des simulations MD [260] ou d'autres systèmes en QENS [261] pour la diffusion translationnelle, mais jamais pour la dynamique interne, ni directement sur un échantillon de type protéine / détergent. De plus, nous avons pu quantifier la dynamique moléculaire d'apo B-100 à l'échelle de la ps-ns.

Avec la méthodologie introduite et employée ici pour les analyses QENS, nous espérons ouvrir la diffusion incohérente de neutrons à plus d'études dynamiques sur des systèmes tels que les protéines membranaires, qui constituent une grande partie du protéome, avec des fonctionnalités variées (reconnaissance, canaux...), mais où la présence de détergent doit souvent être prise en compte.

En parallèle, la structure des LDLs a été étudiée à l'aide de la cryo-EM en «particule seule», ce qui est décrit dans le chapitre 5. Les reconstructions 3D proposées, d'une résolution d'environ 9 Å, ont permis de dériver un modèle actualisé de l'organisation des LDL, où, par exemple, les lamelles centrales se révèlent être des bicouches, par opposition à l'image classique considérant des monocouches. Les cartes 3D présentées ici peuvent être utilisés comme un support pour l'interprétation des courbes SAXS et le développement de modèles associés, ou la création de modèles de départ pour des simulations MD. Cependant, malgré une statistique élevée (environ 700k particules), et des analyses d'images poussées, aucune structure secondaire de la protéine apo B-100 n'a pu être révélée. Seules deux densités supplémentaires ont été identifiées, que nous avons arbitrairement appelées protrusions «frontale» et «latérale». Nous avons associé la protrusion «frontale» au domaine globulaire $NH_2 - \beta\alpha_1$ de la protéine. Quant à la protrusion «latérale», qui n'avait jamais été observée auparavant, nous avons fait l'hypothèse qu'elle correspond aux glycans attachés à ce qui serait le domaine du récepteur de liaison. Enfin, étant donné l'absence de toute structure secondaire, même après des analyses approfondies, nous avons avancé l'hypothèse que la protéine apo B-100 serait intrinsèquement flexible sur la surface-même des LDLs.

À l'heure actuelle, la grande proportion des lipides et la flexibilité intrinsèque d'apo B-100 rendent irréalisable avec les outils actuels la reconstruction 3D à haute résolution des LDLs. En présentant l'état de nos travaux et les différents essais menés, tout en diffusant les données brutes, notamment au travers de la base de données

EMPIAR [279], nous souhaitons initier de nouvelles discussions dans la communauté cryo-EM, et l'émergence de nouveaux algorithmes pour reconstruire de tels complexes.

Dans l'ensemble, les études menées dans le cadre de cette thèse ont permis de dégager une image plus dynamique de la protéine apo B-100. Sous sa forme solubilisée, elle semble à l'échelle de temps la plus lente (ns) être assez contrainte par son environnement (à savoir le détergent NP40), tout en manifestant à l'échelle de la ps une dynamique importante, comparable à celle des lipides (voir le chapitre 4). En association avec la LDL, apo B-100 semble également présenter une flexibilité intrinsèque, qui est indirectement mise en évidence par l'absence de toute structure secondaire dans les cartes cryo-EM, et des zones floues pour certaines régions exposées à la surface, comme le domaine globulaire $NH_2 - \beta\alpha_1$ (Chapitre 5).

Élucider la structure de la protéine apo B-100 est une perspective qui découle naturellement de ces travaux de thèse. Une structure d'apo B-100 aiderait à déchiffrer les reconstructions 3D des LDLs, mais aussi à mieux comprendre certaines fonctionnalités des LDLs, comme leurs propriétés de liaison. Pour ce faire, les conditions de préparation d'apo B-100 doivent encore être optimisées pour la cryo-EM, afin d'éviter l'agrégation observée pendant les premiers tests (voir dans le chapitre 5). En première approche, nous pouvons citer l'ajustement des conditions de la solution, ou l'ajout de tensioactifs, qui sont des pratiques courantes en cryo-EM, comme présenté dans [227]. Des stratégies plus élaborées peuvent également être envisagées pour reconstituer un environnement lipidique minimal pour apo B-100, comme l'utilisation de nanodisques [280] ou de liposomes [271]. Cependant, la préparation des échantillons devra également être optimisée au niveau de la vitrification, avec des essais sur différents types de grilles (par exemple, les grilles recouvertes de carbone ont donné de meilleurs résultats avec les LDLs), ou des paramètres de blotting. Ils seront discutés plus en détail dans le cadre de nouvelles thèses qui débutent à l'Université médicale de Graz (Autriche), et de la venue du Dr Karin Kornmueller dans le groupe du Pr. Judith Peters (Grenoble, France) pendant plusieurs mois en 2022.

Entre-temps, suivant la récente publication d'AlphaFold pour prédire la structure des protéines [281], un modèle théorique pour apo B-100 pourrait être dérivé de sa séquence primaire, et sera également à l'étude. Il viendrait compléter les prédictions de la structure secondaire [68], dont le modèle «pentapartite» reste aujourd'hui une référence, et le dernier modèle homologue du domaine $NH_2 - \beta\alpha_1$ [75]. De plus, parce que cette structure serait directement dérivée de la séquence primaire d'apo B-100, la structure tertiaire calculée serait plus précise que le modèle proposé par Kriško and Etchebest [78]. Toujours en vue des expériences de cryo-EM, un modèle 3D initial aiderait au traitement des images et à l'interprétation des reconstructions d'apo B-100. De plus, si les cartes 3D finales ont une résolution suffisante, un modèle atomique de

la protéine pourrait être construit à partir da la structure théorique. Enfin, un modèle tertiaire est nécessaire pour développer des simulations MD, qui constituent une autre piste potentielle de recherche.

Dans le cas de la LDL entière, les simulations MD précédentes [102, 109] doivent être élargies. Tout d'abord, elles devraient être appliquées à des températures inférieures à la transition du coeur, afin d'étudier son organisation, mais aussi afin d'intégrer les informations provenant des cartes 3D que nous proposons ici. De plus, la dynamique intrinsèque de la protéine à la surface des LDL devrait être prise en compte. Contrairement aux simulations MD proposées dans [109] où apo B-100 repose sur la surface des LDL, il faut considérer des parties enfouies, ou des feuillets β au même niveau que la monocouche de phospholipides. En effet, même dans les micrographes bruts de cryo-EM, nous n'avons pas observé de fortes densités s'étendant hors de la surface des LDL, à l'exception des protrusions. Evidemment, de telles simulations MD sont coûteuses en termes de puissance de calcul, notamment à cause de la grande proportion et variété des lipides composant la LDL. Des simulations à gros grains seraient adaptées pour commencer, ainsi que pour atteindre de longues échelles de temps. Néanmoins, pour avoir une image plus précise de la dynamique d'apo B-100, il pourrait être intéressant de réaliser des simulations MD sur la protéine apo B-100 seule. En outre, elles se révéleraient un excellent complément aux potentielles études cryo-EM d'apo B-100, que nous avons évoqué plus haut. En particulier, si la protéine présente plusieurs conformations, ce qui a déjà été observé par Gantz et al. [29] ou dans nos images EM à coloration négative, nous devons pouvoir les différencier dans l'analyse des images et les reconstructions 3D. Pour cela, les simulations MD sont une méthode de choix. Dans certains cas extrêmes, elles ont même permis de détecter directement dans les tomogrammes bruts les différentes conformations d'une protéine à la surface d'un virus [282].

De manière concomitante aux simulations MD, les études dynamiques comme celles présentées dans le chapitre 4 avec la diffusion incohérente des neutrons restent essentielles, car elles permettent de quantifier directement et expérimentalement les processus de diffusion à l'échelle de la ps-ns. L'étude que nous avons réalisée sur apo B-100 pourrait aider à construire les simulations MD correspondantes. Nous pouvons envisager de poursuivre ces expériences, et de déterminer encore plus précisément la dynamique d'apo B-100. Par exemple, des mesures par spin-echo de neutrons peuvent être considérées pour atteindre des échelles de temps plus longues. Cela nécessiterait cependant quelques ajustements car l'échantillon est lyophilisé, et non en solution. D'un autre côté, pour atteindre la dynamique d'apo B-100 à la surface des LDLs, et au vue de l'évolution des techniques de deutération des dernières années, des parties de la LDL pourraient être deutérées afin de se concentrer sur la protéine, comme les phospholipides (de manière similaires aux premières études par SANS [18]), ou le

cholestérol, qu'il est maintenant possible de deutérer entièrement [283].

Enfin, la caractérisation moléculaire des formes modifiées des LDLs est également cruciale. En prolongeant les mesures précédentes de SAXS, SANS et de diffusion incohérente des neutrons [23, 36], des études de cryo-EM pourraient être envisagées, notamment sur les LDLs riches en triglycérides (TG), dont les premières images ont montré qu'elles avaient une forme plus arrondie et qu'elles n'avaient pas d'arrangement interne [97]. La protéine apo B-100 pourrait être le sujet de questions intéressante, notamment les effets des formes modifiées des LDL sur sa structure, mais aussi sa dynamique. De plus, si jamais la dynamique d'apo B-100 se révélait plus faible sur les TG-LDLs, cela pourrait faciliter la détermination d'une partie de sa structure. Évidemment, le contraire pourrait être tout aussi vrai, avec une plus grande dynamique de l'apo B-100 sur les TG-LDLs, mais l'alignement des images serait pour sa part moins influencé par l'organisation du noyau, puisqu'on s'attend à ce qu'il soit désordonné, comme c'est le cas au-dessus de la transition de phase.

Pour conclure, les résultats présentés dans cette thèse font partie d'un tableau plus vaste. Les informations dynamiques que nous avons déduits sur apo B-100, le modèle 3D actualisé de la LDL, complété ultérieurement par de futurs détails structuraux de la protéine, une meilleure interprétation des courbes SAXS, ou des simulations MD plus élaborées, parmi les nombreuses pistes de recherche envisagées, constituent dans leur ensemble des informations essentielles pour mieux comprendre la fonctionnalité de la LDL, et son lien avec les maladies cardiovasculaires.

Résumé des chapitres (FR)

Ce manuscript de thèse s'articule autour de cinq chapitres, pour lesquels nous fournissons un résumé dans les paragraphes suivants.

Chapitre 1: Les lipoprotéines à basse densité.

Les lipoprotéines à basse densité (LDLs) font partie de la famille des lipoprotéines, qui transportent les acides gras dans notre corps. En particulier, les LDLs ont un rôle essentiel dans le transport du cholestérol. Ce chapitre commence par une présentation des lipoprotéines et un aperçu des voies métaboliques associées, pour ensuite se concentrer sur les LDLs. Le chapitre dresse ensuite un inventaire exhaustif des propriétés des LDLs à l'échelle moléculaire, avec un accent mis sur les études de sa structure et sa dynamique. Il se termine par une vue d'ensemble de la préparation des LDLs à partir du plasma humain, ainsi que la solubilization de la protéine apo B-100, dont les protocoles présentés ont été mis au point par Dr. Karin Kornmueller, de l'Université médicale de Graz (Autriche). La présentation des objectifs de la thèse dans ce contexte global clôture le chapitre.

Chapitre 2: Théorie.

Les deux techniques qui constituent le coeur de la thèse, à savoir la diffusion incohérente de neutrons et la cryo-microscopie électronique (cryo-EM) reposent sur la théorie de la diffusion et de l'interaction lumière-matière. Dans ce chapitre, nous présentons des éléments de théorie, en partant de l'équation de Schrödinger, et en mettant en avant les approximations et hypothèses avancées pour interpréter les données expérimentales. En particulier, en diffusion de neutrons, l'approximation de Born est incontournable, et nécessite de présenter le formalisme quantique qui sous-tend l'interaction des neutrons avec la matière. Le passage à la description classique est également décrit, et brièvement discuté. Du côté des électrons, l'observation directe des images en sortie d'un microscope électronique découle aussi de la diffusion des électrons par l'échantillon. Quelques éléments sur la formation des images par contraste de phase

sont amenés, et les approximations qui constituent le socle des analyses d'images de cryo-EM sont exposés.

Chapitre 3: Instrumentation. Logiciels.

Dans un chapitre plus technique, nous décrivons la production des particules sondes utilisées, neutrons et électrons, jusque leur détection, en passant par les instruments dans lesquels nous les avons respectivement utilisés, spectromètres et microscopes. Dans le cas de la cryo-EM, nous présentons de manière succinte le type d'analyse que nous avons appliqué sur nos images, pour aboutir aux reconstructions 3D présentées plus tard. Dans le cas des deux techniques, nous discutons en quelques mots des potentiels développements qu'elles pourraient connaître à l'avenir.

Chapitre 4: La dynamique d'apo B-100 examinée par diffusion incohérente de neutrons.

La diffusion incohérente de neutrons (EINS et QENS) a été appliquée à l'échantillon d'apo B-100 solubilisée, afin d'en extraire sa dynamique et de pouvoir la comparer à des précédents résultats sur les LDLs entières. Néanmoins, la présence résiduelle du détergent utilisé pour la solubilisation de la protéine, à savoir le Nonidet P-40 (NP40), a dû être prise en compte dans les analyses. Ce chapitre décrit l'approche innovante que nous avons mis en place pour traiter cette difficulté, puis les informations dynamiques que nous avons tiré de l'échantillon d'apo B-100 / NP40.

Chapitre 5: Exploration de l'organisation des LDLs par cryo-microscopie électronique.

Dans ce dernier chapitre, nous explorons la structure des LDLs grâce à la cryo-EM. Nous y décrivons d'abord les divers essais effectuées, en «particule seule» ou en tomographie, sur LDLs ou apo B-100 solubilisée. Ensuite, nous nous concentrons sur les données collectées des LDLs, notamment sur le Titan Krios de l'ESRF (Grenoble, France), et les différentes étapes d'analyses des images. Les classes 2D sont présentées, puis les reconstructions 3D des deux logiciels, RELION et cryoSPARC, sont décrites en détail. Nous discutons aussi des essais d'analyse de variabilité menés, et enfin nous proposons un modèle actualisé de l'organisation des LDLs.

References

[1] H. J. Pownall and A. M. Gotto Jr. Human plasma apolipoproteins in biology and medicine. Structure and function of apolipoproteins, pages 1–32, 1992.

[2] J. P. Segrest, D. W. Garber, C. G. Brouillette, S. C. Harvey, and G. Anantharamaiah. The amphipathic α helix: A multifunctional structural motif in plasma apolipoproteins. In Lipoproteins, Apolipoproteins, and Lipases, pages 303–369. Elsevier, 1994. URL https://doi.org/10.1016/s0065-3233(08)60643-9.

[3] G. Kostner and P. Laggner. Human plasma lipoproteins, 1989.

[4] T. Hevonoja, M. O. Pentikäinen, M. T. Hyvönen, P. T. Kovanen, and M. Ala-Korpela. Structure of low density lipoprotein (LDL) particles: Basis for understanding molecular changes in modified LDL. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 1488(3):189–210, November 2000. URL https://doi.org/10.1016/s1388-1981(00)00123-2.

[5] R. Prassl and P. Laggner. Molecular structure of low density lipoprotein: current status and future challenges. European Biophysics Journal, 38(2):145–158, September 2008. URL https://doi.org/10.1007/s00249-008-0368-y.

[6] T. E. R. F. Collaboration. Major lipids, apolipoproteins, and risk of vascular disease. JAMA, 302(18):1993, November 2009. URL https://doi.org/10.1001/jama.2009.1619.

[7] A. J. Lusis. Atherosclerosis. Nature, 407(6801):233–241, September 2000. URL https://doi.org/10.1038/35025203.

[8] H. Esterbauer, J. Gebicki, H. Puhl, and G. Jürgens. The role of lipid peroxidation and antioxidants in oxidative modification of LDL. Free Radical Biology and Medicine, 13(4): 341–390, October 1992. URL https://doi.org/10.1016/0891-5849(92)90181-f.

[9] B. J. McKeone, J. R. Patsch, and H. J. Pownall. Plasma triglycerides determine low density lipoprotein composition, physical properties, and cell-specific binding in cultured cells. Journal of Clinical Investigation, 91(5):1926–1933, May 1993. URL https://doi.org/10.1172/jci116411.

[10] B. A. Griffin, D. J. Freeman, G. W. Tait, J. Thomson, M. J. Caslake, C. J. Packard, and J. Shepherd. Role of plasma triglyceride in the regulation of plasma low density lipoprotein (LDL) subfractions: relative contribution of small, dense LDL to coronary heart disease risk. Atherosclerosis, 106(2):241–253, April 1994. URL https://doi.org/10.1016/0021-9150(94)90129-5.

[11] B. A. Griffin. Lipoprotein atherogenicity: an overview of current mechanisms. Proceedings of the Nutrition Society, 58(1):163–169, 1999.

[12] C. Packard, M. Caslake, and J. Shepherd. The role of small, dense low density lipoprotein (LDL): a new look. International Journal of Cardiology, 74:S17–S22, June 2000. URL https://doi.org/10.1016/s0167-5273(99)00107-2.

[13] C. J. Packard. Small dense low-density lipoprotein and its role as an independent predictor of cardiovascular disease. Current Opinion in Lipidology, 17(4):412–417, August 2006. URL https://doi.org/10.1097/01.mol.0000236367.42755.c1.

[14] H. Yoshida and R. Kisugi. Mechanisms of LDL oxidation. Clinica Chimica Acta, 411(23-24): 1875–1882, December 2010. URL https://doi.org/10.1016/j.cca.2010.08.038.

[15] D. Atkinson, R. J. Deckelbaum, D. M. Small, and G. G. Shipley. Structure of human plasma low-density lipoproteins: molecular organization of the central core. Proceedings of the National Academy of Sciences, 74(3):1042–1046, March 1977. URL https://doi.org/10.1073/pnas.74.3.1042.

[16] R. J. Deckelbaum, G. G. Shipley, and D. M. Small. Structure and interactions of lipids in human plasma low density lipoproteins. Journal of Biological Chemistry, 252(2):744–754, January 1977. URL https://doi.org/10.1016/s0021-9258(17)32781-3.

[17] K. Muller, P. Laggner, O. Glatter, and G. Kostner. The structure of human-plasma low-density lipoprotein b. an x-ray small-angle scattering study. European Journal of Biochemistry, 82(1): 73–90, January 1978. URL https://doi.org/10.1111/j.1432-1033.1978.tb11998.x.

[18] P. Laggner, G. M. Kostner, U. Rakusch, and D. Worcester. Neutron small angle scattering on selectively deuterated human plasma low density lipoproteins. the location of polar phospholipid headgroups. J. Biol. Chem., 256(22):11832–11839, nov 1981.

[19] A. Johs, M. Hammel, I. Waldner, R. P. May, P. Laggner, and R. Prassl. Modular structure of solubilized human apolipoprotein b-100. Journal of Biological Chemistry, 281(28):19732–19739, July 2006. URL https://doi.org/10.1074/jbc.m601688200.

[20] R. Prassl, M. Pregetter, H. Amenitsch, M. Kriechbaum, R. Schwarzenbacher, J. M. Chapman, and P. Laggner. Low density lipoproteins as circulating fast temperature sensors. PLoS ONE, 3(12):e4079, December 2008. URL https://doi.org/10.1371/journal.pone.0004079.

[21] M. Papi, R. Brunelli, G. Ciasca, A. Maiorana, G. Maulucci, V. Palmieri, T. Parasassi, and M. D. Spirito. Estradiol protective role in atherogenesis through LDL structure modification. Journal of Physics D: Applied Physics, 49(28):285402, June 2016. URL https://doi.org/10.1088/0022-3727/49/28/285402.

[22] S. Maric, T. K. Lind, J. Lyngsø, M. Cárdenas, and J. S. Pedersen. Modeling small-angle x-ray scattering data for low-density lipoproteins: Insights into the fatty core packing and phase transition. ACS Nano, 11(1):1080–1090, January 2017. URL https://doi.org/10.1021/acsnano.6b08089.

[23] B. Lehofer, M. Golub, K. Kornmueller, M. Kriechbaum, N. Martinez, G. Nagy, J. Kohlbrecher, H. Amenitsch, J. Peters, and R. Prassl. High hydrostatic pressure induces a lipid phase transition and molecular rearrangements in low-density lipoprotein nanoparticles. Particle & Particle Systems Characterization, 35(9):1800149, July 2018. URL https://doi.org/10.1002/ppsc.201800149.

[24] R. Prassl, J. M. Chapman, F. Nigon, M. Sara, S. Eschenburg, C. Betzel, A. Saxena, and P. Laggner. Crystallization and preliminary x-ray analysis of a low density lipoprotein from human plasma. Journal of Biological Chemistry, 271(46):28731–28733, November 1996. URL https://doi.org/10.1074/jbc.271.46.28731.

[25] V. Y. Lunin, N. L. Lunina, S. Ritter, I. Frey, A. Berg, K. Diederichs, A. D. Podjarny, A. Urzhumtsev, and M. W. Baumstark. Low-resolution data analysis for low-density lipoprotein particle. Acta Crystallographica Section D Biological Crystallography, 57(1):108–121, January 2001. URL https://doi.org/10.1107/s0907444900014608.

[26] J. E. Chatterton, M. L. Phillips, L. K. Curtiss, R. W. Milne, Y. L. Marcel, and V. N. Schumaker. Mapping apolipoprotein b on the low density lipoprotein surface by immunoelectron microscopy. Journal of Biological Chemistry, 266(9):5955–5962, March 1991. URL https://doi.org/10.1016/s0021-9258(19)67691-x.

[27] J. Spin and D. Atkinson. Cryoelectron microscopy of low density lipoprotein in vitreous ice. Biophysical Journal, 68(5):2115–2123, May 1995. URL https://doi.org/10.1016/s0006-3495(95)80392-9.

[28] E. V. Orlova, M. B. Sherman, W. Chiu, H. Mowri, L. C. Smith, and A. M. Gotto. Three-dimensional structure of low density lipoproteins by electron cryomicroscopy. Proceedings of the National Academy of Sciences, 96(15):8420–8425, July 1999. URL https://doi.org/10.1073/pnas.96.15.8420.

[29] D. L. Gantz, M. T. Walsh, and D. M. Small. Morphology of sodium deoxycholate-solubilized apolipoprotein b-100 using negative stain and vitreous ice electron microscopy. Journal of Lipid Research, 41(9):1464–1472, September 2000. URL https://doi.org/10.1016/s0022-2275(20)33459-3.

[30] G. Ren, G. Rudenko, S. J. Ludtke, J. Deisenhofer, W. Chiu, and H. J. Pownall. Model of human low-density lipoprotein and bound receptor based on CryoEM. Proceedings of the National Academy of Sciences, 107(3):1059–1064, December 2009. URL https://doi.org/10.1073/pnas.0908004107.

[31] V. Kumar, S. J. Butcher, K. Öörni, P. Engelhardt, J. Heikkonen, K. Kaski, M. Ala-Korpela, and P. T. Kovanen. Three-dimensional cryoEM reconstruction of native LDL particles to 16å resolution at physiological body temperature. PLoS ONE, 6(5):e18841, May 2011. URL https://doi.org/10.1371/journal.pone.0018841.

[32] Y. Liu and D. Atkinson. Enhancing the contrast of ApoB to locate the surface components in the 3d density map of human LDL. Journal of Molecular Biology, 405(1):274–283, January 2011. URL https://doi.org/10.1016/j.jmb.2010.10.034.

[33] Y. Liu and D. Atkinson. Immuno-electron cryo-microscopy imaging reveals a looped topology of apoB at the surface of human LDL. Journal of Lipid Research, 52(6):1111–1116, June 2011. URL https://doi.org/10.1194/jlr.m013946.

[34] Y. Liu, D. Luo, and D. Atkinson. Human LDL core cholesterol ester packing: three-dimensional image reconstruction and SAXS simulation studies. Journal of Lipid Research, 52(2):256–262, February 2011. URL https://doi.org/10.1194/jlr.m011569.

[35] C. Mikl, J. Peters, M. Trapp, K. Kornmueller, W. J. Schneider, and R. Prassl. Softness of atherogenic lipoproteins: A comparison of very low density lipoprotein (VLDL) and low density lipoprotein (LDL) using elastic incoherent neutron scattering (EINS). Journal of the American Chemical Society, 133(34):13213–13215, August 2011. URL https://doi.org/10.1021/ja203679g.

[36] M. Golub, B. Lehofer, N. Martinez, J. Ollivier, J. Kohlbrecher, R. Prassl, and J. Peters. High hydrostatic pressure specifically affects molecular dynamics and shape of low-density lipoprotein particles. Scientific Reports, 7(1), April 2017. URL https://doi.org/10.1038/srep46034.

[37] J. Peters, N. Martinez, B. Lehofer, and R. Prassl. Low-density lipoproteins investigated under high hydrostatic pressure by elastic incoherent neutron scattering. The European Physical Journal E, 40(7), July 2017. URL https://doi.org/10.1140/epje/i2017-11558-8.

[38] D. Vance. Biochemistry of lipids, lipoproteins, and membranes. Elsevier, Amsterdam Boston, 2002. ISBN 0-444-51138-5.

[39] M. S. Sekhar, S. Marupuru, B. S. Reddy, S. J. Kurian, and M. Rao. Physiological role of cholesterol in human body. In Dietary Sugar, Salt and Fat in Human Health, pages 453–481. Elsevier, 2020. URL https://doi.org/10.1016/b978-0-12-816918-6.00021-4.

[40] D. A. Bricarello, J. T. Smilowitz, A. M. Zivkovic, J. B. German, and A. N. Parikh. Reconstituted lipoprotein: A versatile class of biologically-inspired nanostructures. ACS Nano, 5(1):42–57, December 2010. URL https://doi.org/10.1021/nn103098m.

[41] M. J. Chapman, P. M. Laplaud, G. Luc, P. Forgez, E. Bruckert, S. Goulinet, and D. Lagrange. Further resolution of the low density lipoprotein spectrum in normal human plasma: physicochemical characteristics of discrete subspecies separated by density gradient ultracentrifugation. Journal of Lipid Research, 29(4):442–458, April 1988. URL https://doi.org/10.1016/s0022-2275(20)38518-7.

[42] R. A. Cox and M. R. García-Palmieri. Cholesterol, triglycerides, and associated lipoproteins, 1990. URL https://www.ncbi.nlm.nih.gov/books/NBK351.

[43] J. P. Kane. Apolipoprotein b: Structural and metabolic heterogeneity. Annual Review of Physiology, 45(1):637–650, October 1983. URL https://doi.org/10.1146/annurev.ph.45.030183.003225.

[44] J. Dixon and H. Ginsberg. Hepatic synthesis of lipoproteins and apolipoproteins. Seminars in Liver Disease, 12(04):364–372, November 1992. URL https://doi.org/10.1055/s-2008-1040406.

[45] A. Zhyvotovska, D. Yusupov, and S. I. McFarlane. Introductory chapter: Overview of lipoprotein metabolism. In Dyslipidemia. IntechOpen, December 2019. URL https://doi.org/10.5772/intechopen.85094.

[46] M. S. Brown and J. L. Goldstein. Receptor-mediated endocytosis: insights from the lipoprotein receptor system. Proceedings of the National Academy of Sciences, 76(7):3330–3337, July 1979. URL https://doi.org/10.1073/pnas.76.7.3330.

[47] J. L. GOLDSTEIN and M. S. BROWN. The LDL receptor and the regulation of cellular cholesterol metabolism. Journal of Cell Science, 1985(Supplement_3):131–137, February 1985. URL https://doi.org/10.1242/jcs.1985.supplement_3.13.

[48] B. F. Voight, G. M. Peloso, M. Orho-Melander, R. Frikke-Schmidt, M. Barbalic, M. K. Jensen, G. Hindy, H. Holm, E. L. Ding, T. Johnson, H. Schunkert, N. J. Samani, R. Clarke, J. C. Hopewell, J. F. Thompson, M. Li, G. Thorleifsson, C. Newton-Cheh, K. Musunuru, J. P. Pirruccello, D. Saleheen, L. Chen, A. F. Stewart, A. Schillert, U. Thorsteinsdottir, G. Thorgeirsson, S. Anand, J. C. Engert, T. Morgan, J. Spertus, M. Stoll, K. Berger, N. Martinelli, D. Girelli, P. P. McKeown, C. C. Patterson, S. E. Epstein, J. Devaney, M.-S. Burnett, V. Mooser, S. Ripatti, I. Surakka, M. S. Nieminen, J. Sinisalo, M.-L. Lokki, M. Perola, A. Havulinna, U. de Faire, B. Gigante, E. Ingelsson, T. Zeller, P. Wild, P. I. W. de Bakker, O. H. Klungel, A.-H. M. van der Zee, B. J. M. Peters, A. de Boer, D. E. Grobbee, P. W. Kamphuisen, V. H. M. Deneer, C. C. Elbers, N. C. Onland-Moret, M. H. Hofker, C. Wijmenga, W. M. Verschuren, J. M. Boer, Y. T. van der Schouw, A. Rasheed, P. Frossard, S. Demissie, C. Willer, R. Do, J. M. Ordovas, G. R. Abecasis, M. Boehnke, K. L. Mohlke, M. J. Daly, C. Guiducci, N. P. Burtt, A. Surti, E. Gonzalez, S. Purcell, S. Gabriel, J. Marrugat, J. Peden, J. Erdmann, P. Diemert, C. Willenborg, I. R. König, M. Fischer, C. Hengstenberg, A. Ziegler, I. Buysschaert, D. Lambrechts, F. V. de Werf, K. A. Fox, N. E. E. Mokhtari, D. Rubin, J. Schrezenmeir, S. Schreiber, A. Schäfer, J. Danesh, S. Blankenberg, R. Roberts, R. McPherson, H. Watkins, A. S. Hall, K. Overvad, E. Rimm, E. Boerwinkle, A. Tybjaerg-Hansen, L. A. Cupples, M. P. Reilly, O. Melander, P. M. Mannucci, D. Ardissino, D. Siscovick, R. Elosua, K. Stefansson, C. J. O'Donnell, V. Salomaa, D. J. Rader, L. Peltonen, S. M. Schwartz, D. Altshuler, and S. Kathiresan. Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study. The Lancet, 380(9841):572–580, August 2012. URL https://doi.org/10.1016/s0140-6736(12)60312-2.

[49] K. Skålén, M. Gustafsson, E. K. Rydberg, L. M. Hultén, O. Wiklund, T. L. Innerarity, and J. Borén. Subendothelial retention of atherogenic lipoproteins in early atherosclerosis. Nature, 417(6890):750–754, June 2002. URL https://doi.org/10.1038/nature00804.

[50] J. Borén and K. J. Williams. The central role of arterial retention of cholesterol-rich apolipoprotein-b-containing lipoproteins in the pathogenesis of atherosclerosis. Current Opinion in Lipidology, 27(5):473–483, October 2016. URL https://doi.org/10.1097/mol.0000000000000330.

[51] Y. V. Bobryshev. Monocyte recruitment and foam cell formation in atherosclerosis. Micron, 37(3):208–222, April 2006. URL https://doi.org/10.1016/j.micron.2005.10.007.

[52] X.-H. Yu, Y.-C. Fu, D.-W. Zhang, K. Yin, and C.-K. Tang. Foam cells in atherosclerosis. Clinica Chimica Acta, 424:245–252, September 2013. URL https://doi.org/10.1016/j.cca.2013.06.006.

[53] R. A. Hegele. Plasma lipoproteins: genetic influences and clinical implications. Nature Reviews Genetics, 10(2):109–121, February 2009. URL https://doi.org/10.1038/nrg2481.

[54] S. Lund-Katz, P. M. Laplaud, M. C. Phillips, and M. J. Chapman. Apolipoprotein b-100 conformation and particle surface charge in human LDL subspecies: implication for LDL receptor interaction. Biochemistry, 37(37):12867–12874, August 1998. URL https://doi.org/10.1021/bi980828m.

[55] F. Ursini, K. J. Davies, M. Maiorino, T. Parasassi, and A. Sevanian. Atherosclerosis: another protein misfolding disease? Trends in Molecular Medicine, 8(8):370–374, August 2002. URL https://doi.org/10.1016/s1471-4914(02)02382-1.

[56] J. L. Sánchez-Quesada, S. Villegas, and J. Ordóñez-Llanos. Electronegative low-density lipoprotein. a link between apolipoprotein b misfolding, lipoprotein aggregation and proteoglycan binding. Current Opinion in Lipidology, 23(5):479–486, October 2012. URL https://doi.org/10.1097/mol.0b013e328357c933.

[57] E. I. O'Leary, Z. Jiang, M.-P. Strub, and J. C. Lee. Effects of phosphatidylcholine membrane fluidity on the conformation and aggregation of n-terminally acetylated α-synuclein. Journal of Biological Chemistry, 293(28):11195–11205, July 2018. URL https://doi.org/10.1074/jbc.ra118.002780.

[58] R. A. Gottlieb and B. M. Babior. Regulation of fas-mediated apoptosis. In Current Topics in Cellular Regulation, pages 69–105. Elsevier, 1997. URL https://doi.org/10.1016/s0070-2137(97)80003-9.

[59] B. Lindshield. Kansas State University Human Nutrition Flexbook, 2021. URL https://courses.lumenlearning.com/suny-nutrition/chapter/2-35-triglycerides/.

[60] C. I. Plc. Avanti Polar Lipids, 2021. URL https://avantilipids.com.

[61] S. H. Chen, C. Y. Yang, P. F. Chen, D. Setzer, M. Tanimura, W. H. Li, A. M. Gotto, and L. Chan. The complete cDNA and amino acid sequence of human apolipoprotein b-100. Journal of Biological Chemistry, 261(28):12918–12921, October 1986. URL https://doi.org/10.1016/s0021-9258(18)69248-8.

[62] V. N. Schumaker, M. L. Phillips, and J. E. Chatterton. Apolipoprotein b and low-density lipoprotein structure: Implications for biosynthesis of triglyceride-rich lipoproteins. In Lipoproteins, Apolipoproteins, and Lipases, pages 205–248. Elsevier, 1994. URL https://doi.org/10.1016/s0065-3233(08)60641-5.

[63] S. S. I. of Bioinformatics. ExPASy ProtParam database, 2021. URL https://web.expasy.org/protparam/. Accession number: APOB_HUMAN (P04114).

[64] S. Duvaud, C. Gabella, F. Lisacek, H. Stockinger, V. Ioannidis, and C. Durinx. Expasy, the swiss bioinformatics resource portal, as designed by its users. Nucleic Acids Research, 49(W1): W216–W227, April 2021. URL https://doi.org/10.1093/nar/gkab225.

[65] C.-Y. Yang, S.-H. Chen, S. H. Gianturco, W. A. Bradley, J. T. Sparrow, M. Tanimura, W.-H. Li, D. A. Sparrow, H. DeLoof, M. Rosseneu, F.-S. Lee, Z.-W. Gu, A. M. Gotto, and L. Chan. Sequence, structure, receptor-binding domains and internal repeats of human apolipoprotein b-100. Nature, 323(6090):738–742, October 1986. URL https://doi.org/10.1038/323738a0.

[66] T. J. Knott, R. J. Pease, L. M. Powell, S. C. Wallis, S. C. Rall, T. L. Innerarity, B. Blackhart, W. H. Taylor, Y. Marcel, R. Milne, D. Johnson, M. Fuller, A. J. Lusis, B. J. McCarthy, R. W. Mahley, B. Levy-Wilson, and J. Scott. Complete protein sequence and identification of structural domains of human apolipoprotein b. Nature, 323(6090):734–738, October 1986. URL https://doi.org/10.1038/323734a0.

[67] J. P. Segrest, M. K. Jones, H. D. Loof, and N. Dashti. Structure of apolipoprotein b-100 in low density lipoproteins. Journal of Lipid Research, 42(9):1346–1367, September 2001. URL https://doi.org/10.1016/s0022-2275(20)30267-4.

[68] J. P. Segrest, M. K. Jones, V. K. Mishra, G. M. Anantharamaiah, and D. W. Garber. apoB-100 has a pentapartite structure composed of three amphipathic alpha-helical domains alternating with two amphipathic beta-strand domains. detection by the computer program LOCATE. Arteriosclerosis and Thrombosis: A Journal of Vascular Biology, 14(10):1674–1685, October 1994. URL https://doi.org/10.1161/01.atv.14.10.1674.

[69] E. Goormaghtigh, J. D. Meutter, B. Vanloo, R. Brassuer, M. Rosseneu, and J.-M. Ruysschaert. Evaluation of the secondary structure of apo b-100 in low-density lipoprotein (LDL) by infrared spectroscopy. Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism, 1006(1): 147–150, November 1989. URL https://doi.org/10.1016/0005-2760(89)90338-x.

[70] T. Anderson, D. Levitt, and L. Banaszak. The structural basis of lipid interactions in lipovitellin, a soluble lipoprotein. Structure, 6(7):895–909, July 1998. URL https://doi.org/10.1016/s0969-2126(98)00091-4.

[71] C. J. Mann, T. A. Anderson, J. Read, S. Chester, G. B. Harrison, S. Köchl, P. J. Ritchie, P. Bradbury, F. S. Hussain, J. Amey, B. Vanloo, M. Rosseneu, R. Infante, J. M. Hancock, D. G. Levitt, L. J. Banaszak, J. Scott, and C. C. Shoulders. The structure of vitellogenin provides a molecular model for the assembly and secretion of atherogenic lipoproteins 1 1edited by a. r. fersht. Journal of Molecular Biology, 285(1):391–408, January 1999. URL https://doi.org/10.1006/jmbi.1998.2298.

[72] H. Herscovitz, A. Derksen, M. T. Walsh, C. J. McKnight, D. L. Gantz, M. Hadzopoulou-Cladaras, V. Zannis, C. Curry, and D. M. Small. The n-terminal 17% of apoB binds tightly and irreversibly to emulsions modeling nascent very low density lipoproteins. Journal of Lipid Research, 42(1):51–59, January 2001. URL https://doi.org/10.1016/s0022-2275(20)32335-x.

[73] N. Dashti, M. Gandhi, X. Liu, X. Lin, and J. P. Segrest. The n-terminal 1000 residues of apolipoprotein b associate with microsomal triglyceride transfer protein to create a lipid transfer pocket required for lipoprotein assembly†. Biochemistry, 41(22):6978–6987, June 2002. URL https://doi.org/10.1021/bi011757l.

[74] Z. G. Jiang, M. Carraway, and C. J. McKnight. Limited proteolysis and biophysical characterization of the lipovitellin homology region in apolipoprotein b. Biochemistry, 44(4): 1163–1173, December 2004. URL https://doi.org/10.1021/bi048286y.

[75] P. E. Richardson, M. Manchekar, N. Dashti, M. K. Jones, A. Beigneux, S. G. Young, S. C. Harvey, and J. P. Segrest. Assembly of lipoprotein particles containing apolipoprotein-b: Structural model for the nascent lipoprotein particle. Biophysical Journal, 88(4):2789–2800, April 2005. URL https://doi.org/10.1529/biophysj.104.046235.

[76] A. Šali and T. L. Blundell. Comparative protein modelling by satisfaction of spatial restraints. Journal of molecular biology, 234(3):779–815, 1993.

[77] R. Prassl and P. Laggner. Lipoprotein structure and dynamics: Low density lipoprotein viewed as a highly dynamic and flexible nanoparticle. In Lipoproteins - Role in Health and Diseases. InTech, October 2012. URL https://doi.org/10.5772/48145.

[78] A. Kriško and C. Etchebest. Theoretical model of human apolipoprotein b100 tertiary structure. Proteins: Structure, Function, and Bioinformatics, 66(2):342–358, November 2006. URL https://doi.org/10.1002/prot.21229.

[79] S. Ritter, I. Frey, K. Diederichs, D. Grathwohl, J. Keul, and M. W. Baumstark. Crystallization and preliminary x-ray diffraction data of two different human low-density lipoprotein (LDL) subfractions. Proteins: Structure, Function, and Genetics, 28(2):293–297, June 1997. URL https://doi.org/10.1002/(sici)1097-0134(199706)28:2<293::aid-prot17>3.0.co;2-d.

[80] M. W. Baumstark, W. Kreutz, A. Berg, I. Frey, and J. Keul. Structure of human low-density lipoprotein subfractions determined by x-ray small-angle scattering. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 1037(1):48–57, January 1990. URL https://doi.org/10.1016/0167-4838(90)90100-t.

[81] R. van Antwerpen, M. La Belle, E. Navratilova, and R. M. Krauss. Structural heterogeneity of apob-containing serum lipoproteins visualized using cryo-electron microscopy. Journal of lipid research, 40(10):1827–1836, 1999.

[82] R. Deckelbaum, G. Shipley, D. Small, R. Lees, and P. George. Thermal transitions in human plasma low density lipoproteins. Science, 190(4212):392–394, October 1975. URL https://doi.org/10.1126/science.170681.

[83] P. Laggner, G. Degovics, K. W. Müller, O. Glatter, O. Kratky, G. Kostner, and A. Holasek. Molecular packing and fluidity of lipids in human serum low density lipoproteins. Hoppe-Seyler´s Zeitschrift für physiologische Chemie, 358(2):771–778, January 1977. URL https://doi.org/10.1515/bchm2.1977.358.2.771.

[84] P. Laggner and G. M. Kostner. Thermotropic changes in the surface structure of lipoprotein b from human-plasma low-density lipoproteins. a spin-label study. European Journal of Biochemistry, 84(1):227–232, March 1978. URL https://doi.org/10.1111/j.1432-1033.1978.tb12160.x.

[85] P. Laggner, G. M. Kostner, G. Degovics, and D. L. Worcester. Structure of the cholesteryl ester core of human plasma low density lipoproteins: selective deuteration and neutron small-angle scattering. Proceedings of the National Academy of Sciences, 81(14):4389–4393, July 1984. URL https://doi.org/10.1073/pnas.81.14.4389.

[86] M. Pregetter, R. Prassl, B. Schuster, M. Kriechbaum, F. Nigon, J. Chapman, and P. Laggner. Microphase separation in low density lipoproteins. Journal of Biological Chemistry, 274(3):1334–1341, January 1999. URL https://doi.org/10.1074/jbc.274.3.1334.

[87] J. E. Chatterton, M. L. Phillips, L. K. Curtiss, R. Milne, J. C. Fruchart, and V. N. Schumaker. Immunoelectron microscopy of low density lipoproteins yields a ribbon and bow model for the conformation of apolipoprotein b on the lipoprotein surface. Journal of Lipid Research, 36(9):2027–2037, September 1995. URL https://doi.org/10.1016/s0022-2275(20)41120-4.

[88] J. Boren, I. Lee, W. Zhu, K. Arnold, S. Taylor, and T. L. Innerarity. Identification of the low density lipoprotein receptor-binding site in apolipoprotein b100 and the modulation of its binding activity by the carboxyl terminus in familial defective apo-b100. Journal of Clinical Investigation, 101(5):1084–1093, March 1998. URL https://doi.org/10.1172/jci1847.

[89] C. L. Oliveira, P. R. Santos, A. M. Monteiro, and A. M. F. Neto. Effect of oxidation on the structure of human low- and high-density lipoproteins. Biophysical Journal, 106(12):2595–2605, June 2014. URL https://doi.org/10.1016/j.bpj.2014.04.049.

[90] R. Brunelli, G. Mei, E. K. Krasnowska, F. Pierucci, L. Zichella, F. Ursini, and T. Parasassi. Estradiol enhances the resistance of LDL to oxidation by stabilizing apoB-100 conformation†. Biochemistry, 39(45):13897–13903, November 2000. URL https://doi.org/10.1021/bi000341p.

[91] R. Brunelli, G. Greco, M. Barteri, E. K. Krasnowska, G. Mei, F. Natella, A. Pala, S. Rotella, F. Ursini, L. Zichella, and T. Parasassi. One site on the apoB-100 specifically binds 17-β-estradiol and regulates the overall structure of LDL. The FASEB Journal, 17(14):1–14, September 2003. URL https://doi.org/10.1096/fj.02-1181fje.

[92] D. Jakubauskas, M. Jansen, J. Lyngsø, Y. Cheng, J. S. Pedersen, and M. Cárdenas. Toward reliable low-density lipoprotein ultrastructure prediction in clinical conditions: A small-angle x-ray scattering study on individuals with normal and high triglyceride serum levels. Nanomedicine: Nanotechnology, Biology and Medicine, 31:102318, January 2021. URL https://doi.org/10.1016/j.nano.2020.102318.

[93] K. Kornmueller. 13th EBSA congress, july 24–28, 2021, vienna, austria. European Biophysics Journal, 50(S1):1–226, July 2021. URL https://doi.org/10.1007/s00249-021-01558-w. Structural variability and flexibility in low-density lipoprotein (O-73) S62.

[94] G. Forte, A. Nichols, and R. Glaeser. Electron microscopy of human serum lipoproteins using negative staining. Chemistry and Physics of Lipids, 2(4):396–408, November 1968. URL https://doi.org/10.1016/0009-3084(68)90013-3.

[95] T. Forte, K. R. Norum, J. A. Glomset, and A. V. Nichols. Plasma lipoproteins in familial lecithin: cholesterol acyltransferase deficiency: structure of low and high density lipoproteins as revealed by electron microscopy. Journal of Clinical Investigation, 50(5):1141–1148, May 1971. URL https://doi.org/10.1172/jci106586.

[96] R. V. Antwerpen and J. C. Gilkey. Cryo-electron microscopy reveals human low density lipoprotein substructure. Journal of Lipid Research, 35(12):2223–2231, December 1994. URL https://doi.org/10.1016/s0022-2275(20)39928-4.

[97] M. B. Sherman, E. V. Orlova, G. L. Decker, W. Chiu, and H. J. Pownall. Structure of triglyceride-rich human low-density lipoproteins according to cryoelectron microscopy. Biochemistry, 42(50):14988–14993, November 2003. URL https://doi.org/10.1021/bi0354738.

[98] Y. Yu, Y.-L. Kuang, D. Lei, X. Zhai, M. Zhang, R. M. Krauss, and G. Ren. Polyhedral 3d structure of human plasma very low density lipoproteins by individual particle cryo-electron tomography1. Journal of Lipid Research, 57(10):1879–1888, October 2016. URL https://doi.org/10.1194/jlr.m070375.

[99] D. Lei, Y. Yu, Y.-L. Kuang, J. Liu, R. M. Krauss, and G. Ren. Single-molecule 3d imaging of human plasma intermediate-density lipoproteins reveals a polyhedral structure. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids, 1864(3):260–270, March 2019. URL https://doi.org/10.1016/j.bbalip.2018.12.004.

[100] L. Wang, D. D. Martin, E. Genter, J. Wang, R. S. McLeod, and D. M. Small. Surface study of apoB1694-1880, a sequence that can anchor apoB to lipoproteins and make it nonexchangeable. Journal of Lipid Research, 50(7):1340–1352, July 2009. URL https://doi.org/10.1194/jlr.m900040-jlr200.

[101] L. Wang, M. T. Walsh, and D. M. Small. Apolipoprotein b is conformationally flexible but anchored at a triolein/water interface: A possible model for lipoprotein surfaces. Proceedings of the National Academy of Sciences, 103(18):6871–6876, April 2006. URL https://doi.org/10.1073/pnas.0602213103.

[102] L. Pan and J. P. Segrest. Computational studies of plasma lipoprotein lipids. Biochimica et Biophysica Acta (BBA) - Biomembranes, 1858(10):2401–2420, October 2016. URL https://doi.org/10.1016/j.bbamem.2016.03.010.

[103] A. Y. Shih, P. L. Freddolino, A. Arkhipov, and K. Schulten. Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations. Journal of Structural Biology, 157 (3):579–592, March 2007. URL https://doi.org/10.1016/j.jsb.2006.08.006.

[104] F. Gu, M. K. Jones, J. Chen, J. C. Patterson, A. Catte, W. G. Jerome, L. Li, and J. P. Segrest. Structures of discoidal high density lipoproteins. Journal of Biological Chemistry, 285 (7):4652–4665, February 2010. URL https://doi.org/10.1074/jbc.m109.069914.

[105] M. K. Jones, A. Catte, L. Li, and J. P. Segrest. Dynamics of activation of lecithin:cholesterol acyltransferase by apolipoprotein a-i. Biochemistry, 48(47):11196–11210, November 2009. URL https://doi.org/10.1021/bi901242k.

[106] A. Hall, J. Repakova, and I. Vattulainen. Modeling of the triglyceride-rich core in lipoprotein particles. The Journal of Physical Chemistry B, 112(44):13772–13782, October 2008. URL https://doi.org/10.1021/jp803950w.

[107] V. V. Chaban and H. Khandelia. Lipid structure in triolein lipid droplets. The Journal of Physical Chemistry B, 118(35):10335–10340, August 2014. URL https://doi.org/10.1021/jp503223z.

[108] V. V. Chaban and H. Khandelia. Distribution of neutral lipids in the lipid droplet core. The Journal of Physical Chemistry B, 118(38):11145–11151, September 2014. URL https://doi.org/10.1021/jp506693d.

[109] T. Murtola, T. A. Vuorela, M. T. Hyvönen, S.-J. Marrink, M. Karttunen, and I. Vattulainen. Low density lipoprotein: structure, dynamics, and interactions of apoB-100 with lipids. Soft Matter, 7(18):8135, 2011. URL https://doi.org/10.1039/c1sm05367a.

[110] K. A. Henzler-Wildman, V. Thai, M. Lei, M. Ott, M. Wolf-Watz, T. Fenn, E. Pozharski, M. A. Wilson, G. A. Petsko, M. Karplus, C. G. Hübner, and D. Kern. Intrinsic motions along an enzymatic reaction trajectory. Nature, 450(7171):838–844, November 2007. URL https://doi.org/10.1038/nature06410.

[111] M. Bée. Quasi-elastic neutron scattering. Adam Hilger, 1988.

[112] G. Hörl, H. Froehlich, U. Ferstl, G. Ledinski, J. Binder, G. Cvirn, T. Stojakovic, M. Trauner, C. Koidl, E. Tafeit, K. Amrein, H. Scharnagl, G. Jürgens, and S. Hallström. Simvastatin efficiently lowers small LDL-IgG immune complex levels: A therapeutic quality beyond the lipid-lowering effect. PLOS ONE, 11(2):e0148210, February 2016. URL https://doi.org/10.1371/journal.pone.0148210.

[113] W. N. Cottingham and D. A. Greenwood. An Introduction to the Standard Model of Particle Physics. Cambridge University Press, 2007. URL https://doi.org/10.1017/cbo9780511791406.

[114] D. Sivia. Elementary Scattering Theory. Oxford University Press, January 2011. URL https://doi.org/10.1093/acprof:oso/9780199228676.001.0001.

[115] H. Schober. An introduction to the theory of nuclear neutron scattering in condensed matter. Journal of Neutron Research, 17(3-4):109–357, 2014. ISSN 1023-8166.

[116] G. L. Squires. Introduction to the theory of thermal neutron scattering. Dover Publications, Mineola, N.Y, 1996. ISBN 0-486-69447-X.

[117] S. W. Lovesey. Theory of neutron scattering from condensed matter. Clarendon Press, Oxford Oxfordshire, 1986. ISBN 0-19-852028-X.

[118] A. J. Dianoux. Neutron data booklet. Old City, Philadelphia, PA, 2003. ISBN 0-9704143-7-4.

[119] D. Jacobson, M. Arif, P. Huffman, and R. Satija. A new neutron imaging facility at bt-6 for the non-destructive analysis of working fuel cells. Technical report, NIST, 2003.

[120] L. Van Hove. Correlations in space and time and born approximation scattering in systems of interacting particles. Physical Review, 95(1):249–262, jul 1954.

[121] A. Rahman, K. S. Singwi, and A. Sjölander. Theory of slow neutron scattering by liquids. i. Physical Review, 126(3):986–996, may 1962.

[122] K. S. Singwi and A. Sjölander. Diffusive motions in water and cold neutron scattering. Physical Review, 119(3):863–871, August 1960. URL https://doi.org/10.1103/physrev.119.863.

[123] D. J. Bicout. Diffusion incohérente des neutrons : modèles analytiques pour la dynamique interne des protéines. Journal de Physique IV (Proceedings), 130:115–132, November 2005. URL https://doi.org/10.1051/jp4:2005130008.

[124] P. L. Hall and D. Ross. Incoherent neutron scattering functions for random jump diffusion in bounded and infinite media. Molecular Physics, 42(3):673–682, February 1981. URL https://doi.org/10.1080/00268978100100521.

[125] F. Volino and A. Dianoux. Neutron incoherent scattering law for diffusion in a potential of spherical symmetry: general formalism and application to diffusion inside a sphere. Molecular Physics, 41(2):271–279, October 1980. URL https://doi.org/10.1080/00268978000102761.

[126] S. R. Al-Ayoubi, P. H. Schummel, A. Cisse, T. Seydel, J. Peters, and R. Winter. Osmolytes modify protein dynamics and function of tetrameric lactate dehydrogenase upon pressurization. Physical Chemistry Chemical Physics, 21(24):12806–12817, 2019. URL https://doi.org/10.1039/c9cp02310k.

[127] V. Arrighi, J. S. Higgins, A. N. Burgess, and W. S. Howells. Rotation of methyl side groups in polymers: A fourier transform approach to quasielastic neutron scattering. 2. polymer blends. Macromolecules, 28(13):4622–4630, June 1995. URL https://doi.org/10.1021/ma00117a037.

[128] F. Gabel, D. Bicout, U. Lehnert, M. Tehei, M. Weik, and G. Zaccai. Protein dynamics studied by neutron scattering. Quarterly Reviews of Biophysics, 35(4):327–367, nov 2002.

[129] M. Bée. Localized and long-range diffusion in condensed matter: state of the art of QENS studies and future prospects. Chemical Physics, 292(2-3):121–141, August 2003. URL https://doi.org/10.1016/s0301-0104(03)00257-x.

[130] S. Khodadadi and A. P. Sokolov. Protein dynamics: from rattling in a cage to structural relaxation. Soft Matter, 11(25):4984–4998, 2015. URL https://doi.org/10.1039/c5sm00636h.

[131] J. C. Smith, P. Tan, L. Petridis, and L. Hong. Dynamic neutron scattering by biological systems. Annual Review of Biophysics, 47(1):335–354, 2018. URL https://doi.org/10.1146/annurev-biophys-070317-033358. PMID: 29561628.

[132] M. Grimaldo, F. Roosen-Runge, F. Zhang, F. Schreiber, and T. Seydel. Dynamics of proteins in solution. Quarterly Reviews of Biophysics, 52, 2019. URL https://doi.org/10.1017/s0033583519000027.

[133] U. Wanderlingh, G. D'Angelo, C. Branca, V. C. Nibali, A. Trimarchi, S. Rifici, D. Finocchiaro, C. Crupi, J. Ollivier, and H. D. Middendorf. Multi-component modeling of quasielastic neutron scattering from phospholipid membranes. The Journal of Chemical Physics, 140(17):174901, May 2014. URL https://doi.org/10.1063/1.4872167.

[134] D. J. Bicout, A. Cisse, T. Matsuo, and J. Peters. The dynamical matryoshka model: 1. incoherent neutron scattering functions for lipid dynamics in bilayers. bioRxiv, 2021. URL https://www.biorxiv.org/content/early/2021/09/24/2021.09.21.461198. https://www.biorxiv.org/content/early/2021/09/24/2021.09.21.461198.

[135] J. Peters and G. R. Kneller. Motional heterogeneity in human acetylcholinesterase revealed by a non-gaussian model for elastic incoherent neutron scattering. The Journal of Chemical Physics, 139(16):165102, October 2013. URL https://doi.org/10.1063/1.4825199.

[136] G. R. Kneller. Franck–condon picture of incoherent neutron scattering. Proceedings of the National Academy of Sciences, 115(43):E10283–E10283, October 2018. URL https://doi.org/10.1073/pnas.1816793115.

[137] J. Frank. Three-dimensional electron microscopy of macromolecular assemblies : visualization of biological molecules in their native state. Oxford University Press, Oxford New York, 2006. ISBN 978-0-19-515096-4.

[138] L. Reimer. Transmission electron microscopy : physics of image formation and microanalysis. Springer, Berlin New York, 1997. ISBN 3-540-62568-2.

[139] M. Vulović, R. B. Ravelli, L. J. van Vliet, A. J. Koster, I. Lazić, U. Lücken, H. Rullgård, O. Öktem, and B. Rieger. Image formation modeling in cryo-electron microscopy. Journal of Structural Biology, 183(1):19–32, July 2013. URL https://doi.org/10.1016/j.jsb.2013.05.008.

[140] P. Koeck and A. Karshikoff. Limitations of the linear and the projection approximations in three-dimensional transmission electron microscopy of fully hydrated proteins. Journal of Microscopy, 259(3):197–209, April 2015. URL https://doi.org/10.1111/jmi.12253.

[141] G. Jensen. Online course : Getting started in cryo-em, 2021. URL https://jensenlab.caltech.edu/courses/.

[142] D. Dynerman. Reconstruction via direct fourier inversion, 2021. URL https://utoronto.zoom.us/rec/share/iUQlpdJvEwJ4Z-GU3bQjJzccyPSJg0iocKhN5RFwyHXMKkuIpR5I3alav3UKfk4T.hA7WqNfNCNPpbqN7. Zoom password : 041\&NOnv+F.

[143] D. B. Williams and C. B. Carter. Transmission Electron Microscopy. Springer US, 1996. URL https://doi.org/10.1007/978-1-4757-2519-3.

[144] K. Murata and M. Wolf. Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. Biochimica et Biophysica Acta (BBA) - General Subjects, 1862(2):324–334, February 2018. URL https://doi.org/10.1016/j.bbagen.2017.07.020.

[145] J. P. Langmore and M. F. Smith. Quantitative energy-filtered electron microscopy of biological molecules in ice. Ultramicroscopy, 46(1-4):349–373, October 1992. URL https://doi.org/10.1016/0304-3991(92)90024-e.

[146] M. J. Peet, R. Henderson, and C. J. Russo. The energy dependence of contrast and damage in electron cryomicroscopy of biological molecules. Ultramicroscopy, 203:125–131, August 2019. URL https://doi.org/10.1016/j.ultramic.2019.02.007.

[147] J. Goodman. Introduction to Fourier optics. Roberts & Co, Englewood, Colo, 2005. ISBN 9780974707723.

[148] R. Bracewell. Strip integration in radio astronomy. Australian Journal of Physics, 9(2):198, 1956. URL https://doi.org/10.1071/ph560198.

[149] F. J. Sigworth. Principles of cryo-EM single-particle image processing. Microscopy, 65(1):57–67, December 2015. URL https://doi.org/10.1093/jmicro/dfv370.

[150] S. H. Scheres. A bayesian view on cryo-EM structure determination. Journal of Molecular Biology, 415(2):406–418, January 2012. URL https://doi.org/10.1016/j.jmb.2011.11.010.

[151] S. J. Ludtke, P. R. Baldwin, and W. Chiu. EMAN: Semiautomated software for high-resolution single-particle reconstructions. Journal of Structural Biology, 128(1):82–97, December 1999. URL https://doi.org/10.1006/jsbi.1999.4174.

[152] N. Grigorieff. Three-dimensional structure of bovine NADH:ubiquinone oxidoreductase (complex i) at 22 å in ice. Journal of Molecular Biology, 277(5):1033–1046, April 1998. URL https://doi.org/10.1006/jmbi.1998.1668.

[153] S. H. Scheres. RELION: Implementation of a bayesian approach to cryo-EM structure determination. Journal of Structural Biology, 180(3):519–530, December 2012. URL https://doi.org/10.1016/j.jsb.2012.09.006.

[154] J. B. Heymann, M. Chagoyen, and D. M. Belnap. Common conventions for interchange and archiving of three-dimensional electron microscopy information in structural biology. Journal of Structural Biology, 151(2):196–207, August 2005. URL https://doi.org/10.1016/j.jsb.2005.06.001.

[155] F. J. Sigworth, P. C. Doerschuk, J.-M. Carazo, and S. H. Scheres. An introduction to maximum-likelihood methods in cryo-EM. In Methods in Enzymology, pages 263–294. Elsevier, 2010. URL https://doi.org/10.1016/s0076-6879(10)82011-7.

[156] S. H. Scheres. Classification of structural heterogeneity by maximum-likelihood methods. In Methods in Enzymology, pages 295–320. Elsevier, 2010. URL https://doi.org/10.1016/s0076-6879(10)82012-9.

[157] E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, and T. E. Ferrin. UCSF ChimeraX : Structure visualization for researchers, educators, and developers. Protein Science, 30(1):70–82, October 2020. URL https://doi.org/10.1002/pro.3943.

[158] T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, and T. E. Ferrin. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Science, 27(1):14–25, September 2017. URL https://doi.org/10.1002/pro.3235.

[159] C. Sorzano, R. Marabini, J. Velázquez-Muriel, J. Bilbao-Castro, S. Scheres, J. Carazo, and A. Pascual-Montano. XMIPP: a new generation of an open-source image processing package for electron microscopy. Journal of Structural Biology, 148(2):194–204, November 2004. URL https://doi.org/10.1016/j.jsb.2004.06.006.

[160] C. Sorzano, E. Ortiz, M. López, and J. Rodrigo. Improved bayesian image denoising based on wavelets with applications to electron microscopy. Pattern Recognition, 39(6):1205–1213, June 2006. URL https://doi.org/10.1016/j.patcog.2005.12.009.

[161] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature Methods, 14(3):290–296, February 2017. URL https://doi.org/10.1038/nmeth.4169.

[162] R. Henderson. The potential and limitations of neutrons, electrons and x-rays for atomic resolution microscopy of unstained biological molecules. Quarterly Reviews of Biophysics, 28 (2):171–193, May 1995. URL https://doi.org/10.1017/s003358350000305x.

[163] J. Baruchel, J. Hodeau, M. Lehmann, J. Regnard, and C. Schlenker. Neutron and synchrotron radiation for condensed matter studies. Springer-Verlag Les Editions de Physique, Berlin New York Ulis, France, 1993. ISBN 2-86883-185-0.

[164] Institut Laue-Langevin, 2021. URL https://www.ill.eu/.

[165] B. P. Schoenborn, D. L. D. Caspar, and O. F. Kammerer. A novel neutron monochromator. Journal of Applied Crystallography, 7(5):508–510, October 1974. URL https://doi.org/10.1107/s0021889874010302.

[166] G. Knoll. Radiation detection and measurement. John Wiley, Hoboken, N.J, 2010. ISBN 9780470131480.

[167] R. M. Moon, T. Riste, and W. C. Koehler. Polarization analysis of thermal-neutron scattering. Physical Review, 181(2):920–931, May 1969. URL https://doi.org/10.1103/physrev.181.920.

[168] A. M. Gaspar, S. Busch, M.-S. Appavou, W. Haeussler, R. Georgii, Y. Su, and W. Doster. Using polarization analysis to separate the coherent and incoherent scattering from protein samples. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1804(1):76–82, January 2010. URL https://doi.org/10.1016/j.bbapap.2009.06.024.

[169] J. Ollivier, H. Mutka, and L. Didier. The new cold neutron time-of-flight spectrometer IN5. Neutron News, 21(2):22–25, April 2010. URL https://doi.org/10.1080/10448631003757573.

[170] N. Francesca, J. Peters, D. Russo, S. Barbieri, C. Chiapponi, A. Cupane, A. Deriu, M. T. D. Bari, E. Farhi, Y. Gerelli, P. Mariani, A. Paciaroni, C. Rivasseau, G. Schirò, and F. Sonvico. IN13 backscattering spectrometer at ILL: Looking for motions in biological macromolecules and organisms. Neutron News, 19(4):14–18, November 2008. URL https://doi.org/10.1080/10448630802474083.

[171] M. Grimaldo, F. Roosen-Runge, N. Jalarvo, M. Zamponi, F. Zanini, M. Hennig, F. Zhang, F. Schreiber, and T. Seydel. High-resolution neutron spectroscopy on protein solution samples. EPJ Web of Conferences, 83:02005, 2015. URL https://doi.org/10.1051/epjconf/20158302005.

[172] M. Appel, B. Frick, and A. Magerl. First results with the neutron backscattering and TOF spectrometer option BATS on IN16b. Physica B: Condensed Matter, 562:6–8, June 2019. URL https://doi.org/10.1016/j.physb.2018.11.062.

[173] S. Perticaroli, G. Ehlers, C. B. Stanley, E. Mamontov, H. O'Neill, Q. Zhang, X. Cheng, D. A. A. Myles, J. Katsaras, and J. D. Nickels. Description of hydration water in protein (green fluorescent protein) solution. Journal of the American Chemical Society, 139(3):1098–1105, November 2016. URL https://doi.org/10.1021/jacs.6b08845.

[174] H. H. Paalman and C. J. Pings. Numerical evaluation of x-ray absorption factors for cylindrical samples and annular sample cells. Journal of Applied Physics, 33(8):2635–2639, August 1962. URL https://doi.org/10.1063/1.1729034.

[175] D. Richard, M. Ferrand, and G. J. Kearley. Analysis and visualisation of neutron-scattering data. Journal of Neutron Research, 4(1):33–39, December 1996. URL https://doi.org/10.1080/10238169608200065.

[176] O. Arnold, J. Bilheux, J. Borreguero, A. Buts, S. Campbell, L. Chapon, M. Doucet, N. Draper, R. F. Leal, M. Gigg, V. Lynch, A. Markvardsen, D. Mikkelson, R. Mikkelson, R. Miller, K. Palmen, P. Parker, G. Passos, T. Perring, P. Peterson, S. Ren, M. Reuter, A. Savici, J. Taylor, R. Taylor, R. Tolchenov, W. Zhou, and J. Zikovsky. Mantid—data analysis and visualization package for neutron scattering and μ SR experiments. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 764:156–166, November 2014. URL https://doi.org/10.1016/j.nima.2014.07.029.

[177] G. Van Rossum and F. L. Drake. Python 3 Reference Manual. CreateSpace, Scotts Valley, CA, 2009. ISBN 1441412697.

[178] R. Ashkar, H. Z. Bilheux, H. Bordallo, R. Briber, D. J. E. Callaway, X. Cheng, X.-Q. Chu, J. E. Curtis, M. Dadmun, P. Fenimore, D. Fushman, F. Gabel, K. Gupta, F. Herberle, F. Heinrich, L. Hong, J. Katsaras, Z. Kelman, E. Kharlampieva, G. R. Kneller, A. Kovalevsky, S. Krueger, P. Langan, R. Lieberman, Y. Liu, M. Losche, E. Lyman, Y. Mao, J. Marino, C. Mattos, F. Meilleur, P. Moody, J. D. Nickels, W. B. O'Dell, H. O'Neill, U. Perez-Salas, J. Peters, L. Petridis, A. P. Sokolov, C. Stanley, N. Wagner, M. Weinrich, K. Weiss, T. Wymore, Y. Zhang, and J. C. Smith. Neutron scattering in the biological sciences: progress and prospects. Acta Crystallographica Section D Structural Biology, 74(12):1129–1168, December 2018. URL https://doi.org/10.1107/s2059798318017503.

[179] M. Haertlein, M. Moulin, J. M. Devos, V. Laux, O. Dunne, and V. T. Forsyth. Biomolecular deuteration for neutron structural biology and dynamics. In Methods in Enzymology, pages 113–157. Elsevier, 2016. URL https://doi.org/10.1016/bs.mie.2015.11.001.

[180] S. Fujiwara, F. Matsumoto, and Y. Yonezawa. Effects of salt concentration on association of the amyloid protofilaments of hen egg white lysozyme studied by time-resolved neutron scattering. Journal of Molecular Biology, 331(1):21–28, August 2003. URL https://doi.org/10.1016/s0022-2836(03)00722-8.

[181] S. Maric, T. K. Lind, M. R. Raida, E. Bengtsson, G. N. Fredrikson, S. Rogers, M. Moulin, M. Haertlein, V. T. Forsyth, M. R. Wenk, T. G. Pomorski, T. Arnebrant, R. Lund, and M. Cárdenas. Time-resolved small-angle neutron scattering as a probe for the dynamics of lipid exchange between human lipoproteins and naturally derived membranes. Scientific Reports, 9 (1), May 2019. URL https://doi.org/10.1038/s41598-019-43713-6.

[182] J. Pieper, A. Buchsteiner, N. A. Dencher, R. E. Lechner, and T. Hauß. Transient protein softening during the working cycle of a molecular machine. Physical Review Letters, 100(22), June 2008. URL https://doi.org/10.1103/physrevlett.100.228103.

[183] J. Pieper. Time-resolved quasielastic neutron scattering studies of native photosystems. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1804(1):83–88, January 2010. URL https://doi.org/10.1016/j.bbapap.2009.09.016.

[184] J. Pieper. The functional role of protein dynamics in photosynthetic reaction centers investigated by elastic and quasielastic neutron scattering. EPJ Web of Conferences, 83:02013, 2015. URL https://doi.org/10.1051/epjconf/20158302013.

[185] C. Beck, M. Grimaldo, F. Roosen-Runge, R. Maier, O. Matsarskaia, M. Braun, B. Sohmen, O. Czakkel, R. Schweins, F. Zhang, T. Seydel, and F. Schreiber. Following protein dynamics in real time during crystallization. Crystal Growth & Design, 19(12):7036–7045, October 2019. URL https://doi.org/10.1021/acs.cgd.9b00858.

[186] H.-J. Cho, J.-K. Hyun, J.-G. Kim, H. S. Jeong, H. N. Park, D.-J. You, and H. S. Jung. Measurement of ice thickness on vitreous ice embedded cryo-EM grids: investigation of optimizing condition for visualizing macromolecules. Journal of Analytical Science and Technology, 4(1), April 2013. URL https://doi.org/10.1186/2093-3371-4-7.

[187] N. C. for Biotechnology Information. Pubchem compound summary for cid 10915, uranyl acetate., 2021. https://pubchem.ncbi.nlm.nih.gov/compound/Uranyl-acetate.

[188] J. Dubochet, J.-J. Chang, R. Freeman, J. Lepault, and A. W. McDowall. Frozen aqueous suspensions. Ultramicroscopy, 10(1-2):55–61, January 1982. URL https://doi.org/10.1016/0304-3991(82)90187-5.

[189] A. W. McDowall, J.-J. Chang, R. Freeman, J. Lepault, C. A. Walter, and J. Dubochet. Electron microscopy of frozen hydrated sections of vitreous ice and vitrified biological samples. Journal of Microscopy, 131(1):1–9, July 1983. URL https://doi.org/10.1111/j.1365-2818.1983.tb04225.x.

[190] K. A. Taylor and R. M. Glaeser. Electron diffraction of frozen, hydrated protein crystals. Science, 186(4168):1036–1037, December 1974. URL https://doi.org/10.1126/science.186.4168.1036.

[191] I. Angert, C. Burmester, C. Dinges, H. Rose, and R. R. Schröder. Elastic and inelastic scattering cross-sections of amorphous layers of carbon and vitrified ice. Ultramicroscopy, 63(3-4):181–192, July 1996. URL https://doi.org/10.1016/0304-3991(96)00036-8.

[192] M. Knoll and E. Ruska. The electron microscope. Z. Phys, 78(5-6):318–339339, 1932.

[193] L. E. Franken, K. Grünewald, E. J. Boekema, and M. C. A. Stuart. A technical introduction to transmission electron microscopy for soft-matter: Imaging, possibilities, choices, and technical developments. Small, 16(14):1906198, April 2020. URL https://doi.org/10.1002/smll.201906198.

[194] G. McMullan, A. R. Faruqi, and R. Henderson. Direct electron detectors. In Methods in Enzymology, pages 1–17. Elsevier, 2016. URL https://doi.org/10.1016/bs.mie.2016.05.056.

[195] Thermo Fischer Scientific, 2021. URL https://www.thermofisher.com.

[196] E. Kandiah, T. Giraud, A. de Maria Antolinos, F. Dobias, G. Effantin, D. Flot, M. Hons, G. Schoehn, J. Susini, O. Svensson, G. A. Leonard, and C. Mueller-Dieckmann. CM01: a facility for cryo-electron microscopy at the european synchrotron. Acta Crystallographica Section D Structural Biology, 75(6):528–535, May 2019. URL https://doi.org/10.1107/s2059798319006880.

[197] M. van Heel and W. Keegstra. IMAGIC: A fast, flexible and friendly image analysis software system. Ultramicroscopy, 7(2):113–129, January 1981. URL https://doi.org/10.1016/0304-3991(81)90001-2.

[198] J. Frank, M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, and A. Leith. SPIDER and WEB: Processing and visualization of images in 3d electron microscopy and related fields. Journal of Structural Biology, 116(1):190–199, January 1996. URL https://doi.org/10.1006/jsbi.1996.0030.

[199] R. Marabini, I. M. Masegosa, C. S. Martin, S. Marco, J. J. Fernandez, L. G. de la Fraga, C. Vaquerizo, and J. M. Carazo. Xmipp: An image processing package for electron microscopy. Journal of Structural Biology, 116(1):237–240, January 1996. URL https://doi.org/10.1006/jsbi.1996.0036.

[200] S. Scheres. Processing of structurally heterogeneous cryo-EM data in RELION. In Methods in Enzymology, pages 125–157. Elsevier, 2016. URL https://doi.org/10.1016/bs.mie.2016.04.012.

[201] A. F. Brilot, J. Z. Chen, A. Cheng, J. Pan, S. C. Harrison, C. S. Potter, B. Carragher, R. Henderson, and N. Grigorieff. Beam-induced motion of vitrified specimen on holey carbon film. Journal of Structural Biology, 177(3):630–637, March 2012. URL https://doi.org/10.1016/j.jsb.2012.02.003.

[202] X. Li, P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard, and Y. Cheng. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. Nature Methods, 10(6):584–590, May 2013. URL https://doi.org/10.1038/nmeth.2472.

[203] S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, and D. A. Agard. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nature Methods, 14(4):331–332, February 2017. URL https://doi.org/10.1038/nmeth.4193.

[204] A. Rohou and N. Grigorieff. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. Journal of Structural Biology, 192(2):216–221, November 2015. URL https://doi.org/10.1016/j.jsb.2015.08.008.

[205] X. Zhang and Z. H. Zhou. Limiting factors in atomic resolution cryo electron microscopy: No simple tricks. Journal of Structural Biology, 175(3):253–263, September 2011. URL https://doi.org/10.1016/j.jsb.2011.05.004.

[206] A. Philippsen, H.-A. Engel, and A. Engel. The contrast-imaging function for tilted specimens. Ultramicroscopy, 107(2-3):202–212, February 2007. URL https://doi.org/10.1016/j.ultramic.2006.07.010.

[207] L. M. Voortman, S. Stallinga, R. H. Schoenmakers, L. J. van Vliet, and B. Rieger. A fast algorithm for computing and correcting the CTF for tilted, thick specimens in TEM. Ultramicroscopy, 111(8):1029–1036, July 2011. URL https://doi.org/10.1016/j.ultramic.2011.03.001.

[208] A. Roseman. FindEM—a fast, efficient program for automatic selection of particles from electron micrographs. Journal of Structural Biology, 145(1-2):91–99, January 2004. URL https://doi.org/10.1016/j.jsb.2003.11.007.

[209] S. H. Scheres. Semi-automated selection of cryo-EM particles in RELION-1.3. Journal of Structural Biology, 189(2):114–122, February 2015. URL https://doi.org/10.1016/j.jsb.2014.11.010.

[210] D. Si, A. Nakamura, R. Tang, H. Guan, J. Hou, A. Firozi, R. Cao, K. Hippe, and M. Zhao. Artificial intelligence advances for de novo molecular structure modeling in cryo-electron microscopy. WIREs Computational Molecular Science, May 2021. URL https://doi.org/10.1002/wcms.1542.

[211] T. Bepler, A. Morin, M. Rapp, J. Brasch, L. Shapiro, A. J. Noble, and B. Berger. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nature Methods, 16(11):1153–1160, October 2019. URL https://doi.org/10.1038/s41592-019-0575-8.

[212] F. Wang, H. Gong, G. Liu, M. Li, C. Yan, T. Xia, X. Li, and J. Zeng. DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM. Journal of Structural Biology, 195(3):325–336, September 2016. URL https://doi.org/10.1016/j.jsb.2016.07.006.

[213] R. Henderson, A. Sali, M. L. Baker, B. Carragher, B. Devkota, K. H. Downing, E. H. Egelman, Z. Feng, J. Frank, N. Grigorieff, W. Jiang, S. J. Ludtke, O. Medalia, P. A. Penczek, P. B. Rosenthal, M. G. Rossmann, M. F. Schmid, G. F. Schröder, A. C. Steven, D. L. Stokes, J. D. Westbrook, W. Wriggers, H. Yang, J. Young, H. M. Berman, W. Chiu, G. J. Kleywegt, and C. L. Lawson. Outcome of the first electron microscopy validation task force meeting. Structure, 20(2):205–214, February 2012. URL https://doi.org/10.1016/j.str.2011.12.014.

[214] S. H. W. Scheres and S. Chen. Prevention of overfitting in cryo-EM structure determination. Nature Methods, 9(9):853–854, July 2012. URL https://doi.org/10.1038/nmeth.2115.

[215] P. B. Rosenthal and R. Henderson. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. Journal of Molecular Biology, 333(4):721–745, October 2003. URL https://doi.org/10.1016/j.jmb.2003.07.013.

[216] P. A. Penczek. Reliable cryo-EM resolution estimation with modified fourier shell correlation. IUCrJ, 7(6):995–1008, September 2020. URL https://doi.org/10.1107/s2052252520011574.

[217] G. Cardone, J. B. Heymann, and A. C. Steven. One number does not fit all: Mapping local variations in resolution in cryo-EM reconstructions. Journal of Structural Biology, 184(2):226–236, November 2013. URL https://doi.org/10.1016/j.jsb.2013.08.002.

[218] A. Kucukelbir, F. J. Sigworth, and H. D. Tagare. Quantifying the local resolution of cryo-EM density maps. Nature Methods, 11(1):63–65, November 2013. URL https://doi.org/10.1038/nmeth.2727.

[219] T. H. D. Nguyen, W. P. Galej, X. chen Bai, C. Oubridge, A. J. Newman, S. H. W. Scheres, and K. Nagai. Cryo-EM structure of the yeast u4/u6.u5 tri-snRNP at 3.7 Å resolution. Nature, 530(7590):298–302, February 2016. URL https://doi.org/10.1038/nature16940.

[220] T. Nakane, D. Kimanius, E. Lindahl, and S. H. Scheres. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. eLife, 7, June 2018. URL https://doi.org/10.7554/elife.36861.

[221] A. Punjani and D. J. Fleet. 3d variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. Journal of Structural Biology, 213(2):107702, June 2021. URL https://doi.org/10.1016/j.jsb.2021.107702.

[222] E. D. Zhong, T. Bepler, B. Berger, and J. H. Davis. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. Nature Methods, 18(2):176–185, February 2021. URL https://doi.org/10.1038/s41592-020-01049-4.

[223] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF chimera?a visualization system for exploratory research and analysis. Journal of Computational Chemistry, 25(13):1605–1612, 2004. URL https://doi.org/10.1002/jcc.20084.

[224] S. Chen, G. McMullan, A. R. Faruqi, G. N. Murshudov, J. M. Short, S. H. Scheres, and R. Henderson. High-resolution noise substitution to measure overfitting and validate resolution in 3d structure determination by single particle electron cryomicroscopy. Ultramicroscopy, 135:24–35, December 2013. URL https://doi.org/10.1016/j.ultramic.2013.06.004.

[225] P. V. Afonine, B. P. Klaholz, N. W. Moriarty, B. K. Poon, O. V. Sobolev, T. C. Terwilliger, P. D. Adams, and A. Urzhumtsev. New tools for the analysis and validation of cryo-EM maps and atomic models. Acta Crystallographica Section D Structural Biology, 74(9):814–840, September 2018. URL https://doi.org/10.1107/s2059798318009324.

[226] K. M. Yip, N. Fischer, E. Paknia, A. Chari, and H. Stark. Atomic-resolution protein structure determination by cryo-EM. Nature, 587(7832):157–161, October 2020. URL https://doi.org/10.1038/s41586-020-2833-4.

[227] B. Carragher, Y. Cheng, A. Frost, R. M. Glaeser, G. C. Lander, E. Nogales, and H.-W. Wang. Current outcomes when optimizing 'standard' sample preparation for single-particle cryo-EM. Journal of Microscopy, 276(1):39–45, October 2019. URL https://doi.org/10.1111/jmi.12834.

[228] R. M. Glaeser. Proteins, interfaces, and cryo-EM grids. Current Opinion in Colloid & Interface Science, 34:1–8, March 2018. URL https://doi.org/10.1016/j.cocis.2017.12.009.

[229] A. J. Noble, H. Wei, V. P. Dandey, Z. Zhang, Y. Z. Tan, C. S. Potter, and B. Carragher. Reducing effects of particle adsorption to the air–water interface in cryo-EM. Nature Methods, 15(10):793–795, September 2018. URL https://doi.org/10.1038/s41592-018-0139-3.

[230] T. Jain, P. Sheehan, J. Crum, B. Carragher, and C. S. Potter. Spotiton: A prototype for an integrated inkjet dispense and vitrification system for cryo-TEM. Journal of Structural Biology, 179(1):68–75, July 2012. URL https://doi.org/10.1016/j.jsb.2012.04.020.

[231] V. P. Dandey, H. Wei, Z. Zhang, Y. Z. Tan, P. Acharya, E. T. Eng, W. J. Rice, P. A. Kahn, C. S. Potter, and B. Carragher. Spotiton: New features and applications. Journal of Structural Biology, 202(2):161–169, May 2018. URL https://doi.org/10.1016/j.jsb.2018.01.002.

[232] V. P. Dandey, W. C. Budell, H. Wei, D. Bobe, K. Maruthi, M. Kopylov, E. T. Eng, P. A. Kahn, J. E. Hinshaw, N. Kundu, C. M. Nimigean, C. Fan, N. Sukomon, S. A. Darst, R. M. Saecker, J. Chen, B. Malone, C. S. Potter, and B. Carragher. Time-resolved cryo-EM using spotiton. Nature Methods, 17(9):897–900, August 2020. URL https://doi.org/10.1038/s41592-020-0925-6.

[233] J. Frank. Advances in the field of single-particle cryo-electron microscopy over the last decade. Nature Protocols, 12(2):209–212, January 2017. URL https://doi.org/10.1038/nprot.2017.004.

[234] J. L. Vilas, N. Tabassum, J. Mota, D. Maluenda, A. Jiménez-Moreno, T. Majtner, J. M. Carazo, S. T. Acton, and C. O. S. Sorzano. Advances in image processing for single-particle analysis by electron cryomicroscopy and challenges ahead. Current Opinion in Structural Biology, 52: 127–145, October 2018. URL https://doi.org/10.1016/j.sbi.2018.11.004.

[235] M. le Maire, P. Champeil, and J. V. Møller. Interaction of membrane proteins and lipids with solubilizing detergents. Biochimica et Biophysica Acta (BBA) - Biomembranes, 1508(1): 86–111, 2000. ISSN 0005-2736. URL https://www.sciencedirect.com/science/article/pii/S0304415700000101. Detergents in Biomembrane Studies.

[236] V. Chaptal, F. Delolme, A. Kilburg, S. Magnard, C. Montigny, M. Picard, C. Prier, L. Monticelli, O. Bornert, M. Agez, S. Ravaud, C. Orelle, R. Wagner, A. Jawhari, I. Broutin, E. Pebay-Peyroula, J.-M. Jault, H. R. Kaback, M. le Maire, and P. Falson. Quantification of detergents complexed with membrane proteins. Scientific Reports, 7(1), February 2017. URL https://doi.org/10.1038/srep41751.

[237] S. J. Perkins and H. Weiss. Low-resolution structural studies of mitochondrial ubiquinol: Cytochrome c reductase in detergent solutions by neutron scattering. Journal of Molecular Biology, 168(4):847–866, August 1983. URL https://doi.org/10.1016/s0022-2836(83)80078-3.

[238] C. Breyton, F. Gabel, M. Lethier, A. Flayhan, G. Durand, J.-M. Jault, C. Juillan-Binard, L. Imbert, M. Moulin, S. Ravaud, M. Härtlein, and C. Ebel. Small angle neutron scattering for the study of solubilised membrane proteins. The European Physical Journal E, 36(7), July 2013. URL https://doi.org/10.1140/epje/i2013-13071-6.

[239] M. Golub, R. Hussein, M. Ibrahim, M. Hecht, D. C. F. Wieland, A. Martel, B. Machado, A. Zouni, and J. Pieper. Solution structure of the detergent–photosystem II core complex investigated by small-angle scattering techniques. The Journal of Physical Chemistry B, 124 (39):8583–8592, August 2020. URL https://doi.org/10.1021/acs.jpcb.0c07169.

[240] S. R. Midtgaard, T. A. Darwish, M. C. Pedersen, P. Huda, A. H. Larsen, G. V. Jensen, S. A. R. Kynde, N. Skar-Gislinge, A. J. Z. Nielsen, C. Olesen, M. Blaise, J. J. Dorosz, T. S. Thorsen, R. Venskutonytė, C. Krintel, J. V. Møller, H. Frielinghaus, E. P. Gilbert, A. Martel, J. S. Kastrup, P. E. Jensen, P. Nissen, and L. Arleth. Invisible detergents for structure determination of membrane proteins by small-angle neutron scattering. The FEBS Journal, 285(2):357–371, December 2017. URL https://doi.org/10.1111/febs.14345.

[241] K. Hiruma-Shimizu, H. Shimizu, G. S. Thompson, A. P. Kalverda, and S. G. Patching. Deuterated detergents for structural and functional studies of membrane proteins: Properties, chemical synthesis and applications. Molecular Membrane Biology, 32(5-8):139–155, November 2015. URL https://doi.org/10.3109/09687688.2015.1125536.

[242] A. Gall, J. Seguin, B. Robert, and M. C. Bellissent-Funel. Membrane proteins in bulk solution can be used for quasi-elastic neutron scattering studies: The case for the photochemical reaction center. Journal of Physical Chemistry B, 106(24):6303–6309, 2002. ISSN 10895647.

[243] U. R. Shrestha, S. M. D. C. Perera, D. Bhowmik, U. Chawla, E. Mamontov, M. F. Brown, and X.-Q. Chu. Quasi-elastic neutron scattering reveals ligand-induced protein dynamics of a g-protein-coupled receptor. The Journal of Physical Chemistry Letters, 7(20):4130–4136, October 2016. URL https://doi.org/10.1021/acs.jpclett.6b01632.

[244] W. Doster, S. Cusack, and W. Petry. Dynamical transition of myoglobin revealed by inelastic neutron scattering. Nature, 337(6209):754–756, February 1989. URL https://doi.org/10.1038/337754a0.

[245] M. Fomina, G. Schirò, and A. Cupane. Hydration dependence of myoglobin dynamics studied with elastic neutron scattering, differential scanning calorimetry and broadband dielectric spectroscopy. Biophysical Chemistry, 185:25–31, January 2014. URL https://doi.org/10.1016/j.bpc.2013.11.004.

[246] H. Frauenfelder, P. W. Fenimore, and R. D. Young. Protein dynamics and function: Insights from the energy landscape and solvent slaving. IUBMB Life, 59(8):506–512, 2007. URL https://doi.org/10.1080/15216540701194113.

[247] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and F. G. Parak. Slaving: Solvent fluctuations dominate protein dynamics and functions. Proceedings of the National Academy of Sciences, 99(25):16047–16051, November 2002. URL https://doi.org/10.1073/pnas.212637899.

[248] A. M. Tsai, D. A. Neumann, and L. N. Bell. Molecular dynamics of solid-state lysozyme as affected by glycerol and water: A neutron scattering study. Biophysical Journal, 79(5): 2728–2732, November 2000. URL https://doi.org/10.1016/s0006-3495(00)76511-8.

[249] G. Caliskan, D. Mechtani, J.-H. Roh, A. M. Kisliuk, A. P. Sokolov, S. Azzam, M. T. Cicerone, S. Lin-Gibson, and I. Peral. Protein and solvent dynamics: How strongly are they coupled? The Journal of Chemical Physics, 121(4):1978–1983, July 2004. URL https://doi.org/10.1063/1.1764491.

[250] A. Paciaroni, E. Cornicchi, A. D. Francesco, M. Marconi, and G. Onori. Conditioning action of the environment on the protein dynamics studied through elastic neutron scattering. European Biophysics Journal, 35(7):591–599, June 2006. URL https://doi.org/10.1007/s00249-006-0073-7.

[251] V. G. Sakai, S. Khodadadi, M. T. Cicerone, J. E. Curtis, A. P. Sokolov, and J. H. Roh. Solvent effects on protein fast dynamics: implications for biopreservation. Soft Matter, 9(22):5336, 2013. URL https://doi.org/10.1039/c3sm50492a.

[252] J. Peters, M. Appel, C. Beck, and A. Cisse. Dynamics of apob-100 and its detergent, 2020. URL https://doi.ill.fr/10.5291/ILL-DATA.INTER-500.

[253] B. Grenier, J. A. G. Almeida, A. Bellisario, L. Canon, A. Cisse, L. Colin, H. L. C. Colocho, N. Coquelle, Y. Correa, V. Crozet, A. D'Angelo, B. Demé, S. Erikson, A. L. Gandarias, L. Girois, M. Gracheva, I. Hoffmann, D. M. Kehlenbeck, R. Kulandaisamy, M. Armando, O. Matsarskaia, M. G. E. Pena, M. G. E. Pena, J. Peters, A. Piccinini, S. Prevost, T. Seydel, P. K. Singh, G. O. Tomas, M. Viljanen, J. Weder, S. Winnall, and B. Zhou. Remote practicals for hercules 2021 session b, 2021. URL https://doi.ill.fr/10.5291/ILL-DATA.TEST-3146.

[254] D. Di Bari, A. Cisse, O. M. T. Giudici, M. Guiral, A. Nidriche, J. Ollivier, A. Paciaroni, J. Peters, and F. Sterpone. Proteome dynamics as a proxy for cellular thermal stability, 2021. URL https://doi.ill.fr/10.5291/ILL-DATA.8-04-885.

[255] M. Newville, T. Stensitzki, D. B. Allen, and A. Ingargiola. Lmfit: Non-linear least-square minimization and curve-fitting for python, 2014. URL https://zenodo.org/record/11813.

[256] D. Sivia, C. Carlile, W. Howells, and S. König. Bayesian analysis of quasielastic neutron scattering data. Physica B, 182:341–348, 1992.

[257] R. Johnson. Computational chemistry comparison and benchmark database, nist standard reference database 101, 2002. URL http://cccbdb.nist.gov/.

[258] M. Holz, S. R. Heil, and A. Sacco. Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate 1h NMR PFG measurements. Physical Chemistry Chemical Physics, 2(20):4740–4742, 2000. URL https://doi.org/10.1039/b005319h.

[259] M. Trapp, M. Tehei, M. Trovaslet, F. Nachon, N. Martinez, M. M. Koza, M. Weik, P. Masson, and J. Peters. Correlation of the dynamics of native human acetylcholinesterase and its inhibited huperzine a counterpart from sub-picoseconds to nanoseconds. Journal of The Royal Society Interface, 11(97):20140372, August 2014. URL https://doi.org/10.1098/rsif.2014.0372.

[260] M. Grimaldo, H. Lopez, C. Beck, F. Roosen-Runge, M. Moulin, J. M. Devos, V. Laux, M. Härtlein, S. D. Vela, R. Schweins, A. Mariani, F. Zhang, J.-L. Barrat, M. Oettel, V. T. Forsyth, T. Seydel, and F. Schreiber. Protein short-time diffusion in a naturally crowded environment. The Journal of Physical Chemistry Letters, 10(8):1709–1715, March 2019. URL https://doi.org/10.1021/acs.jpclett.9b00345.

[261] J. Wolanin, J. Giraud, C. Payre, M. Benoit, C. Antonelli, D. Quemener, I. Tahiri, M. Vandamme, J.-M. Zanotti, and M. Plazanet. Oedometric-like setup for the study of water transport in porous media by quasi-elastic neutron scattering. Review of Scientific Instruments, 92(2):024106, February 2021. URL https://doi.org/10.1063/5.0030297.

[262] M. Tehei, B. Franzetti, K. Wood, F. Gabel, E. Fabiani, M. Jasnin, M. Zamponi, D. Oesterhelt, G. Zaccai, M. Ginzburg, and B.-Z. Ginzburg. Neutron scattering reveals extremely slow cell water in a dead sea organism. Proceedings of the National Academy of Sciences, 104(3):766–771, January 2007. URL https://doi.org/10.1073/pnas.0601639104.

[263] J. Peters, N. Martinez, M. Trovaslet, K. Scannapieco, M. M. Koza, P. Masson, and F. Nachon. Dynamics of human acetylcholinesterase bound to non-covalent and covalent inhibitors shedding light on changes to the water network structure. Physical Chemistry Chemical Physics, 18(18): 12992–13001, 2016. URL https://doi.org/10.1039/c6cp00280c.

[264] G. Zaccai. How soft is a protein? a protein dynamics force constant measured by neutron scattering. Science, 288(5471):1604–1607, jun 2000.

[265] G. Nagy, J. Pieper, S. B. Krumova, L. Kovács, M. Trapp, G. Garab, and J. Peters. Dynamic properties of photosystem II membranes at physiological temperatures characterized by elastic incoherent neutron scattering. increased flexibility associated with the inactivation of the oxygen evolving complex. Photosynthesis Research, 111(1-2):113–124, November 2011. URL https://doi.org/10.1007/s11120-011-9701-x.

[266] U. Lehnert, V. Réat, M. Weik, G. Zaccaï, and C. Pfister. Thermal motions in bacteriorhodopsin at different hydration levels studied by neutron scattering: Correlation with kinetics and light-induced conformational changes. Biophysical Journal, 75(4):1945–1952, October 1998. URL https://doi.org/10.1016/s0006-3495(98)77635-0.

[267] M. Trapp, M. Trovaslet, F. Nachon, M. M. Koza, L. van Eijck, F. Hill, M. Weik, P. Masson, M. Tehei, and J. Peters. Energy landscapes of human acetylcholinesterase and its huperzine a-inhibited counterpart. The Journal of Physical Chemistry B, 116(51):14744–14753, December 2012. URL https://doi.org/10.1021/jp304704h.

[268] I. Ermilova and A. P. Lyubartsev. Cholesterol in phospholipid bilayers: positions and orientations inside membranes with different unsaturation degrees. Soft Matter, 15(1):78–93, 2019. URL https://doi.org/10.1039/c8sm01937a.

[269] N. Thonghin, V. Kargas, J. Clews, and R. C. Ford. Cryo-electron microscopy of membrane proteins. Methods, 147:176–186, September 2018. URL https://doi.org/10.1016/j.ymeth.2018.04.018.

[270] Y. Cheng. Membrane protein structural biology in the era of single particle cryo-EM. Current Opinion in Structural Biology, 52:58–63, October 2018. URL https://doi.org/10.1016/j.sbi.2018.08.008.

[271] X. Yao, X. Fan, and N. Yan. Cryo-EM analysis of a membrane protein embedded in the liposome. Proceedings of the National Academy of Sciences, 117(31):18497–18503, July 2020. URL https://doi.org/10.1073/pnas.2009385117.

[272] S. Tristram-Nagle, Y. Liu, J. Legleiter, and J. F. Nagle. Structure of gel phase DMPC determined by x-ray diffraction. Biophysical Journal, 83(6):3324–3335, December 2002. URL https://doi.org/10.1016/s0006-3495(02)75333-2.

[273] D. J. Benton, A. Nans, L. J. Calder, J. Turner, U. Neu, Y. P. Lin, E. Ketelaars, N. L. Kallewaard, D. Corti, A. Lanzavecchia, S. J. Gamblin, P. B. Rosenthal, and J. J. Skehel. Influenza hemagglutinin membrane anchor. Proceedings of the National Academy of Sciences, 115(40):10112–10117, September 2018. URL https://doi.org/10.1073/pnas.1810927115.

[274] F. Martino, M. Pal, H. Muñoz-Hernández, C. F. Rodríguez, R. Núñez-Ramírez, D. Gil-Carton, G. Degliesposti, J. M. Skehel, S. M. Roe, C. Prodromou, L. H. Pearl, and O. Llorca. RPAP3 provides a flexible scaffold for coupling HSP90 to the human r2tp co-chaperone complex. Nature Communications, 9(1), April 2018. URL https://doi.org/10.1038/s41467-018-03942-1.

[275] S. Bowerman, J. Wereszczynski, and K. Luger. Archaeal chromatin 'slinkies' are inherently dynamic complexes with deflected DNA wrapping pathways. eLife, 10, March 2021. URL https://doi.org/10.7554/elife.65587.

[276] A. von Appen, J. Kosinski, L. Sparks, A. Ori, A. L. DiGuilio, B. Vollmer, M.-T. Mackmull, N. Banterle, L. Parca, P. Kastritis, K. Buczak, S. Mosalaganti, W. Hagen, A. Andres-Pons, E. A. Lemke, P. Bork, W. Antonin, J. S. Glavy, K. H. Bui, and M. Beck. In situ structural analysis of the human nuclear pore complex. Nature, 526(7571):140–143, September 2015. URL https://doi.org/10.1038/nature15381.

[277] B. W. Hoogenboom, L. E. Hough, E. A. Lemke, R. Y. Lim, P. R. Onck, and A. Zilman. Physics of the nuclear pore complex: Theory, modeling and experiment. Physics Reports, 921:1–53, July 2021. URL https://doi.org/10.1016/j.physrep.2021.03.003.

[278] E. D. Zhong, T. Bepler, J. H. Davis, and B. Berger. Reconstructing continuous distributions of 3d protein structure from cryo-em images, 2020.

[279] A. Iudin, P. K. Korir, J. Salavert-Torres, G. J. Kleywegt, and A. Patwardhan. EMPIAR: a public archive for raw electron microscopy image data. Nature Methods, 13(5):387–388, March 2016. URL https://doi.org/10.1038/nmeth.3806.

[280] R. G. Efremov, C. Gatsogiannis, and S. Raunser. Lipid nanodiscs as a tool for high-resolution structure determination of membrane proteins by single-particle cryo-EM. In Methods in Enzymology, pages 1–30. Elsevier, 2017. URL https://doi.org/10.1016/bs.mie.2017.05.007.

[281] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis. Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873):583–589, July 2021. URL https://doi.org/10.1038/s41586-021-03819-2.

[282] B. Turoňová, M. Sikora, C. Schürmann, W. J. H. Hagen, S. Welsch, F. E. C. Blanc, S. von Bülow, M. Gecht, K. Bagola, C. Hörner, G. van Zandbergen, J. Landry, N. T. D. de Azevedo, S. Mosalaganti, A. Schwarz, R. Covino, M. D. Mühlebach, G. Hummer, J. K. Locker, and M. Beck. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. Science, 370(6513):203–208, August 2020. URL https://doi.org/10.1126/science.abd5223.

[283] M. Moulin, G. A. Strohmeier, M. Hirz, K. C. Thompson, A. R. Rennie, R. A. Campbell, H. Pichler, S. Maric, V. T. Forsyth, and M. Haertlein. Perdeuteration of cholesterol for neutron scattering applications using recombinant pichia pastoris. Chemistry and Physics of Lipids, 212:80–87, May 2018. URL https://doi.org/10.1016/j.chemphyslip.2018.01.006.

[284] A. Cisse, T. Matsuo, M. Plazanet, F. Natali, M. M. Koza, J. Ollivier, D. J. Bicout, and J. Peters. The dynamical matryoshka model: 2. a new modeling of local lipid dynamics at the sub-nanosecond timescale in phospholipid membranes. Biochimica et Biophysica Acta (BBA) - Biomembranes, 2021. To be submitted.

[285] T. Matsuo, A. Cisse, M. Plazanet, F. Natali, M. M. Koza, J. Ollivier, D. J. Bicout, , and J. Peters. The dynamical Matryoshka model: 3. diffusive nature of the atomic motions contained in the bicout-model for deciphering local lipid dynamics. Biochimica et Biophysica Acta (BBA) - Biomembranes, 2021. To be submitted.

[286] A. Cisse, J. Peters, G. Lazzara, and L. Chiappisi. PyDSC: a simple tool to treat differential scanning calorimetry data. Journal of Thermal Analysis and Calorimetry, 145 (2):403–409, May 2020. doi: 10.1007/s10973-020-09775-9. URL https://doi.org/10.1007/s10973-020-09775-9.

[287] A. Cisse, A. Marquette, M. Altangerel, J. Peters, and B. Bechinger. Investigation of the action of peptides on lipid membranes. The Journal of Physical Chemistry B, 125(36):10213–10223, August 2021. doi: 10.1021/acs.jpcb.1c06388. URL https://doi.org/10.1021/acs.jpcb.1c06388.

Appendix A

The dynamical Matryoshka model

In parallel of the thesis project about LDLs, I participated in a collaboration between Dr. Dominique Bicout, theoretician at the ILL (theory group), and my thesis director, Pr. Judith Peters, about the introduction of a new theoretical model for probing local lipid dynamics in quasi-elastic neutron scattering (QENS). This work was also joined by Dr. Tatsuhito Matsuo, researcher at the Institute for Quantum Life Science (NIQRST, Tokai, Ibaraki, Japan) and the Japan Proton Accelerator Research Complex (J-PARC, Japan), in exchange for two years in our group.

This joint work gave rise to the dynamical Matryoshka model, which was tested against experimental data on standard phospholipids systems (bilayers and vesicles), measured on several instruments few years ago by Pr. Judith Peters with Dr. Francesca Natali (IN6 data, 2013) and Dr. Marie Plazanet (IN5 data, 2017). Three papers are considered for publication on BBA - Biomembranes:

- The first part, Bicout et al. [134], was submitted and presents the associated formalism and theory on which the dynamical Matryoshka model is relying. It is attached below, and can also be found on bioRxiv.

- The second part, Cisse et al. [284], is close to the submission, and incorporates the application of the model to various phospholipids systems: DMPC bilayers, tail-deuterated or fully protonated, measured in-plane or out-of-plane, vesicles, DMPG bilayers... In that case, the Matryoshka model was applied on the amplitudes of the Lorentzian functions, and thus concerned the geometry of motions.

- The third part, Matsuo et al. [285], is also close to the submission, and extends the model application to the Lorentzian linewidths, so the diffusive properties of motions. Classical models, as well as the Matryoshka model, were tested against the experimental data.

The dynamical Matryoshka model: 1. Incoherent neutron scattering functions for lipid dynamics in bilayers

Dominique J. Bicout[a,d,*], Aline Cisse[b,d], Tatsuhito Matsuo[b,c,d], Judith Peters[b,d,e]

[a] *Univ. Grenoble Alpes, CNRS, Grenoble INP, VetAgro Sup, TIMC, 38000 Grenoble, France*
[b] *Univ. Grenoble Alpes, CNRS, LiPhy, Grenoble, France*
[c] *Institute for Quantum Life Science, NIQRST, Tokai, Ibaraki, 319-1106, Japan*
[d] *Institut Laue-Langevin, 71 Avenue des Martyrs, 38042 Grenoble, France*
[e] *Institut Universitaire de France, France*

Abstract

Fluid lipid bilayers are the building blocks of biological membranes. Although there is a large amount of experimental data using inconsistent quasi-elastic neutron scattering (QENS) techniques to study membranes, very little theoretical works have been developed to study the local dynamics of membranes. The main objective of this work is to build a theoretical framework to study and describe the local dynamics of lipids and derive analytical expressions of inconsistent diffusion functions (ISF) for QENS. As results, we developed the dynamical Matryoshka model which describes the local dynamics of lipid molecules in membrane layers as a nested hierarchical convolution of three motional processes: (i) individual motions described by the vibrational motions of H-atoms; (ii) internal motions including movements of the lipid backbone, head groups and tails, and (iii) molecule movements of the lipid molecule as a whole. The analytical expressions of the ISF associated with these movements are all derived. For use in analyzing the QENS experimental data, we also derived an analytical expression for the aggregate ISF of the Matryoshka model which involves an elastic term plus three inelastic terms of well-separated time scales and whose amplitudes and rates are functions of the lipid motions. And as an illustrative application, we used the aggregated ISF to analyze the experimental QENS data on a lipid sample of multilamellar bilayers of DMPC (1,2-dimyristoyl-*sn*-glycero-3-phosphocholine). It is clear from this analysis that the dynamical Matryoshka model describes very well the experimental data and allow extracting the dynamical parameters of the studied system.

Keywords: lipids, bilayers, local dynamics, quasi-elastic neutron scattering, modeling

1. Introduction

Biological membranes are complex lipid-rich systems that constitute fundamental interfaces and selectively permeable barriers for the compartmentalization that defines cells and

*Corresponding author
Email address: bicout@ill.fr (Dominique J. Bicout)

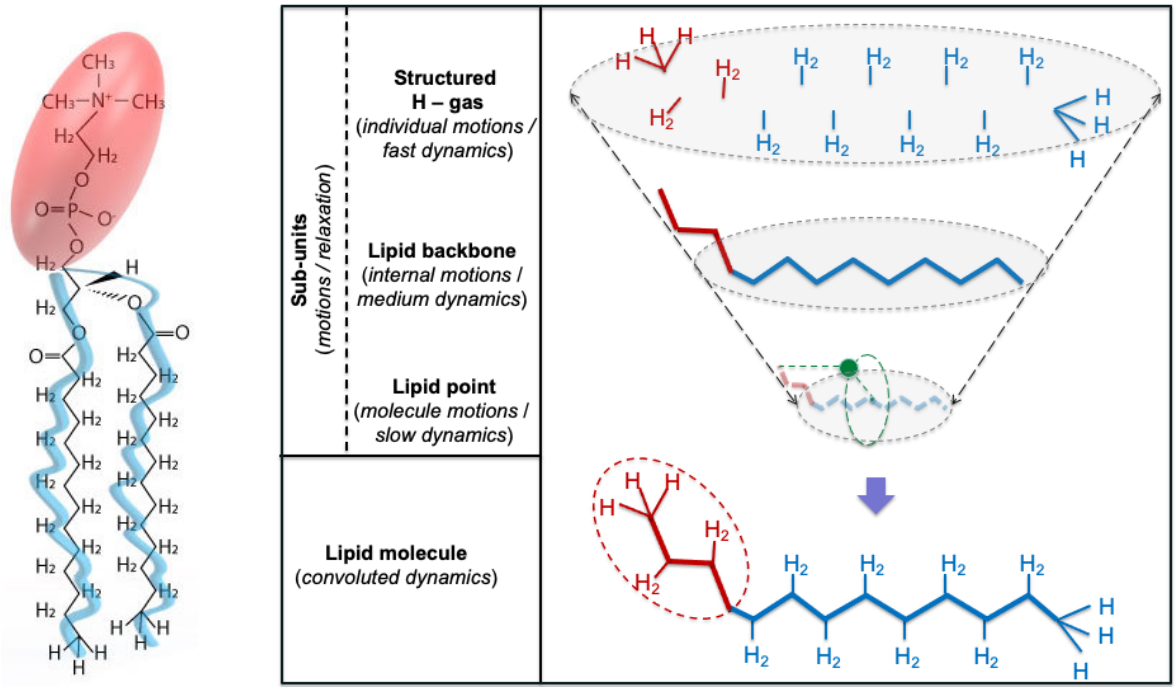
organisms. Biological membranes are composed of lipids, which self-assemble into bilayers, proteins and carbohydrates [1]. In addition to separating the interior from the exterior of cells, for example, membranes are very dynamic systems that host and ensure many essential processes vital for cellular functions, such as the transport of proteins or ions [1, 2]. This dynamics is made up of both the activity at the surface and the movements of the bilayers which give the membrane an important fluid character to ensure its functions. Since lipids are the most abundant constituents of membranes, studies on the dynamics of lipids in membranes are and remain very crucial. Due to their complex structure and dynamics, lipid bilayers are characterized by a hierarchy and heterogeneity of motions over a wide range of time and space. These dynamics include, for example, local movements like lipid rotational, in-and-out of the plane diffusion at very short spatial scales and time scales of pico to nanoseconds, but also collective lipid movements like density fluctuations of short wavelength in pico to nanosecond range and long-wavelength flip-flop or undulation and bending modes of the bilayer in the nano to microsecond range [2, 3, 4]. In this work, we will only deal with short-range movements carried over short time scales (pico to nanoseconds).

Over the years, incoherent quasi-elastic neutron scattering (QENS) has proven to be a key technique for investigations of lipid motions at the pico - to nanosecond time scale and there is a great amount of experimental data that have been accumulated using QENS techniques [5, 6, 7, 8, 9, 10]. However, since the seminal work by Pfeiffer et al. [5], there are very few theoretical works that have been developed to analyze and describe local dynamics of membranes [5, 11]. Along these lines, a study on phospholipid membranes has been proposed to separate the motions over three distinct time scales [12]. Using such models with QENS data allow to retrieve parameters like mean size of the solvent lipid cages, diffusion coefficients lipid rotations or in-plane Brownian motions. Motivated by these experimental investigations and findings, our main objective in this paper is to provide a theoretical framework of a model of membrane layers and to derive analytical expressions of incoherent structure functions (ISF) describing the local dynamics. To this end, we have developed a model that describes a membrane layer as a system of dynamically equivalent lipid molecules, each of which consisting of two connected (via a backbone) bodies (head and tail) undergoing kinds of internal and body motions.

The remaining of the paper is organized as follows. Section 2 consists of three parts: (i) formulation of the dynamical Matryoshka model which describes the hierarchical convolution of movements in the dynamics of a lipid molecule, (ii) detailed description of motions that are included in the model and derivations of analytical expressions of the associated ISFs, and (iii) derivation of an aggregate expression of the global ISF to be used to analyze and fit the experimental data. Technical details of all derivations are described in the Appendices A - C. Section 3 illustrates how the developed theory can be used to analyze the experimental QENS data of a bilayer membrane and extract the parameters of interest. Finally, the main results of the paper are summarized in the Section 4.

2. Model formulation and analytical expressions

The membrane layer is considered as a system made of structurally and dynamically equivalent lipid molecules interacting with each other. Among the whole hierarchy of motions characterizing the membrane dynamics, we are interested in the local motions of lipid molecules as can be studied using techniques like incoherent quasi-elastic neutron scattering, inelastic x-ray scattering, to cite a few [5, 6, 7, 8, 9, 10]. Figure 1A provides an illustration of the types of phospholipid molecules we will be dealing with. To describe the motional processes of an individual lipid molecule, occurring in the potential of mean force generated by the sea of lipid molecules in the membrane layer, we develop the Matryoshka model described below (see Fig. 1B).



A: Lipid molecule **B:** Dynamical Matryoshka model

Figure 1: **A:** Illustration of a phospholipid molecule 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) consisting of a hydrophilic head (in red) and hydrophobic fatty acid tails (in blue) with the distribution of H-atoms. Courtesy from Steph Monfront (ILL). **B: Dynamical Matryoshka model** for the dynamics of the lipid molecule represented as a funnel of three - level convoluted dynamic processes. Lipid tails are represented by an effective single tail for the dynamics. Red and blue colored elements relate to head and tail groups, respectively. H-atoms with C-H bounds (top) are structured along the lipid backbone (middle), with head and tail subunits, mirroring the lipid molecule representation (bottom of the figure). At the bottom of the funnel, the lipid molecule as a whole (rigid body) comes down to a point particle (green sphere) representing the center of inertia (barycenter) of all dynamic H-atoms with respect to the lipid main axis (dashed lipid backbone).

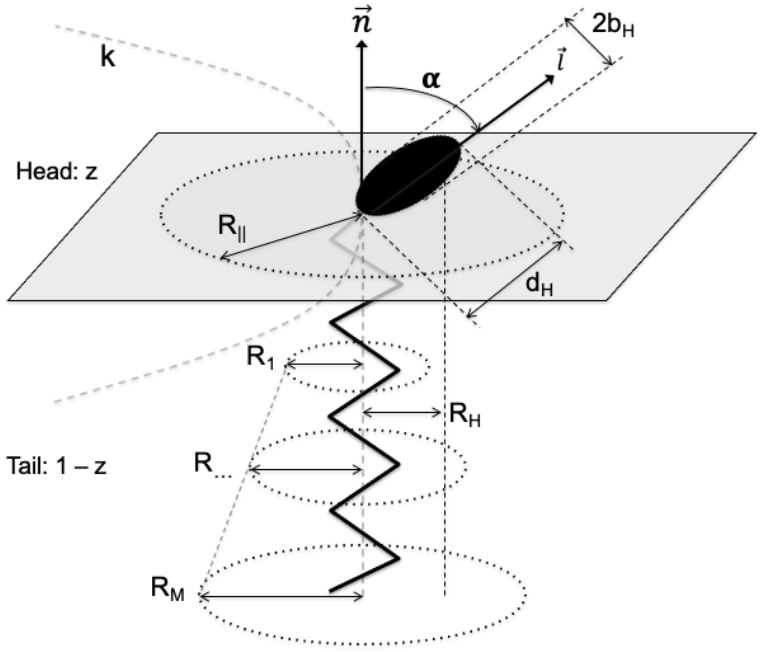


Figure 2: Scheme of a lipid molecule showing the characteristic parameters of motions. The polar head group is represented by the black ellipse (tilted by an angle α) of short and long diameters $2b_{\mathrm{H}}$ and d_{H}, respectively, and the hydrophobic tails are represented by a dynamically effective tail (vertical zigzag black line) of apparent length M. Grey plane separates the head group (with a fraction z of H-atoms) from the tail (with a fraction $1-z$ of H-atoms). The membrane normal (lipid axis) and head axis (tilted by an angle α) are indicated by $\vec{n}$ and $\vec{l}$, respectively. Dashed circle of radius $R_{||}$ in the grey plane represents the effective cage for the in-plane diffusion and the dashed parabola of stiffness k, perpendicular to the membrane plane, indicates in-out of the plane motions of the lipid molecule. Dashed circles of increasing radii, $R_1, \cdots, R_M$, around the effective tail represent areas for the diffusion motions of Hs along the tails and R_{H} is the distance between the lipid axis $\vec{n}$ and the center of inertia of all H-atoms with respect to $\vec{n}$.

2.1. Dynamical Matryoshka model

As illustrated in Fig. 1B, the dynamical Matryoshka model describes the dynamics of a lipid molecule as resulting from the hierarchical convolution of three motional processes (from fastest to slowest motions): (i) *individual motions* of H-atoms forming the cloud of structured H-atoms bound to backbone atoms, (ii) *internal motions* of the lipid backbone or skeleton made up of non-H atoms and H-atoms bound to them, the head and tail subunits, and (iii) *molecule motions* or rigid body motions of the lipid molecule as a whole represented by the movements of the center of inertia all dynamic H-atoms with respect to the lipid main axis. The metaphor of dynamical Matryoshka (nesting dolls) originates from what these motional processes occur in a superimposed and nested way over various (and overlapping) timescales and, therefore, are resolved by zooming in or out over the associated characteristic timescales.

2.2. Dynamical Matryoshka model: Incoherent Structure Function (ISF)

Within the framework of this model and the hypothesis of dynamic independence for different motions, the ISF for the local motions of a lipid molecule can be written as,

$$I(Q,t) = I_{\rm ind}(Q,t) \times I_{\rm int}(Q,t) \times I_{\rm mol}(Q,t)\,, \tag{1}$$

where $I_{\rm ind}(Q,t)$, $I_{\rm int}(Q,t)$ and $I_{\rm mol}(Q,t)$ are the ISFs associated to individual, internal and molecular motions, respectively. In the following, we detail the motional processes of lipids included in the dynamical Matryoshka model with, whenever possible, the associated potentials of mean force, $V(\mathbf{r})$. The elastic incoherent structure factor (EISF) for each motion can be calculated as (see, Table 4),

$$\mathrm{EISF}(Q) = \left| \frac{\int \mathrm{e}^{i\mathbf{Q}.\mathbf{r}} \mathrm{e}^{-\beta V(\mathbf{r})}\, d\mathbf{r}}{\int \mathrm{e}^{-\beta V(\mathbf{r})}\, d\mathbf{r}} \right|^2 \;\; ; \;\; \beta^{-1} = \mathrm{k_B T}\,,\ \text{thermal energy}\,. \tag{2}$$

The relationships between all the dynamical processes are summarized in Table 1 with the details of derivations of IFSs provided in Appendix A and the analytical expressions of the ISFs reported in Table A.7.

Table 1: **Dynamical processes**

Dynamics	**Subunit**	***Motions***	**Process**
Individual	H-atoms	*vibrational Debye-Waller*	Very Fast
Internal	C-H groups (backbone) ⊛ Tail	*jump dynamics* ⊛ *diffusion in circles*	Fast
	C-H groups (backbone) ⊛ Head	*jump dynamics* ⊛ *rotational diffusion* ⊛ *flip-flop*	Intermediate
Molecule	Lipid Molecule	*rotational diffusion*	Intermediate
		in-out of the plane diffusion ⊛ *2d-diffusion in a cage*	Slow

2.2.1. Individual motions

Individual motions relate to the vibrational motions of the cloud of H-atoms bound to the lipid backbone atoms. Such motions are described by a harmonic potential of mean force, $V(r) = r^2/2\langle u^2\rangle$, where $\langle u^2\rangle$ is the ensemble mean-square displacements of H-atoms about their equilibrium positions. These individual movements being relatively fast ($\sim$ 100 meV, [13]) for the timescales that concern us, the associated ISF is reduced to the EISF (i.e., time independent) and, in fact, is factored out in Eq.(1). The ISF for individual motions is therefore given by the Debye-Waller factor as, $I_{\rm ind}(Q,t) = I_{\rm DW}(Q) = A_{\rm DW}(Q)$ (see Table A.7).

2.2.2. Internal motions

Internal motions result from the combination of three motional processes from the lipid backbone, head and tail subunits. The headgroup and tails are considered as independent subunits each with z and $1-z$ fraction of H-atoms, respectively, and the lipid backbone common to both subunits includes all H-atoms (see Fig. 2). The ISF for internal motions writes as,

$$I_{\text{int}}(Q,t) = I_{\text{bkb}}(Q,t)\left[zI_{\text{head}}(Q,t) + (1-z)I_{\text{tail}}(Q,t)\right] \ , \tag{3}$$

where $I_{\text{bkb}}(Q,t)$, $I_{\text{head}}(Q,t)$ and $I_{\text{tail}}(Q,t)$ are the ISFs for motions in the lipid backbone, head and tail.

- *ISF for the Backbone motions,* $I_{\text{bkb}}(Q,t)$: Backbone motions relate to the motions of C-H groups in the lipid molecule. Several models can be used to describe the heterogeneous dynamics of C-H groups depending on whether they are located in the head group or in the tails. For simplicity, we considered that the movements of all the C-H groups could be described using a jump dynamics between two non-equivalent sites distant from d and characteristic relaxation rate Γ_{jump}. Because of heterogeneity, d and Γ_{jump} are ensemble averaged quantities over distributions of site distances and relaxation times, respectively. The potential of mean force for these motions is defined as [14],

$$\frac{\mathrm{e}^{-\beta V(\mathbf{r})}}{\int \mathrm{e}^{-\beta V(\mathbf{r})}\, d\mathbf{r}} = \begin{cases} \phi & ; \quad \mathbf{r} = \mathbf{r}_1 \\ 1-\phi & ; \quad \mathbf{r} = \mathbf{r}_1 + d \\ 0 & ; \quad \text{otherwise} \end{cases} \tag{4}$$

  where $\mathbf{r}_1$ is the position of the low energy site and ϕ is the probability of occupying $\mathbf{r}_1$ (see Table A.7) and the ISF, $I_{\text{bkb}}(Q,t) = I_{\text{jump}}(Q,t)$, is given in Table A.7.

- *ISF for Tail motions,* $I_{\text{tail}}(Q,t)$: Tail motions refer to the movements of the H-atoms in all the tails. In Fig. 2, the tails, which interact with each other, are represented by a dynamically effective tail which indicates the axis of the tails and especially the surfaces within which the movements of the H-atoms take place. The movements of H-atoms in the tails are modeled by classical two-dimensional diffusions, all of an effective diffusion constant D_{tail} within circles of radii R_m in the plane parallel to the membrane at m, where m designates the position of the carbon atoms $C-C$ along the lipid tail from the start to the tail end (with, $1 \leq m \leq M$) and M is the apparent length of the lipid tail. As shown in Fig. 2, as a result of interactions between tails, R_m increases with m as the explorable space increases with m. Unlike the linear increase of Carpentier et al. [15], we use instead, $R_m = R_1\sqrt{m}$, to somehow account for a random walk-like (along the tail axis) positions of carbon atoms to which hydrogens are bound. The potential of mean force for these motions at each m is,

$$V(r,\theta) = \begin{cases} 0 & ; \quad 0 \leq r \leq R_m \ , \ 0 \leq \theta \leq 2\pi \\ \infty & ; \quad \text{otherwise} \end{cases} \tag{5}$$

  and the ISF $I_{\text{tail}}(Q,t)$ is given in Table A.7.

- *ISF for Head motions,* $I_{\text{head}}(Q,t)$: Headgroup motions, involved in the roughness of the membrane surface, include three independent motions (see Fig. 2): the uniaxial rotational diffusion about the head axis $\vec{l}$, the flip-flop or jump dynamics of $\vec{l}$ axis between angles $-\alpha$ and α about the membrane normal $\vec{n}$ axis and the uniaxial rotational diffusion of the $\vec{l}$ axis about the membrane normal $\vec{n}$ axis. However, for our purpose, by symmetry the rotational diffusion about the $\vec{n}$ axis averages out such that only the rotation about $\vec{l}$ and the flip-flop about $\vec{n}$ are retained. The headgroup size b_{H} (see Fig. 2) represents the distance between the head axis and the center of inertia of all H-atoms of the head group with respect to the head axis $\vec{l}$. Derivation of the ISF $I_{\text{head}}(Q,t)$ is detailed in Appendix A.

2.2.3. ISF for molecule motions: $I_{\text{mol}}(Q,t)$

For molecule or rigid body movements, all dynamic H-atoms together perform the same processes regardless of their location in the lipid molecule. Formally, the dynamics of the lipid molecule as a whole can therefore be described by that of the center of inertia of all dynamic H-atoms with respect to the lipid main axis (Fig. 1B). Three types of independent movements are considered: rotational motions around the axis $\vec{n}$ of the lipid molecule, in-out of the membrane plane movements (parallel to $\vec{n}$) and in-plane lateral diffusion of the lipid molecule. The ISF for molecular motions writes as,

$$I_{\text{mol}}(Q,t) = I_{\text{rot}}(Q,t) \times I_{\text{in-out}}(Q,t) \times I_{\text{2d}}(Q,t)\,, \tag{6}$$

where $I_{\text{rot}}(Q,t)$, $I_{\text{in-out}}(Q,t)$ and $I_{\text{2d}}(Q,t)$ are the ISFs for rotational, in-out of the plane and in-plane diffusion motions.

- *ISF for rotational motions,* $I_{\text{rot}}(Q,t)$: For the rotational motions of the lipid molecule about its axis $\vec{n}$, all H-atoms perform exactly the same rotational movement about the lipid axis $\vec{n}$. However, as not all of the H-atoms are located equidistant from the lipid axis, the resulting ISF should be the sum of the individual ISFs. As a first approximation, we can reduce the rotational motions of all H-atoms to that of their center of inertia and describe the rotational motions of the lipid molecule about its axis $\vec{n}$ by a rotational diffusion of diffusion coefficient D_{rot} on a circle of radius R_{H}, where R_{H} is the distance between the lipid axis and the center of inertia of all dynamic H-atoms with respect to the lipid axis $\vec{n}$. The ISF $I_{\text{rot}}(Q,t)$ is given in Table A.7.

- *ISF for in-out of the plane motions,* $I_{\text{in-out}}(Q,t)$: The in-out-of-plane movements, involved in the roughness of the membrane surface, relate on the up and down motions (normal to the membrane plane) around the equilibrium position of the lipid molecule within the membrane layer. Such motions can be described by the one-dimensional (parallel to $\vec{n}$) diffusion within a harmonic potential of force constant k and relaxation time τ. The harmonic potential of mean force is given by, $V(z) = kz^2/2$, where z is the lipid molecule position around its equilibrium position. The ISF $I_{\text{in-out}}(Q,t)$ is given in Table A.7.

- *ISF for in-plane 2d diffusion motions,* $I_{2\mathrm{d}}(Q,t)$: In general, in-plane lateral diffusion of lipid molecules is composed of short range local diffusion in a solvent cage and long range jumps between different sites allowing molecules to travel within the membrane layer. As this work is dealing with spatially short range dynamics, we will only consider short range diffusion in what follows. Thus, in-plane diffusion of lipid molecules describes when molecules exchange places via Brownian motion within a cage. Such motions are described by a two-dimensional diffusion of diffusion constant $D_{||}$ within a circle of radius $R_{||}$, where $R_{||}$ is the effective size of the cage formed by the neighboring lipid molecules. The potential of mean force for such a confined isotropic two-dimensional diffusion is given by,

$$V(r,\theta) = \begin{cases} 0 & ; \quad 0 \leq r \leq R_{||} \\ \infty & ; \quad \text{otherwise} \end{cases} , \quad \forall\, \theta \tag{7}$$

  where r is the position of lipid in the membrane plane. The ISF $I_{2\mathrm{d}}(Q,t)$ is given in Table A.7.

All the 7 movements included in the dynamical Matryoshka model and associated 18 dynamical parameters are summarized in the Table 2.

Table 2: **Local motions and associated parameters of the Matryoshka model.**

Dynamics	***Motions***	**Parameter**	**Parameter definition**
Individual	*Vibrational Debye-Waller*	$\langle u^2 \rangle$	H's mean-square displacements
Internal	*Backbone: jump dynamics*	d	two-site distance
		ϕ	probability of jump events
		Γ_{jump}	jump relaxation rate
	Tails: diffusion in circles	$1-z$	fraction of H's in the tail
		M	length of the effective tail
		R_1	radius of the first circle
		D_{tail}	diffusion coefficient of tail H's
	Head: rotational diffusion+filp-flop	α	head tilt angle
		b_{H}	headgroup H-radius
		τ_{ff}	inter flip-flop mean resting time
		D_{head}	rotational diffusion coefficient
Molecule	*Rotational diffusion*	R_{H}	H-radius of the lipid molecule
		D_{rot}	rotational diffusion coefficient
	In-out of the plane diffusion	k	force constant
		τ	relaxation time
	2d-diffusion in a cage	$R_{||}$	lipid cage radius
		$D_{||}$	molecular diffusion coefficient

2.3. Aggregated processes: 3*-dynamical process approximation*

As outlined above, the ISF of local movements of lipid molecules results from the convolution of several individual ISFs which, in general, are multi-exponential functions of time.

Even when each individual ISF was approximated by a single exponential relaxation, the resulting ISF would still be multi-exponential. In practice, it would be challenging to use such a multi-exponential function in the analyzes of experimental data to extract physical parameters of interest listed in Table 2. The idea is therefore to develop an approximation of the ISF allowing to reduce timescales by coupling them together while keeping the original information (as summarized in Table 2). For that purpose, we follow Wanderlingh et al. [12] and consider how the dynamics of the lipid molecules as described above can be approximatively aggregated into 3-dynamical processes. Under the context of dynamical processes taking place on widely separated time scales (one or few orders of magnitude) as outlined above, the main value of the 3-dynamical process model is to allow fitting the experimental data and, therefore, providing an opportunity for direct connection between analytical models of lipid motions and experimental observations.

The 3-dynamical process model is a phenomenological model in which the position $\mathbf{r}(t)$ at time t of the H-atom in lipid molecules can be split into three independent components, $\mathbf{r}(t) = \mathbf{r}_1(t) + \mathbf{r}_2(t) + \mathbf{r}_3(t)$, where $\mathbf{r}_1(t)$, $\mathbf{r}_2(t)$ and $\mathbf{r}_2(t)$ stand for the slow, intermediate and fast motions, respectively, describing the overall motions of hydrogen atoms in the equilibrium lipid molecule at very different time scales. Each of these motions accounts for a confined dynamics with a quasi-elastic term described by a single exponential decay term. Under the approximation of separate time scales, the aggregated ISF can be written as,

$$I_{\rm agg}(Q,t) = A_{\rm DW}(Q)\Big[B_0 + B_1\,{\rm e}^{-\Gamma_1 t} + B_2\,{\rm e}^{-\Gamma_2 t} + B_3\,{\rm e}^{-\Gamma_3 t}\Big]\ ;\ \Gamma_0 = 0 < \Gamma_1 < \cdots < \Gamma_3\,, \quad (8)$$

where $A_{\rm DW}(Q)$ stands for the Debye-Waller factor, B_0 is the overall EISF, B_i and Γ_i for $i = 1, 2, 3$ are the amplitudes and relaxation rates of the slow, intermediate and fast motions, respectively, such that, $\sum_{i=0}^{3} B_i(Q) = 1$ for all Q. Accordingly, the measurable function in incoherent neutron scattering experiments is,

$$\begin{aligned} S(Q,\omega) &= \mathcal{R}(Q,\omega) \circledast \left\{\frac{1}{\pi}\int_0^{+\infty} dt\, I_{\rm agg}(Q,t)\cos(\omega t)\right\} \\ &= \mathcal{R}(Q,\omega) \circledast \left\{A_{\rm DW}(Q)\sum_{i=0}^{3} B_i(Q)\mathcal{L}_i(Q,\omega)\right\}\ ;\ \mathcal{L}_i(\Gamma_i;Q,\omega) = \frac{1}{\pi}\,\frac{\Gamma_i(Q)}{\omega^2 + \Gamma_i^2(Q)}\,, \end{aligned} \quad (9)$$

where $\mathcal{R}(Q,\omega)$ refers to the experimental resolution function, $B_i(Q)$ and $\Gamma_i(Q)$ are the respective areas and half-widths at half-maximum (HWHM) of the Lorentzian functions $\mathcal{L}_n(Q,\omega)$. It should be stressed here that:

(i) the three time scales slow, intermediate and fast do not necessarily correspond or coincide to and, therefore, are not to be confused with those of individual, internal and collective motions (see Table 1);

(ii) the phenomenological model does not provide which movements are resolved at each timescale.

The purpose of aggregating the $I(Q,t)$, derived above and given in Eq.(1), into 3-dynamical processes is to associate each term in Eq.(8) with a physical meaning and an analytical expression. To this end, we use the mapping, $I(Q,t) \approx I_{\rm agg}(Q,t)$, i.e.,

$$\begin{aligned} & I_{\rm jump}(Q,t)\Big[zI_{\rm head}(Q,t) + (1-z)I_{\rm tail}(Q,t)\Big] I_{\rm rot}(Q,t) I_{\rm in-out}(Q,t) I_{\rm 2d}(Q,t) \\ & \approx B_0 + B_1\,{\rm e}^{-\Gamma_1 t} + B_2\,{\rm e}^{-\Gamma_2 t} + B_3\,{\rm e}^{-\Gamma_3 t}\,, \end{aligned} \tag{10}$$

to derive the expressions of B's and Γ's as functions of physical quantities in $I(Q,t)$. Note that in Eq.(10), B_0 is already known and is equal to $B_0 = I(Q, t\to\infty) = I_{\rm agg}(Q, t\to\infty)$ (i.e., the product of EISF of all local motions), whereas the $B_{i>0}$ and $\Gamma_{i>0}$ remain to be determined. In addition, the number of terms after the approximate sign "$\approx$" in Eq.(10) can be reduced when some of the amplitudes cancel out because the contributing motions turn out to be not observable. The aggregation procedure, based on the hierarchy of relaxation time scales of motional processes in the Matryoshka model, is detailed in Appendix B and the results are summarized in Table 3 in which the expressions of original EISFs and relaxation rates are given in Tables 4 and 5, respectively. Table 3 provides which local motions contribute at each timescale of the 3-dynamical process and how they are combined in the expressions of amplitudes and relaxation rates. Each relaxation rate is not exclusive of a single local motion and the same local motion can therefore contribute to different timescales. The amplitudes result from the combination of several local motions whose relaxation times can be relatively different. From the amplitudes (with the Debye-Waller factoring all), we have the following contributions:

- *Slow motions:* backbone, head, tails, rotational, in-out of the plane and in-plane diffusion;
- *Intermediate motions:* backbone, head, tails and rotational;
- *Fast motions:* backbone and tails.

And, as a function of Q, we have:

- $Q\to 0$ limit: as expected, the scattering function essentially consists of the elastic peak (EISF) with all motions contributing and no Lorentzian functions as, $B_0(Q\to 0)\to 1$ and $B_{i>0}(Q\to 0)\to 0$;
- Moderate Q: as all $B_i \neq 0$, the scattering function consists of the elastic peak plus the three Lorentzian functions with contributions of all motions. Note that slow motions are only observable for these Q-ranges as $B_1(Q)\to 0$ in both $Q\to 0$ and large Q limits;
- Large Q limit: as $B_{i\leq 1}(\text{large}\,Q)\to 0$ and $B_2(\text{large}\,Q) = 1 - B_3(\text{large}\,Q) \approx zA_{\rm bkb}(\text{large}\,Q) \neq 0$, the scattering function reduces to only two Lorentzian functions involving intermediate and fast motions and concerning motions of C-H groups of the backbone.

Table 3: **Amplitudes (areas) and relaxation rates of the 3-dynamical process approximation.** $I_{\rm agg}(Q,t) = A_{\rm DW}(Q)\left[B_0 + B_1\,{\rm e}^{-\Gamma_1 t} + B_2\,{\rm e}^{-\Gamma_2 t} + B_3\,{\rm e}^{-\Gamma_3 t}\right]$

Motions	**Amplitudes (areas)**	
	theoretical$^{(a)}$	**experimental**$^{(b)}$
EISF	$B_0 = [zA_{\rm head} + (1-z)A_{\rm tail}]\,A_{\rm jump}A_{\rm rot}A_{\rm in-out}A_{\rm 2d}$	$A_0 = mB_0 + \varepsilon_0$
Slow	$B_1 = [zA_{\rm head} + (1-z)A_{\rm tail}]\,(1 - A_{\rm in-out}A_{\rm 2d})A_{\rm jump}A_{\rm rot}$	$\left.\begin{array}{l}\\\\\\\end{array}\right\}$ $A_i = mB_i + \dfrac{[1-m-\varepsilon_0]}{3}$
Intermediate	$B_2 = A_{\rm jump}\,[z(1 - A_{\rm head}A_{\rm rot}) + (1-z)A_{\rm tail}(1-A_{\rm rot})]$	
Fast	$B_3 = z(1-A_{\rm jump}) + (1-z)(1-A_{\rm jump}A_{\rm tail})$	
	Relaxation rates	
Slow	$\Gamma_1 = \dfrac{(1-A_{\rm in-out})\Gamma_{\rm in-out} + (1-A_{\rm 2d})\Gamma_{\rm 2d}}{1 - A_{\rm in-out}A_{\rm 2d}}$	
Intermediate	$\Gamma_2 = \dfrac{z(1-A_{\rm head})\Gamma_{\rm head} + [z + (1-z)A_{\rm tail}]\,(1-A_{\rm rot})\Gamma_{\rm rot}}{z(1-A_{\rm head}A_{\rm rot}) + (1-z)A_{\rm tail}(1-A_{\rm rot})}$	
Fast	$\Gamma_3 = \dfrac{(1-A_{\rm jump})\Gamma_{\rm jump} + (1-z)(1-A_{\rm tail})\Gamma_{\rm tail}}{z(1-A_{\rm jump}) + (1-z)(1-A_{\rm jump}A_{\rm tail})}$	

$^{(a)}$ Limiting properties for B_i: $B_0(Q\to 0) = 1$ and $B_{i>0}(Q\to 0) = 0$; $B_i({\rm high}\,Q) \to 0$ for both $i = 0$ and $i = 1$ while $B_2({\rm large}\,Q) = 1 - B_3({\rm large}\,Q) = zA_{\rm jump}({\rm large}\,Q)$ where $A_{\rm jump}({\rm large}\,Q) = A_{\rm bkb}({\rm large}\,Q) \approx 1 - 2\phi(1-\phi)$.

$^{(b)}$ See Appendix C for derivations. m is the fraction of dynamical or mobile H-atoms and the error function $\varepsilon_0(Q)$ accounts for immobile H-atoms, multiple scattering effects observable at low-Q, and others. For simplicity, we will assume that the errors are homogeneously distributed over length scales and, therefore, $\varepsilon_0(Q) = \varepsilon_0$.

Table 4: **Elastic Incoherent Structure Factor (EISF) of dynamical processes.**

Subunit	**EISF**	[a] $(\|\|, \perp)$
H-atom	*vibrational motions: Debye-Waller factor* $A_{\mathrm{DW}}(Q) = \exp\left\{-\frac{Q^2\langle u^2\rangle}{3}\right\}$	(Y,Y)
C-H groups	*jump dynamics between 2 non-equivalent sites distant from d* $A_{\mathrm{jump}}(Q) = 1 - 2\phi(1-\phi)[1 - j_0(Qd)]$	(Y,Y)
Tail	*2d-diffusion of tail H-atoms inside $F(m)$ - distribution of circles of radius R_m* [b] $A_{\mathrm{tail}}(Q) = \sum_{m=1}^{M} F(m)\left[\frac{2J_1(QR_m\sin\gamma)}{QR_m\sin\gamma}\right]^2$	(Y,N)
Head	*rotational diffusion of the headgroup about its axis plus flip-flop motion* [c] $A_{\mathrm{head}}(Q) = \left[\sum_{l=-\infty}^{+\infty} J_{2l}(Qb_{\mathrm{H}}\cos\gamma\sin\alpha) J_l\left(Qb_{\mathrm{H}}\sin\gamma\cos^2\frac{\alpha}{2}\right) J_l\left(Qb_{\mathrm{H}}\sin\gamma\sin^2\frac{\alpha}{2}\right)\right]^2$	(Y,Y)
Molecule	*rotational diffusion of the lipid molecule about the membrane normal axis* $A_{\mathrm{rot}}(Q) = J_0^2(QR_{\mathrm{H}}\sin\gamma)$	(Y,N)
	in-out of the plane 1d-diffusion of the lipid molecule in a harmonic potential $A_{\mathrm{in-out}}(Q) = exp\left\{-(Q\cos\gamma)^2\left(\frac{k_{\mathrm{B}}T}{k}\right)\right\}$	(N,Y)
	in-plane 2d-diffusion of the lipid molecule inside a circle of radius $R_{\|\|}$ $A_{\mathrm{2d}}(Q) = \left[\frac{2J_1(QR_{\|\|}\sin\gamma)}{QR_{\|\|}\sin\gamma}\right]^2$	(Y,N)

[a] **Q** direction: $\perp \Leftrightarrow \gamma = 0$ and $|| \Leftrightarrow \gamma = \frac{\pi}{2}$ refer to **Q** directions perpendicular and parallel to the membrane plane, respectively. "Y=Yes" (i.e., $A_i(Q)$ changes with Q) and "N=No" (i.e., $A_i(Q) = 1$ for all Q) indicate observable and non-observable motions, respectively. Limiting properties: as we are dealing with spatially restricted motions, $A_\alpha(Q=0) = 1$ and $A_\alpha(\mathrm{high}\, Q) \to 0$ for all continuous motions and for the jump process between small number of sites for (backbone) $A_{\mathrm{jump}}(\mathrm{large}\, Q) \neq 0$ and oscillates around, $1 - 2\phi(1-\phi)$.

[b] For simplicity, we use $F(m) = 1/M$ and $R_m = R_1\sqrt{m}$.

[c] We have: $A_{\mathrm{head},\perp}(Q) = J_0^2(Qb_{\mathrm{H}}\sin\alpha)$ and $A_{\mathrm{head},||}(Q) = J_0^2\left(Qb_{\mathrm{H}}\cos^2\frac{\alpha}{2}\right) J_0^2\left(Qb_{\mathrm{H}}\sin^2\frac{\alpha}{2}\right)$

3. Illustrative Example

To compare and test how the theoretical developments outlined above would work when analyzing experimental data, we considered the following neutron scattering experiments. The whole task consists in determining 18 parameters describing the 7 motions of H-atoms in the Matryoshka model (see, Table 2) plus 2 experimental parameters (m and ε_0 in Table 3).

3.1. Quasi-Elastic Neutron Scattering (QENS) experiments and analyses

We used lipid samples of multilamellar bilayers (MLBs) of DMPC (1,2-dimyristoyl-*sn*-glycero-3-phosphocholine) represented in Fig. 1A. The DMPC was purchased from Lipoid (Ludwigshafen Germany) or from Avanti Polar Lipids (Alabaster, USA) and used without further purification. The lipid samples were prepared on 6 Si wafers, and hydrated in D_2O atmosphere within a desiccator at full hydration. They were mounted on flat Aluminum sample holders, gold-coated to avoid sample contamination. The sealing of the cells was done by using Indium wire, and they were weighed before and after the experiment to check for any sample loss.

DMPC samples were scanned on the IN6 time-of-flight spectrometer from ILL (Grenoble, France), with a wavelength of 5.1 Å, corresponding to an energy resolution of about 70 μeV [17]. At this resolution, motions up to around 10 ps are accessible and the attainable Q-range is of, $0.37 \leq Q \leq 2.02\,\mathring{A}^{-1}$. The sample holder was oriented at 135° from the beam to access in-plane motions [8]. QENS scans were performed at three different temperatures, 280 K, 311 K and 340 K, to probe the dynamics below and above the main phase transition of lipid bilayers; DMPC is known to undergo consecutive phase transitions from the gel phase to the ripple phase around 287 K and to the fluid phase around 297 K (corresponding to the physiological state in cells) [18].

Empty cell with and without wafers, as well as Vanadium, were measured for correction and normalization purposes. Raw data were first corrected by the empty cell + 6 wafers contribution, using the Large Array Manipulation Program (LAMP) [19]. The resulting $S(Q,\omega)$ spectra were subsequently analyzed in the range of -10 meV $\leq \Delta E \leq$ 2 meV using IGOR Pro software (WaveMetrics, Lake Oswego, OR, USA). Following the QENS analysis in [12], the model used to fit the spectra is similar to that in Eq.(9) as,

$$S(Q,\omega) = A_{DW}(Q)\Big[\sum_{i=0}^{3} A_i(Q)\mathcal{L}_i(Q,\omega)\Big] \circledast \mathcal{R}(Q,\omega) + \mathcal{B}(Q)\,, \tag{11}$$

where $A_0(Q)$ is the experimental elastic incoherent structure factor (EISF) and the $A_i(Q)$ for $i > 0$ are the experimental areas of the Lorentzian functions $\mathcal{L}_i(Q,\omega)$ of half-widths at half-maximum $\Gamma_i(Q)$ as defined in Section 2.3. The experimentally determined $A_i(Q)$ are related to the areas $B_i(Q)$ in Eqs.(8) and (9) by expressions derived in Appendix C and given in Table 3. $\mathcal{R}(Q,\omega)$ refers to the resolution function and corresponds to the Vanadium measurements directly included in the analysis and $\mathcal{B}(Q)$ is a flat background, that can include the instrument contribution, or fast vibrational motions outside the window.

Table 5: **Relaxation rates of dynamical processes.**

Subunit	**Relaxation rates**[(a)]
C-H groups	*jump dynamics between 2 non-equivalent sites distant from d* $\Gamma_{\rm jump}$
Tail	*2d-diffusion of tail H-atoms inside $F(m)$ - distribution of circles of radius R_m* [(b)] $\Gamma_{\rm tail}(Q) = D_{\rm tail}\left[\sum_{m=1}^{M} F(m)\left(1-A_{\rm tail}(Q)\right)\left[\frac{\left(x_0^1\right)^2}{R_m^2}+Q^2\right]\right] \Big/ \left[\sum_{m=1}^{M} F(m)\left(1-A_{\rm tail}(Q)\right)\right]$
Head	*rotational diffusion of the headgroup about its axis plus flip-flop motion* $\Gamma_{\rm head} = D_{\rm head} + \frac{1}{\tau_{\rm ff}}$
Molecule	*rotational diffusion of the lipid molecule about the membrane normal axis* $\Gamma_{\rm rot} = D_{\rm rot}$ *in-out of the plane 1d-diffusion of the lipid molecule in a harmonic potential* [(c)] $\Gamma_{\rm in-out}(Q) = \left[(Q\cos\gamma)^2\left(\frac{k_{\rm B}T}{k}\right)\frac{1}{\tau}\right] \Big/ \left[1-exp\left\{-(Q\cos\gamma)^2\left(\frac{k_{\rm B}T}{k}\right)\right\}\right]$ *in-plane 2d-diffusion of the lipid molecule inside a circle of radius $R_{\|}$* $\Gamma_{\rm 2d} = \left[\frac{\left(x_0^1\right)^2}{R_{\|}^2}+Q^2\right] D_{\|}$

[(a)] Following Ref.[16], the relaxation rate for each dynamical process is given by, $\frac{1}{\Gamma_\alpha(Q)} = \int_0^\infty dt\left[\frac{I_\alpha(Q,t)-A_\alpha(Q)}{1-A_\alpha(Q)}\right]$ for α = jump, head, tail, rot, in − out, 2d. For the simplicity of calculations, the Γ_α reported in this table are derived at short time limit as, $\Gamma_\alpha(Q) = -\lim_{t\to 0}\frac{d}{dt}\left[\frac{I_\alpha(Q,t)-A_\alpha(Q)}{1-A_\alpha(Q)}\right]$.

[(b)] where $x_0^1 = 1.84118$. Note that $\Gamma_{\rm tail}(Q) = \left(x_0^1\right)^2\frac{D_{\rm tail}}{\langle R_m^2\rangle}$ in the $Q\to 0$ limit and $\Gamma_{\rm tail}(Q) = Q^2 D_{\rm tail}$ in the $Q\to\infty$ limit, where $\langle R_m^2\rangle = (M+1)\,R_1^2/2$ for $F(m) = 1/M$ and $R_m = R_1\sqrt{m}$.

[(c)] Note that $\Gamma_{\rm in-out}(Q) = 1/\tau$ in the $Q\to 0$ limit and $\Gamma_{\rm in-out}(Q) \approx (Q\cos\gamma)^2\left(\frac{k_{\rm B}T}{k}\right)\frac{1}{\tau}$ in the $Q\to\infty$ limit.

The areas retrieved from the QENS analyses were fitted with the A's in Table 3 using the package *lmfit* from Python [20] with Levenberg-Marquardt and Nelder-Mead algorithms. Fitting procedure relies on the simultaneous fitting of the four areas with corresponding functions given in Tables 3 and 4 through a set of shared parameters. In this way, both statistics are improved with more data points included and constraints on parameters are controlled.

3.2. Results

Main results within the framework of this analysis can be summarized as follows:

- Data from QENS experiments are analyzed as illustrated in Fig. 3 where the Eq.(11) is used to fit the data points and extract both the experimental Lorentzian amplitudes (areas) $A_i(Q)$'s and HWHM $\Gamma_i(Q)$'s. Figure 3 shows that the 3-dynamical process model describes very well (with residuals ~ 0) experimental data into an elastic peak ($\delta(\omega)$, a Lorentzian with $\Gamma_0 = 0$, and amplitude $A_0(Q)$) plus three well separated Lorentzians (with, $\Gamma_1 \sim 0.1$ meV, $\Gamma_2 \sim 1$ meV and $\Gamma_3 \sim 10$ meV, for slow, intermediate and fast motions, respectively; there is an order of magnitude between the Γ_i's). Such an analysis is performed for all values of the pair (Q, T) considered in the experiment.

  In the rest of this illustrative example, we will only deal with the $A_i(Q)$'s for a single sample in a given geometry whereas the use of this approach on several samples of lipid systems in different geometries with the analysis including $S(Q,\omega)$, $A_i(Q)$ and $\Gamma_i(Q)$ will be presented and discussed in more details elsewhere [21, 22]. Therefore, the number of parameters to determine is reduced to 11 for the Matryoshka model plus the 2 experimental parameters (m and ε_0). All the parameters are listed in the Table 6 where some are fixed (as extracted from the literature) and others are obtained from the best fit of the model to experimental data.

- Experimental Lorentzian amplitudes (areas) $A_i(Q)$'s extracted as described above in Fig. 3 are represented by data points in Fig. 4. We observe that the $A_i(Q)$ exhibit as a function of Q a behavior exactly as predicted by the model (see the end of the Section 2.3): $A_0(Q) \sim 1$ while $A_{i>0}(Q) \sim 0$ at low $Q \to 0$ as expected for spatially restricted motions, $A_0(Q)$ and $A_1(Q)$ decreases to very low value (closed to zero) in the large Q limit like for continuous motions while $A_2(Q)$ and $A_3(Q)$ increase with Q to a non-zero value reflecting the spatially discrete motions of the backbone.

  To go further in the analysis of these data, expressions of the amplitudes as a function of Q are needed; this is what we have done in deriving the expressions given in Tables 3 and 4. Lines through the data in Fig. 4 represent best fits (with $\chi^2 \sim 1.5$) of $A_i(Q)$'s using expressions in Tables 3 and 4 with the physical parameters of the local motions thus determined and reported in Table 6. This figure clearly demonstrates the relationships between local motions and the amplitudes of the phenomenological 3-dynamical process model.

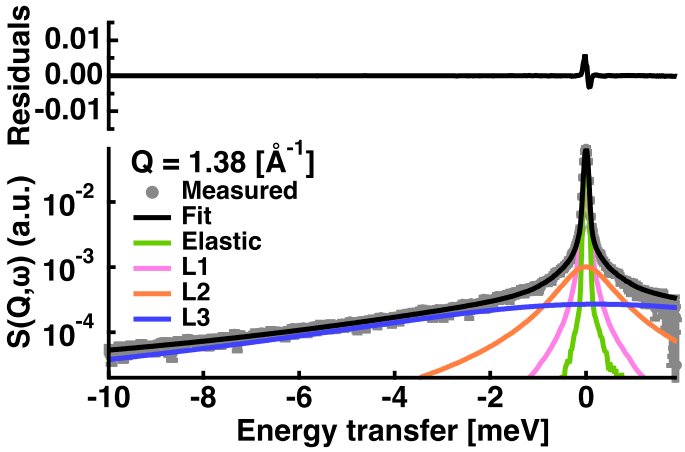


Figure 3: Example of $S(Q = 1.38\ \AA^{-1}, \omega)$ data fitting for DMPC multilamellar bilayers sample measured on IN6 at 135° geometry (in-plane motions) and T = 280 K. Solid black line through data points (grey circles) represents the best fit to the data using Eq.(11) with the resolution function given by the Vanadium measurements at Q = 1.38 $\AA^{-1}$. The fit results from the sum of the elastic peak (green line) and of the three Lorentzian functions for slow (magenta line), intermediate (orange line) and fast (blue line) motions.

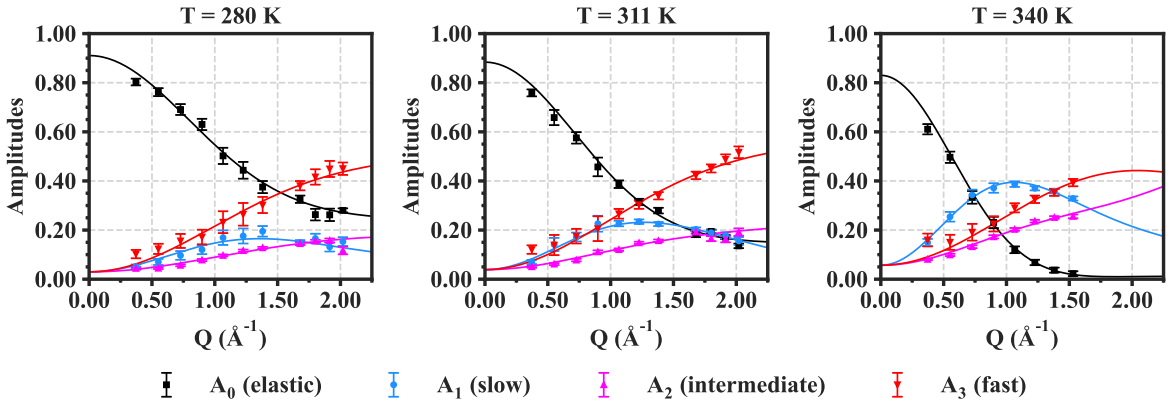


Figure 4: Data fitting for DMPC multilamellar bilayers measured at 135° (in-plane motions). Data points correspond to experimental amplitudes (areas) of the 3-dynamical process as a function of Q for different temperatures, obtained from best fits of QENS data with Eq.(11) as illustrated in Figure 3. Solid lines through the data represent best fits to the data using the expressions of amplitudes given in Tables 3 and 4. Parameters extracted from these fits are reported in Table 6.

Table 6: **Fixed and Fitted parameters for DMPC MLBs.** The sample was measured on IN6 instrument at 135°(in-plane motions or || - geometry).

Id	Parameter	Definition	T = 280K	T = 311K	T = 340K	Units	Reference
1	$\langle u^2 \rangle$	H's mean-squared displacement	0.83 ± 0.05	1.30 ± 0.06	5.06 ± 0.43	Å^2	Fitted
2	d	two-site distance	2.2 ± 0.1	2.2 ± 0.1	2.2 ± 0.1	Å	[23, 24][(a)]
	ϕ	probability of jump events	0.06 ± 0.02	0.07 ± 0.01	0.26 ± 0.02	-	Fitted
3	$1-z$	fraction of H's in the tail	0.75	0.75	0.75	-	[25]
	M	length of the effective tail	14	14	14	-	[25]
	R_1	radius of the first circle	0.43 ± 0.04	0.43 ± 0.02	0.01 ± 0.40	Å	Fitted
4	α	head tilt angle	32.3 ± 0.6	32.3 ± 0.6	32.3 ± 0.6	°	[26, 27]
	b_{H}	headgroup H-radius	1.04 ± 0.04	1.16 ± 0.04	0.66 ± 0.08	Å	Fitted
5	R_{H}	H-radius of the lipid	0.01 ± 0.42	0.46 ± 0.01	0.62 ± 0.01	Å	Fitted
6	k	force constant	No	No	No	N/m	Fitted
7	$R_{\parallel}$	lipid cage radius	1.13 ± 0.08	1.38 ± 0.04	2.06 ± 0.06	Å	Fitted
8	m	% of dynamic H's	67 ± 1	73 ± 2	82 ± 1	-	Fitted
	ε_0	error term	0.24 ± 0.01	0.15 ± 0.01	0.01 ± 0.01	-	Fitted

value $\pm 2\times$standard deviation. No = not observable in || - geometry where **Q** direction parallels the membrane plane (see Talble 4). "**Id**" stands for motions with: 1 = vibrational motions, 2 = jump dynamics of the backbone, 3 = tails: diffusion in circles, 4 = head: rotational diffusion+filp-flop, 5 = rotational diffusion of the lipid molecule, 6 = In-out of the plane diffusion, 7 = in-plane 2d-diffusion, and 8 = experimental parameters.

[(a)] Mean distance between the two sites for the jump diffusion of C-H groups. From Ref.[23]: go to *Home* → *Geometry* → *Experimental* → *Internal coordinates* or *Bond angles*, and look for the HCH to obtain the average angle $\widehat{HCH}$ (https://cccbdb.nist.gov/expangle2x.asp?descript=aHCH&all=0), the experimental length d_{CH} of C-H (https://cccbdb.nist.gov/expbondlengths2x.asp?descript=rCH&all=0) and calculate the distance d as, $d = 2d_{\mathrm{CH}} \sin\left(\widehat{HCH}/2\right)$. See also, Ref.[24], p.88.

- From now on, we will be able to take an interest in local motions. Figure 5 shows the EISF's or amplitudes (except the Debye-Waller factor) of all local motions contributing to the 3-dynamical process. These EISF's are generated by using parameters determined from Fig. 4 into expressions in Table 4. Changes with temperature of the variations in amplitudes as a function of Q reflect changes of parameters with temperature (see Table 6). Here, we did not nor have developed models to predict how these parameters might change with temperature. We can already notice that in general all parameters exhibit an increase with temperature (see Table 6). Indeed, the lipid system gains thermal energy as temperature increases leading to a subsequent enhancement of the dynamics, especially above the main phase transition around $297\,K$. Interestingly, the Matryoshka model already proves to be sensitive enough to allow capturing differences between the gel and liquid phase in such lipid systems.

  In terms of motion amplitudes, observable mouvements have amplitudes $A_i(Q)$ deviating from the horizontal line 1 as a function of Q. The deviation increases when $A_i(Q)$ decreases (i.e., the extent of motions increases) and vice versa, i.e., the area above the amplitude (F), $F_i = (1/Q_{\max}) \int_0^{Q_{\max}} \big[1 - A_i(Q)\big]\, dQ$, increases with the extension or amplitude of motions. More specifically,

  ▷ *Fraction of dynamic H's:* Values of m in Table 6 (given in %) are similar to that found in the literature [28], with m increasing with temperature.

  ▷ *Individual motions:* The Debye-Waller factor is a gaussian function of Q (not shown in Fig. 5) with the mean-square displacements (very similar to that found in literature [28]) increasing with temperature (see Table 6). As a result, $F_{\mathrm{DW}}(Q)$ increases with temperature.

  ▷ *Internal motions*

  → *Backbone motions:* Backbone motions relate to movements of C-H groups described by the jump diffusion between two non-equivalent sites distant of d. As d is fixed (see Table 6), the change in the amplitude of backbone motions with temperature reflects the change in the probability ϕ of jump events that increases with temperature (see Table 6); note that $\phi(T = 340\,K) \approx 0.26$ whereas $\phi_{\max} = 0.5$. Likewise, Fig. 5 shows that the amplitude of backbone motions increases as $F_{\mathrm{jump}}(Q)$ increases with temperature.

  → *Tail motions:* Figure 5 shows that the diffusional motions of tails are clearly observables with almost same amplitudes for temperatures $T = 280\,K$ and $T = 311\,K$ as reflected by values of R_1 in Table 6. And at $T = 340\,K$, the $A_i(Q)$ of tail motions exhibit an increase and become similar to that of backbone motions. This behavior which does not go in the direction of the increase of parameters with temperature finds its origin in the imprecise value of R_1 with large error bars (Table 6). It indicates possible variations of other parameters of the motion with temperature and / or relations between R_1 and others parameters. Further analysis would be needed.

→ *Head motions:* Figure 5 shows that the amplitude of the head motions slightly increases with temperature at low temperature then decreases at higher temperature as reflected by changes in the head size b_{H} (see Table 6). Such a change of b_{H} at higher temperatures is interesting to note, especially since it could be indicative of an interference in the analysis between internal rotational motions of the heads around their axes and the molecule rotational movements of lipid molecules (see *Rotational motions* and Figure 5); both involving the H-atoms in the head group. More analysis would be needed. For example, analyzing amplitudes in $\perp$ geometry, where $\mathbf{Q}$ is perpendicular to the membrane and only the movements of the heads are observable but not those of molecule rotation (see Table 4), could give informative indications on how to disentangle the motions.

▷ *Molecule motions*

→ *Rotational motions:* It appears that the amplitude of the rotational motions of lipid molecules about lipid axis are weak ($A_{\mathrm{rot}}(Q) \sim 1$) at low temperature and increases ($A_{\mathrm{rot}}(Q) < 1$) when the temperature increases thus resulting in an increase of the H-radius R_{H} of lipid molecules with temperature (see Table 6).

→ *In-out the plane motions:* The in-out of the plane motions of lipid molecules are not observable for the $\parallel$ geometry where the neutron scattering vector $\mathbf{Q}$ is parallel to the membrane (see Table 4). In this case, the amplitude remains the horizontal line, $F_{\mathrm{in-out}}(Q) = 0$, for all pairs of (Q, T) as shown in Fig. 5.

→ *In-plane 2d diffusion motions:* The in-plane diffusion of lipid molecules occurs within a cage of radius $R_{||}$ formed by the neighboring lipid molecules. Figure 5 shows that the deviation $F_{\mathrm{2d}}(Q)$ increases as the temperature gets higher, thus indicating that $R_{||}$ increases with the temperature (see Table 6).

Finally, let us underline that some parameters of the model in the Table 6 are not comparable with structure parameters which could have been obtained from diffraction experiments, for example. Indeed, unlike the structural parameters, the model parameters relate to the local movements of the H atoms around their equilibrium positions. By definition, these dynamical parameters are different and will generally be smaller than their structural counterparts. For example, the values of the H-radius, b_{H}, of the head group are slightly smaller than the experimental data ($\sim 2\mathring{A}$) of the head size [26, 29]. Likewise, the values of the smallest radius R_1 for the tail motions are smaller than the experimental values ($\sim 2.2\mathring{A}$) of chain spacings at the start of the tails [26, 29]. As for the H-radius, R_{H}, it can be compared to the span of the lipid molecule given by $D_{\mathrm{H1}} \sin(\alpha)$, where D_{H1} (similar but different from d_{H} in Fig. 2) is the size of the head group and α the tilt angle of the head. We find that R_{H} is smaller than the lipid span for α given in Table 6 and $D_{\mathrm{H1}} \sim 5\mathring{A}$ obtained from diffraction [26, 29].

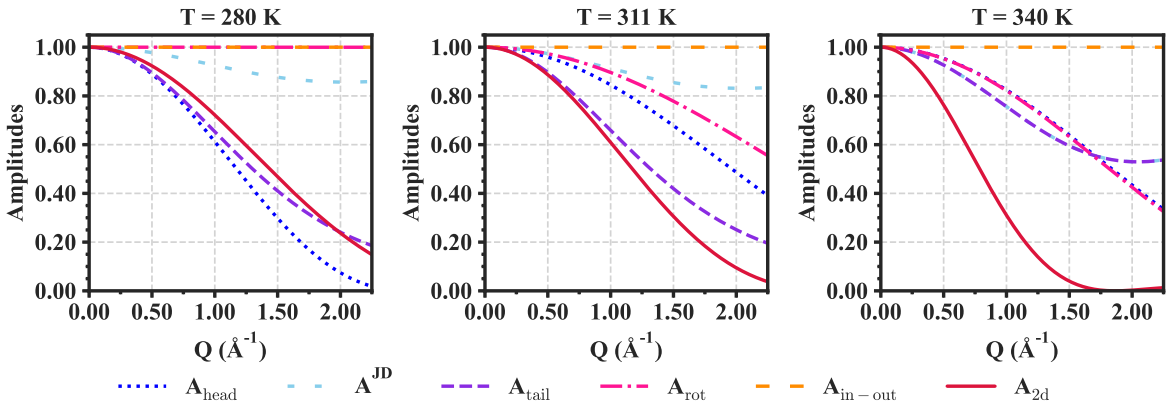


Figure 5: EISF or amplitudes of local motions (contributing to the 3-dynamical process) resolved in DMPC multilamellar bilayers at 135° (in-plane motions) as a function of Q for different temperatures. Lines correspond to expressions in Table 4 using parameters from Table 6.

4. Concluding summary

Our main motivation in developing this work has been to construct a framework for studying and describing local dynamics of lipid molecules in membrane layers. The main results of this work can be summarized as follows:

- We have developed the dynamical Matryoshka model which describes the local dynamics of lipid molecules in a membrane layer as a nested hierarchical convolution of three motional processes (see, Table 1):

  (i) individual motions described by vibrational movements of H-atoms;

  (ii) internal motions including the motions of the lipid backbone described by a jump dynamics between two non-equivalent sites, the head motions described by a rotational diffusion about the head axis plus a flip-flop or jump dynamics of the head tilted angle about the lipid axis, and the motions of the tail described by two-dimensional diffusions of H-atoms within circles of increasing radii along the tails;

  (iii) molecule motions of the lipid molecule as a whole including the rotational diffusion about the lipid axis, in-out of the plane motion and the in-plane local diffusion of the lipid molecule within a cage.

  The model includes seven motions in total (see, Table 2) with the analytical expressions of all associated ISFs provided in Table A.7.

- For the purpose of analyzing QENS experimental data, we have derived an analytical expression for the aggregated ISF of the Matryoshka model which involves an elastic term plus three inelastic terms of well-separated time scales. In doing so, we obtain the relationships between the amplitudes and rates of the aggregated ISF and those of the local movements of the initial model (see, Table 3).

- As a check of the model, we have shown that the theoretical aggregated ISF fit very well the QENS experimental data on a DMPC sample and allow extracting the dynamical parameters of the model.

Although we have shown that the development outlined above already describes very well QENS experiments, it can be complemented with analyzes including both the amplitudes, $A_i(Q)$, and rates, $\Gamma_i(Q)$, for various samples of lipid systems in different geometries. Such work is carried out in [21, 22]. The Matryoshka model provides a framework within which several improvements can be carried out if necessary. For example, for local motions, we have described the motions of the backbone by a jump dynamics between two sites. It is quite possible to replace it with other movements (e.g., jump dynamics between three sites) that the associated EISF (which contributes in the amplitudes of intermediate and fast motions) does not cancel at large Q, as experimental evidence indicates that the movements of the backbone should be spatially discrete. The Matryoshka model outlined above can be further extended to include in the same setting long range jumps for in-plane lipid diffusion and collective undulation modes of lipids. In terms of the time scales discussed here, these motions belong to the class of slow or very slow motions ($\Gamma \sim 1\ \mu$eV). Such a work is in progress. Finally, to be complete in the analysis of QENS data, it might turn out necessary to develop models predicting how the parameters of local motions would change with temperature, for example.

Acknowledgements

The authors thank the Institut Laue-Langevin for the allocation of the beam time to perform the experiments. We also thank Francesca Natali for her help on the experimental aspects of QENS. AC is supported by the JP Aguilar scholarship from Fondation CFM for her PhD thesis.

Data Availability

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

Dominique J. Bicout: Conceptualization, Methodology, Writing - Original Draft, Review & Editing, Supervision. Aline Cisse: Software, Validation, Formal Analysis, Writing - Review & Editing. Tatsuhito Matsuo: Software, Validation, Formal Analysis, Writing - Review & Editing. Judith Peters: Investigation, Writing - Review & Editing, Supervision, Funding acquisition.

Competing Interests:

The authors declare that they have no financial or non-financial competing interests.

Appendix A. Derivations of expressions of ISF

The key quantity in neutron scattering is the incoherent structure function (ISF), defined as,

$$I(Q,t) = \langle e^{i\mathbf{Q}.\mathbf{r(t)}}\, e^{-i\mathbf{Q}.\mathbf{r(0)}} \rangle \tag{A.1}$$

where $\mathbf{Q}$ is the neutron scattering wave vector, $\mathbf{r(t)}$ and $\mathbf{r(0)}$ the hydrogen atom positions at $t = 0$ and time t, respectively, and the sign $\langle \cdots \rangle$ represents an ensemble average over all positions $\mathbf{r(t)}$ and $\mathbf{r(0)}$ in the potential of mean force, $V(\mathbf{r})$. The mathematical expressions of all the ISFs considered in this work are gathered in Table A.7. For the ISFs whose expressions had already been derived elsewhere, we have indicated the associated references, and for the others we derive the expressions in this Appendix. This will be the case for head rotational movements with flipflop and the molecular in-out-of the plane diffusion of the lipid molecule in a harmonic potential.

Table A.7: **Incoherent Structure Functions (ISF) of dynamical processes.**

Subunit	ISF	Mathematical expression
H-atom	[a] $I_{\rm DW}$	*vibrational motions: Debye-Waller factor* $I_{\rm DW}(Q,t) = \exp\left\{-\dfrac{Q^2\langle u^2\rangle}{3}\right\}$
C-H groups	[b] $I_{\rm jump}$	*jump dynamics between 2 non-equivalent sites distant from d* $I_{\rm jump}(Q,t) = A_{\rm jump}(Q) + [1 - A_{\rm jump}(Q)]\,\exp\{-\Gamma_{\rm jump}\lvert t\rvert\}\ ,\ A_{\rm jump}(Q) = 1 - 2\phi(1-\phi)[1 - j_0(Qd)]$
Tail	[c] $I_{\rm tail}$	*2d-diffusion of tail H-atoms inside $F(m)$ - distribution of circles of radius R_m* $I_{\rm tail}(Q,t) = \sum_{m=1}^{M} F(m)\left\{A_0^0(y_m) + \sum_{n=1}^{+\infty} A_n^0(y_m)\,\exp\left\{-\left(x_n^0\right)^2\left(D_{\rm tail}/R_m^2\right)\lvert t\rvert\right\}\right.$ $\left. + 2\sum_{l=1}^{+\infty}\sum_{n=0}^{+\infty} A_n^l(y_m)\,\exp\left\{-\left(x_n^l\right)^2\left(D_{\rm tail}/R_m^2\right)\lvert t\rvert\right\}\right\}$ $A_0^0(y_m) = \left[\dfrac{2J_1(y_m)}{y_m}\right]^2\ ;\ A_n^l(y_m) = \dfrac{4\left(x_n^l\right)^2}{\left(x_n^l\right)^2 - l^2}.\left[\dfrac{y_m J_{l+1}(y_m) - lJ_l(y_m)}{y_m^2 - \left(x_n^l\right)^2}\right]^2\ ;\ y_m = QR_m\sin\gamma$
Head	[d] $I_{\rm head}$	*rotational diffusion of the headgroup about its axis plus flip-flop motion* $I_{\rm head}(Q,t) = \sum_{n=-\infty}^{+\infty} A_n(Q)\,\exp\left\{-\left[n^2 D_{\rm head} + \dfrac{1-(-1)^n}{2\tau_{\rm ff}}\right]\lvert t\rvert\right\}$ $A_n(Q) = \left[\sum_{l=-\infty}^{+\infty} J_{n-2l}(Qb_{\rm H}\cos\gamma\sin\alpha) J_l\left(Qb_{\rm H}\sin\gamma\cos^2\dfrac{\alpha}{2}\right) J_l\left(Qb_{\rm H}\sin\gamma\sin^2\dfrac{\alpha}{2}\right)\right]^2$
Molecule	[e] $I_{\rm rot}$	*rotational diffusion of the lipid molecule about the membrane normal axis* $I_{\rm rot}(QR_{\rm H}\sin\gamma,t) = J_0^2(QR_{\rm H}\sin\gamma) + 2\sum_{n=1}^{+\infty} J_n^2(QR_{\rm H}\sin\gamma)\,\exp\{-D_{\rm rot}n^2\lvert t\rvert\}$
	[f] $I_{\rm in-out}$	*in-out of the plane 1d-diffusion of the lipid molecule in a harmonic potential* $I_{\rm in-out}(Q\cos\gamma,t) = exp\left\{-(Q\cos\gamma)^2\left(\dfrac{k_{\rm B}T}{k}\right)\left[1 - \mathrm{e}^{-\lvert t\rvert/\tau}\right]\right\}$
	[g] $I_{\rm 2d}$	*in-plane 2d-diffusion of the lipid molecule inside a circle of radius $R_{\lvert\rvert}$* $I_{\rm 2d}(QR_{\lvert\rvert}\sin\gamma,t) = A_0^0(y) + \sum_{n=1}^{+\infty} A_n^0(y)\,\exp\left\{-\left(x_n^0\right)^2\left(D_{\lvert\rvert}/R_{\lvert\rvert}^2\right)\lvert t\rvert\right\} + 2\sum_{l=1}^{+\infty}\sum_{n=0}^{+\infty} A_n^l(y)\,\exp\left\{-\left(x_n^l\right)^2\left(D_{\lvert\rvert}/R_{\lvert\rvert}^2\right)\lvert t\rvert\right\}$ $A_0^0(y) = \left[\dfrac{2J_1(y)}{y}\right]^2\ ;\ A_n^l(y) = \dfrac{4\left(x_n^l\right)^2}{\left(x_n^l\right)^2 - l^2}.\left[\dfrac{yJ_{l+1}(y) - lJ_l(y)}{y^2 - \left(x_n^l\right)^2}\right]^2\ ;\ y = QR_{\lvert\rvert}\sin\gamma$

[a] corresponds to the long time limit as the relaxation of vibrational motions is very fast and out of the time scales considered here. $\langle u^2 \rangle$ is the mean square amplitude of these very fast motions.

[b] $j_0(\cdots)$ is the spherical Bessel function of firts kind and order 0, $\phi = \tau_1/(\tau_1+\tau_2) = 1/\left[1+\exp(-\Delta G/k_{\mathrm{B}}T)\right]$ and $\Gamma_{jump} = \tau_1^{-1} + \tau_2^{-1}$ where τ_1 and τ_2 are the mean residence times in each site and ΔG the standard energy variation of the transition between the two sites [14].

[c] See [g] for the derivation. D_{tail} is the diffusion constant. The simplest distribution law is, $F(m) = 1/M$, where M is the apparent chain length of the lipid tail, and the circle-radius $R_m = R_1\sqrt{m}$. Carpentier et al. [15] rather used a two-parameters model for the circle-radius, $R_m = \left(\dfrac{m-2}{M-2}\right)(R_M - R_2) + R_2$.

[d] b_{H} is the H-radius of the head group (i.e., the distance between the head axis and the center of inertia of all H-atoms of the head group with respect to the head axis $\vec{l}$), α the tilt angle between the membrane normal axis and head axis (see Fig.2), D_{head} the head rotational diffusion constant and τ_{flip} the mean resting time between two successive flip-flops. $J_n(\cdots)$ is the cylindrical Bessel function of first kind and order n. See the derivation in Sec. Appendix A.1.

[e] See Eq.(13) in Ref.[30]. The size of the cross section of the lipid molecule is given by, R_{H} (i.e., the distance between the lipid axis and the center of inertia of all dynamic H-atoms with respect to the lipid axis), γ the angle between $\mathbf{Q}$ and the membrane normal axis and D_{rot} the rotational diffusion constant.

[f] k is the force constant, τ the relaxation time, $k_{\mathrm{B}}T$ the thermal energy with k_{B} the Boltzmann constant and T temperature. See the derivation in Appendix A.2.

[g] See the derivation in Ref.[31]. $J_l(\cdots)$ is the cylindrical Bessel function of first kind and order l and x_n^l is the $(n+1)$th root of, $\dfrac{d}{dx}[J_l(x)] = 0$; these roots are tabulated in [32], e.g., $x_n^0 = 3.83170, 7.01558, 10.17346$ for $n = 1, 2, 3$ and $x_n^1 = 1.84118, 5.33144, 8.53632$ for $n = 0, 1, 2$, where x_0^1 is the smallest non-zero root. $D_{||}$ is the diffusion constant.

Appendix A.1. Rotational diffusions and flip-flop of headgroup

We will consider the head group as an ellipsoid of circular cross-section of diameter $2b_{\mathrm{H}}$ and directing vector $\vec{l}$ (see Fig.2), where b_{H} is the distance between the head axis and the center of inertia of all H-atoms of the head group with respect to the head axis $\vec{l}$ and $\vec{l}$ is given in the orthonormal frame $\widehat{xyz}$ of basis $(\vec{i},\vec{j},\vec{k})$ by,

$$\vec{l} = \begin{pmatrix} \sin\alpha\cos\eta \\ \sin\alpha\sin\eta \\ \cos\alpha \end{pmatrix} . \tag{A.2}$$

The angles are defined in Fig.A.6.

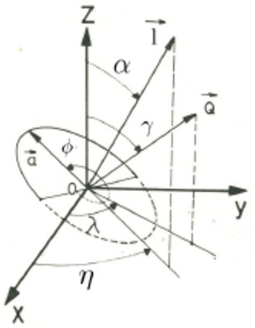


Figure A.6: Definition of angles for the rotations about $\vec{n}$ (parallell to z) and $\vec{l}$ directions and the flip-flop motion of $\vec{l}$ about $\vec{n}$. Adaptation from [30].

That thus makes it possible to define a local orthonormal coordinate system $(\vec{e}_x,\vec{e}_y,\vec{e}_z)$ associated with the head group as follows,

$$\begin{cases} \vec{e}_x = (\cos\alpha\cos\eta)\,\vec{i} + (\cos\alpha\sin\eta)\,\vec{j} - (\sin\alpha)\,\vec{k} \\ \vec{e}_y = -(\sin\eta)\,\vec{i} + (\cos\eta)\,\vec{j} \\ \vec{e}_z = \vec{l} = (\sin\alpha\cos\eta)\,\vec{i} + (\sin\alpha\sin\eta)\,\vec{j} + (\cos\alpha)\,\vec{k} \end{cases} \tag{A.3}$$

from which the matrix of passage from $(\vec{e}_x,\vec{e}_y,\vec{e}_z)$ to $(\vec{i},\vec{j},\vec{k})$ can be determined as,

$$P = [\vec{e}_x,\vec{e}_y,\vec{e}_z] = \begin{pmatrix} \cos\alpha\cos\eta & -\sin\eta & \sin\alpha\cos\eta \\ \cos\alpha\sin\eta & \cos\eta & \sin\alpha\sin\eta \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix} \tag{A.4}$$

Note that $P = R_z(\eta)R_y(\alpha)$ results from the operation of two rotations where $R_y(\alpha)$ describes the rotation of an angle α about $y-$axis and $R_z(\eta)$ the rotation of an angle η about $z-$axis. The head group position vector $\vec{b}$ in the local head group frame $(\vec{e}_x,\vec{e}_y,\vec{e}_z)$ is given by,

$$\vec{b} = b_{\mathrm{H}} \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} = b_{\mathrm{H}}\left[(\cos\phi)\,\vec{e}_x + (\sin\phi)\,\vec{e}_y\right] \tag{A.5}$$

where ϕ is the angle between $\vec{b}$ and $\vec{e}_x$. Thus, the head group position vector $\mathbf{r}$ in the frame $(\vec{i},\vec{j},\vec{k})$ is obtained using the passage matrix as,

$$\begin{aligned}\mathbf{r} &= P\vec{b} = b_{\mathrm{H}} \begin{pmatrix} \cos\alpha\cos\eta & -\sin\eta & \sin\alpha\cos\eta \\ \cos\alpha\sin\eta & \cos\eta & \sin\alpha\sin\eta \\ -\sin\alpha & 0 & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} \\ &= b_{\mathrm{H}} \begin{pmatrix} \cos\alpha\cos\eta\cos\phi - \sin\eta\sin\phi \\ \cos\alpha\sin\eta\cos\phi + \cos\eta\sin\phi \\ -\sin\alpha\cos\phi \end{pmatrix} \end{aligned} \tag{A.6}$$

Note that, $\mathbf{r}\cdot\vec{l} = 0$ as expected since $\mathbf{r} \perp \vec{l}$.

Let the $z-$axis be the normal to the membrane, the scattering wave vector $\mathbf{Q}$ is given by,

$$\mathbf{Q} = Q \begin{pmatrix} \sin\gamma\cos\lambda \\ \sin\gamma\sin\lambda \\ \cos\gamma \end{pmatrix} \tag{A.7}$$

where γ is the angle between $\vec{k}$ and $\mathbf{Q}$ and λ the angle between the projection of $\mathbf{Q}$ onto the $x-y$ plan and $\vec{i}$. Now, the scalar product entering in the calculation of the ISF writes as,

$$\begin{aligned}\mathbf{Q}\cdot\mathbf{r} &= Qb_{\mathrm{H}} \begin{pmatrix} \sin\gamma\cos\lambda \\ \sin\gamma\sin\lambda \\ \cos\gamma \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha\cos\eta\cos\phi - \sin\eta\sin\phi \\ \cos\alpha\sin\eta\cos\phi + \cos\eta\sin\phi \\ -\sin\alpha\cos\phi \end{pmatrix} \\ &= Qb_{\mathrm{H}} \left[\sin\gamma\cos\phi\cos\alpha\left(\cos\lambda\cos\eta + \sin\lambda\sin\eta\right) + \sin\gamma\sin\phi\left(\sin\lambda\cos\eta - \cos\lambda\sin\eta\right)\right. \\ &\quad \left. - \cos\gamma\sin\alpha\cos\phi\right] \\ &= Qb_{\mathrm{H}} \left[\sin\gamma\cos\phi\cos\alpha\cos(\lambda-\eta) + \sin\gamma\sin\phi\sin(\lambda-\eta) - \cos\gamma\sin\alpha\cos\phi\right] \\ &= Qb_{\mathrm{H}} \left\{\sin\gamma\left[\cos^2\left(\frac{\alpha}{2}\right)\cos(\phi-\lambda+\eta) - \sin^2\left(\frac{\alpha}{2}\right)\cos(\phi+\lambda-\eta)\right] - \cos\gamma\sin\alpha\cos\phi\right\} \\ &= Qb_{\mathrm{H}} f(\alpha,\phi,\lambda,\gamma,\eta) \end{aligned} \tag{A.8}$$

The head group performs two independent motions:

- Uniaxial rotational diffusion of angle ϕ about the $\vec{l}-$axis and diffusion constant D_{head} described by the Green's function:

$$G_{\mathrm{rot}}(\phi,t|\phi_0) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} \exp\left\{in(\phi-\phi_0)\right\}\cdot\exp\left\{-D_{\mathrm{head}}n^2|t|\right\} \tag{A.9}$$

  with the equilibrium distribution, $P_{\mathrm{eq,rot}}(\phi) = \lim_{t\to\infty} G_{\mathrm{rot}}(\phi,t|\phi_0) = 1/(2\pi)$,

- Flip-flop motion of the $\vec{l}-$axis between angles $\theta = -\alpha$ and $\theta = \alpha$ and residence time τ_{ff} described by the jump diffusion with the Green's function:

$$G_{\mathrm{ff}}(\theta,t|\theta_0) = \frac{\delta_{\theta,\alpha}}{2}\left[u\,\delta_{\theta_0,\alpha} + v\,\delta_{\theta_0,-\alpha}\right] + \frac{\delta_{\theta,-\alpha}}{2}\left[u\,\delta_{\theta_0,-\alpha} + v\,\delta_{\theta_0,\alpha}\right] \tag{A.10}$$

where,

$$\left.\begin{matrix} u \\ v \end{matrix}\right\} = 1 \pm \exp\left\{-\frac{|t|}{\tau_{\rm ff}}\right\} \tag{A.11}$$

with the equilibrium distribution, $P_{\rm eq,ff}(\theta) = \lim_{t\to\infty} G_{\rm ff}(\theta, t|\theta_0) = 1/2$.

For experimental configurations where the direction of the scattering vector $\mathbf{Q}$ with respect to the membrane normal is set constant (i.e., the angle γ is set constant for all λ), the ISF is calculated as,

$$\begin{aligned} I(Q,t) &= \sum_{\theta_0=-\alpha}^{\alpha} P_{\rm eq,ff}(\theta_0) \sum_{\theta=-\alpha}^{\alpha} G_{\rm ff}(\theta,t|\theta_0) \int_0^{2\pi}\frac{d\eta_0}{2\pi}\int_0^{2\pi}\frac{d\eta}{2\pi}\int_0^{2\pi} d\phi_0\, P_{\rm eq,rot}(\phi_0) \\ &\quad \int_0^{2\pi} d\phi\, G_{\rm rot}(\phi,t|\phi_0)\exp\left\{iQb_{\rm H}\left[f(\theta,\phi,\lambda,\gamma,\eta) - f(\theta_0,\phi_0,\lambda,\gamma,\eta_0)\right]\right\} \qquad \text{(A.12)} \\ &= \sum_{\theta_0=-\alpha}^{\alpha} P_{\rm eq,ff}(\theta_0) \sum_{\theta=-\alpha}^{\alpha} G_{\rm ff}(\theta,t|\theta_0) \sum_{n=-\infty}^{\infty} B_n(\theta,\gamma)B_n^*(\theta_0,\gamma)\,\exp\left\{-D_{\rm head}n^2|t|\right\}\ , \end{aligned}$$

where,

$$B_n(\theta,\gamma) = \int_0^{2\pi}\frac{d\eta}{2\pi}\int_0^{2\pi}\frac{d\phi}{2\pi}\exp\left\{i\left[Qb\, f(\theta,\phi,\lambda,\gamma,\eta) + n\phi\right]\right\} = \frac{1}{2\pi}\int_0^{2\pi} d\eta\, C_n(\theta,\lambda,\gamma,\eta) \tag{A.13}$$

Using $G_{\rm ff}(\theta,t|\theta_0)$ in Eq.(A.10) back into $I(Q,t)$, we have:

$$\begin{aligned} I(Q,t) &= \sum_{n=-\infty}^{\infty}\left\{\frac{u}{4}\left[B_n(\alpha,\gamma)B_n^*(\alpha,\gamma) + B_n(-\alpha,\gamma)B_n^*(-\alpha,\gamma)\right]\right. \\ &\quad \left.\frac{v}{4}\left[B_n(\alpha,\gamma)B_n^*(-\alpha,\gamma) + B_n(-\alpha,\gamma)B_n^*(\alpha,\gamma)\right]\right\}\exp\left\{-D_{\rm head}n^2|t|\right\}\ . \qquad \text{(A.14)} \end{aligned}$$

We have,

$$\begin{aligned} C_n(\alpha,\lambda,\gamma,\eta) &= \frac{1}{2\pi}\int_0^{2\pi} d\phi\exp\left\{i\left[Qb_{\rm H}\, f(\alpha,\phi,\lambda,\gamma,\eta) + n\phi\right]\right\} \qquad \text{(A.15)} \\ &= \frac{1}{2\pi}\int_0^{2\pi} d\phi\exp\left\{in\phi\right\}\exp\left\{iz_1\cos\phi\right\}\exp\left\{iz_2\cos(\phi-\mu)\right\}\exp\left\{iz_2\cos(\phi+\mu)\right\}\ , \end{aligned}$$

where,

$$\begin{cases} \mu = \lambda - \eta \\ z_1 = -Qb_{\rm H}\cos\gamma\sin\alpha \\ z_2 = Qb_{\rm H}\sin\gamma\cos^2\left(\frac{\alpha}{2}\right) \\ z_3 = Qb_{\rm H}\sin\gamma\sin^2\left(\frac{\alpha}{2}\right) \end{cases} \tag{A.16}$$

Now, using the following relations,

$$\begin{cases} \exp\{\pm iz\cos\psi\} = \displaystyle\sum_{m=-\infty}^{\infty} i^{\pm m}\exp\{\mp im\psi\}\,J_m(z)\ ;\ \text{Jacobi-Anger expansion} \\ \\ \displaystyle\int_0^{2\pi} d\phi\exp\{\pm ik\phi\} = 2\pi\,\delta_{k,0} \end{cases} \tag{A.17}$$

back into Eq;(A.15), we obtain:

$$C_n(\alpha,\lambda,\gamma,\eta) = i^n\sum_{m=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} J_m(z_1)J_l(z_2)J_{n-m-l}(z_3)\,\exp\{i(2l+m-n)\mu\}\ , \tag{A.18}$$

where $J_m(\cdots)$ is the Bessel function of first kind and order m. It follows that,

$$\begin{aligned} B_n(\alpha,\gamma) &= \int_0^{2\pi}\frac{d\eta}{2\pi}\,C_n(\alpha,\lambda,\gamma,\eta) = \int_0^{2\pi}\frac{d\mu}{2\pi}\,C_n(\alpha,\lambda,\gamma,\eta) \\ &= \frac{i^n}{2\pi}\sum_{m=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} J_m(z_1)J_l(z_2)J_{n-m-l}(z_3)\int_0^{2\pi} d\mu\,\exp\{i(2l+m-n)\mu\} \\ &= i^n\sum_{l=-\infty}^{\infty} J_{n-2l}(z_1)J_l(z_2)J_l(z_3)\,. \end{aligned} \tag{A.19}$$

Using the relation for the Bessel functions, $J_n(-x) = (-1)^n J_n(x)$, we obtain the relation: $B_n(\alpha,\gamma) = (-1)^n B_n(-\alpha,\gamma)$. Now, using this relation relation in Eq.(A.14), we finally obtain,

$$I(Q,t) = \sum_{n=-\infty}^{\infty}\left[\frac{u+(-1)^n v}{2}\right]|B_n(-\alpha,\gamma)|^2\exp\left\{-D_{\text{head}}n^2|t|\right\} \tag{A.20}$$

$$= \sum_{n=-\infty}^{\infty}\left[\sum_{l=-\infty}^{+\infty} J_{n-2l}(-z_1)J_l(z_2t)J_l(z_3)\right]^2\exp\left\{-\left[n^2D_{\text{head}}+\frac{1-(-1)^n}{2\tau_{\text{ff}}}\right]|t|\right\} \tag{A.21}$$

Appendix A.2. 1d Diffusion in a Harmonic Potential

Let the $z-$axis coincides with the normal to the membrane, we consider that each lipid molecule, of coordinate z, is undergoing as a whole a 1d diffusion parallel to $z-$axis in a harmonic potential of mean force, $V(z) = kz^2/2$, of force constant k and relaxation time τ. The Green's function, $G(z,t|z_0)$, describing such motions for the lipid molecule is given by,

$$G(z,t|z_0) = \frac{\exp\left\{-\dfrac{\left(z-z_0\mathrm{e}^{-t/\tau}\right)^2}{2\sigma_0^2\left[1-\mathrm{e}^{-2t/\tau}\right]}\right\}}{\sqrt{2\pi\sigma_0^2\left[1-\mathrm{e}^{-2t/\tau}\right]}} \tag{A.22}$$

with the equilibrium distribution of lipid molecule positions given by,

$$P_{\rm eq}(z) \;=\; \lim_{t\to\infty} G(z,t|z_0) = \frac{\exp\left\{-\dfrac{z^2}{2\sigma_0^2}\right\}}{\sqrt{2\pi\sigma_0^2}}\;,\; \sigma_0^2 = \frac{k_{\rm B}T}{k} \tag{A.23}$$

Denoting by γ the angle between the $z-$axis and the scattering $\mathbf{Q}$, the ISF can be written as,

$$\begin{aligned} I(Q,t) &= \langle e^{i\mathbf{Q}.\mathbf{r(t)}}\, e^{-i\mathbf{Q}.\mathbf{r(0)}}\rangle = \langle e^{iQ\cos\gamma z(t)}\, e^{-iQ\cos\gamma z(0)}\rangle \\ &= \int_{-\infty}^{+\infty} dz_0 \int_{-\infty}^{+\infty} dz\, \mathrm{e}^{iQ\cos\gamma z} G(z,t|z_0)\mathrm{e}^{-iQ\cos\gamma z_0} P_{\rm eq}(z_0) \\ &= \int_{-\infty}^{+\infty} dz_0 B(z_0) P_{\rm eq}(z_0) \end{aligned} \tag{A.24}$$

where,

$$\begin{cases} B(z_0) = \int_{-\infty}^{+\infty} dz\, \mathrm{e}^{iQ\cos\gamma z} G(z,t|z_0) = \mathrm{e}^{iyz_0\mathrm{e}^{-t/\tau} - (ya)^2/4} \\ \\ y = Q\cos\gamma\;;\; a^2 = 2\sigma_0^2\left[1-\mathrm{e}^{-2t/\tau}\right] \end{cases} \tag{A.25}$$

Then,

$$\begin{aligned} I(Q,t) &= \mathrm{e}^{-(ya)^2/4} \int_{-\infty}^{+\infty} dz_0 \mathrm{e}^{-iyz_0\left[1-\mathrm{e}^{-t/\tau}\right]} P_{\rm eq}(z_0) \\ &= \exp\left\{-\frac{(ya)^2}{4} - \frac{\left[y\sigma_0\left(1-\mathrm{e}^{-t/\tau}\right)\right]^2}{2}\right\} \end{aligned} \tag{A.26}$$

Finally, we obtain:

$$I(Q,t) = \exp\left\{-(Q\cos\gamma)^2\left(\frac{k_{\rm B}T}{k}\right)\left[1-\mathrm{e}^{-t/\tau}\right]\right\} \tag{A.27}$$

Appendix B. Derivation of aggregated expressions of B's amplitudes and Γ's rates

Starting from a time dependent ISF, the aim of the aggregation is in general to derive an approximate expression of ISF as an expansion of only few relaxation functions, $E_k(Q,t)$ such that, $E_k(Q,t=0)=1$ and $E_k(Q,t\to\infty)=0$. For our purpose, we deal with single exponential relaxation function, $E_k(Q,t) = \exp\{-\Gamma_k t\}$, where $\Gamma_k(Q)$ is the relaxation rate. The ISF to derive an aggregation is given in the right-hand-side of Eq.(10) as,

$$I = \left[zI_{\rm jump}I_{\rm head} + (1-z)I_{\rm jump}I_{\rm tail}\right]\, I_{\rm rot}I_{\rm in-out}I_{\rm 2d} \tag{B.1}$$

where $I_\alpha(Q,t)$ are the ISF's of local motions described in the Matryoshka model and given as [16],

$$I_\alpha(Q,t) = A_\alpha(Q) + \Big[1 - A_\alpha(Q)\Big]\, C_\alpha(Q,t)\,;\ \alpha = \text{jump, head, tail, rot, in}-\text{out, 2d} \quad \text{(B.2)}$$

where $A_\alpha(Q)$ is the amplitude and $C_\alpha(Q,t)$ the relaxation function such that, $C_\alpha(Q,t=0)=1$ and $C_\alpha(Q,t\to\infty)=0$. As, in general, $C_\alpha(Q,t)$ is a multi-exponential function of time with decay rates functions of Q, the first step toward the aggregation is to derive a single exponential approximation of $C_\alpha(Q,t)$ as [16],

$$C_\alpha(Q,t) \approx E_\alpha(Q,t) = \exp\{-\Gamma_\alpha(Q)\,t\} \ \text{ with } \ \frac{1}{\Gamma_\alpha(Q)} = \int_0^\infty dt\, C_\alpha(Q,t) \quad \text{(B.3)}$$

where $\Gamma_\alpha(Q)$ is the $Q-$dependent relaxation rate. Next, the goal is to derive the amplitudes and relaxation rates such that Eq.(B.1) can be rewritten as,

$$\begin{cases} I_{\text{agg}}(Q,t) = \displaystyle\sum_{i=0}^{3} B_i(Q)\, E_i(Q,t) \\ \\ \text{with: } i = (0=\text{EISF}), (1=\text{slow}), (2=\text{intermediate}), (3=\text{fast})\ , \end{cases} \quad \text{(B.4)}$$

where $B_i(Q)$ is the amplitude and $E_i(Q,t)$ the relaxation function with relaxation rates such that, $\Gamma_0 = 0 < \Gamma_{\text{slow}} < \Gamma_{\text{intermediate}} < \Gamma_{\text{fast}}$. To proceed, expanding Eq.(B.1) yields,

$$\begin{cases} I = \ \Big[z A_{\text{jump}} A_{\text{head}} + z A_{\text{jump}}(1-A_{\text{head}}) E_{\text{head}} + z(1-A_{\text{jump}}) A_{\text{head}} E_{\text{jump}} \\ \quad + z(1-A_{\text{jump}})(1-A_{\text{head}}) E_{\text{jump}} E_{\text{head}} + (1-z) A_{\text{jump}} A_{\text{tail}} \\ \quad + (1-z) A_{\text{jump}}(1-A_{\text{tail}}) E_{\text{tail}} + (1-z)(1-A_{\text{jump}}) A_{\text{tail}} E_{\text{jump}} \\ \quad + (1-z)(1-A_{\text{jump}})(1-A_{\text{tail}}) E_{\text{jump}} E_{\text{tail}} \Big] \times \Big[A_{\text{rot}} A_{\text{in}-\text{out}} A_{\text{2d}} \\ \quad + (1-A_{\text{rot}}) A_{\text{in}-\text{out}} A_{\text{2d}} E_{\text{rot}} + A_{\text{rot}}(1-A_{\text{in}-\text{out}}) A_{\text{2d}} E_{\text{in}-\text{out}} \\ \quad + A_{\text{rot}} A_{\text{in}-\text{out}}(1-A_{\text{2d}}) E_{\text{2d}} + (1-A_{\text{rot}})(1-A_{\text{in}-\text{out}}) A_{\text{2d}} E_{\text{rot}} E_{\text{in}-\text{out}} \\ \quad + (1-A_{\text{rot}}) A_{\text{in}-\text{out}}(1-A_{\text{2d}}) E_{\text{rot}} E_{\text{2d}} + A_{\text{rot}}(1-A_{\text{in}-\text{out}})(1-A_{\text{2d}}) E_{\text{in}-\text{out}} E_{\text{2d}} \\ \quad + (1-A_{\text{rot}})(1-A_{\text{in}-\text{out}})(1-A_{\text{2d}}) E_{\text{rot}} E_{\text{in}-\text{out}} E_{\text{2d}} \Big] \end{cases} \quad \text{(B.5)}$$

Next, we use the following rules,

- Hierarchy of relaxations: the hierarchy of relaxation time scales of motional processes in the Matryoshka model indicates that: $\Gamma_{\text{jump}} \sim \Gamma_{\text{tail}} > \Gamma_{\text{head}} \sim \Gamma_{\text{rot}} > \Gamma_{\text{in}-\text{out}} \sim \Gamma_{\text{2d}}$, implying that, $E_{\text{jump}} \sim E_{\text{tail}} < E_{\text{head}} \sim E_{\text{rot}} < E_{\text{in}-\text{out}} \sim E_{\text{2d}}$ for all t;
- Operation rules: for timescales $\Gamma_i t \sim 1$, we have:
  - ▷ Summation: $aE_i + bE_j \approx \begin{cases} aE_i + b & , \ \Gamma_i > \Gamma_j \\ aE_i & , \ \Gamma_i < \Gamma_j \end{cases}$

$$\triangleright \text{ Product: } E_i \times E_j \approx \begin{cases} E_i & , \quad \Gamma_i > \Gamma_j \\ 0 & , \quad \Gamma_i < \Gamma_j \end{cases}$$

to rearrange Eq.(B.5) into 4 terms and derive the aggregated approximation in Eq.(B.4) where,

- **EISF:** collects all time independent terms in Eq.(B.5).

  Taking the $t \to \infty$ limit (i.e., setting all $E_\alpha(Q, t \to \infty) = 0$) in Eq.(B.5) gives,

$$B_0 = \Big[z A_{\text{head}} + (1-z) A_{\text{tail}} \Big] A_{\text{jump}} A_{\text{rot}} A_{\text{in-out}} A_{\text{2d}} \tag{B.6}$$

- **Slow motions:** collect all terms involving $E_{\text{in-out}}$ and E_{2d} in Eq.(B.5).

$$\begin{aligned} B_{\text{slow}}\, E_{\text{slow}} \;=\; & \Big\{ A_{\text{jump}} A_{\text{rot}} \Big[z A_{\text{head}} + (1-z) A_{\text{tail}} \Big] + T_1 \Big\} \Big\{ T_2 + A_{\text{in-out}} (1 - A_{\text{2d}}) E_{\text{2d}} \\ & + (1 - A_{\text{in-out}}) A_{\text{2d}} E_{\text{in-out}} + (1 - A_{\text{in-out}})(1 - A_{\text{2d}}) E_{\text{in-out}} E_{\text{2d}} \Big\} \end{aligned} \tag{B.7}$$

  in which T_1 groups terms involving E_{jump}, E_{tail} and E_{head} and T_2 terms involving the products of E_{rot} with $E_{\text{in-out}}$ and E_{2d} (i.e., terms $\approx E_{\text{rot}}$). For timescales $\Gamma_{\text{2d}} t \sim \Gamma_{\text{in-out}} t \sim 1$, both T_1 and T_2 vanish as they relax faster to zero and, therefore, will be omitted. Now, taking the $t \to 0$ limit in Eq.(B.7) (without T_1 and T_2) gives,

$$B_{\text{slow}} = \Big[z A_{\text{head}} + (1-z) A_{\text{tail}} \Big] (1 - A_{\text{in-out}} A_{\text{2d}}) A_{\text{jump}} A_{\text{rot}} \,. \tag{B.8}$$

  And, taking the $t \to 0$ limit in the time derivative of Eq.(B.7) gives,

$$\Gamma_{\text{slow}} = \frac{(1 - A_{\text{in-out}}) \Gamma_{\text{in-out}} + (1 - A_{\text{2d}}) \Gamma_{\text{2d}}}{1 - A_{\text{in-out}} A_{\text{2d}}} \tag{B.9}$$

- **Intermediate motions:** collect all terms involving E_{head} and E_{rot} in Eq.(B.5).

  For timescales $\Gamma_{\text{head}} t \sim \Gamma_{\text{rot}} t \sim 1$, we use the rules above, $E_{\text{in-out}} \sim E_{\text{2d}} \approx 1$ and $E_{\text{jump}} \sim E_{\text{tail}} \approx 0$ to first reduces Eq.(B.5) to,

$$\begin{aligned} B_{\text{intermediate}}\, E_{\text{intermediate}} = \; & \Big[z A_{\text{jump}} A_{\text{head}} + z A_{\text{jump}} (1 - A_{\text{head}}) E_{\text{head}} \\ & + z (1 - A_{\text{jump}})(1 - A_{\text{head}}) E_{\text{jump}} E_{\text{head}} + (1-z) A_{\text{jump}} A_{\text{tail}} \Big] \times \\ & \Big[A_{\text{rot}} + (1 - A_{\text{rot}}) A_{\text{in-out}} A_{\text{2d}} E_{\text{rot}} + \\ & + (1 - A_{\text{rot}})(1 - A_{\text{in-out}}) A_{\text{2d}} E_{\text{rot}} E_{\text{in-out}} \\ & + (1 - A_{\text{rot}}) A_{\text{in-out}} (1 - A_{\text{2d}}) E_{\text{rot}} E_{\text{2d}} \\ & + (1 - A_{\text{rot}})(1 - A_{\text{in-out}})(1 - A_{\text{2d}}) E_{\text{rot}} E_{\text{in-out}} E_{\text{2d}} \Big] \\ & - \text{time independent terms} \end{aligned} \tag{B.10}$$

and next, $E_{\rm rot}E_{\rm in-out} \approx E_{\rm rot}$, $E_{\rm rot}E_{\rm 2d} \approx E_{\rm rot}$, $E_{\rm rot}E_{\rm in-out}E_{\rm 2d} \approx E_{\rm rot}$ and $E_{\rm jump}E_{\rm head} \approx 0$ to,

$$\begin{aligned} B_{\rm intermediate}\, E_{\rm intermediate} \;=\; & \Big\{ z(1-A_{\rm head})A_{\rm rot}E_{\rm head} + z(1-A_{\rm head})(1-A_{\rm rot})E_{\rm head}E_{\rm rot} \\ & + \left[zA_{\rm head} + (1-z)A_{\rm tail} \right](1-A_{\rm rot})E_{\rm rot} \Big\} A_{\rm jump} \end{aligned} \tag{B.11}$$

Taking the $t \to 0$ limit in Eq.(B.11) gives,

$$B_{\rm intermediate} = \Big\{ z(1-A_{\rm head}A_{\rm rot}) + (1-z)A_{\rm tail}(1-A_{\rm rot}) \Big\} A_{\rm jump}\,. \tag{B.12}$$

And, taking the $t \to 0$ limit in the time derivative of Eq.(B.11) gives,

$$\Gamma_{\rm intermediate} = \frac{z(1-A_{\rm head})\Gamma_{\rm head} + \left[z + (1-z)A_{\rm tail} \right](1-A_{\rm rot})\Gamma_{\rm rot}}{z(1-A_{\rm head}A_{\rm rot}) + (1-z)A_{\rm tail}(1-A_{\rm rot})} \tag{B.13}$$

- **Fast motions:** collect all terms involving $E_{\rm jump}$ and $E_{\rm tail}$ in Eq.(B.5).

  For timescales $\Gamma_{\rm jump}t \sim \Gamma_{\rm tail}t \sim 1$, we use the rules above, $E_{\rm in-out} \sim E_{\rm 2d} \sim E_{\rm rot} \sim E_{\rm head} \approx 1$ to reduces Eq.(B.5) to,

$$\begin{aligned} B_{\rm fast}\, E_{\rm fast} = \; & z(1-A_{\rm jump})A_{\rm head}E_{\rm jump} \\ & + z(1-A_{\rm jump})(1-A_{\rm head})E_{\rm jump}E_{\rm head} + (1-z)A_{\rm jump}A_{\rm tail} \\ & + (1-z)A_{\rm jump}(1-A_{\rm tail})E_{\rm tail} + (1-z)(1-A_{\rm jump})A_{\rm tail}E_{\rm jump} \\ & + (1-z)(1-A_{\rm jump})(1-A_{\rm tail})E_{\rm jump}E_{\rm tail} \end{aligned} \tag{B.14}$$

  or,

$$\begin{aligned} B_{\rm fast}\, E_{\rm fast} \;=\; & \Big[zA_{\rm head} + (1-z)A_{\rm tail} \Big](1-A_{\rm jump})E_{\rm jump} + (1-z)A_{\rm jump}(1-A_{\rm tail})E_{\rm tail} \\ & + \Big[z(1-A_{\rm head})E_{\rm head} + (1-z)(1-A_{\rm tail})E_{\rm tail} \Big](1-A_{\rm jump})E_{\rm jump} \end{aligned} \tag{B.15}$$

  Taking the $t \to 0$ limit in Eq.(B.15) gives,

$$B_{\rm fast} = (1-A_{\rm jump}) + (1-z)A_{\rm jump}(1-A_{\rm tail}) = z(1-A_{\rm jump}) + (1-z)(1-A_{\rm jump}A_{\rm tail}) \tag{B.16}$$

  And, taking the $t \to 0$ limit in the time derivative of Eq.(B.15) gives,

$$\Gamma_{\rm fast} = \frac{(1-A_{\rm jump})\Gamma_{\rm jump} + (1-z)(1-A_{\rm tail})\Gamma_{\rm tail}}{B_{\rm fast}} \tag{B.17}$$

Appendix C. Relation between experimental *A*'s and theoretical *B*'s amplitudes

Given that the theoretical aggregated ISF in Eq.(B.4) writes as follows,

$$I_{\rm agg}(Q,t) = \sum_{i=0}^{n} B_i(Q)\, E_i(Q,t) \;\;;\;\; \sum_{i=0}^{n} B_i(Q) = 1 \;\;;\;\; B_i(Q=0) = \delta_{i,0}\,, \tag{C.1}$$

we assume that the experimental ISF can also be written in the same way as,

$$I_{\rm exp}(Q,t) = \sum_{i=0}^{n} A_i(Q)\, E_i(Q,t) \;\; ; \;\; \sum_{i=0}^{n} A_i(Q) = 1\,, \tag{C.2}$$

with the same $E_i(Q,t)$ but different amplitudes $A_i(Q)$ and such that, $0 < A_i(Q=0) < 1\,, \forall i$. The correspondence between $I_{\rm agg}(Q,t)$ and $I_{\rm exp}(Q,t)$ can be written as,

$$I_{\rm exp}(Q,t) \;\; = \;\; m I_{\rm agg}(Q,t) + \underbrace{\sum_{i=0}^{n} \varepsilon_i(Q)\, E_i(Q,t)}_{\text{error terms}}\,, \tag{C.3}$$

where m is the fraction of observable mobile H-atoms and $\varepsilon_i(Q)$'s the errors accounting for the fraction of immobile H-atoms and other experimental errors like multiple scatterings, etc. Collecting in Eq.(C.3) terms under the same relaxation function $E_i(Q,t)$, we obtain the general relationship between $B(Q)$'s and $A(Q)$'s as,

$$A_i(Q) = m B_i(Q) + \varepsilon_i(Q) \;\; ; \;\; \forall i\,. \tag{C.4}$$

From the point of view of the analysis of experimental data, the expression in Eq.(C.4) involves $n+2$ unknowns to determine: m and $n+1$ functions $\varepsilon_i(Q)$. Therefore, in the absence of any information we use the closure relation (obtained by construction) satisfied by the unknowns and assume that the errors of the quasi-elastic terms ($i > 0$) are all identical, i.e.,

$$\left\{ \begin{array}{l} m + \displaystyle\sum_{i=0}^{n} \varepsilon_i(Q) = 1 \\ \\ \varepsilon_i(Q) = \varepsilon(Q) \qquad ; \;\; i > 0 \end{array} \right. \implies \varepsilon_i(Q) = \frac{[1 - m - \varepsilon_0(Q)]}{n} \;\; ; \; i > 0\,, \tag{C.5}$$

thus reducing the number of unknowns from $n+2$ to 2. Using this back in Eq.(C.4), we obtain,

$$\left\{ \begin{array}{l} A_0(Q) = m B_0(Q) + \varepsilon_0(Q) \\ \\ A_i(Q) = m B_i(Q) + \dfrac{[1 - m - \varepsilon_0(Q)]}{n} \;\; ; \;\; i > 0 \end{array} \right. \tag{C.6}$$

where the remaining unknowns are m and $\varepsilon_0(Q)$.

References

[1] M. Luckey. *Membrane structural biology: with biochemical and biophysical foundations.* Cambridge University Press, New York, 2008.

[2] R. Lipowsky and L. Sackmann. *Structure and Dynamics of Membranes.* North Holland, 1995.

[3] S. König, T. M. Bayerl, G. Coddens, D. Richter, and E. Sackmann. Hydration dependence of chain dynamics and local diffusion in l-alphadipalmitoylphosphtidylcholine multilayers studied by incoherent quasielastic neutron scattering. *Biophys. J.*, 68:1871–1880, 1995.

[4] E. Lindahl and O. Edholm. Mesoscopic undulations and thickness fluctuations in lipid bilayers from molecular dynamics simulations. *Biophys. J.*, 79:426–433, 2000.

[5] W. Pfeiffer, Th. Henkel, E. Sackmann, W. Knoll, and D. Richter. Local dynamics of lipid bilayers studied by incoherent quasi-elastic neutron scattering. *Europhys. Lett.*, 8(2):201–206, 1989. URL https://doi.org/10.1209/0295-5075/8/2/016.

[6] U. Wanderlingh, G. DAngelo, V. Conti Nibali, M. Gonzalez, C. Crupi, and C. Mondelli. Influence of gramicidin on the dynamics of dmpc studied by incoherent elastic neutron scattering. . *Phys.: Condens. Matter*, 20:104214, 2008.

[7] M. C. Rheinstädter, C. Ollinger, G. Fragneto, F. Demmel, and T. Salditt. Collective dynamics of lipid membranes studied by inelastic neutron scattering. *Phys. Rev. Lett.*, 93:108107, 2004.

[8] F. Natali, C. Castellano, D. Pozzi, and A. Congiu Castellano. Dynamic properties of an oriented lipid/DNA complex studied by neutron scattering. *Biophysical Journal*, 88(2):1081–1090, February 2005. doi: 10.1529/biophysj.104.042788. URL https://doi.org/10.1529/biophysj.104.042788.

[9] M. Trapp, T. Gutberlet, F. Juranyi, T. Unruh, B. Demé, M. Tehei, and J. Peters. Hydration dependent studies of highly aligned multilayer lipid membranes by neutron scattering. *J. Chem. Phys.*, 133:164505, 2010.

[10] S. Gupta and G. J. Schneider. Modeling the dynamics of phospholipids in the fluid phase of liposomes. *Soft Matter*, 16:3245–3256, 2020.

[11] S. König, W. Pfeiffer, T. Bayerl, D. Richter, and E. Sackmann. Molecular dynamics of lipid bilayers studied by incoherent quasi-elastic neutron scattering. *J. Phys. II France*, 2:1589–1615, 1992. URL https://doi.org/10.1051/jp2:1992100.

[12] U. Wanderlingh, G. DAngelo, C. Branca, V. Conti Nibali, A. Trimarchi, S. Rifici, D. Finocchiaro, C. Crupi, J. Ollivier, and H. D. Middendorf. Multi-component modeling of quasielastic neutron scattering from phospholipid membranes. *J. Chem. Phys.*, 140:174901, 2014. URL https://doi.org/10.1063/1.4872167.

[13] Q. Berrod, K. Lagrené, J. Ollivier, and J.-M. Zanotti. Inelastic and quasi-elastic neutron scattering. application to soft-matter. *EPJ Web of Conferences*, 188:05001, 2018. URL https://doi.org/10.1051/epjconf/201818805001.

[14] D. J. Bicout. Diffusion incohrente des neutrons : modles analytiques pour la dynamique interne des protines. *J. Phys. IV France*, 130:115–132, 2005. URL https://doi.org/10.1051/jp4:2005130008.

[15] L. Carpentier, M. Be, A. M. Giroud-Godquin, P. Maldivi, and J. C. Marchon. Alkyl chain motions in columnar mesophases: A quasielastic neutron scattering study of dicopper tetrapalmitate. *Mol. Phys.*, 68(6):1367–1378, 1989.

[16] D. J. Bicout. Incoherent neutron scattering functions for diffusion inside two concentric spheres. *Phys. Rev. E*, 62:261–271, Jul 2000. doi: 10.1103/PhysRevE.62.261. URL https://link.aps.org/doi/10.1103/PhysRevE.62.261.

[17] In6: Cold neutron time-focusing time-of-flight spectrometer. URL http://www.ill.eu/instruments-support/instruments-groups/instruments/in6/description/instrument-layout/.

[18] J. Peters, J. Marion, F. Natali, E. Kats, and D. J. Bicout. The dynamical transition of lipid multilamellar bilayers as a matter of cooperativity. *J. Phys. Chem. B*, 121(28):6860–6868, 2017. doi: 10.1021/acs.jpcb.7b05167. URL https://doi.org/10.1021/acs.jpcb.7b05167.

[19] D. Richard, M. Ferrand, and G. J. Kearley. Analysis and visualisation of neutron-scattering data. *Journal of Neutron Research*, 4(1):33–39, December 1996. doi: 10.1080/10238169608200065. URL https://doi.org/10.1080/10238169608200065.

[20] Matthew Newville, Till Stensitzki, Daniel B. Allen, Antonino Ingargiola, and Andrew Nelson. *Lmfit: Non-linear least-square minimization and curve-fitting for python.* June 2016. URL https://ui.adsabs.harvard.edu/abs/2016ascl.soft06014N.

[21] A. Cissé and al. In preparation. 2021.
[22] T. Matsuo and al. In preparation. 2021.
[23] Computational chemistry comparison and benchmark database, Release 21, August 2020. URL https://cccbdb.nist.gov/.
[24] Wiebke Knoll. *Le Rôle de la myéline dans les maladies dégénératives*. Theses, Université de Grenoble, September 2012. URL https://tel.archives-ouvertes.fr/tel-00744455.
[25] Avanti polar lipids website, Copyright 2021 Croda International Plc. URL avantilipids.com.
[26] Stephanie Tristram-Nagle, Yufeng Liu, Justin Legleiter, and John F. Nagle. Structure of gel phase DMPC determined by x-ray diffraction. *Biophysical Journal*, 83(6):3324–3335, December 2002. doi: 10.1016/s0006-3495(02)75333-2. URL https://doi.org/10.1016/s0006-3495(02)75333-2.
[27] Anthony Watts, Karl Harlos, and Derek Marsh. Charge-induced tilt in ordered-phase phosphatidylglycerol bilayers evidence from x-ray diffraction. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 645(1):91–96, June 1981. doi: 10.1016/0005-2736(81)90515-0. URL https://doi.org/10.1016/0005-2736(81)90515-0.
[28] Bachir Aoun, Eric Pellegrini, Marcus Trapp, Francesca Natali, Laura Cantù, Paola Brocca, Yuri Gerelli, Bruno Demé, Michael Marek Koza, Mark Johnson, and Judith Peters. Direct comparison of elastic incoherent neutron scattering experiments with molecular dynamics simulations of dmpc phase transitions. *The European Physical Journal E*, 39(4), apr 2016. doi: 10.1140/epje/i2016-16048-y. URL https://doi.org/10.1140/epje/i2016-16048-y.
[29] Wiebke Knoll, Judith Peters, Petri Kursula, Yuri Gerelli, Jacques Ollivier, Bruno Demé, Mark Telling, Ewout Kemner, and Francesca Natali. Structural and dynamical properties of reconstituted myelin sheaths in the presence of myelin proteins MBP and p2 studied by neutron scattering. *Soft Matter*, 10 (3):519–529, 2014. doi: 10.1039/c3sm51393a. URL https://doi.org/10.1039/c3sm51393a.
[30] A. J. Dianoux, F. Volino, and H. Hervet. Incoherent scattering law for neutron quasi-elastic scattering in liquid crystals. *Mol. Phys.*, 30(4):1181–1194, 1975.
[31] A. J. Dianoux, M. Pineri, and F. Volino. Incoherent scattering law for restricted diffusion inside a volume with an anisotropic shape: Application to the problem of water absorbed in nafion membranes. *Mol. Phys.*, 46(1):129–137, 1982.
[32] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1972.

Appendix B

Additional publications

I was involved as well in other projects:

- In 2019, I analyzed QENS data for the group of Pr. Roland Winter (TU Dortmund University, Germany). The system was the tetrameric lactate dehydrogenase protein, studied with different osmolytes, under high-hydrostatic pressure [126].
- During my second internship at ILL, and at the beginning of my thesis, I wrote with Dr. Leonardo Chiappisi (Large Scale Structure group, ILL) a Python script for the analysis of Differential Scanning Calorimetry (DSC) data, which we called *pyDSC*, and which was released online on the Github platform. The associated paper was published in the Journal of Thermal Analysis and Calorimetry [286].
- All along my thesis, I participated in a collaboration between the group of Pr. Burkhard Bechinger (Université de Strasbourg, France), including Dr. Arnaud Marquette, and my thesis director, Pr. Judith Peters, who was co-responsible of IN13 spectrometer during their experiments. I performed the measurements with them, carried out additional DSC experiments on their lipid systems along with one intern of our group (Munkhtuguldur Altangerel), and analyzed the data. This work was published in the Journal of Physical Chemistry B [287].

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2019, **21**, 12806

Osmolytes modify protein dynamics and function of tetrameric lactate dehydrogenase upon pressurization†

Samy R. Al-Ayoubi,[a] Paul Hendrik Schummel,[a] Aline Cisse,[bc] Tilo Seydel,[b] Judith Peters[bc] and Roland Winter*[a]

Received 24th April 2019,
Accepted 28th May 2019

DOI: 10.1039/c9cp02310k

rsc.li/pccp

We present a study of the combined effects of natural cosolvents (TMAO, glycine, urea) and pressure on the activity of the tetrameric enzyme lactate dehydrogenase (LDH). To this end, high-pressure stopped-flow methodology in concert with fast UV/Vis spectroscopic detection of product formation was applied. To reveal possible pressure effects on the stability and dynamics of the enzyme, FTIR spectroscopic and neutron scattering measurements were carried out. In neat buffer solution, the catalytic turnover number of the enzyme, k_{cat}, increases up to 1000 bar, the pressure range where dissociation of the tetrameric species to dimers sets in. Accordingly, we obtain a negative activation volume, $\Delta V^{\#} = -45.3$ mL mol^{-1}. Further, the enzyme substrate complex has a larger volume compared to the enzyme and substrate in the unbound state. The neutron scattering data show that changes in the fast internal dynamics of the enzyme are not responsible for the increase of k_{cat} upon compression. Whereas the magnitude of k_{cat} is similar in the presence of the osmolytes, the pressure of deactivation is modulated by the addition of cosolvents. TMAO and glycine increase the pressure of deactivation, and in accordance with the observed stabilizing effect both cosolvents exhibit against denaturation and/or dissociation of proteins. While urea does not markedly affect the magnitude of the Michaelis constant, K_M, both 1 M TMAO and 1 M glycine exhibit smaller K_M values of about 0.07 mM and 0.05 mM below about 1 kbar. Such positive effect on the substrate affinity could be rationalized by the effect the two cosolutes impose on the thermodynamic activities of the reactants, which reflect changes in water-mediated intermolecular interactions. Our data show that the intracellular milieu, *i.e.*, the solution conditions that have evolved, may be sufficient to maintain enzymatic activity under extreme environmental conditions, including the whole pressure range encountered on Earth.

Introduction

Modulation of enzymatic activity by high hydrostatic pressure (HHP) has become a promising trait of engineering enzymatic activity in recent years, though the underlying mechanisms are still poorly understood.[1–16] The use of HHP can have several advantages: (i) given that the rate of an enzymatic reaction is often limited by the thermostability of the corresponding enzyme, superimposing pressure-induced thermostabilization of an enzyme with an accelerated substrate conversion at increased temperatures can lead to an enhanced overall reaction rate; (ii) the catalytic turnover number can also be directly enhanced by HHP (*e.g.*, by selection of a different conformational substate of the enzyme) if the activation volume associated with the reaction is negative; (iii) based on the same thermodynamic principle, HHP can alter the substrate specificity and stereoselectivity of an enzyme by favoring the conversion of a substrate to a product with a smaller volume; (iv) as pressure can weaken particular intermolecular interactions, HHP may also increase conformational flexibility in the active site, thereby improving reaction rates, as sufficient conformational flexibility is often required for the activation of enzymes.

Though some data are available regarding the effect of pressure on the activity of monomeric enzymes, little is known about the effect of pressure on the activity of oligomeric ones, which constitute a large family of enzymes.[17] However, one of the effects upon compression of oligomeric enzymes, their pressure-induced dissociation, is a well-known phenomenon and has been observed in several cases.[18–27] In general, most oligomeric proteins dissociate between 1 and 2 kbar due to imperfect interfacial packing of the monomeric units. What is

[a] *Physical Chemistry I – Biophysical Chemistry, Faculty of Chemistry and Chemical Biology, TU Dortmund University, Otto-Hahn-Str. 4a, 44227 Dortmund, Germany. E-mail: roland.winter@tu-dortmund.de*
[b] *Institut Laue Langevin, F-38042 Grenoble Cedex 9, France*
[c] *University Grenoble Alpes, LiPhy, F-38000 Grenoble, France*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9cp02310k

largely terra incognita, however, is the combined effect of pressure and cosolvents on the pressure stability, dynamics and activity of oligomeric enzymes.

In this study, the tetrameric enzyme lactate dehydrogenase (LDH) has been in the focus. This enzyme catalyzes the interconversion of pyruvate to lactate with nicotinamide adenine dinucleotide (NADH) as co-substrate during the anaerobic glycolysis. The catalytic mechanism involves the binding of pyruvate to the LDH–NADH complex, closing of the active site by a mobile loop, hydride transfer from NADH to the carbonyl carbon atom of pyruvate, opening of the loop and finally product release.[28,29] The single steps have been studied using temperature-jump relaxation methods.[30–33] Pyruvate does not bind to the enzyme in the absence of NADH.[34,35] MD simulations revealed that the tetrameric conformation is necessary to maintain the geometry of the active site and to prevent penetration of water into the reaction center.[36] In the apo-enzyme and the LDH–NADH-complex, the active site is open to the solvent while in the presence of the substrate the active site is closed by the mobile loop.[37] Loop closure is also the rate determining step of the catalysis reaction.[37] Inhibitors of LDH are potential anticancer drugs.[38,39] The pressure-induced dissociation of LDH is well studied and sets in at approximately 1000 bar at room temperature, which is accompanied by a loss of enzymatic activity.[40–44] Small-angle X-ray scattering (SAXS) measurements revealed that dissociation of the tetramer leads to dimer formation.[44]

To reveal the effects of pressure and substrate binding on the dynamic properties of LDH, elastic incoherent neutron scattering (EINS) and quasi-elastic neutron scattering (QENS) measurements were carried out in the presence of the enzyme's cosubstrate NADH, the substrate analogue oxamate, and in the presence of the naturally occurring osmolyte glycine as an important representative of a compatible cosolvent. The pressure variable has been applied here not only to tune the oligomeric state of the protein, but also because pressure is able to modulate the activity of the enzyme.

In general, the activity of multimeric enzymes depends on their state of oligomerization. Cosolvents, such as natural osmolytes, may influence both the oligomerization state as well as kinetic properties of the enzyme reaction.[7,9,45] Hence, we studied also the pressure dependent enzymatic activity in the presence of selected osmolytes. In general, osmolytes like the amino acid glycine and the methylamine trimethylamine-*N*-oxide (TMAO) are synthesized by marine organisms to counteract stress factors like hydrostatic pressure (in deep sea trenches, pressures up to 1000 bar are encountered[46]). For comparison, we also studied the destabilizing cellular osmolyte urea.[47–52] We included all three cosolvents in our enzyme activity studies, EINS and QENS studies have been carried out for one of the osmolytes, glycine, next to the neat buffer data for restricted beam time reasons. Glycine is the main osmolyte in shallow water invertebrates, while the TMAO content of teleosts increases linearly with increasing pressure.[49] TMAO is thought to be excluded from the protein surface, hence stabilizing proteins against unfolding *via* an excluded volume mechanism. Conversely, urea interacts preferentially with the protein backbone, thereby destabilizing proteins.[51] Differently, glycine was also found to interact directly with the protein but has nevertheless a stabilizing effect.[53] The effects of TMAO and urea on the sub-ns dynamics of a monomeric protein, lysozyme, at high pressure have been investigated previously by our lab using EINS.[54] While 2 M TMAO reduces the protein's fast (sub-ns) internal dynamics, 2 M urea did not show a marked effect on the fast internal dynamics of the protein.

Experimental section

Sample preparation

Lactate dehydrogenase (LDH) from rabbit muscle was purchased from Merck (Darmstadt, Germany). Urea, TMAO, glycine, TRIS, DTT, EDTA and D_2O were purchased from Sigma Aldrich (Darmstadt, Germany). Sodium oxamate and NADH was obtained from Alpha Aesar (Darmstadt, Germany) and sodium pyruvate from Carl Roth (Karlsruhe, Germany). All measurements were performed in pressure-stable Tris–HCl buffer at pH 7.6. To prevent the oxidation of sulfhydryl groups at elevated pressure, the buffer contained 10 mM DTT and 1 mM EDTA.[41,42] Neutron scattering and FTIR spectroscopy measurements were carried out in D_2O based buffers to avoid contamination of the sample signal by the solution's signal. Regarding neutron scattering, the incoherent neutron cross section of H-atoms is indeed much higher than that of all other atoms and permits to highlight parts of the sample. Inversely, H/D exchange of the solution allows to hide its contribution. The pD was adjusted by adding 0.4 to the pH meter reading.[55]

Neutron scattering experiments

Neutron scattering experiments were carried out on the two backscattering spectrometers IN13 and IN16B of the Institut Laue-Langevin (ILL, Grenoble, France).[56,57] IN13 is a thermal backscattering spectrometer using a wavelength of 2.23 Å. The time-window is about 100 ps and it covers a spatial range of 1–30 Å. IN16B was used with a Si(111) backscattering monochromator and analyzer crystals (elastic scattering wavelength of 6.27 Å), employing a Phase Space Transformation chopper to increase the neutron flux at the sample position[58] and a Doppler drive to scan the incident energy in a range of $\pm$30 μeV relative to the elastic energy. Each spectrometer has its characteristic energy resolution. The 8 μeV resolution of IN13 corresponds to a time window of about 0.1 ns, whereas IN16B gives access to motions within a 1 ns time scale. Accordingly, these instruments can be used as motion filters, as local elastic vibrations and rotations of amino acid side chains can, for instance, be probed on IN13, whereas IN16B provides information on diffusive motions. When the sample is probed in solution as in the present case, global diffusion of the whole protein can also become visible on the latter instrument.

The high-pressure equipment was developed by the Service for Advanced Neutron Environment (SANE) of the ILL for measurements on biological samples in solution.[59] It consists

of high-pressure sample cells, built using a high-tensile aluminum alloy which is mainly transparent for neutrons and withstanding pressure loads up to 6 kbar. The pressure stick maintaining the cell is placed inside the cryostat controlling temperature. As each cryostat has a cold point around 77 K, the liquid transmitting the hydrostatic pressure has to be heated and isolated by a secondary vacuum from the surrounding. Fluorinert was used as pressure transmitting medium, which was separated from the sample by a specific separator device.[60] All measurements were performed at 25 °C and each pressure point was measured for at least 6 h on IN13 and for 30 min on IN16B, which has a much higher neutron flux. To treat the data extracted from elastic scattering, the signal of the empty cell and of the buffer at the corresponding pressure was subtracted from the scattering intensity of the sample. Additionally, normalization to a vanadium standard was performed to correct for geometric and detector efficiency effects. The complete data reduction was carried out using the LAMP software available at ILL.[61]

Elastic measurements allow the determination of the apparent atomic mean-squared displacements (MSD), $\langle u^2 \rangle$, from the elastic scattering function, S, using the Gaussian approximation, which assumes that the atomic nuclei undergo harmonic motions around their equilibrium positions:[62]

$$S(Q, 0 \pm \Delta E) \approx S_0 \exp\left(\frac{-Q^2\langle u^2\rangle}{3}\right) \quad (1)$$

Here, S_0 is the incident neutron flux, Q the momentum transferred from the incident to the outgoing neutron wave vector, and ΔE is the instrumental energy resolution. The MSD values are obtained for each pressure point from the slope of the logarithm of the scattered intensities plotted *versus* Q^2 according to

$$\langle u^2 \rangle \approx -3\frac{\partial \ln S(Q, 0 \pm \Delta E)}{\partial Q^2}. \quad (2)$$

QENS measurements were performed on the instrument IN16B of the ILL. They provide information about the incoherent scattering function as a function of the momentum transfer Q and the energy transfer $\hbar\omega$. The model scattering function for fitting the IN16B data was implemented as[63,64]

$$S(Q,\omega) = R \otimes \{\beta[A_0 L(\gamma,\omega) + (1 - A_0)L(\gamma + \Gamma,\omega)] + \beta_{D_2O} L(\gamma_{D_2O},\omega) + \beta_\delta \delta(\omega)\}. \quad (3)$$

Therein, $R = R(Q,\omega)$ denotes the spectrometer resolution function, $\beta = \beta(Q)$ a scalar scaling factor, $\beta_{D_2O}(Q)$ the amplitude of the solvent (D_2O) diffusive contribution, and $A_0(Q)$ the elastic incoherent structure factor (EISF). The symbols L represent Lorentzian functions, and $\gamma(Q)$, $\Gamma(Q)$, and $\gamma_{D_2O}(Q)$ denote the linewidths associated with the centre-of-mass diffusion of the proteins in the solvent, the internal protein diffusion, and the solvent diffusion, respectively. The additional elastic contribution δ with amplitude β_δ accounts for an imperfect subtraction of the high-pressure sample cell.

The IN16B data were reduced using Mantid[65] and subsequently fitted using Python3, notably employing scipy.optimize.curve_fit. The spectrometer resolution R was modelled by a sum of several Gaussian functions, and the convolution of the model scattering function with this resolution function was carried out analytically by Voigt functions implemented using the real part of the Faddeeva function from scipy.special. The IN16B spectra were fitted in a global fit approach for all Q-values at once. The linewidth associated with the solvent diffusion, $\gamma_{D_2O}(Q)$, was interpolated from the IN6 measurements presented in ref. 64. The centre-of-mass diffusion of the protein and solvent were modelled by Brownian diffusion, so the associated linewidths are expressed as:

$$\gamma(Q) = DQ^2 \quad (4)$$

with D being the respective global diffusion coefficient.

Finally, the internal protein diffusion is represented by a confined Brownian diffusion model that takes also into account the restrictions induced by pressure application. The original model of diffusion in a sphere, proposed by Volino and Dianoux,[66] assumes an infinite sum within the expression of the scattering function taking also into account a series of rotational diffusion terms. The model used here for fitting $S(Q,\omega)$ is restricted to only one Lorentzian for the internal dynamics as data quality is not sufficient to distinguish more of them. In the global fit, the corresponding linewidths Γ are fitted as a function of Q. The form of $\Gamma(Q)$ provides indications of the kind of motion present, but to extract characteristic parameters, a model is required for $\Gamma(Q)$. A heuristic approach was used here in the data analysis, taking into account the extreme values of the theoretical $\Gamma(Q)$ in the model proposed by Volino and Dianoux:[66]

$$\Gamma(Q) = \frac{1}{1 + e^{+10[Q^2 - (\pi/a)^2]}} \cdot \frac{4.33 D_1}{a^2} + \frac{1}{1 + e^{-10[Q^2 - (\pi/a)^2]}} D_1 Q^2, \quad (5)$$

with D_1 being the internal diffusion coefficient of the protein, and a the radius of the confinement. In the global fit approach, D, D_1 and a are free parameters, whose final values are determined by the best fits in the least-square definition.

Furthermore, we calculated theoretical global diffusion coefficients for monomers and tetramers of the protein based on the crystal structures of LDH (PDB code 5QNB) using HYDROPRO,[67] in which the protein molecules are assumed to be rigid. Such approach permits to disentangle the amount of monomers and tetramers at the different pressure values. To this end, the theoretical values for the hydrodynamic volume fraction, φ_t, have been calculated as explained in C. Beck *et al.*,[68] $\varphi_t = \varphi(R_h/R)^3$, which rescales the theoretical values to the known dry protein volume fraction, φ. Therein, the hydrodynamic radius of the protein, R_h, was calculated from the HYDROPRO dilute-limit translational diffusion coefficient, D_{t0}, according to the Stokes–Einstein equation,

$$R_h = \frac{k_B T}{6\pi\eta D_{t0}}, \quad (6)$$

where k_B is Boltzmann's constant, T the temperature in Kelvin, η the viscosity of the solvent, and the radius of the dry protein

radius, R, was calculated from the molecular mass M_w of the protein and its specific volume, V_s, according to

$$R = \left(\frac{3V_s M_w}{4\pi N_A}\right)^{1/3}, \quad (7)$$

where N_A is the Avogadro constant.

Measurements of the enzymatic activity

For the activity assays, a NADH concentration of 0.6 mM was used. The enzyme concentration, c_{E0}, was 0.46 nM. The concentration was determined UV-spectroscopically using a Nanovue Plus Spectrometer (Biochrom, Cambridge, UK). The activity measurements were performed using a PerkinElmer Lambda 25 spectrometer. Pressure dependent measurements were performed in a custom-built high-pressure cell equipped with sapphire windows. The initial reaction velocity, v_0, was determined by a linear fit of the time dependent development of the absorbance at 340 nm, which is due to interconversion of NADH to NAD^+. The data were processed according to the mechanism of Michaelis–Menten:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P \quad (8)$$

E characterizes the free enzyme, S the substrate, ES the enzyme–substrate complex and P the products of the reaction, $k_2 = k_{cat}$ characterizes the turnover number, $K_M = (k_{-1} + k_{-2})/k_1$ is the Michaelis constant, an inverse measure of substrate affinity. The activation volume, $\Delta V^{\#}$, was determined from the pressure dependence of the initial reaction velocity, v_0, according to:[14]

$$\frac{d\ln(v_0/v_{0,1bar})}{dp} = -\frac{\Delta V^{\#}}{RT} \quad (9)$$

Fourier-transform infrared (FTIR) spectroscopy

Prior to the FTIR measurements, the protein was dissolved in D_2O and equilibrated for several hours to exchange labile H-atoms and lyophilized overnight. The protein concentration was 3 wt%. FTIR spectroscopic measurements were conducted using a Nicolet Magna 550 FTIR spectrometer (Thermo Fisher Scientific, Waltham, MA) and a Diacell Vivo DAC diamond anvil cell (Almax EasyLab, Diksmuide, Belgium. Pressure was applied by the pneumatic PACE 5000 system (GE Sensing, Frankfurt, Germany)). The IR spectra were background-subtracted and the amide I′ band was normalized to its area. The subbands for the analysis of the secondary structure elements were derived from the 2nd derivatives of the spectra as well as from the corresponding Fourier self-deconvoluted spectra.[45,69]

Results and discussion

Elastic incoherent neutron scattering (EINS) results

The elastic incoherent neutron scattering data are representative of small and local motions within the sample, typically vibrations and rotations of the side chains of the amino acids. Fig. S1 (ESI†) shows the EINS intensities as $\ln S$ *vs.* Q^2 plots in the neutron momentum transfer range from Q^2 = 1.3 $Å^{-2}$ to 3.5 $Å^{-2}$ at various pressure points as well as linear fits according to the Gaussian approximation. This Q^2 range was chosen to fulfill the validity of the approximation assuming a spherical geometry $\left(\sqrt{\langle u^2\rangle Q^2} \approx 2.8\right)$ and to reduce the influence of diffusive motions of H_2O or HDO impurities arising from HD exchange processes at small Q-values.[70]

The pressure dependence of the mean-squared displacement, MSD, is shown in Fig. 1. Upon compression, the MSD of the apo-enzyme remains essentially constant up to 1000 bar. A further increase of the pressure leads to increased MSD values, which are followed by a drastic decrease of the MSD beyond 2000 bar. As mentioned above, the enzyme dissociates at approximately 1000 bar. Hence, the observed increase of the MSD is most likely attributable to the subunit dissociation, as such dissociation leads to a higher flexibility which is counteracting the general pressure effect provoking reduced motions. The FTIR spectroscopic measurements shown in Fig. S2 (ESI†) reveal that this dissociation is accompanied by changes of the protein's secondary structure, namely a reduction of the α-helical content accompanied by the formation of more β-sheet structures. Such minor conformational changes upon subunit dissociation of multimeric proteins has been observed in several cases and denoted conformational drift by Weber *et al.*[22,23,40] The dissociation leads to the hydration of the intersubunit area and induces water penetration into the active site of the enzyme, thereby reducing the overall volume of the system, which is the cause of the pressure-induced dissociation process. This hydration change might not only induce the variation of the protein's secondary structure but also the increase of the fast internal protein dynamics. The observed decrease of the MSD beyond 2000 bar might be due to previously observed aggregation effects under pressure.[42,44] However, owing to the absence of any IR-band characteristic for protein aggregation (Fig. S2, ESI†), *e.g.*, bands at 1615 and 1680 cm^{-1} due to the formation of intermolecular β-sheets,[71,72] such scenario cannot be confirmed. We found the secondary structure of LDH to be essentially unaffected by pressure application between 2000 and 7000 bar. Consequently, the observed changes are rather due to the intrinsic elastic effect of pressure, which leads to a sizeable reduction of molecular motions due to increased volume restriction at these high pressures beyond 2 kbar. This is also the pressure range, where the structure of the water and physico-chemical properties of H_2O change markedly (*e.g.*, the isothermal compressibility and diffusivity decrease and the viscosity increases substantially).[73,74]

In the presence of NADH, the MSDs compared to those of the apo-enzyme seem to be slightly reduced. Calculating the p-value with a Student bilateral test, we find however for most of the pressure values that the p-value is higher than the critical p-value of 0.32 assuming a 1σ confidence interval, which means that the variation in the MSD between LDH and LDH + NADH is not significant within 1σ. The increase of the MSD at ~1000 bar is also seen in the complex of NADH and LDH. Interestingly, as shown in Fig. 2B, the fast dynamics of the

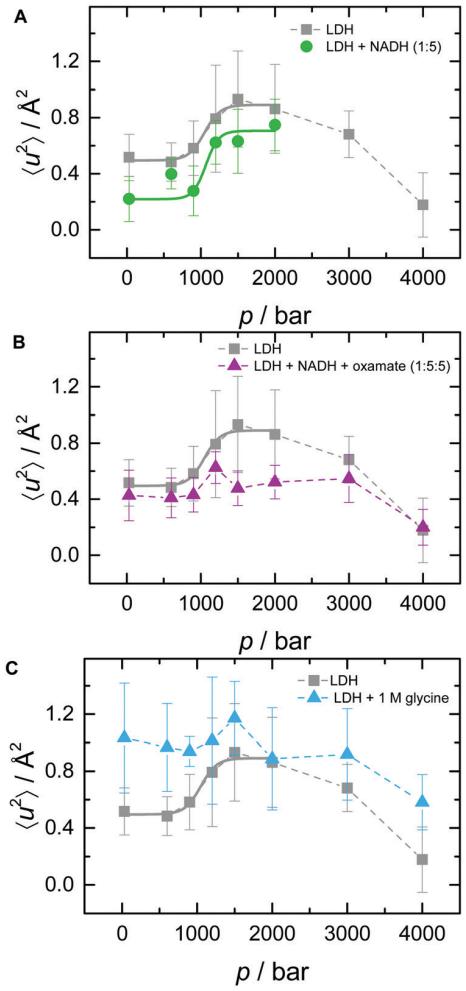


Fig. 1 Pressure dependence of the mean-squared displacement, MSD, of 90 mg mL^{-1} LDH (apo-LDH) (A), the complex between LDH and NADH, the complex between LDH, NADH and the substrate analogue oxamate (B), and of apo-LDH in the presence of 1 M glycine (C). The MSD results are from IN13. The error bars were evaluated through a least square fitting method. The sigmoidal curves and dashed lines are shown as guide to the eyes only.

ternary complex of the enzyme with both substrates, NADH and oxamate, is essentially unaffected by pressure application up to ~3000 bar, where the decrease of the MSD sets in as well. Slightly reduced MSD values compared to the apo-enzyme are visible. The binding of oxamate to the LDH–NADH-complex induces closure of the active site *via* a rearrangement of the mobile loop region. This closed conformation by substrate binding seems to stabilize the protein against pressure-induced dissociation. Remarkably, 1 M glycine increases the MSD significantly. Concerning the variation of the MSD below 1500 bar, the *p*-value is higher than the critical *p*-value 0.32, which indicates that it is not significant within the 1σ confidence interval. On the other hand, the *p*-value corresponding to the variation of the MSD above 1500 bar is smaller than 0.32, so that the decrease of the MSD observed in the Fig. 1 is meaningful.

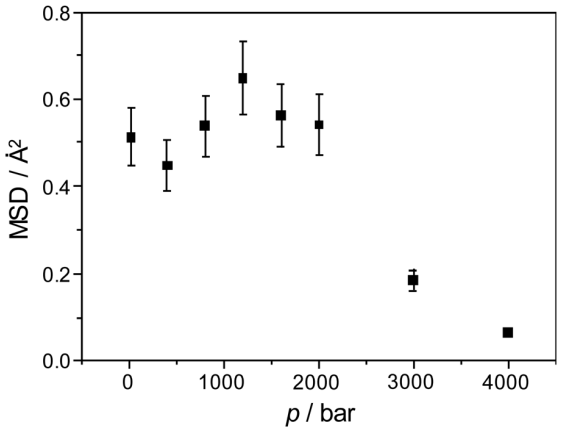


Fig. 2 Pressure dependence of the MSD, $\langle u^2 \rangle$, of all hydrogen atoms of 90 mg mL^{-1} LDH upon increasing pressure measured on IN16B (T = 25 °C).

This decrease is associated with the general effect of pressure, which decreases the volume accessible for motions (according to Le Châtelier's principle). Hence, the cosolute glycine seems to impose slightly larger fast internal protein fluctuations, but also seems to stabilize the protein tetramer against pressure-induced dissociation.

Quasi-elastic incoherent neutron scattering results

Quasi-elastic neutron scattering (QENS) data give access to the geometry of motions and permit to identify the presence of different kinds of movements and to extract associated dynamical parameters. For beam time limitations, we investigated the apo-enzyme, only. Fig. S3 (ESI†) presents an example of the fit of the total quasi-elastic spectrum, from which the different contributions could be extracted. First, we analyzed the elastic scattering contribution to the overall spectrum (Fig. 2). This contribution is similar to the results obtained on IN13, but on a much longer time scale of about 1 ns. The elastic results resemble the IN13 MSD data with a maximum at about 1200 bar when increasing pressure. At 3 and 4 kbar, the MSDs are drastically reduced, leading almost to a 'freezing' of the internal motions on that timescale at highest pressures.

Following eqn (3), we are able to separate the characteristics of the global diffusion, corresponding to a combination of translational and rotational (tumbling) movements of the whole protein within the solution, and the internal diffusion of atoms or small molecular subgroups of the protein. Fig. 3a presents the parameter extracted for the global diffusion, *i.e.*, the linewidth γ of the corresponding Lorentzian as a function of Q^2. The linewidth shows a linear behavior in Q^2, indicating an essentially Fickian-type of diffusion of the protein, and tends to zero at the origin, pointing to no confinement within this movement, what is expected for free Brownian diffusion of a particle in solution. From the slope of this curve, the diffusion coefficient, D, can be calculated (eqn (4)), which is depicted in Fig. 3b as a function of pressure. The results of the fits were very stable for this parameter, leading to small error bars. Interestingly, the pressure dependence of this global center-of-mass

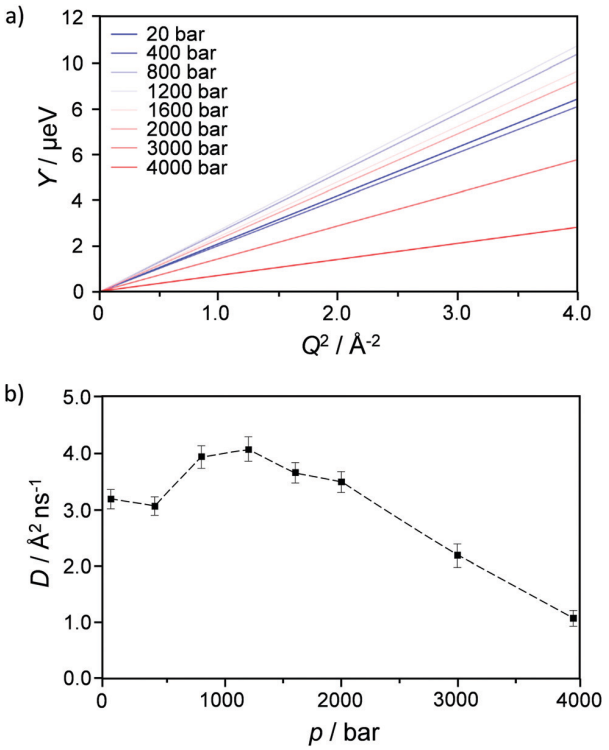


Fig. 3 (a) Linewidth $\gamma(Q)$ of the global diffusion (Lorentzian) of LDH for selected pressures as obtained from the global fit results. (b) Short-time (ns) center-of-mass translational diffusion coefficient, D, as a function of pressure at T = 25 °C.

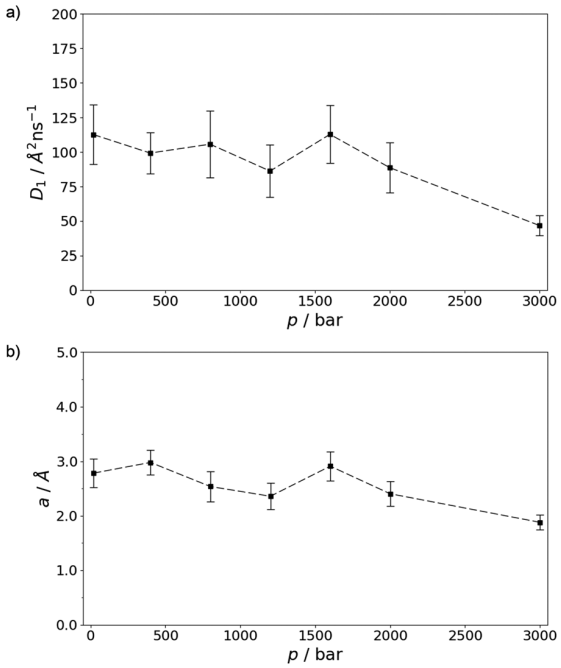


Fig. 4 (a) Internal diffusion coefficient, D_1, and (b) radius of confinement, a, of LDH at 25 °C as a function of pressure. Data are shown up to 3 kbar, only, as fits at 4 kbar became unstable.

diffusion coefficient is similar to that of the MSD, increasing up to the subunit dissociation around 1 kbar and decreasing beyond that pressure. This result can in fact be rationalized invoking dissociation of the tetrameric protein around 1 kbar, and subsequent drastic increase of the solvent viscosity at high pressures (*cf.* eqn (6)), which increases particularly above 2 kbar[73,74] and leads to the observed marked decrease of the global translational diffusion coefficient. Analysis of the pressure dependence of the translational diffusion coefficient from the QENS data and comparison with modelling of the hydrodynamic radius of LDH seems to be consistent with a dissociation event of tetramers or larger assemblies into smaller assemblies – presumably dimers – (see Fig. S4, ESI†) in terms of the observed hydrodynamic properties. We note that association or dissociation events may be transient in time, such that this consistency drawn from the observable hydrodynamic properties does not provide static average information. Comparing the results with those of Fig. 3 on the restricted internal protein motions, it appears that the pressure dependence of the fast motions on the local scale is correlated with the pressure dependence of the global translational motion of the protein.

Fig. S5 (ESI†) shows the linewidth Γ of the internal diffusion, which is characterized by a linear behavior at higher Q^2 values and finite values at low Q, indicating confinement in a spherical volume. The radius of confinement, a, can be extracted from the inflexion point of Γ around 1.2 Å^{-2} and the internal diffusion coefficient, D_1, through the fit using eqn (5) (see Fig. 4); the parameter a presents values slightly below 3 Å, typical for the low degree of freedom of atoms attached to the densely packed amino acids' side chains. Both quantities, D_1 and a, crease slightly around 3 kbar only, as expected upon strong compression.

Enzymatic activity of the LDH catalyzed reaction

Representative Michaelis–Menten plots of the enzymatic activity (initial reaction velocity normalized to the enzyme concentration, v_0/c_{LDH}) at selected pressures of the LDH catalyzed reaction are shown in Fig. 5. The initial velocity of the enzyme reaction is given by v_0 = d[P]/dt, where [P] is the concentration of the product, *i.e.* the rate of product formation, as calculated from the slope of the linear fit of the time dependent absorbance data at 340 nm.

We used a large excess of substrate that is saturating the enzyme, which also avoids significant back reaction. The kinetic parameters $K_M = (k_{-1} + k_2)/k_1$ (Michaelis constant) and $k_{cat} = k_2$ (rate constant of the catalysis or turnover number) were obtained by linear regression analysis of the Michaelis–Mention plots. Owing to the rather complex two-substrate enzyme-catalyzed reaction, the kinetic parameters derived are model dependent and hence may be viewed as apparent ones. In the literature, turnover numbers, k_{cat}, of 190 s^{-1} for a bacterial LDH (pH 6, 25 °C)[29] and 350 s^{-1} for human LDH were reported.[75] These values are of comparable magnitude to our results with k_{cat} values of about 400 s^{-1}. As shown in Fig. 6, k_{cat} increases with increasing pressure up to 400 bar. Upon further pressure application, k_{cat}

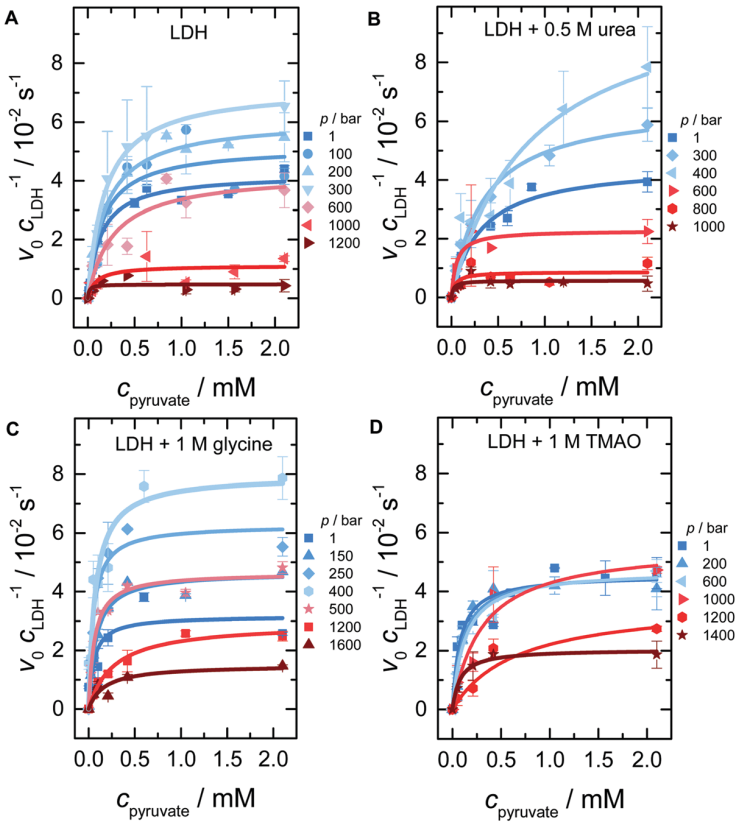


Fig. 5 Michaelis–Menten plots (initial reaction velocity normalized to the enzyme concentration, v_0/c_{LDH}) of the LDH catalyzed interconversion of pyruvate to lactate at ambient and elevated pressure in the absence (A) and presence of the osmolytes urea (B), glycine (C) and TMAO (D). The error bars represent mean values of at least three independent kinetic measurements. The solid lines represent fits according to the Michaelis–Menten formalism.

decreases again to the initial value, until deactivation occurs at about 1000 bar. Molecular dynamics simulations showed that dissociation changes the active site's geometry of LDH, which explains the deactivation by pressure-induced subunit dissociation.[36]

As can be seen in Fig. 6, the magnitude of the catalytic turnover number k_{cat} is similar to the neat buffer data in the presence of the osmolytes, the pressure dependence is affected by the presence of osmolytes, in particular for TMAO, however. Remarkably, the increase of k_{cat} is not observed in the presence of 1 M TMAO and remains essentially constant over a large pressure range up to about 1200 bar. Glycine and urea exhibit a similar pressure dependence of k_{cat} like the neat buffer data.

In the absence of osmolytes, deactivation of the enzyme sets in at approximately 1000 bar, which is the pressure range were dissociation of the protein occurs. The pressure of deactivation is modulated by the addition of cosolvents. While 0.5 M urea decreases the pressure of deactivation to $\sim$600 bar, 1 M TMAO and glycine increase the pressure of deactivation to about 1200 bar, in accordance with the observed stabilizing effect both cosolvents have shown against denaturation and/or dissociation of proteins.[7,52,53]

In neat buffer solution, K_M increases upon pressure application from about 0.14 mM up to 0.34 mM at 800 bar, indicating a reduced affinity of the substrate with increasing pressure (Fig. 6). In the presence of 0.5 M urea, a slight increase of K_M might be implied, which would be in accordance with the fact that urea was found to be a competitive inhibitor of pyruvate in teleost, elasmobranch and mammalian LDH.[76–78] In the presence of TMAO and glycine, K_M remains essentially constant up to 800 bar and increases at higher pressure. While urea does not markedly affect the magnitude of K_M, both, 1 TMAO and 1 M glycine exhibit smaller K_M values of about 0.07 mM and 0.05 mM below about 1000 bar, respectively. Previously, 250 mM TMAO was also found to decrease K_M of tNADH LDH from a deep-sea grenadier fish, while 250 mM glycine had no effect.[48]

In general, the Michaelis constant K_M may be influenced by various factors induced by the addition of cosolutes, such as (i) by minor conformational changes of the binding pocket (or selection of a different conformational substate), (ii) by compaction or partial unfolding of the whole protein, (iii) by increased or reduced protein dynamics imposed by the solvation shell ('slaving effect'), or (iv) by changes in the thermodynamic activity (activity coefficient) of the substrate (and/or enzyme). TMAO has been observed to show activating and deactivating effects on enzyme kinetics.[79] Distinct effects of glycine were also reported for the enzymes xylanase

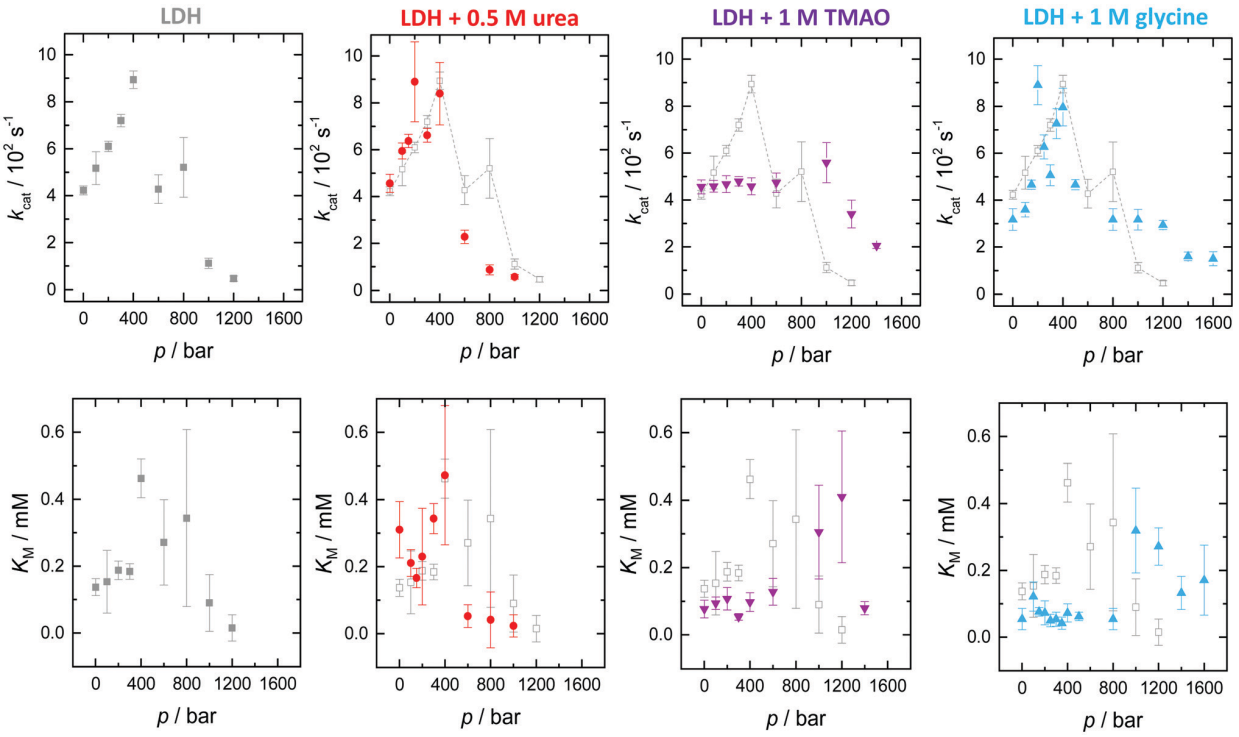


Fig. 6 Pressure dependence of the kinetic parameters k_{cat} and K_M in the absence and presence of the osmolytes urea, TMAO and glycine.

(increase of k_{cat} and K_M)[80] and α-chymotrypsin (increase of k_{cat}).[8] These effects were generally attributed to the zwitterionic nature of glycine at physiological pH, which allows interaction with multiple parts of the protein *via* its positive and negative charged regions.[41] However, it was recently shown that cosolute-induced changes of K_M can also be merely attributed to changes of the activity coefficient of the substrate rather than to structural changes of the enzyme itself.[52,81]

High-pressure enzymology enables the determination of volume changes in the course of the catalyzed reaction.[2,14] At low substrate concentrations, v_0 is proportional to k_{cat}/K_M and the activation volume includes volume changes due to substrate binding.[8,14,20,82] If the enzyme is saturated, *i.e.* at high substrate (pyruvate) concentrations, v_0 is proportional to k_{cat} and the activation volume determined is the difference between the volume of the transition state, $ES^{\#}$, and the ES complex. Consequently, the concentration dependence of the activation volume enables estimation of the volume change ΔV_{ES} induced by the binding of the substrate. In general, the activation volume is caused by changes of the packing density and/or changes of the hydration in the active site.[2,14] The overall volume profile of the LDH-catalyzed reaction is depicted in Fig. 7. The activation volumes at high and low substrate concentration were determined according to eqn (3) using $k = k_{cat}$ and $k = k_{cat}/K_M$, respectively, and are shown in Fig. 7 for the enzyme in neat buffer solution at $T = 25$ °C. We obtain an activation volume, $\Delta V^{\#}$, of -45.3 ± 1.3 mL mol^{-1} from the data at high substrate concentrations. A negative activation volume of about -18 mL mol^{-1} was described previously for

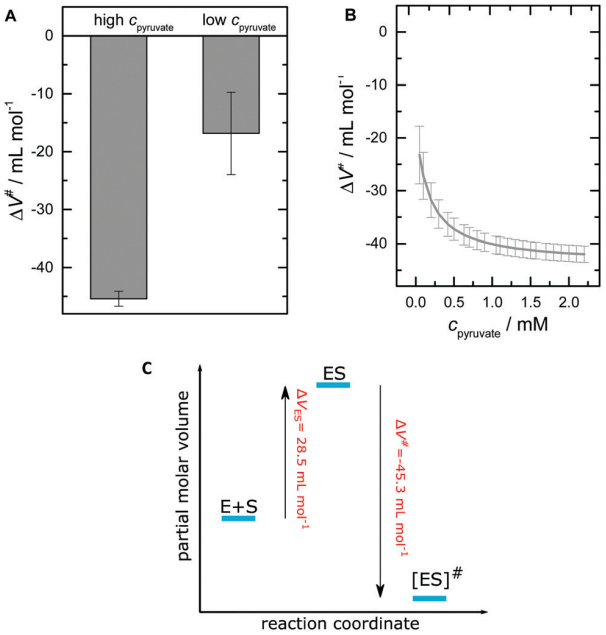


Fig. 7 (A) Activation volume, $\Delta V^{\#}$, of the LDH catalyzed interconversion of pyruvate for low and high substrate concentrations. (B) $\Delta V^{\#}$ as a function of substrate concentration for neat buffer conditions. (C) Volume profile of the LDH catalyzed reaction.

a different reaction condition (24 mM TRIS buffer, pH 8, pressure range from 1 to 612 atm, $T = 15$ °C).[83] At low substrate concentrations (Fig. 7B), the volume change is less negative

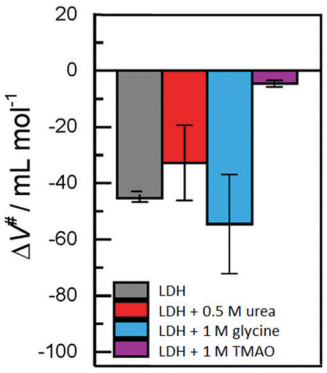


Fig. 8 Activation volume, $\Delta V^{\#}$, of the LDH catalyzed interconversion of pyruvate for high substrate concentrations in the absence and presence of several osmolytes.

and we obtain ΔV_{ES} = 28.5 mL mol^{-1}, indicating that the enzyme substrate complex has a larger volume compared to enzyme and substrate in the unbound state. The overall volume profile of the LDH-catalyzed reaction is depicted in Fig. 7C.

Fig. 8 depicts the effect of the osmolytes on the activation volume of the catalytic step. No drastic change of $\Delta V^{\#}$ is observed in the presence of the osmolytes glycine and urea. In the presence of 1 M TMAO, the activation volume is less negative, however ($\Delta V^{\#} = -4.5 \pm 1.1$ mL mol^{-1}). A similar effect has been observed upon addition of salts.[83]

Summary and conclusions

We set out to study the combined effects of pressure and cosolutes (TMAO, glycine, urea) on the activity of the tetrameric enzyme lactate dehydrogenase (LDH). To this end, high-pressure stopped-flow methodology in concert with fast UV/Vis spectroscopic detection of product formation was applied. To reveal possible pressure effects on the stability and dynamics of the enzyme, FTIR spectroscopic measurements, elastic incoherent neutron scattering (EINS) and quasi-elastic neutron scattering (QENS) measurements were carried out as well.

In neat buffer solution, the catalytic turnover number, k_{cat}, increases up to 1000 bar, the pressure range where dissociation of the tetrameric species to dimers sets in. Accordingly, we obtain a negative activation volume, $\Delta V^{\#} = -45.3 \pm 1.3$ mL mol^{-1}. Further, we could show that the enzyme substrate complex has a larger volume compared to enzyme and substrate in the unbound state, *i.e.*, the enzyme–substrate (ES) complex is destabilized by pressure (ΔV_{ES} = 28.5 mL mol^{-1}).

Whereas the magnitude of the catalytic turnover number k_{cat} is similar to the neat buffer data in the presence of the osmolytes, the pressure of deactivation is modulated by the addition of cosolvents. TMAO and glycine increase the pressure of deactivation to about 1200 bar, in accordance with the observed stabilizing effect both cosolvents have shown against denaturation and/or dissociation of proteins.[7,53] According to the high-pressure neutron scattering and FTIR spectroscopic data, no marked unfolding of the protein takes place even up to the multi-kbar regime, rather dissociation which is accompanied by a minor conformational drift between about 1000 and 3000 bar, only.

Interestingly, the pressure dependence of k_{cat} seems to vanish largely in the case of 1 M TMAO, which is reflected in a less negative $\Delta V^{\#}$ value ($\Delta V^{\#} = -4.5$ mL mol^{-1}). This may be due to the strong excluded volume effect TMAO imposes on the system. TMAO has been found to interact unfavorably with the peptide backbone. A strong force for minimizing exposure of the polypeptide backbone to this solute may result in a smaller volume of the ES complex, thereby leading to the less negative $\Delta V^{\#}$ value, rendering k_{cat} less pressure sensitive. This would also be consistent with the view of a TMAO-induced selection of a more compact conformational substate of the protein.[51,52] In fact, LDH has been shown to sample various conformational substates.[30] Such scenario would also be consistent with the decrease of the MSD upon addition of TMAO on LDH,[30] as also observed for the monomeric protein lysozyme, which results in a compaction (higher rigidity of the protein backbone) of the protein.[54]

Recently, internal motions of proteins have attracted considerable interest because of their potential role in enzyme function. Enzyme molecules undergo a range of internal motions that are also affected by the surrounding environment, including osmolytes, by temperature and by pressure. For example, slow reaction-promoting motions associated with the transition state of the protein could be enslaved by faster motions of surface residues affected by the surrounding solvent, *i.e.*, by thermodynamic coupling with the solvent.[84] The pressure dependence of the mean-squared displacement, MSD, shows that the MSD of the enzyme in neat buffer solution remains essentially constant up to 1000 bar, *i.e.*, a marked change in the internal dynamics by pressure does not seem to be responsible for the increase of k_{cat} upon compression, it is due to the smaller volume of the transition state (ES$^{\#}$). Remarkably, 1 M glycine increases the MSD significantly, but does not change k_{cat}. Within the accuracy of the data, no significant pressure effect on the MSD is observed up to 3000 bar in this case. Hence, the 1 M glycine solution seems to impose larger fast internal protein fluctuations, which have no significant effect on the turnover number. On the other hand, favorable enthalpic interactions of the zwitterionic glycine molecules with the protein interface renders the protein more stable against pressure-induced dissociation. Conversely, as expected, 0.5 M urea leads to a decrease of the dissociation pressure and hence a drastic decrease of k_{cat} already beyond $\sim$400 bar. This effect is due to the well-known effect of urea to destabilize proteins by direct interaction with the protein's sidechains and backbone. The dissociated state provides greater surface exposure of the protein and hence more favorable interactions with the cosolute.[51,52]

While urea does not markedly affect the magnitude of K_M, both, 1 TMAO and 1 M glycine exhibit smaller K_M values of about 0.07 mM and respectively 0.05 mM below about 1000 bar, *i.e.* below dissociation of the tetrameric enzyme. Such positive

effect on the substrate affinity could be rationalized by the effect of the two cosolutes on the thermodynamic activities of the reactants ($a_i = f_i c_i$, f_i = activity coefficient, c_i = concentration of species i), which enter the equilibrium constant of ES formation. The activity coefficients reflect changes in water-mediated intermolecular interactions, such as of the osmolyte with the substrate. For example, TMAO (and probably also glycine), by repulsive interactions with the substrate (corresponding to activity coefficients $f_i > 0$), seems to favor a greater affinity of the substrate to enzyme because the bound state restricts exposure of the substrate to solvent. That is, TMAO by itself would increase the activity above that in water by shifting the equilibrium toward the more compact ES form. In contrast, urea, with its different chemical character allowing easy formation of H-bonds, may weakly interact with the substrate, promotes more expanded protein conformations and thus also substrate dissociation from the enzyme. The interaction of urea, TMAO and glycine with the substrate alone may be sufficient to explain the tendency of the osmolytes to induce opposing effects on the substrate-K_M. Of note, a marked increase in catalytic efficiency (and decrease of K_M) upon addition of glycine has been observed for the α-chymotrypsin-catalyzed hydrolysis reactions.[85] Glycine, being a zwitterion, has the ability to interact with the reactants *via* both its negative and positive charged groups, and due to its rather small size it is not sterically hindered from weak binding to multiple parts of the protein.

To conclude, the data presented may also shed light on the significance of natural cosolutes (osmolytes) of different nature in the biology of adaptation. They are not only able to increase the thermodynamic stability of the enzyme, they are also able to change the energy landscape and the ability to sample conformational substates that are relevant for enzyme catalysis. Changes in the intracellular environment, *i.e.*, the solution conditions that have evolved, not changes in the structure of the enzyme *via* mutational changes, may be sufficient to modulate and control enzymatic activity under extreme environmental conditions, such as high pressure in the whole pressure range encountered by organisms on Earth ($p_{max} \approx 1000$ bar). Providing an appropriate intracellular solution might be sufficient to protect the organism also against other environmental stresses, such as dehydration, low or high temperature, or the intracellular presence of the metabolic waste product urea.[86–88]

Data accessibility

For the IN16B neutron scattering data, the DOI is https://doi.ill.fr/10.5291/ILL-DATA.8-04-809, for the IN13 neutron scattering data, the DOI is http://doi.ill.fr/10.5291/ILL-DATA.8-04-762.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

R. W. acknowledges funding from the Cluster of Excellence RESOLV (Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy, EXC-2033, Project number 39067787) and from the DFG Research Unit FOR 1979. R. W. and J. P. acknowledge also financial support from the German Academic Exchange Service (DAAD) (travel grant PROCOPE 57386797).

References

1 M. J. Eisenmenger and J. I. Reyes-De-Corcuera, *Enzyme Microb. Technol.*, 2009, **45**, 331–347.
2 *High Pressure Bioscience – Basic Concepts, Applications and Frontiers*, ed. K. Akasaka and H. Matsuki, Springer Netherlands, 2015.
3 P. Masson, in *Trends in High Pressure Bioscience and Biotechnology*, ed. R. Hayashi, Elsevier Science B.V., 2002, pp. 177–187.
4 E. Morild, *Adv. Protein Chem.*, 1981, **34**, 93–166.
5 V. V. Mozhaev, R. Lange, E. V. Kudryashova and C. Balny, *Biotechnol. Bioeng.*, 1996, **52**, 320–331.
6 S. Kara, W. S. Long, M. Berheide, S. Peper, B. Niemeyer and A. Liese, *J. Biotechnol.*, 2011, **152**, 87–92.
7 K. Julius, S. R. Al-Ayoubi, M. Paulus, M. Tolan and R. Winter, *Phys. Chem. Chem. Phys.*, 2018, **20**, 7093–7104.
8 M. W. Jaworek, V. Schuabb and R. Winter, *Chem. Commun.*, 2018, **54**, 5696–5699.
9 T. Q. Luong and R. Winter, *Phys. Chem. Chem. Phys.*, 2015, **17**, 23273–23278.
10 T. Q. Luong, N. Erwin, M. Neumann, A. Schmidt, C. Loos, V. Schmidt, M. Fändrich and R. Winter, *Angew. Chem., Int. Ed.*, 2016, **55**, 12412–12416.
11 M. Berheide, S. Peper, S. Kara, W. S. Long, S. Schenkel, M. Pohl, B. Niemeyer and A. Liese, *Biotechnol. Bioeng.*, 2010, **106**, 18–26.
12 K. Akasaka, H. Nagahata, A. Maeno and K. Sasaki, *Biophysics*, 2008, **4**, 29–32.
13 E. Decaneto, S. Suladze, C. Rosin, M. Havenith, W. Lubitz and R. Winter, *Biophys. J.*, 2015, **109**, 2371–2381.
14 C. Czeslik, T. Q. Luong and R. Winter, *MRS Bull.*, 2017, **42**, 738–742.
15 E. Ohmae, C. Murakami, S. Tate, K. Gekko, K. Hata, K. Akasaka and C. Kato, *Biochim. Biophys. Acta, Proteins Proteomics*, 2012, **1824**, 511–519.
16 C. Schuabb, N. Kumar, S. Pataraia, D. Marx and R. Winter, *Nat. Commun.*, 2017, **8**, 14661.
17 J. Marion, M. Trovaslet, N. Martinez, P. Masson, R. Schweins, F. Nachon, M. Trapp and J. Peters, *Phys. Chem. Chem. Phys.*, 2015, **17**, 3157–3163.
18 M. Gross and R. Jaenicke, *Eur. J. Biochem.*, 1994, **221**, 617–630.
19 J. L. Silva, D. Foguel and C. A. Royer, *Trends Biochem. Sci.*, 2001, **26**, 612–618.
20 B. B. Boonyaratanakornkit, C. B. Park and D. S. Clark, *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, 2002, **1595**, 235–249.

21 G. Weber, *Biochemistry*, 1986, **25**, 3626–3631.
22 J. L. Silva and G. Weber, *Annu. Rev. Phys. Chem.*, 1993, **44**, 89–113.
23 J. L. Silva, D. Foguel, A. T. Da Poian and P. E. Prevelige, *Curr. Opin. Struct. Biol.*, 1996, **6**, 166–175.
24 X. Peng, J. Jonas and J. L. Silva, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 1776–1780.
25 V. V. Mozhaev, K. Heremans, J. Frank, P. Masson and C. Balny, *Proteins*, 1996, **24**, 81–91.
26 X. Peng, J. Jonas and J. L. Silva, *Biochemistry*, 1994, **33**, 8323–8329.
27 D. Knorr, V. Heinz and R. Buckow, *Biochim. Biophys. Acta, Proteins Proteomics*, 2006, **1764**, 619–631.
28 K. Swiderek, I. Tuñón, S. Martí and V. Moliner, *ACS Catal.*, 2015, **5**, 1172–1185.
29 P. Kędzierski, K. Moreton, A. R. Clarke and J. J. Holbrook, *Biochemistry*, 2001, **40**, 7247–7252.
30 N. Zhadin and R. Callender, *Biochemistry*, 2011, **50**, 1582–1589.
31 S. McClendon, N. Zhadin and R. Callender, *Biophys. J.*, 2005, **89**, 2024–2032.
32 L. Qiu, M. Gulotta and R. Callender, *Biophys. J.*, 2007, **93**, 1677–1686.
33 R. Callender and R. B. Dyer, *Chem. Rev.*, 2006, **106**, 3031–3042.
34 J. J. Holbrook and H. Gutfreund, *FEBS Lett.*, 1973, **31**, 157–169.
35 J. R. E. T. Pineda, R. Callender and S. D. Schwartz, *Biophys. J.*, 2007, **93**, 1474–1483.
36 R. K. Schmidt and J. E. Gready, *J. Mol. Model.*, 1999, **5**, 153–168.
37 A. R. Clarke, T. Atkinson and J. J. Holbrook, *Trends Biochem. Sci.*, 1989, **14**, 101–105.
38 F. Farabegoli, M. Vettraino, M. Manerba, L. Fiume, M. Roberti and G. Di Stefano, *Eur. J. Pharm. Sci.*, 2012, **47**, 729–738.
39 Z. Zhao, F. Han, S. Yang, J. Wu and W. Zhan, *Cancer Lett.*, 2015, **358**, 17–26.
40 L. King and G. Weber, *Biophys. J.*, 1986, **49**, 72–73.
41 B. C. Schade, R. Rudolph, H. D. Luedemann and R. Jaenicke, *Biochemistry*, 1980, **19**, 1121–1126.
42 G. Schmid, H.-D. Ludemann and R. Jaenicke, *Eur. J. Biochem.*, 1979, **97**, 407–413.
43 L. King and G. Weber, *Biochemistry*, 1986, **25**, 3637–3640.
44 T. Fujisawa, M. Kato and Y. Inoko, *Biochemistry*, 1999, **38**, 6411–6418.
45 R. Winter, D. Lopes, S. Grudzielanek and K. Vogtt, *J. Non-Equilib. Thermodyn.*, 2007, **32**, 41–97.
46 I. Daniel, P. Oger and R. Winter, *Chem. Soc. Rev.*, 2006, **35**, 858–875.
47 P. H. Yancey and J. F. Siebenaller, *J. Exp. Biol.*, 2015, **218**, 1880–1896.
48 P. H. Yancey, *J. Exp. Biol.*, 2005, **208**, 2819–2830.
49 P. H. Yancey, W. R. Blake and J. Conley, *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.*, 2002, **133**, 667–676.
50 M. A. Schroer, Y. Zhai, D. C. F. Wieland, C. J. Sahle, J. Nase, M. Paulus, M. Tolan and R. Winter, *Angew. Chem., Int. Ed.*, 2011, **50**, 11413–11416.
51 D. R. Canchi and A. E. García, *Annu. Rev. Phys. Chem.*, 2013, **64**, 273–293.
52 M. Gao, C. Held, S. Patra, L. Arns, G. Sadowski and R. Winter, *ChemPhysChem*, 2017, **18**, 2951–2972.
53 L. Arns, V. Schuabb, S. Meichsner, M. Berghaus and R. Winter, *Z. Phys. Chem.*, 2017, **232**, 615–634.
54 S. R. Al-Ayoubi, P. H. Schummel, M. Golub, J. Peters and R. Winter, *Phys. Chem. Chem. Phys.*, 2017, **19**, 14230–14237.
55 A. K. Covington, M. Paabo, R. A. Robinson and R. G. Bates, *Anal. Chem.*, 1968, **40**, 700–706.
56 N. Francesca, J. Peters, D. Russo, S. Barbieri, C. Chiapponi, A. Cupane, A. Deriu, M. T. Di Bari, E. Farhi, Y. Gerelli, P. Mariani, A. Paciaroni, C. Rivasseau, G. Schirò and F. Sonvico, *Neutron News*, 2008, **19**, 14–18.
57 B. Frick, E. Mamontov, L. van Eijck and T. Seydel, *Z. Phys. Chem.*, 2010, **224**, 33–60.
58 M. Hennig, B. Frick and T. Seydel, *J. Appl. Crystallogr.*, 2011, **44**, 467–472.
59 J. Peters, M. Trapp, D. Hughes, S. Rowe, B. Demé, J.-L. Laborier, C. Payre, J.-P. Gonzales, S. Baudoin, N. Belkhier and E. Lelièvre-Berna, *High Press. Res.*, 2012, **32**, 97–102.
60 E. Lelièvre-Berna, B. Demé, J. Gonthier, J.-P. Gonzales, J. Maurice, Y. Memphis, C. Payre, P. Oger, J. Peters and S. Vial, *J. Neutron Res.*, 2017, **19**, 77–84.
61 D. Richard, M. Ferrand and G. J. Kearley, *J. Neutron Res.*, 1996, **4**, 33–39.
62 A. Rahman, K. S. Singwi and A. Sjölander, *Phys. Rev.*, 1962, **126**, 986–996.
63 M. Grimaldo, F. Roosen-Runge, F. Zhang, T. Seydel and F. Schreiber, *J. Phys. Chem. B*, 2014, **118**, 7203–7209.
64 M. Grimaldo, F. Roosen-Runge, N. Jalarvo, M. Zamponi, F. Zanini, M. Hennig, F. Zhang, F. Schreiber and T. Seydel, *EPJ Web Conf.*, 2015, **83**, 02005.
65 O. Arnold, *et al.*, *Nucl. Instrum. Methods Phys. Res., Sect. A*, 2014, **764**, 156–166.
66 F. Volino and A. J. Dianoux, *Mol. Phys.*, 1980, **41**, 271–279.
67 A. Ortega, D. Amorós and J. García de la Torre, *Biophys. J.*, 2011, **101**, 892–898.
68 C. Beck, M. Grimaldo, F. Roosen-Runge, M. K. Braun, F. Zhang, F. Schreiber and T. Seydel, *J. Phys. Chem. B*, 2018, **122**, 8343–8350.
69 G. Panick, R. Malessa, R. Winter, G. Rapp, K. J. Frye and C. Royer, *J. Mol. Biol.*, 1998, **275**, 389–402.
70 G. Zaccai, *Science*, 2000, **288**, 1604–1607.
71 F. Meersman and K. Heremans, *Biochemistry*, 2003, **42**, 14234–14241.
72 Y. Cordeiro, J. Kraineva, M. C. Suarez, A. G. Tempesta, J. W. Kelly, J. L. Silva, R. Winter and D. Foguel, *Biophys. J.*, 2006, **91**, 957–967.
73 R. Ludwig, *Angew. Chem., Int. Ed.*, 2001, **40**, 1808–1827.
74 H. Weingärtner, E. U. Franck, G. Wiegand, N. Dahmen, F. H. Frimmel, B. C. Gordalla, K. Johannsen, R. S. Summers, W. Höll, M. Jekel, R. Gimbel, R. Rautenbach and W. H. Glaze, in *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH, Weinheim, 1996, pp. 1–101.
75 A. Cortes, D. C. Emery, D. J. Halsall, R. M. Jackson, A. R. Clarke and J. John Holbrook, *Protein Sci.*, 1992, **1**, 892–901.

76 P. H. Yancey and G. N. Somero, *J. Comp. Physiol., B*, 1978, **125**, 135–141.
77 K. Rajagopalan, I. Fridovich and P. Handler, *J. Biol. Chem.*, 1961, **236**, 1059–1065.
78 R. L. Joseph, D. C. Watts and E. Baldwin, *Biochem. J.*, 1963, **87**, 416–422.
79 T. Mashino and I. Fridovich, *Arch. Biochem. Biophys.*, 1987, **258**, 356–360.
80 V. Vathipadiekal, A. Verma and M. Rao, *Biol. Chem.*, 2007, **388**, 61–65.
81 A. Wangler, R. Canales, C. Held, T. Q. Luong, R. Winter, D. H. Zaitsau, S. P. Verevkin and G. Sadowski, *Phys. Chem. Chem. Phys.*, 2018, **20**, 11317–11326.
82 P. Butz, K. O. Greulich and H. Ludwig, *Biochemistry*, 1988, **27**, 1556–1563.
83 P. S. Low and G. N. Somero, *Proc. Natl. Acad. Sci. U. S. A.*, 1975, **72**, 3014–3018.
84 P. K. Agarwal, *Biochemistry*, 2019, **58**, 438–449.
85 M. W. Jaworek, V. Schuabb and R. Winter, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1347–1354.
86 I. Baskakov, A. Wang and D. W. Bolen, *Biopyhs. J.*, 1998, **74**, 2666–2673.
87 P. H. Schummel, M. W. Jaworek, C. Rosin, J. Högg and R. Winter, *Phys. Chem. Chem. Phys.*, 2018, **20**, 28400–28411.
88 P. H. Schummel, C. Anders, M. W. Jaworek and R. Winter, *ChemPhysChem*, 2019, **20**, 1098–1109.

Journal of Thermal Analysis and Calorimetry (2021) 145:403–409
https://doi.org/10.1007/s10973-020-09775-9

PyDSC: a simple tool to treat differential scanning calorimetry data

Aline Cisse[1,2] · Judith Peters[1,2] · Giuseppe Lazzara[3] · Leonardo Chiappisi[1,4]

Received: 12 December 2019 / Accepted: 30 April 2020 / Published online: 16 May 2020

Abstract
Herein, we describe an open-source, Python-based, script to treat the output of differential scanning calorimetry (DSC) experiments, called *pyDSC*, available free of charge for download at https://github.com/leonardo-chiappisi/pyDSC under a GNU General Public License v3.0. The main aim of this program is to provide the community with a simple program to analyze raw DSC data. Key features include the correction from spurious signals, and, most importantly, the baseline is computed with a robust, physically consistent approach. We also show that the baseline correction routine implemented in the script is significantly more reproducible than different standard ones proposed by proprietary instrument control software provided with the microcalorimeter used in this work. Finally, the program can be easily applied to large amount of data, improving the reliability and reproducibility of DSC experiments.

Keywords DSC · Baseline correction · Python · Phase transition · Protein conformation · Polymer stability

Introduction

Differential scanning calorimetry (DSC) is a powerful thermo-analytical technique which detects heat changes associated with physical and chemical transformation in biological and non-biological samples. Due to the simplicity of the technique, the relatively low-cost of the apparatus, the ease of data analysis, DSC found wide application in very diverse fields of both an academic and industrial research activities. Several excellent reviews covering the use of DSC can be found in the literature. With focus on colloidal and biophysical science, we address the reader to some covering topics such as protein conformation [16, 18, 31], lipid phase transitions [13], (bio) polymer stability [26], and on mixed systems [17, 25, 30].

✉ Leonardo Chiappisi
chiappisil@ill.eu

1 Institut Max von Laue - Paul Langevin, 71 avenue des Martyrs, 38042 Grenoble, France

2 Laboratoire Interdisciplinaire de Physique, Univ. Grenoble Alpes, CNRS, 140 Rue de la Physique, 38000 Grenoble, France

3 Dipartimento di Fisica e Chimica, Università degli Studi di Palermo, Viale delle Scienze pad 17, 90128 Palermo, Italy

4 Stranski Laboratorium für Physikalische und Theoretische Chemie, Institut für Chemie, Technische Universität Berlin, Strasse des 17. Juni 124, 10623 Berlin, Germany

In many cases, a DSC experiment is used to extract the enthalpy change ΔH of the studied process and the temperature of the transition. This analysis is straightforward, especially when the data exhibit a good signal-to-noise ratio. However, a DSC curve usually contains an abundance of information, which can be extracted by an in-depth analysis [7, 11, 12, 20, 24, 28, 32].

For every, even simple, analysis of a DSC curve, a correct evaluation of the baseline represents a critical step, whose importance is a matter of discussion since the origins of the technique [4, 14, 22, 27]. However, the baseline correction is mostly performed in a semiautomatic procedure by the proprietary software provided with the majority of commercial calorimeters. In fact, the correction consists in a spline interpolation between two arbitrary chosen regions before and after the transition region. While the procedure delivers correct results in the majority of cases, it is prone to artifacts—in particular, when the signal is close to the limits of the investigated temperature range and/or if a relevant heat capacity change occurs before and after the transition. In addition, most correction procedures are not based on a physical consistent model.

The baseline of a DSC curve reflects the difference between the heat capacity of the sample and of the reference. Provided spurious contributions to the signal from the empty cells and from a mismatch of buffer amount in sample and reference, besides the asymmetries of the measuring system, are taken into account, the baseline reflects the apparent heat

capacity of the investigated compound. In the approximation of a two step transition, e.g., from a state A to a state B, the apparent heat capacity as a function of temperature T ($C_p(T)$) is given by:

$$C_P(T) = \alpha(T)C_P^B + (1-\alpha)C_P^A \quad (1)$$

with $\alpha(T)$ being the degree of conversion of the process as a function of temperature; C_P^A and C_P^B are the apparent heat capacities of the compound in state A and B, respectively. C_P^A and C_P^B will be considered to be temperature independent within the investigated range in this work.

A physical consistent baseline subtraction, associated with a correction of residual signals, is simple steps which may significantly improve the data quality and the amount of information which can be extracted from the experiment. Herein, we describe a simple, freely accessible and modifiable tool, *pyDSC*, to treat DSC experiments. The program, which will be constantly updated, can be freely downloaded from the GitHub repository https://github.com/leonardo-chiappisi/pyDSC. At the same link, the up-to-date manual of the program can be found.

In the following, we provide a few experimental notes on a typical DSC experiment, a concise description of the operations performed by *pyDSC*, together with some examples.

Materials and methods

The DSC curves used to describe the functionalities of pyDSC were recorded on a micro-DSC III from Setaram, France. The instrument is provided with the SetSoft2000 software suite. To test the program, we have investigated a 8 mass% solution of the EO_{26}–PO_{40}–EO_{26} triblock copolymer Pluronic P85, bought from Sigma-Aldrich. The solutions were prepared using water of Millipore grade. Furthermore, *pyDSC* was applied to analyze DSC curves recorded on a Nano DSC, on a Multicell DSC and on a DSC 2920 from TA-Instruments.

Experimental

In order to work with the apparent heat capacity of the investigated compound, the DSC raw data need to be corrected for the contribution of the empty cells and of the buffer. Accordingly, three sets of measurements can be provided to the script:

- The standard sample versus buffer run. For a proper normalization of the data, information such as the mass of sample and the concentration of the investigated compound needs to be known. Moreover, to minimize the contribution of the buffer to the baseline, the mass of buffer in the sample cell and in the reference cell should be as close as possible. We define the difference of mass of buffer in the sample cell and in the reference cell as Δm_{BS}.
- The buffer versus buffer run. This run is used to determine the heat capacity of the buffer solution and will be used to correct the DSC data for the contribution of the buffer to the recorded signal. In contrast to the sample versus buffer run, where the cells are filled with possibly the exact amount of sample, in this run the amount of buffer in the two cells should differ, in order to get a good signal-to-noise. We recommend that the difference in mass between the two buffer cells, Δm_{BB}, is about 10% of the mass of buffer in the cell.
- The empty cell versus empty cell run. This run is used to correct both the sample versus buffer and the buffer versus buffer run with the contribution arising from the different heat capacities of the empty cells and electronic asymmetries. In order to be effective, we recommend to use always the same reference and sample cells for all the measurements.

Program description

The script is written in Python 3, it makes use of the SciPy [19] and matplotlib [15] packages and is platform independent. Several operations on the data are performed, provided the required information is available. The following operations are executed on the raw data files:

1. *Data averaging and resizing* The DSC raw data of the sample and, if provided, of the empty cell and buffer-buffer runs are read by the program and the points which fall outside the temperature region of interest are discarded. The reference measurements are averaged. If requested by the user, in order to facilitate any further data treatment, the number of points is reduced by a factor N, defined by the user.
2. *Heat flow conversion in heat capacity* The heat flow is converted into a heat capacity upon division by the heating rate. The heating rate is either provided by the user, or it is determined from the raw data file if the time information is contained. The raw heat flows are plotted in the rawdata.pdf file while the heat flows normalized by the heating rate and the sample amount are plotted in corrected_data.pdf.

3. *Empty cell and buffer correction* If the data file corresponding to the empty cell/empty cell and/or buffer/buffer runs are provided, the raw data are corrected for both contributions. In particular, the empty cell/empty cell run is used to take into account any non-constant contribution arising from the empty cells or the instrument. The buffer/buffer run is used to take into account any contribution arising from a mismatch in amount of buffer contained in the reference and in the sample cell. In detail, the apparent heat capacity of the sample run Cp_s, is corrected for the apparent heat capacity of the empty cell Cp_{EC} and for the apparent heat capacity of the buffer/buffer run Cp_{bb} as:

$$\overline{Cp_s} = Cp_s - Cp_{EC} - \frac{\Delta m_{BS}(Cp_{bb} - Cp_{EC})}{\Delta m_{BB}} \quad (2)$$

with Δm_{BS} and Δm_{BB} being the mass difference between the buffer contained in the sample and in the reference in the sample run, and between the buffers in the buffer/buffer run. If no empty cell/empty cell or buffer/buffer runs are provided, this correction will not be performed and the non-corrected sample run will be further analyzed. In most of the cases, Δm_{BS} should be so small that the last term might be negligible. For instruments working with non-removable sample cells with constant filling volume, i.e., $\Delta m_{BB} = 0$, by loading in *pyDSC* a buffer-buffer titration instead of the empty cell/empty cell run.

4. *Normalization of apparent heat flow* The apparent heat flow of the sample run is finally normalized by the amount of solute. In particular, the apparent heat capacity is normalized either by the mass of solute, resulting in a Cp_s given in J K^{-1} g^{-1}, or by the moles of solute, if the molecular weight is known, thus resulting in a Cp_s given in J K^{-1} mol^{-1}.

5. *Baseline correction* The core task carried out by the program is to compute a physically meaningful baseline, as described earlier in the text. The baseline is constructed following an iterative procedure proposed elsewhere [22, 29]. In detail, the baseline $CP_{bl}(T)$ is defined as follows:

$$CP_{bl}(T) = (1 - \alpha)CP_{bl}^{pre}(T) + \alpha \cdot CP_{bl}^{post}(T) \quad (3)$$

with $CP_{bl}^{pre}(T)$ and $CP_{bl}^{post}(T)$ are the baselines determined in the regions at lower and higher temperatures than the peak, respectively. α is the degree of conversion, defined as:

$$\alpha(T) = \frac{\int_0^T CP_s(T) - CP_{bl}(T)dT}{\int_0^\infty CP_s(T) - CP_{bl}(T)dT} \quad (4)$$

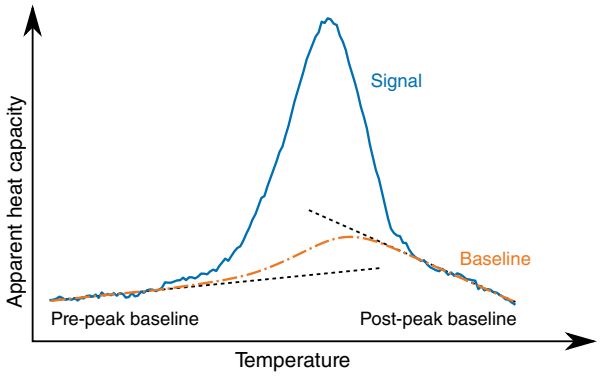


Fig. 1 Schematic diagram describing the baseline construction. The region before and after the peak are used to calculate the corresponding baselines. See text for further details

where $CP_s(T)$ is the apparent heat capacity difference between sample and reference, after normalization and eventual correction. $CP_{bl}^{pre}(T)$ and $CP_{bl}^{post}(T)$ represent the temperature-dependent difference of heat capacity of sample and reference, eventually corrected from the contribution from the empty cells and solvent mismatch. They are assumed to exhibit a linear dependence from temperature. In Fig. 1, the baseline construction is schematically depicted.

At this stage of the program, the enthalpy change of the process, computed as $\Delta H = \int_0^\infty CP_s(T) - CP_{bl}(T)\,dT$, is evaluated and its uncertainty is estimated. The uncertainty in ΔH is estimated from the stochastic fluctuations of the signal in the regions where $CP_{bl}^{pre}(T)$ and $CP_{bl}^{post}(T)$ are determined. By definition, in these regions no processes take place, and the signal fluctuates around an apparent heat capacity of zero J K^{-1} is used to determine the standard deviation of the heat capacity. The heat exchanged during the examined process is computed by integration using the trapezoidal rule, and its uncertainty by standard error propagation.

Program usage

pyDSC performs all the aforementioned steps on the DSC raw data provided by the user, located in the folder `raw-data`. In addition, the user will provide all the necessary information to perform the corrections within the `Files.txt` and `Input_params.txt` text files. Full details are provided in the user manual and exemplary files at https://github.com/leonardo-chiappisi/DSC_correction. Finally, the pyDSC script can be executed within a terminal by invoking the Python3 interpreter as:

```
python3 pyDSC.py
```

In addition to an output ASCII file for each sample run, the script will generate five output figures: Alpha.pdf, Baseline_data.pdf, Corrected_data.pdf, Final_data.pdf, and Rawdata.pdf.

In `Rawdata.pdf`, the recorded heat flows of the different runs are reported as a function of temperature.

In `Corrected_data.pdf`, the sample run heat capacity, after binning and normalization by sample amount, is reported as a function of temperature.

In `Baseline_data.pdf`, the heat capacity of each sample run, after binning, normalization by sample amount, and empty cell and buffer correction, is reported, together with the calculated baseline as a function of temperature. Moreover, the regions used for the determination of the baseline before and after the peak are highlighted.

In `Final_data.pdf`, the heat capacity of each sample run, after binning, normalization by sample amount, empty cell and buffer correction, and baseline subtraction, is reported as a function of temperature.

Finally, in `alpha.pdf`, the degree of conversion determined for each sample run is reported as a function of temperature.

The different plots allow the user to determine whether the correction procedures executed during each step are performed correctly and keep control over data treatment process.

In addition to the five figures, one ASCII file for each sample run is created. The file contains a summary of the information used to treat the data in the header, in addition to the computed values of enthalpy change and change in heat capacity at the peak temperature. Moreover, the file contains four data columns: temperature, the normalized, corrected, and binned heat capacity with baseline subtracted, without baseline subtraction, the baseline, and the degree of conversion of the process.

Results and discussion

The script is tested using a semidiluted solution of the triblock copolymer poly-(ethyleneoxide)–poly-(propyleneoxide)–poly-(ethyleneoxide) Pluronic P85 (EO_{26}–PO_{40}–EO_{26}). DSC was extensively employed to reveal the self-assembly behavior of this class of polymers in solutions [1–3, 10, 21], and the system was chosen due to its good signal-to-noise ratio, non-negligible baseline dependence from the temperature, and the large availability of reference data to compare with.

The proprietary software commonly used to perform a first analysis of the microcalorimetric data provide different routines for the baseline subtraction and the blank correction, which typically consists in a simple subtraction of the raw heat flow data. For what concerns the baseline evaluation and subtraction, different algorithms are available, depending on the vendor and on the software version. However, in all cases, the baseline under the transition peak is evaluated from, somehow arbitrary chosen, regions prior and after the process signal.

To test the robustness of the *pyDSC* script, the enthalpy of micellization recorded for a solution of the triblock copolymer EO_{26}–PO_{40}–EO_{26} was evaluated using *pyDSC* and two different algorithms provided by the Setaram data analysis software: "linear"—which uses a line from the starting to the end point arbitrarily selected by the user—and "curve"—which generates a spline curve from the selected initial/final points that are selected by user—using different regions for the baseline evaluation. For the sake of the comparison, we also report data analyzed with *pyDSC* without the correction for the buffer and empty cell contributions. The results are shown in Table 1 and two clear results stand out: the arbitrary choice of the baseline region has a non-negligible effect on the obtained values of ΔH_m, with a variation of approx. 2 and 6% for the linear and the curve algorithm, respectively, versus only 0.7% found with *pyDSC*. The effect is even smaller when the contribution of the solvent and empty cells are taken into account. Furthermore, trending values are observed with a systematic variation of the obtained enthalpy change for linear integration. The DSC

Table 1 Micellization enthalpy ΔH_m of a EO_{26}–PO_{40}–EO_{26} triblock copolymer in aqueous solution, given in kJ mol^{-1}

Peak temperature interval	Linear	Curve	pyDSC	pyDSC[a]
18–50 °C	233.4	222.9	216.0 ± 0.1	214.7 ± 0.1
17–51 °C	237.5	221.3	216.7 ± 0.1	217.0 ± 0.1
16–52 °C	242.4	250.5	217.3 ± 0.1	218.0 ± 0.1
15–53 °C	245.6	242.7	217.4 ± 0.1	217.5 ± 0.1
Average/kJ mol^{-1}	239.7	234.3	216.8	216.8
Standard deviation kJ mol^{-1}	5.4	14.5	0.7	1.5

All values refer to the same heating curve, analyzed in the temperature interval between 10 and 60 °C, using different, arbitrarily chosen temperature intervals to define the transition peak. The ΔH_m values were obtained using the "linear" and "curve" algorithm provided by the Setaram control software SetSoft2000 and by *pyDSC*, described in this work

a: Data without empty cell and buffer correction

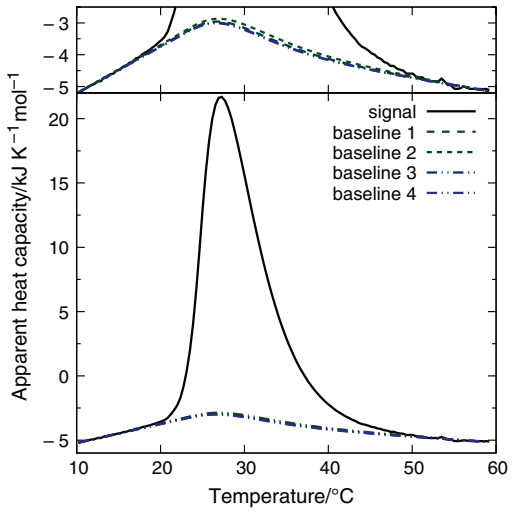


Fig. 2 Apparent heat capacity recorded for a 8 mass% aqueous solution of EO_{26}–PO_{40}–EO_{26} during a heating ramp at 0.5 °C min^{-1}, with the corresponding baselines, determined assuming a peak temperature interval of 18–50 °C (baseline 1), 17–51 °C (baseline 2), 16–52 °C (baseline 3), 15–53 °C (baseline 4). The top insert shows a magnification of the baseline region

curves with the corresponding four baselines are shown in Fig. 2, further demonstrating that the arbitrary choice of the integration region as only a minimal effect on the results.

pyDSC was also applied to correct the calorimetric output from DSC experiments performed on very diverse systems on other instruments. In particular, in Fig. 3, *pyDSC* was applied for the correction of data stemming from the micro-DSC III from Setaram and from the DSC 2920, nano DSC, and Multicell DSC from TA-Instruments. The curves refer to the melting of isopropylpalmitate in a mixture of surfactant and alcohols [8], to the melting of Indium, to the thermal transition within self-assembled fibrils [5], and to the thermal denaturation of a protein [23] (See figure caption and corresponding references for further details).

Finally, it is useful to recall that the trustworthiness of the microcalorimetric data does not depend only on the data correction and baseline subtraction procedures, but a regular instrument maintenance and calibration are equally essential [6, 9].

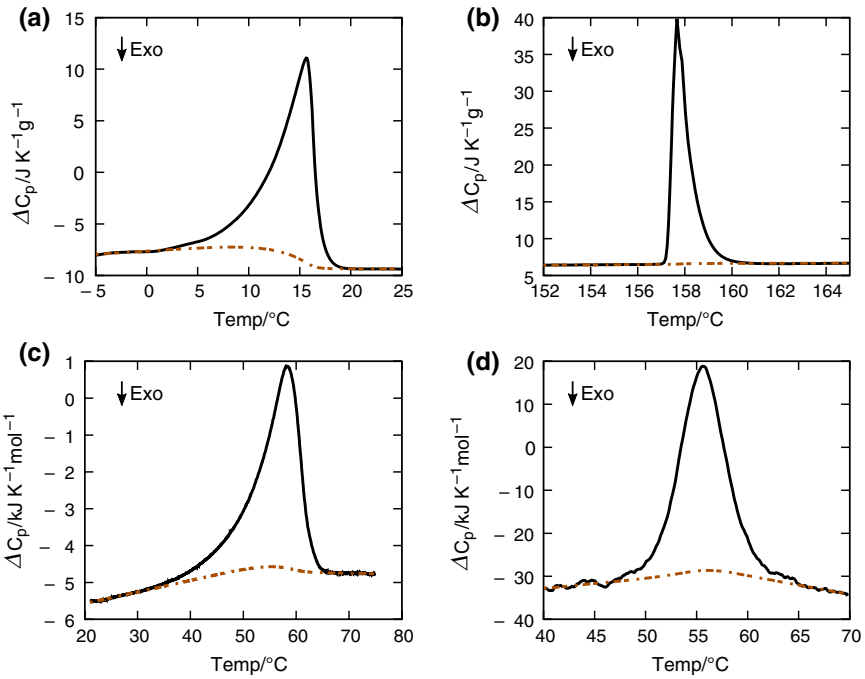


Fig. 3 Raw heat capacities (full lines) and calculated baselines (dashed lines) for different DSC curves. In particular, **a** the melting of Isopropylpalmitate in the presence of butanol and a polyoxyethylene glycol oleic ether nonionic surfactant, recorded on a multi-cell DSC from TA-Instruments (see Ref. [8] for full details on the system); **b** the melting of Indium, recorded on a DSC 2920 from TA-Instruments; **c** the melting of fibrils made from the supramolecular polymerization of a π-donor π-acceptor monomer, recorded on a Setaram micro-DSC III (see Ref. [5] for full details on the system); **d** thermal denaturation of the orange carotenoid protein recorded on a nano DSC from TA-Instruments (see Ref. [23] for full details on the system)

Conclusions

In this contribution, we present *pyDSC* a simple, python-based script to treat differential scanning calorimetry data. The software is platform independent and freely usable. It offers a simple but robust and physically meaningful approach to evaluate the baseline in DSC experiments. Moreover, the raw DSC data can be corrected for the contribution of the empty cells and the buffer, and their size reduced without loss of quality.

We anticipate that all these features will increase the robustness and reproducibility of the output of DSC experiments across very different communities. In addition, we commit ourselves to keep the software up to date, and ask the scientific community for feedback and suggestions for further implementations. The source code is available free of charge under a GNU General Public License v3.0 (GPLv3).

Acknowledgements Open Access funding provided by Projekt DEAL. LC acknowledges the TU-Berlin and the ILL for postdoctoral funding though a three-year cooperation agreement. AC acknowledges the CFM foundation for funding her doctoral thesis. The partnership for soft condensed matter (PSCM) at the ILL Grenoble is acknowledged for providing the differential scanning calorimeter (DSC).

References

1. Alexandridis P, Hatton TA. Poly(ethylene oxide)–poly(propylene oxide)–poly(ethylene oxide) block-copolymer surfactants in aqueous-solutions and at interfaces—thermodynamics, structure, dynamics, and modeling. Colloids Surf A Physicochem Eng Asp. 1995;96(1–2):1–46.
2. Alexandridis P, Holzwarth JF, Hatton TA. Micellization of poly(ethylene oxide)–poly(propylene oxide)–poly(ethylene oxide) triblock copolymers in aqueous solutions: thermodynamics of copolymer association. Macromolecules. 1994;27(9):2414–25.
3. Barba AA, D'Amore M, Grassi M, Chirico S, Lamberti G, Titomanlio G. Investigation of Pluronic F127-water solutions phase transitions by DSC and dielectric spectroscopy. J Appl Polym Sci. 2009;114(2):688–95.
4. Brennan WP, Miller B, Whitwell JC. An improved method of analyzing curves in differential scanning calorimetry. Ind Eng Chem Fundam. 1969;8(2):314–8.
5. Cera L, Chiappisi L, Böttcher C, Schulz A, Schoder S, Gradzielski M, Schalley CA. PolyWhips: directional particle transport by gradient-directed growth and stiffening of supramolecular assemblies. Adv Mater. 2017;29(8):1604430.
6. Cerdeiriña CA, Míguez JA, Carballo E, Tovar CA, De La Puente E, Romaní L. Highly precise determination of the heat capacity of liquids by DSC: calibration and measurement. Thermochim Acta. 2000;347(1–2):37–44.
7. Chiappisi L, Lazzara G, Gradzielski M, Milioto S. Quantitative description of temperature induced self-aggregation thermograms determined by differential scanning calorimetry. Langmuir. 2012;28(51):17609–16.
8. Chiappisi L, Noirez L, Gradzielski M. A journey through the phase diagram of a pharmaceutically relevant microemulsion system. J Colloid Interface Sci. 2016;473:52–9.
9. Della Gatta G, Richardson MJ, Sarge SM, Stølen S. Standards, calibration, and guidelines in microcalorimetry. Part 2. Calibration standards for differential scanning calorimetry* (IUPAC Technical Report). Pure Appl Chem. 2006;78(7):1455–76.
10. Dintcheva NT, Catalano G, Arrigo R, Morici E, Cavallaro G, Lazzara G, Bruno M. Pluronic nanoparticles as anti-oxidant carriers for polymers. Polym Degrad Stabil. 2016;134:194–201.
11. Dlubek G, Pompe G, Pionteck J, Janke A, Kilburn D. Differential scanning calorimetry (DSC) for interdiffusion studies in PCV/PnBMA blend: a quantitative analysis. Macromol Chem Phys. 2003;204(9):1234–44.
12. Fessas D, Iametti S, Schiraldi A, Bonomi F. Thermal unfolding of monomeric and dimeric β-lactoglobulins. Eur J Biochem. 2001;268(20):5439–48.
13. Heerklotz H. The microcalorimetry of lipid membranes. J Phys Condens Matter. 2004;16(15):R441–67.
14. Höhne GWH, Hemminger WF, Flammersheim HJ. Differential scanning calorimetry. 2nd ed. Heidelberg: Springer; 2003.
15. Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng. 2007;9(3):90–5.
16. Ibarra-Molero B, Naganathan AN, Sanchez-Ruiz JM, Muñoz V. Modern analysis of protein folding by differential scanning calorimetry. In: Feig AL, editor. Methods in enzymology—calorimetry. Amstedam: Academic Press/Elsevier; 2016. p. 281–318.
17. Jelesarov I, Bosshard HR. Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. J Mol Recognit. 1999;12(1):3–18.
18. Johnson CM. Differential scanning calorimetry as a tool for protein folding and stability. Arch Biochem Biophys. 2013;531(1–2):100–9.
19. Jones E, Oliphant T, Paterson P. SciPy: open source scientific tools for Python
20. Leharne S. The physical chemistry of high-sensitivity differential scanning calorimetry of biopolymers. ChemTexts. 2017;3(1):1–12.
21. Lopes JR, Loh W. Investigation of self-assembly and micelle polarity for a wide range of ethylene oxidepropylene oxideethylene oxide block copolymers in water. Langmuir. 1998;14(4):750–6.
22. Malakhov DV, Abou Khatwa MK. Constructing a self-consistent integral baseline by using cubic splines. J Therm Anal Calorim. 2007;87(2):595–9.
23. Moldenhauer M, Sluchanko NN, Tavraz NN, Junghans C, Buhrke D, Willoweit M, Chiappisi L, Schmitt FJ, Vukojević V, Shirshin EA, Ponomarev VY, Paschenko VZ, Gradzielski M, Maksimov EG, Friedrich T. Interaction of the signaling state analog and the apoprotein form of the orange carotenoid protein with the fluorescence recovery protein. Photosynth Res. 2018;135(1–3):125–39.
24. Paterson IF, Armstrong JK, Chowdhry BBZ, Leharne SA. Thermodynamic model fitting of the calorimetric output obtained for aqueous solutions of oxyethyleneoxypropyleneoxyethylene triblock copolymers. Langmuir. 1997;13(8):2219–26.

25. Prenner E, Chiu M. Differential scanning calorimetry: an invaluable tool for a detailed thermodynamic characterization of macromolecules and their interactions. J Pharm Bioallied Sci. 2011;3(1):39.
26. Privalov PL, Dragan AI. Microcalorimetry of biological macromolecules. Biophys Chem. 2007;126(1–3):16–24.
27. Sandu C, Singh RK. Modeling differential scanning calorimetry. Thermochim Acta. 1990;159:267–98.
28. Schulthess T, Schönfeld HJ, Seelig J. Thermal unfolding of apolipoprotein A—1. Evaluation of methods and models. Biochemistry. 2015;54(19):3063–75.
29. Van der Plaats G. A theoretical evaluation of a heat-flow differential scanning calorimeter. Thermochim Acta. 1984;12:77–82.
30. Weber PC, Salemme FR. Applications of calorimetric methods to drug discovery and the study of protein interactions. Curr Opin Struct Biol. 2003;13(1):115–21.
31. Yang Z, Brouillette CG. A guide to differential scanning calorimetry of membrane and soluble proteins in detergents; 2016. pp. 319–58
32. Zehender F, Ziegler A, Schönfeld HJ, Seelig J. Thermodynamics of protein self-association and unfolding. The case of apolipoprotein A—I. Biochemistry. 2012;51(6):1269–80.

THE JOURNAL OF PHYSICAL CHEMISTRY B
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCB

Article

Investigation of the Action of Peptides on Lipid Membranes

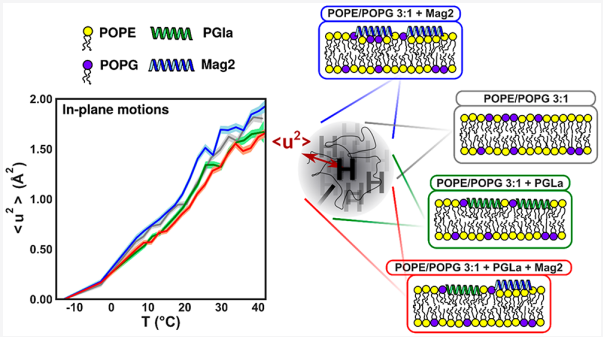
Aline Cisse,[#] Arnaud Marquette,[#] Munkhtuguldur Altangerel, Judith Peters,[*] and Burkhard Bechinger

Cite This: https://doi.org/10.1021/acs.jpcb.1c06388

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Calorimetric and incoherent neutron scattering methods were employed to investigate the action of magainin 2 and PGLa peptides on the phase behavior and molecular dynamics of lipids mimicking cytoplasmic membranes of Gram-negative bacteria. The impact of the peptides, tested individually and cooperatively by differential scanning calorimetry, presented a broadened peak, sometimes with a second shoulder, depicting the phase transition temperature around 21 °C. Neutron scattering revealed a small but significant variation of the membrane dynamics due to the peptides in both in-plane and out-of-plane directions. Although we did not find a clear hint for synergy in the interplay of the two peptides, the calorimetric and neutron data give compatible results in terms of a decrease of the enthalpy due to the presence of the peptides, which destabilize the membrane. The dynamics in the two directions was differentiated when the individual peptides were added to the membranes, but the impact was smaller when both peptides were added together.



INTRODUCTION

While during the past decades many pathogenic microorganisms have developed resistance against a multitude of conventional antibiotics,[1] the number of new antimicrobial compounds entering the market has decreased consistently.[2] In combination, the panel of therapeutic options to treat infections caused by multidrug-resistant bacteria has dropped considerably. Therefore, new compounds and strategies are heavily sought after, and clinical trials using cationic antimicrobial peptides or mimetics thereof have already shown good promise for the development of new classes of antibiotics.[3,4] Several living species produce antimicrobial peptides as a first line of defense against infections working as sentinels of the innate immune system. Notably, pathogenic bacteria struggle to develop resistance against their mode of action that is thought to target the physicochemical properties of the cell membrane, unlike conventional antibiotics that interfere with the metabolic pathways.[5,6] The knowledge about the interactions between antimicrobial peptides and membranes is still elusive, because it is a multicomponent and complex process[7,8] and might be different for each type of peptide.

Magainins and PGLa were among the first peptides where the antimicrobial activity was tested and described,[9] and their mechanism of action was investigated in great detail.[7,8] These peptides are about 25 residues in length, carry an overall positive charge, and they adopt amphipathic α-helical conformations when interacting with membranes.[10−12] Enantiomers, retromers, or retroenantiomers of magainins exhibit the same high antibiotic and pore-forming activities of the parent L-peptide, supporting the idea that the antibiotic activity is primarily related to their interactions with cell membranes rather than through specific, chiral receptor recognition.[13]

Magainin 2 and PGLa are naturally stored together in the skin granules of *Xenopus laevis* frogs for rapid release upon infection[14] and have shown good evidence to selectively impair the bacterial plasma membrane.[15,16] Interestingly, the two peptide systems have shown synergistic effects with an enhancement of about an order of magnitude in antimicrobial assays, as well as in fluorescent dye-release experiments using lipid-only model membranes.[17−19] The mechanistic basis of membrane disruption and antibacterial activities of the peptide cocktails remains unclear[18,20,21] and has recently been reviewed and investigated in refs 8, 22, and 23.

Many biophysical investigations are indicative that the amphipathic helices of this and many related linear cationic peptides preferentially orient parallel to the bilayer surface.[15,24] This interfacial positioning results in the destabilization of the lipid bilayer, thereby causing leakage of cellular contents, an increase of electric conductance and collapse of electrochemical gradients across the bilayer, giving a good explanation for the killing of microorganisms (reviewed in, e.g., refs 23, 25, and 26). A detailed analysis of the peptide topologies using

Received: July 17, 2021
Revised: August 23, 2021

solid-state NMR investigations show that magainin 2 adopts a very stable alignment parallel to the membrane surface in the presence or absence of PGLa,[27−30] while the interactions of PGLa are more sensitive to details of the membrane composition. In the presence of unsaturated lipids, including membranes made of *E. coli* lipid extracts, PGLa orients along the interface of liquid crystalline bilayers under all conditions tested so far (tilt angles close to 90°). In contrast, in fully saturated lipid bilayers the tilt angle is dependent on the peptide-to-lipid ratio, the presence of magainin, the hydration level, and the composition of the lipid bilayer.[28−31] Recently, a negative intrinsic curvature of the PE lipids was correlated to the synergistic enhancement of dye release activities from model membrane vesicles when both peptides are present.[22] The presence of the PE lipid was found essential to promote the formation of higher order peptide−lipid supramolecular arrangements that foster the membrane association of the peptides when added as a mixture.[32]

The 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine/1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1′-rac-glycerol) POPE/POPG (3:1 mol/mol) lipid mixture is considered a good model membrane mimetic of the composition of the cytoplasmic membranes of Gram-negative bacteria.[22] This membrane composition was used in several recent studies investigating the synergistic interaction of PGLa and magainin at a peptide-to-lipid ratio of ≈2 mol %, where the synergistic effect is high and the detergent action of individual molecules remains moderate.[22,32] Under these conditions, both peptides adopt an in-plane orientation (see Figure 1), similar to the topology also observed in *E. coli* lipid extracts, disturbing the membrane in the vicinity of the peptides and locally inducing its thinning.[30] Interestingly, under these conditions, the formation of peptide−lipid supramolecular assemblies has also been observed by optical techniques[20,32] as well as by MD simulations,[33] in agreement with previous studies suggesting the (transient) formation of parallel heterodimers[18,34] and elongated assemblies stabilized by electrostatic and hydrophobic interactions.[33,35] Such supramolecular arrangements are a key component to understanding the synergistic activities of PGLa and magainin 2.[32] The heterodimers could even be involved in cross-linking apposing bilayers.[20,36]

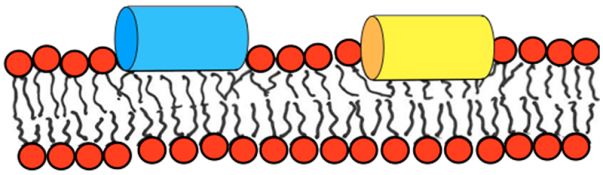


Figure 1. Schematic drawing of the in-planar arrangements of magainin 2 (blue) and PGLa (yellow) within membranes.

Here, we studied the lipid movements by testing their in-plane and out-of-plane dynamics with respect to the membrane surface. To this end, oriented multilamellar membranes made of POPE/POPG (3:1 mol/mol) supplemented with 2% of PGLa and magainin 2 either alone or at equimolar mixtures were investigated. Elastic Incoherent Neutron Scattering (EINS) is a well-adapted technique for such purposes, as neutrons represent a nondamaging way to probe the samples. The wavelengths and energies used here match well the typical binding energies found in biological systems. We combined this approach with measurements of multilamellar vesicles of the same composition by Differential Scanning Calorimetry (DSC). DSC does not permit to distinguish in-plane and out-of-plane movements, thus, we used spherical vesicles with many layers, which are easier to produce. In addition, DSC is more adapted to measurements of vesicles in suspension instead of oriented membranes on wafers. These experiments allowed us to shed light on the impact of the peptides on the lipid phase transitions and on thermodynamic quantities, which reversely also influence molecular dynamics.

METHODS

Differential Scanning Calorimetry. DSC probes heat changes within a solution and permits to extract various thermodynamic parameters linked to the transition, such as the transition temperature T_d, the changes of enthalpy ΔH, entropy ΔS, and heat capacity ΔC_p. The specific heat capacities are normalized to a theoretical baseline, built with the degree of reaction method.[37] We call $f(T)$ the equation of the baseline, where $f_{pre}(T) = a_{pre} \times T + b_{pre}$ corresponds to the linear fit of the data before the transition peak and $f_{post}(T) = a_{post} \times T + b_{post}$ corresponds to the linear fit of the data after the transition peak. With $H(T)$ being the entropy, calculated as in ref 37, the function $\alpha(T) = H(T)/\Delta H$ is the degree of reaction. In that case, the equation of the baseline is written as $f(T) = [1 - \alpha(T)] \times f_{pre}(T) + \alpha(T) \times f_{post}(T)$.

Elastic Incoherent Neutron Scattering. EINS is a technique giving access to the molecular motions of hydrogen atoms, as the incoherent neutron scattering cross section of hydrogen is much higher than for any other type of atoms commonly found in biomolecules. In a neutron scattering experiment, the neutron structure function $S(\vec{Q},\hbar\omega)$ is probed, with $\vec{Q} = \vec{k}_{incident} - \vec{k}_{scattered}$ being the momentum transfer, and $\hbar\omega$ is the energy transfer between the incident neutron beam and the sample. Within the Born approximation and neglecting multiple scattering, $S(\vec{Q},\hbar\omega)$ is directly linked to the classical Van Hove pair correlation function $g(\vec{r},t)$ through a Fourier transform in time and space[38] and enables to retrieve information about the sample structure (coherent contribution) or dynamics (incoherent contribution).

In an EINS experiment, the elastic incoherent neutron structure function is linked to the normalized measured intensity by the following expression:[39]

$$S_{inc}^{el}(Q) \approx I(Q,\tau) = \sum_{\alpha}^{N} x_\alpha \exp\left\{-\frac{1}{3}Q^2\langle r_\alpha^2\rangle[1 - C_\alpha(\tau)]\right\} \quad (1)$$

with x_α being the proportion of particles having the same dynamics and accounting for the polydispersity. Q is the modulus of the momentum transfer, $\langle r_\alpha^2\rangle$ is the equilibrium mean-square displacement (MSD) of particle α, and $C_\alpha(\tau)$ is the stationary position relaxation function in the time frame of the instrumental resolution time τ.

If we assume on the one hand that all atoms have similar dynamics, the polydispersity $x_\alpha = 1$, and on the other hand that the motions of these atoms around their equilibrium positions follow a Gaussian distribution, a hypothesis also known as the Gaussian approximation,[40] the total MSD, $\langle R^2(T)\rangle$, representing the average dynamics of hydrogen atoms under the time resolution τ, can then be obtained through the following:

$$\langle R^2(T)\rangle = -3\frac{d[\ln\{S_{inc}^{el}(Q)\}]}{dQ^2}\bigg|_{Q=0} \quad (2)$$

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

By extracting the MSD for each measured temperature, we can have insight into the dynamics upon phase transition when scanning the sample. In the case of the lipid gel-phase transition, it is commonly characterized by a steep increase of the MSD around the transition temperature T_d, preceded and followed by a linear increase.[41] To describe the process quantitatively, we use a two-state model developed by Bicout and Zaccai, which relates the sample flexibility to the thermodynamics of the phase transition through:[39]

$$\langle R^2(T)\rangle = [1-\phi(T)]\frac{k_B T}{k_1} + \phi(T)\frac{k_B T}{k_2} \quad (3)$$

$$\text{with} \quad \phi(T) = \frac{1}{1+\exp\left(\frac{\Delta H - T\Delta S}{k_B T}\right)} \quad (4)$$

where k_B is the Boltzmann constant, k_1 and k_2 are the respective average force constants of the sample before and after the phase transition describing the median rigidity of the bonds between atoms, and ΔH and ΔS are the excesses of enthalpy and entropy upon the transition. The probability of finding an atom below the phase transition at a lower temperature is given by $(1-\phi)$, where ϕ is defined by eq 4 and ϕ is the probability to find it above. One usually defines a transition temperature T_d as the temperature at which ΔG is zero, that is, when a compensation between enthalpy and entropy occurs, such that $[1-\phi(T_d)] = \phi(T_d) = 0.5$. As defined in ref 39, the apparent force constant k_3 of the system at any temperature was calculated at a reference temperature of T_r = 37 °C, the highest temperature reached here:

$$\frac{1}{k_3} = \frac{[1-\phi(T_r)]}{k_1}\left\{1 - \frac{\Delta H}{k_B T_r}\phi(T_r)\right\} + \frac{\phi(T_r)}{k_2}\left\{1 + \frac{\Delta H}{k_B T_r}[1-\phi(T_r)]\right\} \quad (5)$$

The model further assumes small and large cages for the motions taking place in the sample. These motions can be vibrational, rotations of, for example, hydrogens from methyl groups, or translational, like the whole diffusion of a lipid inside the membrane. Depending on the phase of the lipids, these motions are confined in small cages in the gel state or of larger scale in the liquid disordered phase.

When the temperature increases, a certain proportion of hydrogen atoms passes from small cages to large cages in terms of dynamics; this proportion is represented by $\phi(T)$, and the corresponding force constant k_2. Reversely, $[1-\phi(T)]$ is the proportion of atoms moving in small cages and is described by the force constant k_1.

Materials. High purity D_2O (99.9%) was purchased from Merck (Darmstadt, Germany). The two phospholipids, 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1′-rac-glycerol) sodium salt (POPG), were purchased from Avanti Polar Lipids (Alabaster, AL). POPE has a main phase transition temperature of 25 °C, whereas the one of POPG occurs at −2 °C.

Peptide Synthesis and Purification. The magainin 2 (GIGKF LHSAK KFGKA FVGEI MNS, prepared with a carboxyterminal free acid) and PGLa (GMASK AGAIA GKIAK VALKA L-NH_2) peptides were prepared by solid-phase peptide synthesis on a Millipore 9050 (Millipore, Burlington, MA) automatic peptide synthesizer and its standard Fmoc chemistry. The peptides were purified (>90%) by semipreparative HPLC, as described previously for other antimicrobial peptides.[42] The identity of the products was confirmed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. The peptides were lyophilized and the trifluoroacetic acid counterions were exchanged in an excess of 5% acetate (v/v) in water and then lyophilized again.[43]

In all experiments, the following four samples were studied: (1) POPE/POPG 3:1 (mol/mol); (2) POPE/POPG 3:1 (mol/mol) + 2 mol % PGLa; (3) POPE/POPG 3:1 (mol/mol) + 2 mol % magainin 2; (4) POPE/POPG 3:1 (mol/mol) + 1 mol % PGLa + 1 mol % magainin 2 (total peptide concentration 2 mol %).

Multilamellar Vesicles for DSC. Suspensions of multilamellar vesicles (MLVs) were prepared as follows: 2 mg of POPE/POPG (3:1 mol/mol) in chloroform/methanol 2:1 v/v were mixed with aliquots of peptide solutions to obtain a peptide-to-lipid ratio of 2 mol %. The solution was then dried at room temperature under nitrogen flow before being placed in a vacuum chamber ($P \approx 10^{-3}$ Torr) for 2 h. Prior to the DSC measurements, the preparation was hydrated with H_2O to a final concentration of 5 mg/mL, corresponding to an excess of water and therefore full hydration and vortexed for about 1 h. The use of H_2O here does not influence the behavior around the phase transition.

Oriented Lipid Bilayers for Neutron Scattering. To achieve a peptide-to-lipid ratio of 2 mol % peptide, powder was codissolved with 200 mg of phospholipid (POPE/POPG 3:1 mol/mol) in organic solvent (chloroform and methanol). The homogeneous solution was deposited onto small-size (30 × 40 × 0.28 mm^3) crystalline silicon wafers (from Si-Mat Silicon Materials, Kaufering, Germany) and dried at room temperature under a nitrogen flow before being placed in a vacuum chamber ($P \approx 10^{-3}$ Torr) for about 2 h. The samples were then equilibrated in a hydration chamber (100% relative humidity of D_2O) at 40 °C for at least 24 h. Diffraction measurements on the high resolution diffractometer D16 at ILL[44] showed *d*-spacing of the membrane stacks of 53.4 Å for the pure lipids and in between 54.1 and 54.8 Å with the peptides incorporated. A short calculation[45,46] provides then a hydration of 13 molecules of water per lipid for the pure lipids and 14−15 water molecules per lipid for the complex samples, thus, all samples were close to full hydration. Another simple calculation shows that the peptides were in contact with about 11 lipid headgroups. The wafers where stacked on top of each other and placed in slab-shaped aluminum sample holders (six wafers per samples). Aliquots of 50 μL of D_2O were added inside each of the sample holders to ensure full hydration, before being closed and sealed using wires made of indium. The weight of each sample was monitored before and after the EINS measurements to verify the stability of the hydration levels of the oriented bilayers.

Differential Scanning Calorimetry. Scans of MLV samples, at a concentration of 5 mg/mL in H_2O buffer, were measured against H_2O on a 6300 Nano DSC III. Successive heating and cooling cycles were performed at a scan rate of 1 °C/min between 0 and 40 °C, with 10 min equilibration between heating and cooling cycles. Measurements were all done at 4 bar pressure to avoid water fusion and breaking of the capillaries. At least three heating−cooling cycles were performed for each sample to overcome the annealing effect of lipid membranes, so that the first cycle was never considered in our analyses.

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

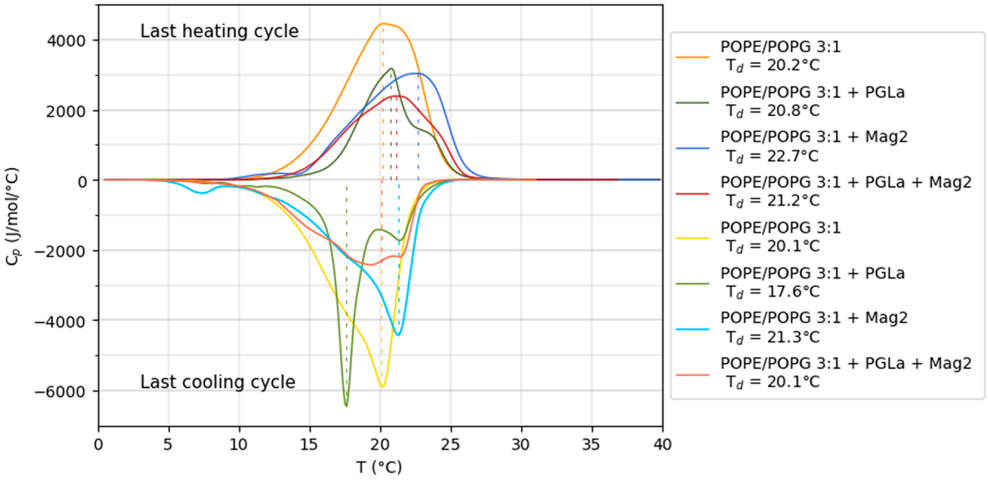


Figure 2. Comparison of the heat capacities after correction by buffer and empty cell and normalization by the baseline. Only the last heating and cooling cycles are shown here.

Table 1. Thermodynamic Parameters Averaged on All Cycles (First One Excluded)[a]

	heating				cooling			
	POPE/POPG	POPE/POPG + PGLa	POPE/POPG + Mag2	POPE/POPG + PGLa + Mag2	POPE/POPG	POPE/POPG + PGLa	POPE/POPG + Mag2	POPE/POPG + PGLa + Mag2
T_d (°C)	20.2	20.8	22.7	21.2	20.1	17.6	21.2	19.8
ΔH (kJ/mol)	29.6 (3)	14.9 (1)	22.4 (1)	17.4 (5)	30.7 (4)	19.3 (1)	24.4 (1)	19.4 (3)
ΔC_p (J/mol/°C)	−458 (23)	−861 (4)	−754 (11)	−351 (14)	276 (14)	478 (1)	744 (12)	160 (16)

[a]The error on the last digit is shown between brackets.

The data were normalized by the sample mass and heat rate and baseline-subtracted where the baseline was computed by the degree of reaction method,[37] using *pyDSC* script, as described in ref 47.

Elastic Incoherent Neutron Scattering. EINS measurements were performed on the thermal backscattering spectrometer IN13 (ILL, Grenoble, France),[48] using a wavelength of λ = 2.23 Å. The instrument has a high momentum transfer range, 0.2 < Q < 4.9 Å^{-1}, and a full-width energy resolution of 8 μeV, meaning that motions between 1 and 30 Å and up to 100 ps are accessible. Elastic fixed window scans were recorded. The points at −13 and −3 °C were measured separately for 2 h each, then continuous temperature ramps were performed from 7 to 43 °C, with heat rates between 30 and 40 m°C/min, depending on the sample. The measurements were recorded every 10 min.

As the amount of peptide was low and the peptides small (ca. 20 amino acids with a length of about 3 nm, comparable to the typical length of about 3.7 nm of a lipid), the direct contribution of the peptide dynamics to the signal is negligible. We rather probed here the lipid dynamics impacted or not by the peptides.

Each sample was measured at two different geometries, at 135° and 45°, which is the angle between the flat aluminum sample holder and the incident neutron beam. The 135° geometry corresponds to the case where the momentum transfer vector $\vec{Q}$ is mainly parallel to the surface of the membranes deposited on the Si wafers, thereby focusing on the in-plane motions. While in the 45° geometry, the momentum transfer vector $\vec{Q}$ is mainly perpendicular to the surface and thereby focuses on the out-of-plane motions.[49]

Scans of an empty Al cell with six clean Si wafers were performed in both geometries at 27 °C. The data from all samples were corrected by the latter contribution and normalized with the corresponding measurements at −13 °C, using the Large Array Manipulation Program (LAMP) available at the ILL.[50] The error bars come from Poisson statistics applied to the raw data.

The MSD were retrieved from the corrected and normalized data within the framework of the Gaussian approximation,[40] following eq 2, and then fitted with the Bicout-Zaccai model,[39] according to eqs 3 and 4. This permitted the extraction of the force constants and thermodynamic parameters ΔH and ΔS, using the *lmfit* package from Python.[51]

RESULTS

Differential Scanning Calorimetry. DSC on the four different samples showed very broad transitions of over more than 10 °C, with a maximum peak between 20 and 23 °C for the heating cycle and a splitting into a double peak for some of the curves. The maximum peak corresponds to a value in between the transition temperatures of the two individual lipids (POPE has a main phase transition temperature of 25 °C, POPG of −2 °C), but closer to the one of POPE, which contributes here three times more than POPG. Such findings are in agreement with results reported by Pozo Navas et al.[52] who did a systematic study in dependence of the ratio of POPE and POPG. They always found a single peak, but the transition temperature changed continuously in between the individual transition temperatures.

As shown in Figure 2 and the transition temperatures displayed in Table 1, a hysteresis is found in the heating−

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

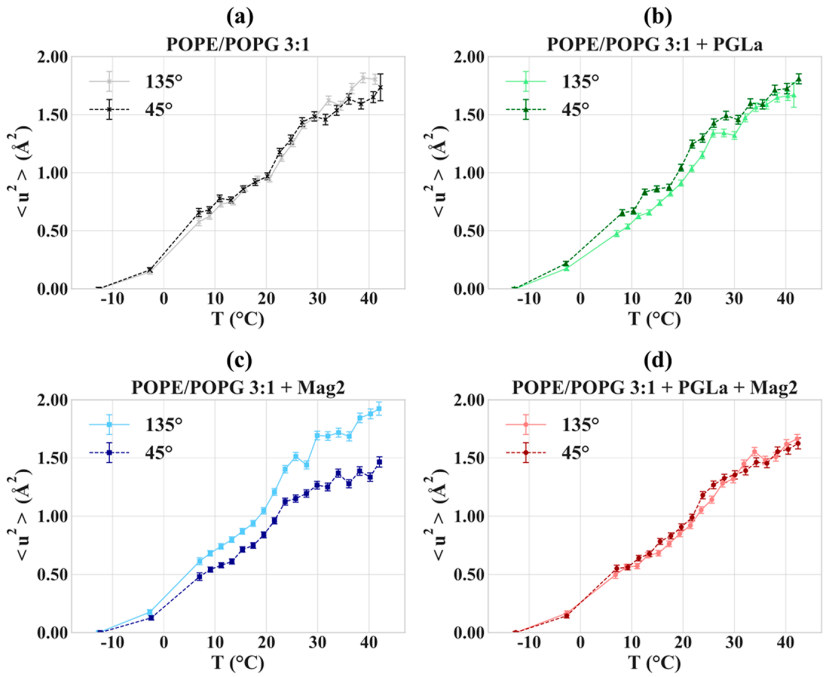


Figure 3. Comparison of the MSD at 135°and 45° for each sample.

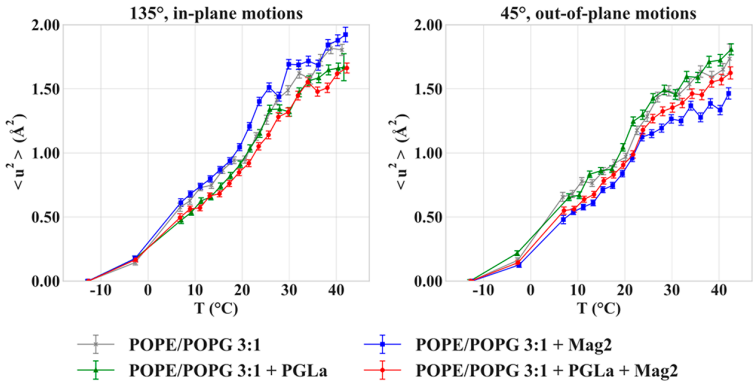


Figure 4. Mean-square displacements calculated from eq 2 for each sample and orientation as a function of the temperature. Lines are guide for the eye.

cooling cycle, with a lower transition temperature T_d when cooling the membranes. Such behavior is known for lipid membranes[53] and was already observed in the presence of such types of lipids and peptides.[36]

The data analysis of the heat capacities C_p according to ref 47 prevails certain trends. We fitted the integral of the whole transition, as the approach is based on a two-state model. We also tried to fit two transitions when two peaks were visible, but it resulted in an overfitting with unstable results. The curves corresponding to the pure lipids present the biggest surfaces and, therefore, higher excesses of enthalpy compared to the complex samples. ^{15}N solid-state NMR spectroscopy indicates an alignment of the peptides parallel to the membrane surface, while molecular modeling calculations are indicative of a partitioning of the amphipathic helices into the membrane interface.[30,33] In this manner, the peptides push the lipid head groups apart, creating extra space in the bilayer interior and thereby a decrease in the order parameter of the fatty acyl chain packing as evident from ^{2}H solid-state NMR spectroscopy investigations of deuterated lipid fatty acyl chains.[30] As a result, the cooperativity of the gel−liquid transition becomes less pronounced.

Both peptides show the tendency to induce shoulders in the peaks, indicating more complex phase transitions, but PGLa has the most pronounced effect. Molecular modeling calculations are indicative that the hydrophobic surface of the amphipathic PGLa exhibits a much larger hydrophobic angle concomitant with a deeper penetration into the membrane when compared to magainin 2.[19,33] When heating the sample, the absolute change of heat capacity is slightly decreased when compared to pure lipids, meaning that the membrane can store less energy in the liquid phase. On the contrary, it is observed upon cooling because in the presence of PGLa the membrane ponds more energy in the gel phase.

Elastic Incoherent Neutron Scattering. Here, we compared the MSD of the H atoms in all four samples

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

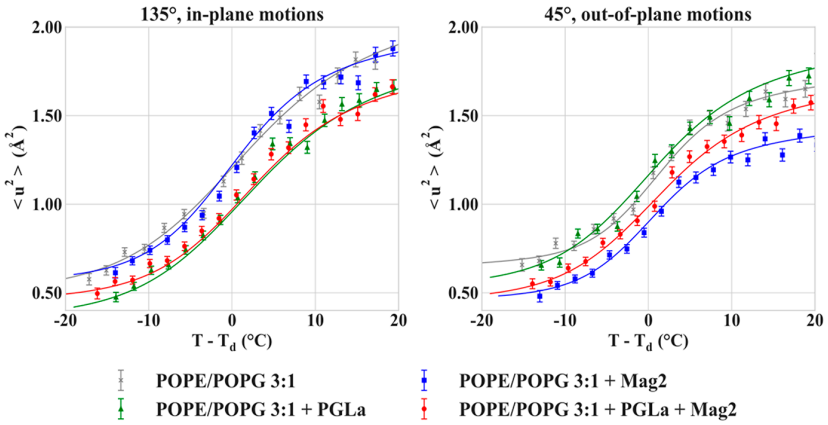


Figure 5. Fit results along with corresponding data for each sample and orientation as a function of $T - T_d$, with T_d being the value for the transition temperature determined by the fit and displayed in Table 2.

Table 2. Fit Results and Derived Parameters (the Error on the Last Digit Is Shown between Brackets)[a]

	POPE/POPG		POPE/POPG + PGLa		POPE/POPG + Mag2		POPE/POPG + PGLa + Mag2	
	45°	135°	45°	135°	45°	135°	45°	135°
k_1 (N/m)	0.57 (3)	0.72 (9)	0.68 (10)	1.0 (2)	0.80 (6)	0.65 (6)	0.80 (7)	0.80 (6)
k_2 (N/m)	0.26 (1)	0.22 (2)	0.24 (1)	0.25 (1)	0.31 (1)	0.23 (1)	0.27 (1)	0.26 (1)
k_3 (37 °C) (N/m)	0.12 (2)	0.049 (5)	0.088 (9)	0.064 (7)	0.17 (2)	0.096 (9)	0.091 (9)	0.062 (7)
ΔH (kJ/mol)	181 (27)	107 (22)	128 (30)	111 (17)	170 (24)	145 (23)	131 (17)	130 (16)
ΔS (J/mol/°C)	612 (91)	360 (74)	436 (101)	376 (57)	580 (83)	494 (77)	445 (56)	437 (52)
ΔG (37 °C) (J/mol)	−8720 (95)	−4600 (78)	−7160 (106)	−5560 (60)	−9500 (87)	−8140 (81)	−6950 (59)	−5470 (55)
T_d (°C)	22 (1)	24 (2)	21 (2)	21 (1)	20 (1)	21 (1)	21 (1)	23 (1)
χ^2_{red}	1.33	2.13	2.98	1.48	1.53	2.52	1.06	1.23

[a]To calculate $\Delta G = \Delta H - T\,\Delta S$, the temperature must be converted to values in Kelvin.

measured in the in-plane (135°) and out-of-plane (45°) directions (see Figure 3). As oriented membranes require a good degree of alignment, we tested it on the high resolution diffractometer D16 at ILL.[44] The corresponding rocking curves are shown in the Supporting Information (SI, Figure S1 and Table S1). The full width half-maximum (fwhm) of the peaks was found between 1.5 and 3.1° as the various wafers were not perfectly parallel to each other, but it is fully sufficient for the purpose of our study.

As membranes are either gel-like (below T_d) or fluid (above T_d) within the temperature range considered here, the motions within the plane of the membrane are requesting less energy than the movements where lipids partially leave the plane. For pure lipid bilayers and at high temperatures, in-plane motions are more pronounced when compared to out-of-plane ones (Figure 3a). The differences are even larger for magainin 2 (Figure 3c) and inverse for PGLa (Figure 3b), for example, the in-plane motions being below the out-of-plane motions in that case, in contrast to the other samples. At low temperature, the differences remain for magainin 2 and PGLa, but disappear for pure membranes and the peptide mix. The most remarkable effect relies certainly in the observation that the membranes with the two peptides included present almost the same behavior in- and out-of-plane, as if the two peptides together provided a more homogeneous behavior to the membrane in different directions (Figure 3d).

Figure 4 directly compares the MSD of all four samples for in-plane and out-of-plane motions, respectively. Here the differences are more striking, especially for the sample including magainin 2 (blue line and symbols), for which the in-plane motions are highest and the out-of-plane motions lowest when compared to the other samples. The sample without peptides shows the expected behavior with movements in-plane favored over the out-of-plane motions (see Figure 3a). When compared to the sample without peptides (a), PGLa damps the motions in-plane, but leaves out-of-plane motions almost unaffected (b). Overall, for out-of-plane motions, the sample in the presence of both peptides has the MSD the most independent from the orientation (see Figure 4), pointing toward a counterbalancing dynamic effect, which aims at smearing out the opposite impacts caused by the two peptides.

We fit the curves with the model described by eqs 3 and 4 at temperatures above 7 °C (to exclude the high increase in MSD due to water melting in between −3 and 7 °C, where the two-state model does not apply; see Figure 5). They present the typical sigmoidal forms with very similar slopes before and after the transition.[41] The main phase transition appears again very broad when measured by neutron scattering spreading over 15−20 °C. The curves are shifted in absolute height, and the individual peptides exchange their role in the two directions, magainin 2 conferring more and PGLa less dynamics to the membranes in the in-plane direction, whereas PGLa provides more and magainin 2 less dynamics to the membranes in the out-of-plane direction.

The fit results are summarized in Table 2.

The fits converged satisfactorily, resulting in χ^2 values ≤ 3. The extracted transition temperature T_d was between 21 and 24 °C for all samples and orientations, close to but not identical with the main phase transition temperature of the majority fraction lipid POPE. The force constants k_1 were all

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

higher than k_2, reflecting that the membrane becomes more flexible in the liquid-crystalline phase compared to the gel phase and in agreement with earlier findings for DMPC membranes.[54] k_3, which was calculated at the reference temperature 37 °C, is even significantly lower than k_2.

DISCUSSION

The investigations presented here demonstrate the impact of two linear cationic antimicrobial peptides, namely, magainin 2 and PGLa, when added alone or as a synergistic mixture, on membrane dynamics. We used a mixture of zwitterionic and anionic phospholipids POPE and POPG at a 3:1 molar ratio to mimic the composition of the cytoplasmic membranes of Gram-negative bacteria. The membranes were fully hydrated, and the samples were scanned in the temperature range from −253 to 37 °C. The samples did not present significant differences in MSD at low temperatures (data below −13 °C, not shown here), but interesting features were present above the melting point of water. Through the MSD, elastic incoherent neutron scattering permits to extract information on the average molecular dynamics of H atoms and the molecular subgroups to which they are bound within the preferential in-plane or out-of-plane directions. Whereas the MSD show that, in general, in-plane motions are privileged over out-of-plane motions, especially in the liquid-crystalline phase, the individual peptides can have differentiated effects.

Both peptides are cationic amphipathic and partition into the membrane interface at stable orientations parallel to the surface, concomitant with considerable membrane disordering and thinning of the lipid bilayer.[30] However, PGLa is more hydrophobic and therefore partitions deeper into the bilayer interface,[19,33] and differences in the distribution of charges can be observed. As a consequence, magainin exhibits very persistent in-plane tilt angles close to 90° under most conditions investigated, suggesting that the peptide forms a stable interaction network with the lipids, thereby damping out-of-plane motions.[29] In contrast, PGLa tends to respond to changes in membrane composition and to adopt many different alignments.[28,31,43] These observations correlate with PGLa suppressing in-plane motions and favoring the out-of-plane ones, whereas in the presence of magainin 2, the opposite trends are observed.

Interestingly, the pure membrane and the synergistic peptide mixture reconstituted into the membrane present similar dynamics on the 100 ps time scale investigated here, with the peptides merely damping the motions to result in slightly lower MSD. Indeed, previous investigations indicate that, in POPE/POPG 3:1, the peptide mixtures tend to pack more densely and to exhibit correlated diffusion.[32,36] These observations are suggestive that the peptides are enriched in areas leaving other lipid regions depleted from the influence of peptides. Because the cationic peptides interact preferentially with the anionic lipids,[55] such phase separation processes probably also involve some demixing of the lipids.

Globally, the impact of peptides on lipid membrane dynamics does not seem to follow a generic scheme, which becomes obvious when comparing magainin 2 and PGLa (Figures 3 and 4). Other authors reported also various effects, depending on specific conditions such as peptide concentration or charges: Sharma et al. compared the dynamics of DMPC membranes altered by two antimicrobial peptides at low concentrations of 0.5 mol %,[56] because otherwise they could induce pore formation in the membranes. At low concentration, they are situated parallel to the surface similar to the peptides investigated here. The peptide with a net charge of −1, alamethicin, suppressed the pretransition in the membrane, whereas the peptide with +5 net surface charge, melittin, suppressed both pre- and main phase transitions, disrupting the regular phase behavior in that way. In our study, we observed the main phase transition in all four samples, even if the signal of it is rather weak in elastic neutron scattering. Sharma et al. reported a decrease in dynamics on addition of alamethicin. On the contrary, melittin increases the dynamics below the phase transition and decreases it above, making the DMPC structure less ordered below T_d and less disordered above. Barrett et al. found an increase in dynamics far away from the phase transition in the presence of a peptide, but a decrease close to the transition.[57] Knoll et al.[58] investigated lipids mimicking a myelin membrane in the presence of two of the myelin proteins, which had the synergistic effect of stiffening the membrane, therefore, reducing the motions.

The application of a two-state model, describing the thermodynamics of the phase transition from the gel to the liquid-crystalline state proposed by Bicout and Zaccai,[39] permitted to determine thermodynamic parameters and force constants, describing the averaged resilience of the membrane (see Figure 5 and Table 2). An open question remained regarding the comparability with thermodynamic results extracted from DSC, as both methods probe very different time scales. However, in biological systems, strong correlations of molecular dynamics over time scales covering many decades are increasingly observed (see, for example, literature in refs 59−61). To our best knowledge, the two methods were applied here for the first time simultaneously on membrane-associated peptides. We found that the transition was broad for all samples, as also seen with DSC, and not affected by the presence of the peptides. The main phase transition temperature was almost the same in all cases; only the absolute values of the MSD were impacted to different degrees. The force constants validated that the flexibility of the membranes was increased with temperature (the force constants are then decreased) and even higher at 37 °C, well above the transition temperature.

Furthermore, the model gives access to the excesses of enthalpy (ΔH), entropy (ΔS), and free energy (ΔG) of the phase transition by differentiating the two directions in-plane and out-of-plane of the membrane. Generally, the enthalpy and entropy differences are lower for the in-plane motions than for the out-of-plane ones, in agreement with the higher energy required for motions perpendicular to the membrane. The only exception to this rule is the sample incorporating the two peptides, for which ΔH and ΔS are the same within error bars in both directions, as if the peptides had the effect of averaging over the angles due to their locations.

Overall, force constants are in good agreement with values determined earlier for DMPC, but the thermodynamic quantities determined here are higher by about 1 order of magnitude.[62] However, we probed the thermodynamics of oriented membranes, whereas in the earlier study we concentrated on multilamellar vesicles (MLVs). The supported bilayers are often found stiffer than MLVs, mainly because the hydration is more homogeneous in the latter case.[63]

Whereas DSC and neutron scattering yield very similar results for the transition temperatures and the widths of the transition, the thermodynamic parameters found by DSC are much smaller than those determined by neutron scattering. We

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

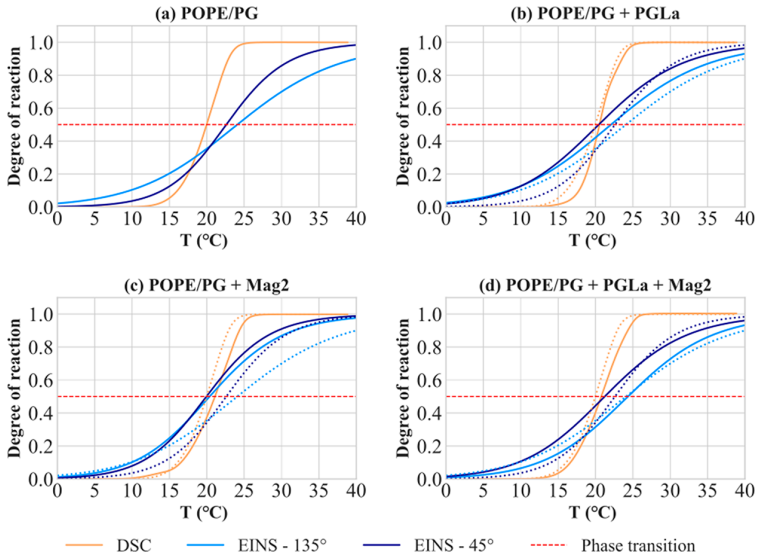


Figure 6. Comparison of the degrees of reaction α and ϕ, derived respectively from DSC measurements and EINS results, for the four samples. The dotted lines correspond to the pure POPE/PG degrees of reaction of Figure 6a.

speculate that such findings are mainly due to differences in concentration and hydration required by the two techniques. ΔH and ΔC_p are sensitive to the energy needed for the transition, which depends on the sample structuration and state. Not only are the samples prepared differently (MLVs against flat supported lipid bilayers, so that the differentiation of the directions was not possible) and are investigated at a much different concentration, the DSC measurements are also much faster than neutron experiments (a few hours against days). Most importantly, both techniques give access to molecular processes acting on very different time scales (s to min for DSC and sub-ns for EINS), so that the results cannot be directly compared. However, more and more hints can be found in the literature that molecular dynamics in biological systems are correlated over many decades in time (see literature in refs 59−61). Consequently, we find similar trends with both techniques. The peptides included individually and in combination decrease the thermodynamic parameters compared to the pure membrane, certainly due to a lower order of the chains.

Furthermore, the two approaches permit to compare the degrees of cooperativity in dynamics recorded for the transition. They correspond to the function $\phi(T)$ defined in eq 4 and to the function $\alpha(T)$ defined in the Methods on DSC measurements and are represented in Figure 6. Clearly, the degree of reaction or dynamical cooperativity is much higher when measured with DSC, for example, when the reaction time observed is of the order of seconds, the reaction appears much narrower. It is lower in case of EINS, where sub-ns time scales are probed, and is broadened over 20 or more degrees. Again, one remarkable difference of ϕ between the pure lipid membrane and in the presence of one or two peptides is that the degree of reaction for out-of-plane (dark blue in Figure 6) is higher above 20 °C, meaning that the liquid phase is reached in the out-of-plane direction before the in-plane case. It is probably due to the packing of lipids and especially their tails; to reach the liquid phase in the in-plane direction, it necessitates more thermal energy. The differences in the two directions are more marked for the pure membranes or in the presence of both peptides.

CONCLUSIONS

We investigated the synergistic effect of two antimicrobial peptides on a membrane of POPE/POPG phospholipids by EINS and DSC. Neutron scattering allows to separate the motions oriented mainly in-plane and out-of-plane, whereas MLVs were used for the DSC measurements. The pure membrane was compared to samples with one or the two peptides incorporated. From previous studies it is known that the peptide magainin 2 is situated at the membrane surface, and parallel to it. PGLa is oriented in a similar way but exhibits a higher tendency to slightly tilt toward a transmembrane orientation in fully saturated membranes, whatever the membrane curvature and molecular shape.[28−31] Although the synergistic action of both peptides on the bacterial membrane cannot be directly correlated with in-plane and out-of-plane molecular motions, small but significant differences in the MSD of the samples are apparent. The pure membranes and the membrane with the two peptides included present mainly homogeneous motions in both directions, whereas the membranes with only one peptide depict more variations, especially magainin 2 has a pronounced effect.

Systematic trends between the thermodynamic results obtained from DSC and neutron scattering are observed, pointing toward the importance of these parameters for the antimicrobial actions of the peptides. We followed the approach suggested by Bicout and Zaccai[39] to extract thermodynamic parameters from neutron data. Whereas in previous work these values were never directly compared to DSC measurements, the two techniques probe thermodynamic effects on very different time scales, and they allowed the extraction of the common behavior of the samples.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.1c06388.

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Rocking curves of all four samples used for neutron scattering and the extracted mosaic spreads (PDF)

AUTHOR INFORMATION

Corresponding Author

Judith Peters − *Univ. Grenoble-Alpes, CNRS, LiPhy, 38000 Grenoble, France; Institut Laue-Langevin, 38000 Grenoble, France; Institut Universitaire de France, 75231 Paris, France;* orcid.org/0000-0001-5151-7710; Email: jpeters@ill.fr

Authors

Aline Cisse − *Univ. Grenoble-Alpes, CNRS, LiPhy, 38000 Grenoble, France; Institut Laue-Langevin, 38000 Grenoble, France;* orcid.org/0000-0002-2193-0843

Arnaud Marquette − *University of Strasbourg/CNRS, Chemistry Institute, Membrane Biophysics and NMR, UMR7177 Strasbourg, France*

Munkhtuguldur Altangerel − *Univ. Grenoble-Alpes, CNRS, LiPhy, 38000 Grenoble, France; Institut Laue-Langevin, 38000 Grenoble, France*

Burkhard Bechinger − *University of Strasbourg/CNRS, Chemistry Institute, Membrane Biophysics and NMR, UMR7177 Strasbourg, France; Institut Universitaire de France, 75231 Paris, France;* orcid.org/0000-0001-5719-6073

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.1c06388

Author Contributions

#These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank H. Guillou for the permission to use its DSC instrument, T. Matsuo and X. Qu for their help with the DSC measurements, and G. Pabst for his most valuable discussion. Furthermore, we are grateful to B. Demé who permitted and helped to measure the rocking curves on D16. We thank the foundation JP Aguilar for financing the doctoral thesis of A. Cisse. Data are available at DOIs 10.5291/ILL-DATA.8-02-817 and 10.5291/ILL-DATA.8-02-856. The financial contributions of the Agence Nationale de la Recherche (Projects MemPepSyn 14-CE34-0001-01, Biosupramol 17-CE18-0033-3, Naturalarsenal 19-AMRB-0004-02, and the LabEx Chemistry of Complex Systems 10-LABX-0026_CSC), the University of Strasbourg, the CNRS, the Région Grand-Est, and the RTRA International Center of Frontier Research in Chemistry are gratefully acknowledged. B.B. and J.P. are grateful to the Institut Universitaire de France for providing additional time to be dedicated to research.

REFERENCES

(1) Doi, Y.; Bonomo, R. A.; Hooper, D. C.; Kaye, K. S.; Johnson, J. R.; Clancy, C. J.; Thaden, J. T.; Stryjewski, M. E.; van Duin, D. Gram-Negative Bacterial Infections: Research Priorities, Accomplishments, and Future Directions of the Antibacterial Resistance Leadership Group. *Clin. Infect. Dis.* **2017**, *64*, S30−S35.

(2) Lohner, K. Development of Novel Antimicrobial Agents: Emerging Strategies. *J. Antimicrob. Chemother.* **2003**, *51*, 1321.

(3) Domingues, M. M.; Santos, N. C.; Castanho, M. A. Antimicrobial peptide rBPI21: a translational overview from bench to clinical studies. *Curr. Protein Pept. Sci.* **2012**, *13* (7), 611−9.

(4) Hancock, R. E.; Haney, E. F.; Gill, E. E. The immunology of host defence peptides: beyond antimicrobial activity. *Nat. Rev. Immunol.* **2016**, *16*, 321−334.

(5) Perron, G. G.; Zasloff, M.; Bell, G. Experimental evolution of resistance to an antimicrobial peptide. *Proc. R. Soc. London, Ser. B* **2006**, *273* (1583), 251−6.

(6) Lazzaro, B. P.; Zasloff, M.; Rolff, J. Antimicrobial peptides: Application informed by evolution. *Science* **2020**, *368* (6490), eaau5480.

(7) Aisenbrey, C.; Marquette, A.; Bechinger, B. The Mechanisms of Action of Cationic Antimicrobial Peptides Refined by Novel Concepts from Biophysical Investigations. *Adv. Exp. Med. Biol.* **2019**, *1117*, 33−64.

(8) Marquette, A.; Bechinger, B. Biophysical Investigations Elucidating the Mechanisms of Action of Antimicrobial Peptides and Their Synergism. *Biomolecules* **2018**, *8* (2), 18.

(9) Zasloff, M. Magainins, a class of antimicrobial peptides from Xenopus skin: isolation, characterization of two active forms, and partial cDNA sequence of a precursor. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (15), 5449−53.

(10) Gesell, J.; Zasloff, M.; Opella, S. J. Two-dimensional 1H NMR experiments show that the 23-residue magainin antibiotic peptide is an alpha-helix in dodecylphosphocholine micelles, sodium dodecylsulfate micelles, and trifluoroethanol/water solution. *J. Biomol. NMR* **1997**, *9* (2), 127−35.

(11) Bechinger, B.; Zasloff, M.; Opella, S. J. Structure and dynamics of the antibiotic peptide PGLa in membranes by solution and solid-state nuclear magnetic resonance spectroscopy. *Biophys. J.* **1998**, *74*, 981−987.

(12) Haney, E. F.; Hunter, H. N.; Matsuzaki, K.; Vogel, H. J. Solution NMR studies of amphibian antimicrobial peptides: linking structure to function? *Biochim. Biophys. Acta, Biomembr.* **2009**, *1788* (8), 1639−55.

(13) Wade, D.; Boman, A.; Wahlin, B.; Drain, C. M.; Andreu, D.; Boman, H. G.; Merrifield, R. B. All-D amino acid-containing channel-forming antibiotic peptides. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (12), 4761−5.

(14) Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **2002**, *415* (6870), 389−95.

(15) Bechinger, B. The SMART model: Soft Membranes Adapt and Respond, also Transiently, in the presence of antimicrobial peptides. *J. Pept. Sci.* **2015**, *21* (5), 346−55.

(16) Lee, D. K.; Bhunia, A.; Kotler, S. A.; Ramamoorthy, A. Detergent-type membrane fragmentation by MSI-78, MSI-367, MSI-594, and MSI-843 antimicrobial peptides and inhibition by cholesterol: a solid-state nuclear magnetic resonance study. *Biochemistry* **2015**, *54* (10), 1897−907.

(17) Westerhoff, H. V.; Zasloff, M.; Rosner, J. L.; Hendler, R. W.; Waal, A.; Gomes, A. V.; Jongsma, A. P. M.; Riethorst, A.; Juretic, D. Functional synergism of the magainins PGLa and magainin-2 in Escherichia coli, tumor cells and liposomes. *Eur. J. Biochem.* **1995**, *228*, 257−264.

(18) Matsuzaki, K.; Mitani, Y.; Akada, K. Y.; Murase, O.; Yoneyama, S.; Zasloff, M.; Miyajima, K. Mechanism of synergism between antimicrobial peptides magainin 2 and PGLa. *Biochemistry* **1998**, *37* (43), 15144−53.

(19) Glattard, E.; Salnikov, E. S.; Aisenbrey, C.; Bechinger, B. Investigations of the synergistic enhancement of antimicrobial activity in mixtures of magainin 2 and PGLa. *Biophys. Chem.* **2016**, *210*, 35−44.

(20) Marquette, A.; Salnikov, E. S.; Glattard, E.; Aisenbrey, C.; Bechinger, B. Magainin 2-PGLa Interactions in Membranes - Two Peptides that Exhibit Synergistic Enhancement of Antimicrobial Activity. *Curr. Top. Med. Chem.* **2015**, *16* (1), 65−75.

(21) Zerweck, J.; Strandberg, E.; Kukharenko, O.; Reichert, J.; Burck, J.; Wadhwani, P.; Ulrich, A. S. Molecular mechanism of

synergy between the antimicrobial peptides PGLa and magainin 2. *Sci. Rep.* **2017**, *7* (1), 13153.
(22) Leber, R.; Pachler, M.; Kabelka, I.; Svoboda, I.; Enkoller, D.; Vacha, R.; Lohner, K.; Pabst, G. Synergism of Antimicrobial Frog Peptides Couples to Membrane Intrinsic Curvature Strain. *Biophys. J.* **2018**, *114* (8), 1945−1954.
(23) Bechinger, B.; Juhl, D. W.; Glattard, E.; Aisenbrey, C. Revealing the mechanisms of synergistic action of magainin antimicrobial peptides. *Frontiers in Medical Technology, section Pharmaceutical Innovation* **2020**, *2*, 615494.
(24) Perrin, B. S., Jr.; Tian, Y.; Fu, R.; Grant, C. V.; Chekmenev, E. Y.; Wieczorek, W. E.; Dao, A. E.; Hayden, R. M.; Burzynski, C. M.; Venable, R. M.; Sharma, M.; Opella, S. J.; Pastor, R. W.; Cotten, M. L. High-resolution structures and orientations of antimicrobial peptides piscidin 1 and piscidin 3 in fluid bilayers reveal tilting, kinking, and bilayer immersion. *J. Am. Chem. Soc.* **2014**, *136* (9), 3491−504.
(25) Bechinger, B. The structure, dynamics and orientation of antimicrobial peptides in membranes by multidimensional solid-state NMR spectroscopy. *Biochim. Biophys. Acta, Biomembr.* **1999**, *1462*, 157−183.
(26) Shai, Y. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochim. Biophys. Acta, Biomembr.* **1999**, *1462*, 55−70.
(27) Bechinger, B. Insights into the mechanisms of action of host defence peptides from biophysical and structural investigations. *J. Pept. Sci.* **2011**, *17* (5), 306−14.
(28) Strandberg, E.; Zerweck, J.; Wadhwani, P.; Ulrich, A. S. Synergistic insertion of antimicrobial magainin-family peptides in membranes depends on the lipid spontaneous curvature. *Biophys. J.* **2013**, *104* (6), L9−11.
(29) Salnikov, E. S.; Bechinger, B. Lipid-controlled peptide topology and interactions in bilayers: structural insights into the synergistic enhancement of the antimicrobial activities of PGLa and magainin 2. *Biophys. J.* **2011**, *100* (6), 1473−80.
(30) Harmouche, N.; Bechinger, B. Lipid-Mediated Interactions between the Antimicrobial Peptides Magainin 2 and PGLa in Bilayers. *Biophys. J.* **2018**, *115* (6), 1033−1044.
(31) Tremouilhac, P.; Strandberg, E.; Wadhwani, P.; Ulrich, A. S. Conditions affecting the re-alignment of the antimicrobial peptide PGLa in membranes as monitored by solid state 2H-NMR. *Biochim. Biophys. Acta, Biomembr.* **2006**, *1758* (9), 1330−42.
(32) Aisenbrey, C.; Amaro, M.; Pospisil, P.; Hof, M.; Bechinger, B. Highly synergistic antimicrobial activity of magainin 2 and PGLa peptides is rooted in the formation of supramolecular complexes with lipids. *Sci. Rep.* **2020**, *10* (1), 11652.
(33) Pachler, M.; Kabelka, I.; Appavou, M. S.; Lohner, K.; Vacha, R.; Pabst, G. Magainin 2 and PGLa in Bacterial Membrane Mimics I: Peptide-Peptide and Lipid-Peptide Interactions. *Biophys. J.* **2019**, *117* (10), 1858−1869.
(34) Hara, T.; Mitani, Y.; Tanaka, K.; Uematsu, N.; Takakura, A.; Tachi, T.; Kodama, H.; Kondo, M.; Mori, H.; Otaka, A.; Nobutaka, F.; Matsuzaki, K. Heterodimer formation between the antimicrobial peptides magainin 2 and PGLa in lipid bilayers: a cross-linking study. *Biochemistry* **2001**, *40* (41), 12395−9.
(35) Juhl, D. W.; Glattard, E.; Lointier, M.; Bampilis, P.; Bechinger, B. The Reversible Non-covalent Aggregation Into Fibers of PGLa and Magainin 2 Preserves Their Antimicrobial Activity and Synergism. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 526459.
(36) Kabelka, I.; Pachler, M.; Prevost, S.; Letofsky-Papst, I.; Lohner, K.; Pabst, G.; Vacha, R. Magainin 2 and PGLa in Bacterial Membrane Mimics II: Membrane Fusion and Sponge Phase Formation. *Biophys. J.* **2020**, *118* (3), 612−623.
(37) Höhne, H.; Hemminger, G. W. H.; Flammersheim, W. F. *Differential Scanning Calorimetry*, 2nd ed.; Springer: Heidelberg, 2003.
(38) van Hove, L. Correlations in Space and Time and Born Approximation Scattering in Systems of Interacting Particles. *Phys. Rev.* **1954**, *95*, 249−262.
(39) Bicout, D. J.; Zaccai, G. Protein flexibility from the dynamical transition: a force constant analysis. *Biophys. J.* **2001**, *80* (3), 1115−23.
(40) Rahman, A.; Singwi, K. S.; Sjolander, A. Theory of Slow Neutron Scattering by Liquids 0.1. *Phys. Rev.* **1962**, *126* (3), 986−996.
(41) Trapp, M.; Marion, J.; Tehei, M.; Deme, B.; Gutberlet, T.; Peters, J. High hydrostatic pressure effects investigated by neutron scattering on lipid multilamellar vesicles. *Phys. Chem. Chem. Phys.* **2013**, *15* (48), 20951−20956.
(42) Verly, R. M.; de Moraes, C. M.; Resende, J. M.; Aisenbrey, C.; Bemquerer, M. P.; Pilo-Veloso, D.; Valente, A. P.; Almeida, F. C.; Bechinger, B. Structure and membrane interactions of the antibiotic peptide dermadistinctin K by multidimensional solution and oriented ^{15}N and ^{31}P solid-state NMR spectroscopy. *Biophys. J.* **2009**, *96* (6), 2194−203.
(43) Aisenbrey, C.; Michalek, M.; Salnikov, E. S.; Bechinger, B. Solid-state NMR approaches to study protein structure and protein-lipid interactions. *Methods Mol. Biol.* **2013**, *974*, 357−87.
(44) Cristiglio, V.; Giroud, B.; Didier, L.; Demé, B. D16 is back to business: More neutrons, more space, more fun. *Neutron News* **2015**, *26*, 22−24.
(45) Pandit, K. R.; Klauda, J. B. Membrane models of E. coli containing cyclic moieties in the aliphatic lipid chain. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818* (5), 1205−10.
(46) Salvador-Castell, M.; Golub, M.; Erwin, N.; Deme, B.; Brooks, N. J.; Winter, R.; Peters, J.; Oger, P. M. Characterisation of a synthetic Archeal membrane reveals a possible new adaptation route to extreme conditions. *Commun. Biol.* **2021**, *4* (1), 653.
(47) Cisse, A.; Peters, J.; Lazzara, G.; Chiappisi, L. PyDSC: a simple tool to treat differential scanning calorimetry data. *J. Therm. Anal. Calorim.* **2021**, *145*, 403.
(48) Francesca, N.; Peters, J.; Russo, D.; Barbieri, S.; Chiapponi, C.; Cupane, A.; Deriu, A.; Di Bari, M. T.; Farhi, E.; Gerelli, Y.; Mariani, P.; Paciaroni, A.; Rivasseau, C.; Schiro, G.; Sonvico, F. IN13 Backscattering Spectrometer at ILL: Looking for Motions in Biological Macromolecules and Organisms. *Neutron News* **2008**, *19* (4), 14−18.
(49) Knoll, W.; Peters, J.; Kursula, P.; Gerelli, Y.; Natali, F. Influence of myelin proteins on the structure and dynamics of a model membrane with emphasis on the low temperature regime. *J. Chem. Phys.* **2014**, *141* (20), 205101.
(50) Richard, D.; Ferrand, M.; Kearley, G. J. Analysis and visualisation of neutron-scattering data. *J. Neutron Res.* **1996**, *4* (1), 33−39.
(51) Newville, M.; Stensitzki, T.; Allen, D. B.; Ingargiola, A. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python. *Zenodo* **2014**, na.
(52) Pozo Navas, B.; Lohner, K.; Deutsch, G.; Sevcsik, E.; Riske, K. A.; Dimova, R.; Garidel, P.; Pabst, G. Composition dependence of vesicle morphology and mixing properties in a bacterial model membrane system. *Biochim. Biophys. Acta, Biomembr.* **2005**, *1716* (1), 40−8.
(53) D'Angelo, G.; Wanderlingh, U.; Nibali, V. C.; Crupi, C.; Corsaro, C.; Di Marco, G. Physical study of dynamics in fully hydrated phospholipid bilayers. *Philos. Mag.* **2008**, *88* (33−35), 4033−4046.
(54) Peters, J.; Marion, J.; Natali, F.; Kats, E.; Bicout, D. J. The Dynamical Transition of Lipid Multilamellar Bilayers as a Matter of Cooperativity. *J. Phys. Chem. B* **2017**, *121* (28), 6860−6868.
(55) Mason, A. J.; Martinez, A.; Glaubitz, C.; Danos, O.; Kichler, A.; Bechinger, B. The antibiotic and DNA-transfecting peptide LAH4 selectively associates with, and disorders, anionic lipids in mixed membranes. *FASEB J.* **2006**, *20* (2), 320−2.
(56) Sharma, V. K.; Mamontov, E.; Tyagi, M.; Qian, S.; Rai, D. K.; Urban, V. S. Dynamical and Phase Behavior of a Phospholipid Membrane Altered by an Antimicrobial Peptide at Low Concentration. *J. Phys. Chem. Lett.* **2016**, *7* (13), 2394−401.

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

(57) Barrett, M. A.; Zheng, S. B.; Toppozini, L. A.; Alsop, R. J.; Dies, H.; Wang, A. L.; Jago, N.; Moore, M.; Rheinstadter, M. C. Solubility of cholesterol in lipid membranes and the formation of immiscible cholesterol plaques at high cholesterol concentrations. *Soft Matter* **2013**, *9* (39), 9342−9351.
(58) Knoll, W.; Peters, J.; Kursula, P.; Gerelli, Y.; Ollivier, J.; Deme, B.; Telling, M.; Kemner, E.; Natali, F. Structural and dynamical properties of reconstituted myelin sheaths in the presence of myelin proteins MBP and P2 studied by neutron scattering. *Soft Matter* **2014**, *10* (3), 519−29.
(59) Hu, X.; Hong, L.; Smith, M. D.; Neusius, T.; Cheng, X.; Smith, J. C. The dynamics of single protein molecules is non-equilibrium and self-similar over thirteen decades in time. *Nat. Phys.* **2016**, *12*, 171.
(60) Trapp, M.; Tehei, M.; Trovaslet, M.; Nachon, F.; Martinez, N.; Koza, M. M.; Weik, M.; Masson, P.; Peters, J. Correlation of the dynamics of native human acetylcholinesterase and its inhibited huperzine A counterpart from sub-picoseconds to nanoseconds. *J. R. Soc., Interface* **2014**, *11*, na.
(61) Kneller, G. R.; Hinsen, K. Fractional Brownian dynamics in proteins. *J. Chem. Phys.* **2004**, *121* (20), 10278−83.
(62) Peters, J.; Marion, J.; Becher, F. J.; Trapp, M.; Gutberlet, T.; Bicout, D. J.; Heimburg, T. Thermodynamics of lipid multi-lamellar vesicles in presence of sterols at high hydrostatic pressure. *Sci. Rep.* **2017**, *7* (1), 15339.
(63) Aoun, B.; Pellegrini, E.; Trapp, M.; Natali, F.; Cantu, L.; Brocca, P.; Gerelli, Y.; Deme, B.; Koza, M. M.; Johnson, M.; Peters, J. Direct comparison of elastic incoherent neutron scattering experiments with molecular dynamics simulations of DMPC phase transitions. *Eur. Phys. J. E: Soft Matter Biol. Phys.* **2016**, *39* (4), 1−10.

https://doi.org/10.1021/acs.jpcb.1c06388
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Appendix C

Portfolio Label RES (FR)

(EN): The thesis is associated to the Label RES (Research and Higher Education) of the University Grenoble-Alpes. Within this framework, teachings were carried out (∼100 hours TD equivalent over the three years), and specific trainings in pedagogy were followed. The RES Label was validated through the production of a portfolio (in french), attached below, which synthesizes my teaching experience, and presents a reflection on my teaching practices.

(FR): La thèse est accompagnée du Label RES (Recherche et Enseignement Supérieur) de l'Université Grenoble-Alpes. Dans ce cadre, des enseignements ont été effectuées (∼100 heures équivalent TD sur les trois ans), et des formations spécifiques à la pédagogie ont été suivies. Le Label RES a été validé au travers de la production d'un portfolio, joint ci-dessous, qui synthétise mon expérience d'enseignement, et présente une démarche de réflexion sur mes pratiques pédagogiques.

Portfolio Label RES

Tribulations pédagogiques

Aline Cissé

Université Grenoble-Alpes, Ecole doctorale de Physique

Référente Label : Irène Ventrillard
Directrice de thèse : Judith Peters

Table des matières

Remerciements

Avant toute chose, je tiens à remercier toutes les équipes pédagogiques mentionnées pour m'avoir intégrée, accepté mes propositions, mes questions ou mes remarques, pour m'avoir laissé une grande liberté quant à ma façon d'enseigner, et enfin pour toutes les discussions, les conseils, et les retours prodigués tout au long de ces deux ans d'enseignement.

1 Introduction

Commencer à enseigner comporte plusieurs défis ; on peut se sentir intimidé en face de notre première classe, ne pas se croire à la hauteur, ou batailler pour trouver le temps par rapport aux travaux de thèse. Mais un des plus imposants est bien celui-ci : désapprendre tous les préconçus accumulés au cours de plus de quinze ans de scolarité, puis apprendre ce que peut être enseigner, et tenter de s'en approcher du mieux possible. Tout comme l'acte de recherche est une remise en cause constante de ses acquis, de ses savoirs, au gré des expériences et des hypothèses que nous élaborons, il en va de même pour l'enseignement.

En trois années de thèse, je n'ai évidemment pas abouti à une recette miracle, et d'ailleurs il n'en existe pas. Néanmoins, j'en ai retiré des leçons importantes, modifié certaines de mes pratiques, et constaté les progrès directement en classe. Une des grandes problématiques que j'avais concernait la motivation et l'attention des étudiants, que ce soit pendant le cours ou en-dehors. J'ai commencé par orienter cette question sur les méthodes d'évaluation. Étant donné que la plupart des efforts fournis par les étudiants se concentrent sur la note, je souhaitais détourner un peu cette forme d'apprentissage «bachoté». Dans cette optique-là, j'en ai appris plus sur la grille critériée, et j'ai eu l'occasion de la mettre en œuvre dans plusieurs types de cours.

La question de l'attention est également au centre de la conception d'un cours, à égal avec le contenu en soi. J'en ai pris conscience lors de plusieurs formations, et de manière décuplée lors d'enseignements à distance, où il n'est plus possible de se cantonner à un cours purement magistral. Intimement lié à cette problématique, celle de l'animation de groupes est essentielle. Dans mes derniers enseignements, j'ai ainsi exploré ces deux côtés en adaptant mon cours en séquences, et en m'appuyant sur des jeux-cadres, comme ceux de Thiagi [1]. Ces derniers m'ont permis d'étendre le travail en groupe, qui m'a toujours paru comme indispensable, et d'ouvrir de nouvelles pistes.

Mon parcours, agrémenté des formations du Label, des nombreuses discussions avec mes collègues et les équipes pédagogiques, ainsi que par des éléments de littérature, est une succession de déconstructions, notamment dans la manière frontale d'enseigner, et de nouvelles constructions. Au-delà de la motivation, même si la question est toujours présente, j'ai réalisé que l'autonomie des étudiants était un point essentiel à travailler, et qu'il avait été sous-estimé jusqu'ici. Nous vivons à une époque où les savoirs sont disponibles en ligne, sous tous les médias possibles ; livres, cours, vidéos, encyclopédies en ligne... Ce n'est plus tant enseigner le contenu tel quel qui importe, mais donner une façon de l'appréhender, une indépendance dans les apprentissages, et au-delà développer l'esprit critique.

Ce portfolio sera organisé autour des trois axes que j'ai mentionnés auparavant ; l'évaluation des apprentissages, qui a été ma première approche, puis la conception d'un cours et l'animation de groupes, qui m'ont permis d'explorer des nouvelles voies. Les annexes contiennent des détails sur les différents enseignements dont j'ai eu la responsabilité. Dans une dernière partie, je mentionnerai les perspectives qui découlent de ces trois années d'expérience en enseignement, en gardant à l'esprit tout ce qu'il me reste à apprendre, à lire, à expérimenter, si je veux poursuivre dans cette voie.

Première partie

Axe 1 : Évaluer des apprentissages

Lorsque j'ai commencé à enseigner, je sortais tout juste de Master, et j'avais pu expérimenter l'effet délétère que pouvait avoir les examens sur la façon d'apprendre. A savoir, les efforts vont se concentrer sur l'obtention de la note, puisque c'est cela qui conditionne notre avancée. Sauf que l'examen terminal est parfois en inadéquation avec les véritables objectifs du cours, ou même avec ce qui est fait pendant les séances. La plupart des étudiants se retrouvent donc dans un cas de figure que Rolland Viau dénommait stratégie du minimax : «le minimum d'effort pour le maximum de points» [9].

Généralement, la faute est entièrement mise sur le compte des étudiants, qui ne seraient pas assez motivés, voire fainéants. La réalité est plus complexe que ça. Lors de plusieurs séances du Label, ou encore dans une formation sur les évaluations [6], la notion d'alignement pédagogique a été évoquée. Introduite par Biggs [3], puis Leclercq [8], cet alignement concerne la triple concordance entre les objectifs du cours, les activités mises en place durant les séances, et l'évaluation. Pour qu'un cours garde en cohérence, il est donc nécessaire d'avoir en tête ces trois notions. Un désalignement peut ainsi mener aux problèmes observés aujourd'hui, à savoir une démotivation, et un manque d'intérêt pour le cours. On l'entend d'ailleurs assez souvent de la part des étudiants ; qu'ils ne savent pas «à quoi ça sert», ou qu'ils demandent sans arrêt «ce qu'il y aura à l'examen». Si les objectifs ne sont pas explicités, la perception de valeur, telle qu'énoncée par Viau [9]devient faible, et le seul objectif que les étudiants se donneront alors sera de décrocher la moyenne et de passer le semestre. Plusieurs solutions nous ont été présentées lors du Label, et je me suis concentrée sur deux outils :

— La grille critériée ;
— L'évaluation formative, notamment au travers de QCM.

1 Utiliser la grille critériée

Entre mes premiers TPs et les derniers cours, je suis passée d'une grille d'évaluation purement informative, et pas nécessairement partagée avec les étudiants, à une grille d'objectifs par niveaux, qui permet aux étudiants de se positionner tout le long du cours, jusqu'à l'examen terminal. Le stage d'Autrans [7], et surtout les TPs d'optique donnés à l'IUT 1 Mesures Physiques (voir Annexe A4), m'ont permis de réellement prendre la mesure de comment s'écrit une grille critériée, et comment elle est utilisée pour optimiser l'apprentissage.

1.1 Évaluer les travaux pratiques

Lorsque je suis arrivée à l'IUT 1, l'équipe pédagogique d'Irène Ventrillard continuait à s'interroger sur la manière d'évaluer les étudiants en travaux pratiques. La nouveauté de cette année consistait à décorréler la partie rédaction de compte-rendus des apports scientifiques et pratiques. Ainsi, les étudiants devaient rendre à chaque fin de séance un compte-rendu, mais ils n'étaient évalués que sur la rédaction. Les compétences expérimentales devaient quant à elles être évaluées séparément dans une dernière séance d'examen pratique.

Plusieurs raisons soutenaient cette méthode ; pendant une séance de TP, les étudiants sont encore en situation d'apprentissage, et souvent le format du compte-rendu est négligé.

Au premier semestre, ils avaient donc eu des cours sur la rédaction de compte-rendu, dont les principaux éléments étaient rappelés dans des fiches disponibles en salle de TP. D'un autre côté, chaque compétence relative à la rédaction, par exemple les graphiques ou les incertitudes, était explicitée dans une grille critériée

en quatre niveaux (voir la grille complète en suivant ce lien). Chaque semaine, lorsqu'on leur rendait leur compte-rendu, ils pouvaient donc se positionner par rapport à cette grille, et savoir comment s'améliorer.

Etant donné que les TPs ont été interrompus par le confinement du printemps 2020, nous n'avons pas pu mener l'intégralité de ce qui était prévu. Après discussion au sein de l'équipe pédagogique, les étudiants devaient donc reprendre leur premier compte-rendu, en suivant la grille et les retours que nous leur avions déjà donnés lors de la seconde séance. Une fois évalués, nous leur avons renvoyé leur note et où leur CR se positionnait sur la grille. Un exemple de retour que j'ai donné est montré dans la figure 1.

Thème	N0	N1	N2	N3	Coef	
Organisation générale				1	3	Mise en forme parfaite !
Introduction				1	1	Introduction très bien étoffée, bravo.
Description disp exp				1	1	Bel effort pour les schémas, continuez comme ça pour vos futurs CR / rapports de stage etc.
Protocole			1		3	Attention à ne pas mélanger protocoles et explications. Les étapes étaient là, mais il aurait fallu les mettre à part (protocole = recette).
Tableaux / Unités					0	
Graphiques					0	
Décrire résultats			1		1	Présentation via « Hypothèses » et « Observations » très bien, continuez comme ça.
Confronter avec théorie			1		1	Même remarque. Il manque juste une discussion des limites / difficultés lors des observations pendant le TP, et une explication de ces possibles limites / difficultés (notamment pour le cas différencier lentille convergente / divergente).
Conclusion			1		1	Conclusion très bien étoffée, bravo.
Incertitudes					0	

Figure 1 – Retour donné sur un compte-rendu de TP en suivant la grille critériée pour la rédaction. Les poids de chaque thématique varient selon les TPs, par exemple le premier TP ne fait pas intervenir d'incertitudes, mais met l'accent sur la rédaction de protocole. Ces coefficients sont les mêmes pour tous les groupes de TPs et ont été décidés par Irène Ventrillard lors de la rédaction des énoncés.

Même si la note n'a finalement pas compté, le but était que les étudiants ressortent de ces TPs en sachant rédiger un compte-rendu. L'étudiant à qui j'avais envoyé ce retour m'avait d'ailleurs répondu par mail :

> « Un très grand merci pour votre retour encourageant ! Je tiendrai compte de vos remarques pour mes prochain CR et rapports. »

Si nous avions suivi une approche traditionnelle, probablement aurions-nous eu plus de mécontentement du fait que le travail n'était finalement pas noté. Mais puisque nous avons été transparents sur les objectifs du TP dés le début, ainsi que pendant le confinement, où nous avions communiqué avec les étudiants en ligne (salons Discord notamment), ils en ont retiré une meilleure perception de valeur, ainsi que de compétence, au sens de Viau [9]. Les travaux rendus étaient pour la majorité bons, avec une moyenne de 12, ils avaient fait l'effort de réécrire leur CR, et ils avaient conscience d'avoir acquis des compétences, à plus long terme que la simple note.

1.2 Vers un meilleur alignement pédagogique en CM/TD

En suivant cet exemple, j'ai voulu adapter cette méthode à l'évaluation du cours sur les séries et transformées de Fourier, que j'avais commencé à encadrer l'année d'avant (voir Annexe A2). Cet enseignement appartient à une UE de mathématiques pour la physique, gérée par Elisabeth Charlaix, qui en assure la première partie traitant de l'algèbre linéaire. En général, cette UE a toujours présenté des difficultés pour la majorité des étudiants. Une des causes est le faible niveau en mathématiques, qui provient même d'avant le baccalauréat, et un manque d'entrainement en calculs. La question de la motivation y est cruciale, car cet enseignement nécessite un travail régulier, voire répétitif, en-dehors des séances. De plus, les étudiants ont souvent du mal à voir l'intérêt du cours dans leur poursuite d'études, alors que ces bases sont essentielles.

La première année où j'avais enseigné cette partie, j'avais fait l'erreur de ne pas distribuer de grille critériée, même si j'en avais utilisé pour les devoirs maison ou l'examen terminal (voir barème à ce lien). Ainsi les étudiants avaient pu avoir des bons retours, mais ils n'avaient pas eu de vision dés le début sur les objectifs du cours, un des trois points essentiels de l'alignement pédagogique selon Biggs [3] et Leclercq [8].

J'ai voulu corriger cela la deuxième année, et j'ai ainsi repensé le cours/TD et ses objectifs, puis résumé cela dans une grille critériée rassemblant des compétences générales, des éléments correspondant aux séries ou aux transformées de Fourier (disponible ici). Chaque exercice de TD indiquait ensuite quelles compétences étaient mises en oeuvre, comme montré dans ce TD en annexe. De cette façon, les étudiants pouvaient se positionner au fur et à mesure du semestre, mais aussi comprendre dans quel but on faisait tel ou tel exercice pendant les séances. Par ailleurs, j'avais couplé cela à des versions Moodle de certaines questions, ce qui leur donnait la possibilité de s'entrainer eux-mêmes sur les exercices non corrigés en cours, et où ils estimaient devoir s'améliorer.

Lors du questionnaire d'auto-évaluation, alors que ma partie ne couvre que la moitié du semestre, qu'elle comportait déjà quatre TDs, dont la moitié des exercices étaient laissés à faire en autonomie, les seize étudiants ayant répondu ont tous jugé qu'il aurait fallu « un peu plus » d'exercices et d'applications. D'un autre côté, lorsque je les ai interrogés sur leur degré d'implication personnelle (participation en cours, travail à la maison...), 69% l'ont jugé « assez important », contre 25% à la considérer « peu important », et 6% « très important », mais 0% à avoir répondu « pas important du tout ». J'avais donc une petite fraction d'étudiants qui s'investissaient énormément, ce que je ressentais par mail où j'avais des questions sur les TDs non corrigés, et tout de même deux tiers des étudiants qui s'impliquaient plus que le minimum.

En résumé, une grille critériée en lien fort avec les TDs et des exercices en ligne, ont poussé à un travail plus régulier, et en-dehors des séances. Plus important, sur tous les cours que j'avais encadrés en deux ans, ou tous ceux que j'avais pu connaître en tant qu'étudiante, c'est le premier pour lequel je n'ai pas entendu la fameuse question « Qu'est-ce qu'on aura à l'examen ? ». Ça peut paraître anecdotique, mais présenter clairement ce que j'attendais d'eux, montrer que les TDs et le cours sont en conformation avec ces objectifs, leur ont donné plus confiance envers les activités proposées, et surtout en l'évaluation finale. Ils savaient qu'ils auront un retour rapide s'ils travaillaient régulièrement, et leur motivation en a été plus importante. D'une année à l'autre sur cet enseignement de mathématiques, la moyenne est ainsi passée de 9 à 12 à l'examen terminal, pour un sujet de difficulté similaire, grâce entre autres à un meilleur alignement pédagogique.

2 Mettre en avant l'évaluation formative

Dans la plupart des cours où j'ai enseigné, l'évaluation officielle ne tenait qu'à un examen terminal, ce qui limitait les retours que l'on pouvait donner aux étudiants. Lors du stage d'Autrans [7], ou de la formation « Evaluer les apprentissages des étudiants » [6], l'accent a été mis sur l'évaluation dite formative. Effectuée tout au long du cours, elle permet de suivre la progression et de donner des retours immédiats. Plusieurs formats existent, et j'en ai utilisé principalement trois : le devoir-maison (DM), le questionnaire à choix multiples (QCM), et les possibilités d'exercices en ligne offertes par Moodle.

2.1 Le devoir-maison ou favoriser le travail en-dehors des cours

Lors de mes vacations en CM/TD de mathématiques pour la physique, j'ai organisé les deux années des devoirs-maisons. Cela me permettait de jauger ce qu'ils avaient compris du cours, de voir où les principales erreurs se faisaient, et de pouvoir vite adapter le cours et les TDs. Je notais ces devoirs sur deux points, qui étaient ajoutés en bonus à la note de ma partie à l'examen terminal.

De cette expérience, j'avais constaté plusieurs choses ; d'abord le temps de correction était non négligeable. D'une année à l'autre, j'ai ainsi réduit le rendu de DM chaque semaine, à un seul à faire pendant les vacances de la Toussaint sur un exercice plus long (application de préférence), et laisser les QCM pour les retours sur calculs plus courts. Ensuite, il était évident que les étudiants travaillaient en groupe, c'est

pourquoi j'ai limité l'impact de la note de DM. De plus, la deuxième année, j'avais fait le retour sur le DM grâce à la grille critériée, afin qu'ils aient une idée de comment elle était utilisée, et pour qu'ils se positionnent sur les attendus à la moitié du cours.

En général, les DMs étaient bien accueillis de la part des étudiants, qui en demandaient même parfois plus. Un élève m'a même proposé dans le questionnaire d'auto-évaluation de donner des DMs d'une semaine à l'autre, mais avec une chance sur deux de devoir le rendre. Cette idée pourrait être reprise à l'avenir, avec peut-être plusieurs exercices au choix tout le long du cours, en plus de l'application, afin de donner plus de perception de « controlabilité », au sens de Viau [9].

2.2 Les QCM pour un retour rapide

Lorsque j'ai commencé à enseigner, je m'étais promise de ne jamais évaluer sur des QCM. Après deux années, plusieurs formations, et le passage au distanciel, j'ai commencé à y voir un intérêt, non pas en tant qu'évaluation certificative, mais bien pour prendre le pouls de la classe à un instant T. Je reviendrai sur l'apport des QCM sur d'autres aspects du cours, notamment pour animer des groupes.

Ainsi, dans l'enseignement des mathématiques pour la physique, j'ai utilisé la deuxième année des QCM Moodle et Kahoot pour la partie sur les transformées de Fourier, par exemple la convolution ou les opérations. Certaines questions correspondaient même à certaines parties du TD. Cela m'a permis de voir rapidement où les erreurs se faisaient systématiquement, et de pouvoir cibler certaines explications.

J'avais commencé ce cours en insistant auprès des étudiants que je voulais qu'ils fassent des erreurs pendant les séances, pour justement savoir comment s'améliorer et ne pas les faire à l'examen. Néanmoins, l'erreur a un certain statut qu'il est difficile de dépasser pour les étudiants, comme la plupart des cours la considère comme pénalisante. Utiliser les QCM, et notamment sous forme de jeux-cadres comme avec Kahoot, a permis de dépasser ça, et de pouvoir diagnostiquer les erreurs, comme Astolfi le présente dans un de ses ouvrages [2].

Evidemment, la part des réponses au hasard n'est pas négligeable, et a même été pointé du doigt par certains étudiants lors du questionnaire d'auto-évaluation. Néanmoins, le format du QCM a en général été apprécié des étudiants. A la question de quels éléments les avaient le plus aidé à suivre le cours, les QCM revenaient dans au moins 40% des réponses.

A l'avenir, je souhaiterais pousser plus loin dans cette direction, ajouter des solutions générales implicites si possibles, ou de trouver des applications qui permettent d'écrire des réponses (à l'instar d'une «ardoise numérique mais anonyme»), afin de regrouper les réponses au tableau rapidement. Une autre idée serait d'envoyer les étudiants expliquer telle ou telle erreur. Même s'ils ont eu bon, comprendre pourquoi leurs collègues font certaines erreurs constitue une étape supplémentaire dans l'apprentissage.

2.3 Moodle et la possibilité d'une banque d'exercices en ligne

Comme je l'avais expliqué précédemment, pendant l'automne 2020 j'avais converti certains de mes TDs en version Moodle en ligne, où les étudiants pouvaient entrer leurs réponses et rapidement savoir s'ils avaient bon ou pas, accompagnés d'indications où corriger.

En général, les étudiants avaient donné de bons retours sur cette idée, notamment « pour éviter de rester bloqué », mais pointaient du doigt les nombreux bugs liés au formattage trop strict des questionnaires Moodle.

En parallèle, la responsable de l'UE, Elisabeth Charlaix, a poussé au développement des questions Stack, et souhaiterait continuer dans cette direction. L'objectif serait de créer des banques de questions en ligne pour que les étudiants travaillent plus, et gagnent en autonomie concernant les mathématiques en physique.

Deuxième partie

Axe 2 : Concevoir un enseignement

Dans deux cours (voir Annexes A2 et A3), j'ai eu la responsabilité, en partie ou totale, de la conception d'un enseignement. Par là, j'entends le choix partiel ou total des activités proposées, de certains contenus du cours, et de comment découper et organiser tous ces éléments. J'ai affronté plusieurs problématiques, et je vais mettre en parallèle ces deux enseignements ;

1. L'un en mathématiques pour la physique pour une vingtaine d'étudiants en Licence 3 (automne 2019 et 2020, voir Annexe A2), où j'avais quasiment une liberté totale. Ce cours a dû être adapté en distanciel en 2020 ;
2. Le second en opto-électronique pour environ quatre-vingt étudiants en deuxième année d'IUT (début 2020), où le support existait déjà mais permettait quelques adaptations (voir Annexe A3).

1 Aborder le contenu

Dans les deux cas, j'ai dû passer énormément de temps sur le contenu, car les disciplines étaient complètement différentes de mon sujet de thèse. Je travaille en biophysique en utilisant des neutrons (spectroscopie), des électrons (microscopie) ou des méthodes thermodynamiques, et je ne fais ainsi ni des mathématiques aussi fondamentales (même si j'applique les séries et transformées de Fourier au quotidien), et encore moins de l'opto-électronique. Enseigner dans ces disciplines représentait donc un premier défi dans la manière d'aborder le contenu quand je n'étais pas moi-même spécialiste. La communication avec les responsables de cours a été essentielle en ce sens, notamment pour comprendre les enjeux et le contexte, et structurer l'enseignement à partir de là.

Dans le cas des mathématiques, l'objectif est de donner à des étudiants en fin de licence physique-chimie les outils nécessaires pour d'autres disciplines, comme l'optique, où les séries et transformées de Fourier sont essentielles. Le contexte était similaire à celui que j'avais moi-même connu en licence, et après discussion avec Elisabeth Charlaix, il a été décidé de mettre en avant la maîtrise des calculs, et de travailler sur des applications.

Pour gérer le temps au mieux (je n'avais que 13.5 heures en tout), et me conformer à ces objectifs, j'ai pris la décision de ne pas forcément exposer les démonstrations, qui se trouvent cependant dans le support de cours initial. J'ai conscience que ces choix ne sont pas à prendre à la légère, et que réduire toute une discipline à un rôle purement utilitaire n'est pas anodin. Je l'ai donc bien spécifié aux étudiants dés le début du cours, ainsi que pendant les exposés, où j'indiquais où trouver les démonstrations, et comment les amorcer eux-mêmes s'ils le souhaitaient.

Dans le cas de l'opto-électronique, le cours était fortement lié au TD, dont j'encadrais un groupe, et à des TPs. Il était donc nécessaire de garder la cohérence de l'ensemble, et Guillermo Martin m'a expliqué que la modulation électro-optique constituait le coeur du cours, ainsi que des TDs, et des TPs y étaient consacrés. J'ai donc gardé le plan de cours, et mis l'accent sur cette partie, qui comportait par ailleurs des parties plus compliquées pour des étudiants de deuxième année, mettant en œuvre des tenseurs et des matrices.

2 Travailler le support de cours

Une des questions qui revient souvent est celle du support de cours, notamment le choix entre le tableau ou l'utilisation de diapositives. Dans les deux cours, j'ai eu à expérimenter les différents types de support.

D'un côté, le tableau m'a facilité les schémas, ou les démonstrations faisant intervenir des équations. Il permet aussi de rebondir plus facilement aux questions des étudiants. Lors du passage au distanciel à l'automne 2020, j'ai tenu à en garder un, et j'ai utilisé une version minimaliste couplée à une Speechicam (voir figure 2).

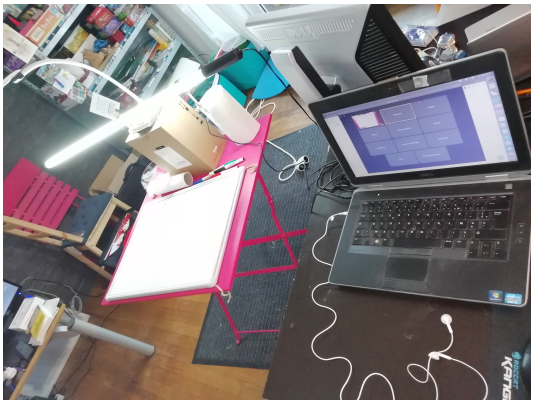


Figure 2 – Adaptation du tableau en distanciel à l'aide d'une Speechicam.

D'un autre côté, les diapositives permettent de projeter sur un plus grand écran, et d'être accessibles à tous les étudiants dans un amphithéâtre, ce qui était le cas en opto-électronique, et constituaient le support de cours initial envoyé par Guillermo Martin.

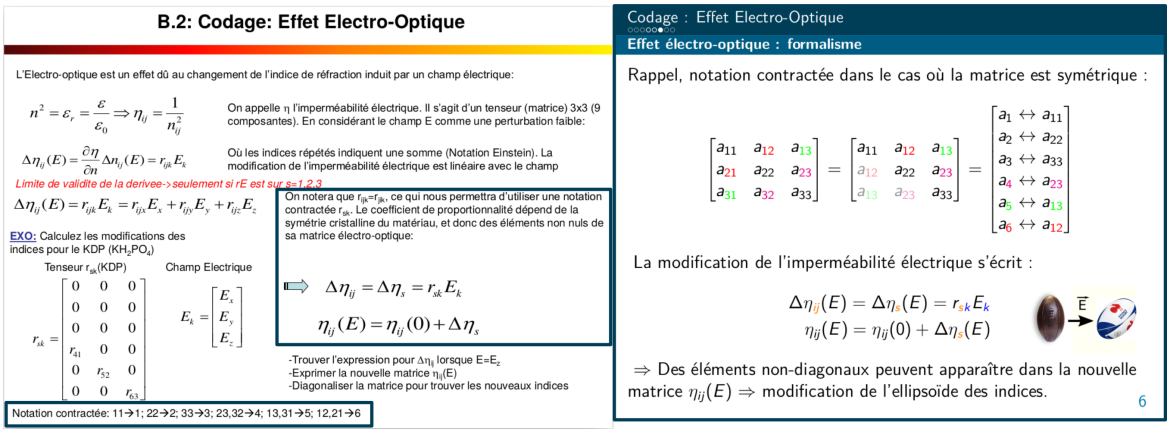


Figure 3 – Adaptation d'une diapositive en opto-électronique. A gauche, le support initial, dont les parties encadrées en bleu sont montrées dans leur nouvelle version à droite. Version agrandie en Annexe A3.

Néanmoins, j'ai pu constater après le premier amphithéâtre que quand je couplais les diapositives au tableau, si c'était trop chargé, les étudiants recopiaient plus que m'écoutaient. Ils pouvaient rater certains schémas ou explications. J'ai donc choisi d'adapter les diapositives, notamment concernant les parties importantes. J'ai privilégié des slides plus épurées, avec les informations apparaissant pas à pas, suivant ce que je présentais. Concernant les tenseurs et matrices, j'ai tenté de rendre ça plus visuel, afin de faciliter la compréhension, par exemple des différents indices. Un exemple d'adaptation est montré dans la figure 3.

J'ai conscience qu'en ce temps-là, je me suis posée en tant que facilitatrice d'apprentissage, et j'ai conservé une manière frontale d'exposer le cours. Lors du deuxième amphi, l'adaptation des slides a été bien perçue, et lorsque j'avais demandé s'ils trouvaient le cours plus clair, la majorité avait répondu oui (questionnaire léger de type « minute paper » comme présenté lors d'Autrans [7]). Cependant, ce qu'ils avaient surtout retenu était le fait d'avoir travaillé pendant l'amphi, et j'aurais dû mettre l'accent dessus dés le début avec des apprentissages actifs, comme préconisé par l'ouvrage de Markus Brauer [4], ou mentionné lors du stage

d'Autrans [7]. Cela mériterait d'être discuté à l'avenir avec Guillermo Martin si l'occasion se présentait.

Enfin, avec le passage au distanciel en 2020, j'ai été obligée d'adapter le cours sur les transformées de Fourier. Le tableau ne pouvait pas suffire à tout inclure, et je l'ai conservé pour les schémas et les interactions avec les étudiants. Mais pour le contenu principal, il était nécessaire de l'adapter en diapositives. J'ai repris les principes que j'avais suivis en opto-électronique : slides épurées, avec les informations principales, apparition progressive, et rendre les équations plus visuelles (voir exemple dans la figure 4).

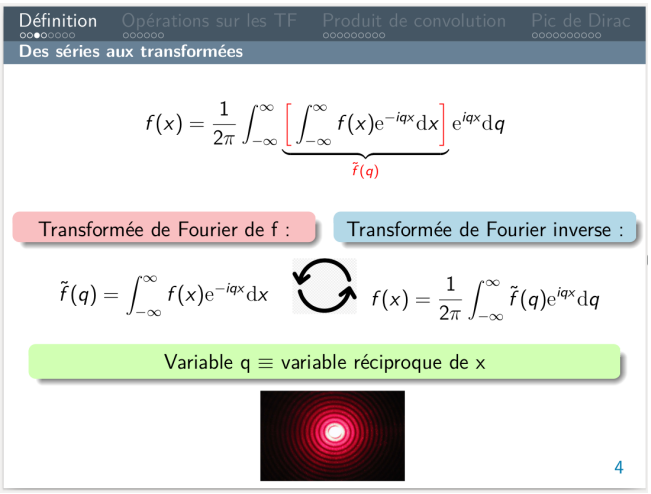


Figure 4 – Exemple d'une diapositive du cours de transformées de Fourier à l'automne 2020.

Lors du questionnaire d'auto-évaluation, certains étudiants ont mis en avant « la clarté des diapos », ou le fait que le cours était « bien structuré [...] Et il n'y avait aussi pas des tonnes et des tonnes d'informations, juste celles essentielles ».

3 Découper le cours en séquences

C'est après une formation sur l'enseignement donnée par l'ADEA [5], en octobre 2020, que j'ai pleinement pris conscience de la différence et de l'équilibre à trouver entre contenu, processus pédagogiques et sens. Après avoir mis l'accent sur le contenu, ce que j'ai présenté dans les deux précédentes sections, j'ai réalisé que je devais contrebalancer par plus d'activités, et surtout que je devais revoir l'organisation même de mes cours.

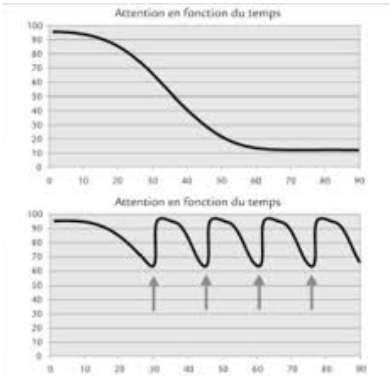


Figure 5 – Graphique de l'attention en fonction du temps (adaptée de [4]). Les flèches indiquent les moments où les changement d'activités recaptent l'attention des étudiants.

La figure 5 (haut) ci-dessous montre de manière directe que l'attention de n'importe quel étudiant – et tout simplement tout être humain – baisse au bout de quinze minutes. Prendre conscience de ça, et mettre en

parallèle avec les deux enseignements donnés fin 2019 et début 2020, m'ont vraiment poussé à aller au-delà d'un cours uniquement frontal. La formation de l'ADEA [5], ainsi que le livre de Brauer [4],m'ont permis d'envisager un enseignement non plus comme un « cours magistral », mais comme une réelle « mise en scène ». Une idée est donc de découper son cours en séquences, de trente minutes environ, alternant exposés et activités, pour recapter l'attention, comme montré dans la figure 5 (bas).

Cette question de l'attention devenait par ailleurs cruciale avec le passage au distanciel. Tous les problèmes sont exacerbés lors d'un cours à distance. C'est encore plus difficile de se concentrer sur un écran que lors d'un véritable cours, on peut facilement faire autre chose, et les interactions sont diminuées. Je ne parle même pas des nombreux problèmes liés à la précarité numérique, du fait que certains étudiants partagent un espace de travail réduit avec plusieurs autres personnes (famille, colocataires ou conjoint.e), ou encore des états anxieux et dépressifs qu'ils peuvent traverser. Il était clair pour moi qu'en aucun cas je voulais motiver les étudiants de manière punitive, par exemple en les forçant à allumer leur webcam.

Lors du cours de mathématiques pour la physique de 2020, la première chose que j'ai faite a été de revoir chaque séance, et de les découper en séquences. Un extrait de planning écrit pour une séance de trois heures est donné dans la figure 6 ci-dessous.

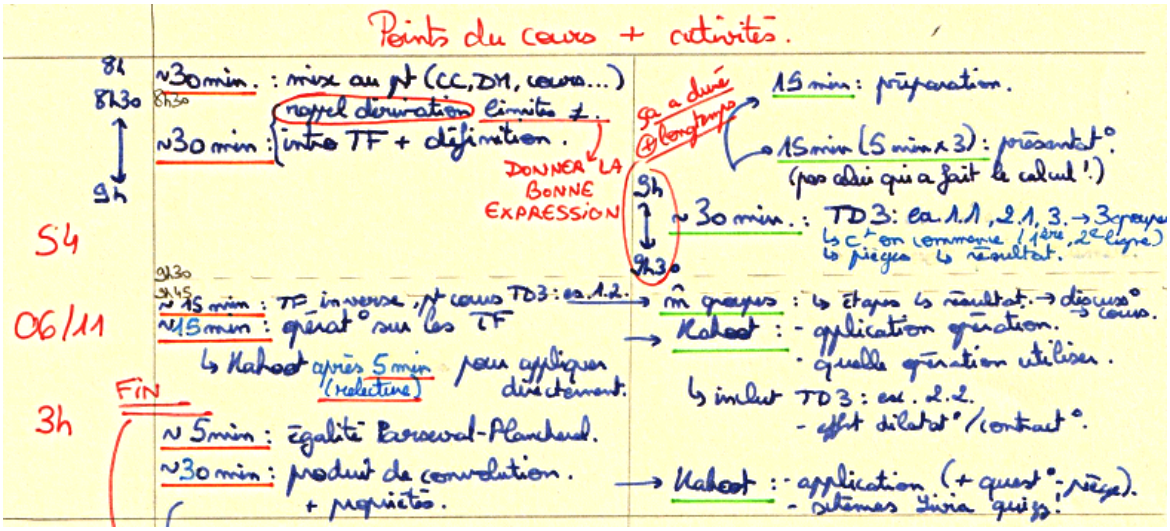


Figure 6 – Extrait d'un découpage d'une séance en séquences.

Les activités proposées seront détaillées dans l'axe sur l'animations de groupes. Mais c'est vraiment leur association avec un rythme imposé par les séquences qui ont favorisé une participation active. A la question de quels éléments avaient le plus aidé à suivre ce cours, un étudiant avait mis en avant :

> « Il y avait plusieurs "phases" dans une séance de cours ce qui évite de se lacer d'une lecture de diapo "classique". Autrement dit on faisait pas 1h30 de diapo, il y a des parties interactives, évidemment des moments ou on fait du cours mais ce n'était pas barbant car on ne faisait pas que ça. Et cela donne envie de revenir la semaine d'apres pour participer au cours (ou essayer de battre M*** sur Kahoot) [...] » (sic) (les astérisques ont été ajoutés pour masquer le nom).

Les étudiants étaient aussi beaucoup plus réactifs pendant le cours, et aux points qu'ils faudrait améliorer, un élève avait même écrit : « on manque d'un tout petit peu de temps, 1h de cours en plus pour pouvoir bien finir les derniers exercices » (sic)

Alors que le cours durait déjà trois heures, et se déroulait un vendredi matin dés 8h. Introduire des « stimulations diverses », pour citer une réponse d'un étudiant, les a aidés à suivre le cours, à participer, et j'ai constaté la différence avec mes précédents cours.

Troisième partie

Axe 3 : Animer des groupes.

J'ai eu l'occasion d'enseigner à des groupes de tailles diverses, de petits groupes d'une dizaine d'étudiants, à des amphithéâtres de quatre-vingt élèves, et pour différents types de cours, travaux pratiques, dirigés ou cours magistral. La dynamique n'était jamais identique, et d'ailleurs même d'une année à l'autre, cela pouvait changer.

La question de l'animation de groupes est donc celle pour laquelle j'ai dû faire le plus preuve d'adaptation et d'improvisation. Néanmoins, c'est aussi le côté le plus humain, même le plus amusant dans certains cas, et finalement celui duquel j'ai retiré énormément de leçons.

1 Gérer l'espace, favoriser les échanges

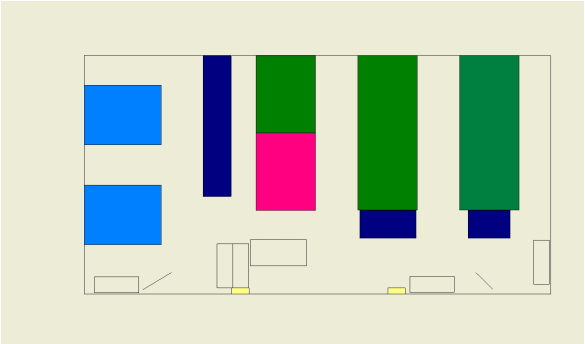
La gestion de l'espace s'est révélé assez importante concernant les travaux pratiques. J'ai encadré des TPs mettant en œuvre des bancs optiques dans deux cas différents ; les TPs photographie au DLST (voir Annexe A1), et des TPs d'optique à l'IUT (voir Annexe A4).

D'un côté, au DLST, la salle est tout en longueur, et les paillasses alignées les unes derrière les autres. Dans cette configuration, ceux qui étaient tout au fond m'entendaient mal, et ne voyaient pas le tableau. Dés que j'allais voir un groupe, ceux à l'autre extrêmité n'avait « plus de contact », et je passais mon temps à faire des aller-retours entre les huit groupes.

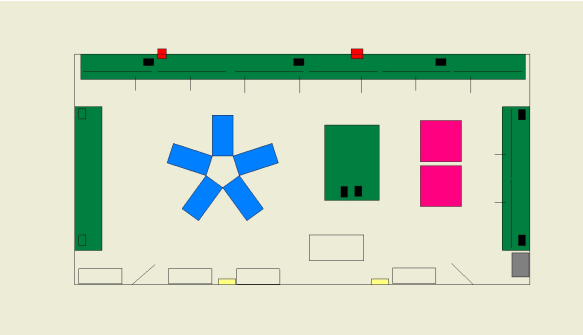
Lors de la première séance, j'ai vite remarqué que le groupe arrivé en dernier, appelons-le le groupe A, s'est retrouvé isolé au fond, et les deux étudiants avaient plus de difficultés à se mettre à travailler, même si je passais régulièrement. Lors de la deuxième séance, j'ai envoyé au fond le groupe qui avait bien avancé le TP précédent, en leur expliquant les conditions. Très vite, au cours de l'heure qui a suivi, j'ai pu voir que le groupe A avançait mieux, qu'ils avaient plus de contact avec leurs collègues et moi. En parallèle, le groupe du fond, qui avançait plus vite, n'était pas non plus isolé car leurs collègues venaient d'eux-mêmes demander conseil.
Cette simple adaptation a favorisé les échanges entre étudiants, ce qui a aussi permis d'équilibrer les avancements, sans que je sois constamment derrière eux. Par ailleurs, mettre en confiance le groupe A les a poussés à s'investir plus, et ils ont présenté un des meilleurs oraux de projet à la fin du TP.

D'un autre côté, à l'IUT 1 Mesures Physiques, Aurélie Potnisky, en charge de plusieurs salles de TP, de leur gestion et du matériel, a organisé la salle d'optique d'une manière plus centralisée. La figure 7 montre les plans de la salle avant et après aménagement. On peut voir dans la nouvelle version que les paillasses sont toutes en contact les uns avec les autres, on peut passer aisément d'un groupe à l'autre, sans laisser non plus d'étudiants isolés. Par ailleurs, le bureau et le tableau sont visibles de la plupart des tables, et assez d'espace reste pour regrouper les étudiants autour. Cela aide grandement à l'organisation des mini-oraux, qui nécessitent la projection de transparents, qu'Irène Ventrillard met en place pour ces TPs.
J'ai vraiment senti la différence dans cette configuration, je ne courais pas pour voir chaque groupe, je pouvais recentraliser l'attention des étudiants si nécessaire (notamment pour les oraux), et aucun groupe n'était isolé.

Ces deux exemples montrent que la participation active des étudiants, qui est nécessaire en TP, se réfléchit bien en amont, ou nécessite des adaptations rapides. L'implication des techniciens dans l'organisation de la salle (si la salle le permet évidemment), et dans l'équipe pédagogique en général, est ainsi un véritable atout.



Plan de la salle de TP initiale



Plan de la salle de TP finale

Figure 7 – Aménagement de la salle de TP d'optique par l'équipe pédagogique et technique de l'IUT 1 Mesures Physiques. Les paillasses avec bancs optiques sont représentées en vert, celles avec goniomètre en bleu, et le tableau ainsi que le bureau se trouvent au sud des deux plans. (Courtoisie d'Aurélie Potnisky).

2 Encourager le travail en groupe

Selon la pyramide de l'apprentissage, que j'ai découverte dans l'ouvrage de Brauer [4], représentée dans la figure 8 ci-dessous, les méthodes qui permettent le plus grand taux de mémorisation sont celles qui impliquent de discuter en groupe et d'enseigner aux autres. J'en ai moi-même fait l'expérience durant tout mon cursus, et je considère une grande partie de ma réussite due aux travaux de groupe que j'ai pu faire, et à toutes les fois où j'ai aidé des collègues.

Il est donc indispensable d'encourager les travaux en groupe. En plus de la dimension pédagogique, ils renferment une dimension éthique et humaine : favoriser la collaboration au-delà du « chacun pour soi », même si nous finissons par être en compétition d'une manière ou d'une autre (concours, postes...).

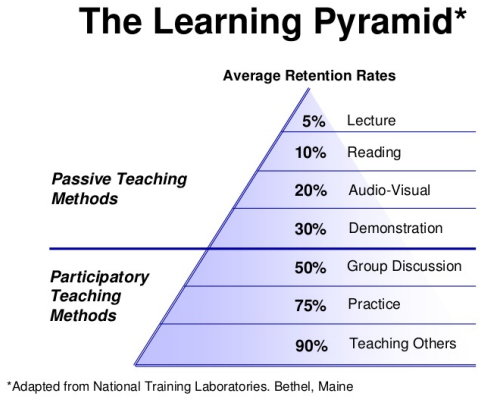
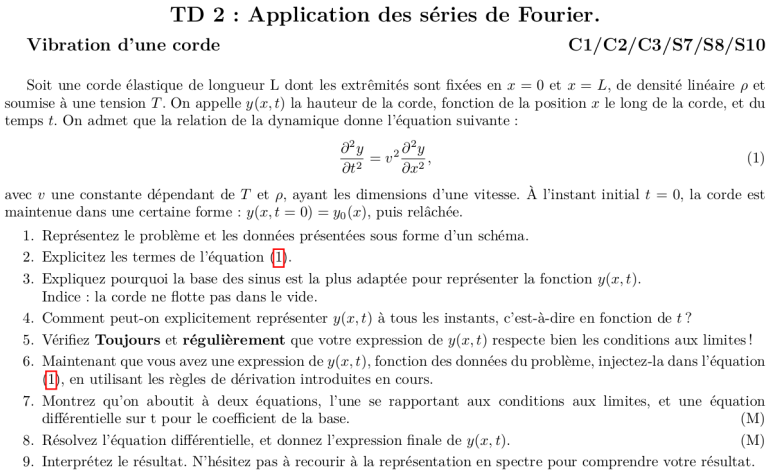


Figure 8 – Pyramide de l'apprentissage, affichant le taux de mémorisation selon différentes techniques d'enseignement. Source : National Training Laboratories, Bethel (Maine, USA).

Dans les premiers TDs que j'ai encadrés, en mathématiques pour la physique (Annexe A2) et en opto-électronique (Annexe A3), je voyais que le travail en groupe s'organisait toujours plus ou moins naturellement. Néanmoins, les groupes avançaient de manière inégale, je finissais toujours par corriger au tableau pour rester dans les temps, et parce que les étudiants le réclamaient.

Lors de la formation donnée par l'ADEA [5], cette question avait été abordée, et la formatrice nous avait présenté, et mis en application avec nous, les jeux-cadres de Thiagi (voir par exemple la thiagipedia en ligne [1]). En 2020, en plus d'adapter le cours de mathématiques en séquences, j'ai revu la manière d'encadrer les TDs, et d'inclure ces jeux cadres pour les travaux en groupe.

Le premier exemple est celui d'un exercice d'application (voir figure 9 ci-dessous). J'ai séparé la classe en quatre groupes (ils étaient trois étudiants par groupe). Comme les quatre premières questions étaient indépendantes, chaque groupe a dû se concentrer sur une question pendant cinq à dix minutes. Ensuite, un des étudiants devait présenter la réponse de son groupe au tableau, et répondre aux éventuelles questions de ses collègues.



TD 2 : Application des séries de Fourier.

Vibration d'une corde **C1/C2/C3/S7/S8/S10**

Soit une corde élastique de longueur L dont les extrêmités sont fixées en $x = 0$ et $x = L$, de densité linéaire ρ et soumise à une tension T. On appelle $y(x,t)$ la hauteur de la corde, fonction de la position x le long de la corde, et du temps t. On admet que la relation de la dynamique donne l'équation suivante :

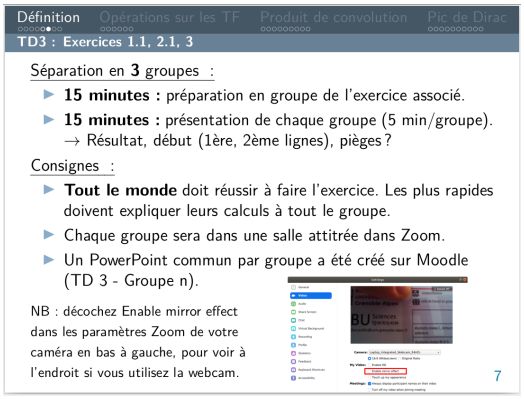
$$\frac{\partial^2 y}{\partial t^2} = v^2 \frac{\partial^2 y}{\partial x^2}, \quad (1)$$

avec v une constante dépendant de T et ρ, ayant les dimensions d'une vitesse. À l'instant initial $t = 0$, la corde est maintenue dans une certaine forme : $y(x, t = 0) = y_0(x)$, puis relâchée.

1. Représentez le problème et les données présentées sous forme d'un schéma.
2. Explicitez les termes de l'équation (1).
3. Expliquez pourquoi la base des sinus est la plus adaptée pour représenter la fonction $y(x,t)$.
   Indice : la corde ne flotte pas dans le vide.
4. Comment peut-on explicitement représenter $y(x,t)$ à tous les instants, c'est-à-dire en fonction de t ?
5. Vérifiez **Toujours** et **régulièrement** que votre expression de $y(x,t)$ respecte bien les conditions aux limites !
6. Maintenant que vous avez une expression de $y(x,t)$, fonction des données du problème, injectez-la dans l'équation (1), en utilisant les règles de dérivation introduites en cours.
7. Montrez qu'on aboutit à deux équations, l'une se rapportant aux conditions aux limites, et une équation différentielle sur t pour le coefficient de la base. (M)
8. Résolvez l'équation différentielle, et donnez l'expression finale de $y(x,t)$. (M)
9. Interprétez le résultat. N'hésitez pas à recourir à la représentation en spectre pour comprendre votre résultat.

Figure 9 – Exercice d'application en mathématiques pour la physique.

Un second exemple concerne plusieurs exercices de calculs de transformées de Fourier, qui appliquaient directement le cours, mais sur différentes fonctions. Nous étions cette fois-ci en distanciel, et j'ai séparé en trois groupes répartis dans des salles sur Zoom. La figure 10a ci-dessous montre les consignes données, et la figure 10b un exemple de slide d'un groupe.



Définition Opérations sur les TF Produit de convolution Pic de Dirac

TD3 : Exercices 1.1, 2.1, 3

Séparation en **3** groupes :

- **15 minutes :** préparation en groupe de l'exercice associé.
- **15 minutes :** présentation de chaque groupe (5 min/groupe).
  → Résultat, début (1ère, 2ème lignes), pièges ?

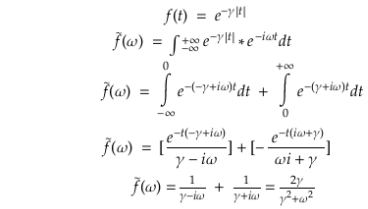
Consignes :

- **Tout le monde** doit réussir à faire l'exercice. Les plus rapides doivent expliquer leurs calculs à tout le groupe.
- Chaque groupe sera dans une salle attitrée dans Zoom.
- Un PowerPoint commun par groupe a été créé sur Moodle (TD 3 - Groupe n).

NB : décochez Enable mirror effect dans les paramètres Zoom de votre caméra en bas à gauche, pour voir à l'endroit si vous utilisez la webcam.

7

(a) Consignes



$$f(t) = e^{-\gamma|t|}$$
$$\tilde{f}(\omega) = \int_{-\infty}^{+\infty} e^{-\gamma|t|} * e^{-i\omega t} dt$$
$$\tilde{f}(\omega) = \int_{-\infty}^{0} e^{-(-\gamma+i\omega)t} dt + \int_{0}^{+\infty} e^{-(\gamma+i\omega)t} dt$$
$$\tilde{f}(\omega) = [\frac{e^{-t(-\gamma+i\omega)}}{\gamma - i\omega}] + [-\frac{e^{-t(i\omega+\gamma)}}{\omega i + \gamma}]$$
$$\tilde{f}(\omega) = \frac{1}{\gamma - i\omega} + \frac{1}{\gamma + i\omega} = \frac{2\gamma}{\gamma^2 + \omega^2}$$

(b) Exemple d'une slide d'un groupe

Figure 10 – Adaptation d'exercices en distanciel et travaux de groupe.

Que ce soit en présentiel ou distanciel, les étudiants s'impliquaient plus, et gagnaient en confiance. Je leur avais dit à l'oral que de préférence, celui qui avait calculé le plus vite ne devait pas être la personne qui devait présenter. Cela forçait les élèves avec plus de facilités à expliquer leur raisonnement, à leurs camarades de pouvoir comprendre, pour à leur tour présenter aux autres groupes. Dans la pyramide figure 8, cela signifiait que chaque étudiant passait par les trois méthodes actives d'apprentissage : discussion en groupe, pratique et présentation aux autres.

De plus, en encourageant à travailler par petits groupes à distance, la moitié des étudiants activaient leur webcam dans les salles, branchaient leur micro et participaient. Quand je passais de salle en salle, je pouvais donc directement communiquer avec eux, et le cours/TD était moins frontal.

A la question de quels éléments les avaient le plus aidé à suivre, dans 50% des réponses, les étudiants mentionnaient les travaux en groupe, et dans 30% des cas, ils mettaient en avant le côté interactif.

J'avais également posé la question suivante : « Est-ce que vous avez le sentiment de sortir de ces six séances de cours en ayant appris quelque chose ? Cela englobe autant des éléments de théorie, que des savoir-faire (calculs...) ». 87.5% ont répondu « Oui, totalement », contre 12.5% « Un peu ». Aucun n'a choisi « Pas vraiment » ou « Pas du tout ».

Cela s'est d'ailleurs ressenti dans les copies d'examen, où la moyenne était de 12. Pour une compétence travaillée en groupe (par exemple le calcul de transformées de Fourier), plus de la moitié des étudiants avaient le niveau maximal, c'est-à-dire qu'ils savaient en calculer une, et sans erreurs.

3 Dynamiser avec des jeux

Certains jeux-cadres encouragent le travail en groupe, mais d'autres, sans forcément se faire par groupe, peuvent apporter un côté ludique, non négligeable dans la motivation des étudiants.

Pour dynamiser les cours à distance, et prolonger l'apport des QCM dans les apprentissages, j'ai créé des QCM Kahoot qui reprenaient des éléments de cours et même des questions de TDs, concernant les opérations sur les transformées de Fourier, ou la convolution (voir planning figure 6).

Je leur laissais deux à trois minutes par questions, et je faisais en sorte de mettre en distracteurs des erreurs communes qu'ils pouvaient faire, ou que j'avais pu voir dans des copies des années précédentes. La part laissée au hasard restait cependant importante, il aurait fallu pouvoir inclure des solutions générale implicites pour aller plus loin (peut-être dans la version pro ou d'autres logiciels).

Néanmoins, le côté ludique a permis plus de spontanéité dans les réponses aux questions. Certaines étaient des questions de TD, et si je les avais posées en temps normal, les étudiants auraient mis plus de temps à réfléchir, ou n'auraient même pas osé écrire quelque chose sur leur feuille, de peur de se tromper. Alors que dans un format jeu, ils étaient moins hésitants, et cela m'a permis de rapidement diagnostiquer des erreurs et leur permettre de se corriger.

La moitié des étudiants ont d'ailleurs mis ces quizz en avant quand je leur avais demandé ce qu'ils avaient préféré dans cet enseignement. Ils les associaient à un « côté interactif », « d'apprendre plus facilement », ou « de voir tout de suite quelles formules sont importantes ».

Enfin, dans un genre différent, mais qui a très bien fonctionné, j'avais lancé par binômes un jeu des questions en fin de cours, dont le résumé peut être trouvé dans la figure 11 ci-dessous.

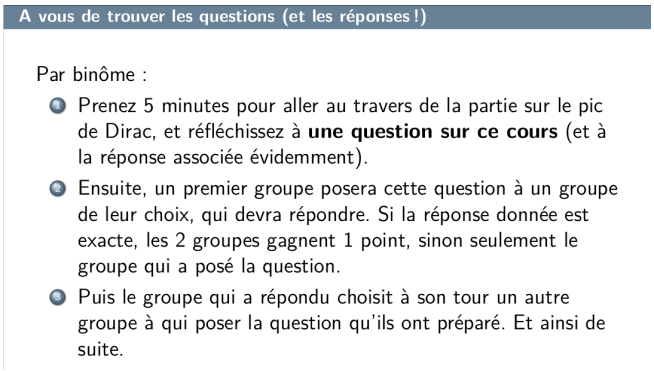


Figure 11 – Jeu des questions en fin de cours.

J'avais demandé à la moitié des groupes une question sur les séries de Fourier, et l'autre moitié sur les transformées. De cette façon, ils pouvaient survoler de nouveau le cours, et se concentrer sur des parties

puisqu'ils devaient créer une question. Des exemples de question sont montrés ci-dessous :

— Effet de la dérivation dans l'espace direct sur la TF ?

— Y a t-il une différence entre la dérivation des séries de cosinus et de sinus ? Et si oui, laquelle ?

— TF de la fonction $\frac{\mathrm{d}}{\mathrm{d}x}[f(-(x-a))]$

Sous l'effet de la mini-compétition avec le comptage des points, j'ai eu la surprise de trouver des questions plus compliquées, et pour lesquelles les autres binômes répondaient rapidement.

Pour conclure sur cet axe concernant l'animation de groupes, remettre au centre les étudiants, et développer des activités centrées sur leur apprentissage, que ce soit via les travaux de groupe, ou les jeux-cadres, a vraiment été un tournant dans ma manière d'enseigner.

Il m'a fallu plusieurs formations, le stage d'Autrans, plusieurs heures d'enseignements, la lecture d'un ouvrage, avant de déconstruire la conception « classique » que j'avais d'un cours. J'ai conscience que ça a été long, mais lorsque j'ai enfin mis en place ces activités dans mon dernier cours, j'ai eu des bonnes surprises de la part des étudiants. Ils étaient plus motivés, apprenaient mieux, et de leur propre point de vue, à la question si ces activités les avaient aidés à mieux intégrer le cours et les TDs, 94% d'entre eux ont répondu « Tout à fait », contre 6% « Assez ».

Quand je leur ai demandés s'ils avaient d'autres idées d'activités, qu'ils auraient vues dans d'autres cours, un étudiant a répondu :

> « Dans les autres matières les profs ne proposent pas vraiment d'activité, ou si parfois du travail en groupe mais ça reste rare (cela peut aussi se comprendre par le besoin d'avancer plus vite en TD..) donc je n'ai aucune suggestion je pense pas qu'on ait manqué d'activité. »

J'ai souligné la phrase car c'est exactement ce que je pensais avant. Néanmoins, même si nous n'avons pas bouclé tous les TDs, en réalité les étudiants ont avancé ensemble plus vite que ce que je pensais, et certains ont travaillé en autonomie d'eux-mêmes en-dehors du cours. En termes de gain d'apprentissage, cela vaut bien plus de mille corrections au tableau.

4 Perspectives

Au terme de ces deux ans et demi d'enseignements, constituant cent heures de vacations, que ce soit en travaux pratiques, dirigés, ou cours magistral, et des formations du Label et d'Adum, j'en ressors avec une vision différente de celle que j'avais au départ, ainsi que plus de bagage et d'éléments à travailler.

D'abord, j'ai appris à encore plus apprécier l'enseignement, et à même y voir des formes d'amusement, notamment dans la « mise en scène » des cours (découpage par séquences), ou les jeux-cadres, sans perdre néanmoins de vue l'objectif d'apprentissage (bien au contraire !). Lors du stage d'Autrans, nous avions eu une grande table ronde le dernier jour, et un collègue avait dit que « peut-être qu'après quelques années on peut s'ennuyer ». En réalité, si on prend un peu moins au sérieux son enseignement, parce qu'au fond c'est de ça qu'il s'agit, et qu'on imagine des activités, qu'on s'adapte aux étudiants, chaque année ne sera jamais la même.

Si on se centre sur l'humain, jamais une classe ne sera similaire, puisqu'on n'aura jamais les mêmes groupes, alors qu'en recherche, vous êtes sûrs chaque année de retrouver les mêmes experts (je caricature mais c'est pour montrer à quel point l'enseignement peut être caricaturé aussi).

D'un autre côté, et c'est ce que défend Markus Brauer dans son livre [4], adapter son enseignement ne fait pas forcément perdre plus de temps. Le temps de préparation est beaucoup plus long, évidemment, mais ensuite on gagne beaucoup de temps. Par exemple dans les corrections de copies, en utilisant une grille critériée. Ou pendant les cours et TDs, lorsqu'on fait travailler les étudiants ; au bout d'un moment, nous nous retrouvons même à un rôle de modérateur, ce sont les élèves qui présentent au tableau et font le travail. Enfin, et non négligeable, avec un bon alignement pédagogique, je n'ai plus eu à répondre mille fois aux éternelles questions de « à quoi ça sert » et « qu'est-ce qu'il y aura à l'examen » (promis).

Finalement, cette expérience d'enseignement m'aura permis de déconstruire les schémas traditionnels et la vision frontale que j'avais d'un cours. Même la manière dont j'avais envisagé la problématique de la motivation a changé, en passant d'un « Les étudiants ne sont pas motivés, et ne veulent pas travailler, comment je les pousse ? », à « Si les étudiants veulent travailler, quelles sont les meilleures conditions pour qu'ils apprennent le mieux, et qui vont concentrer leur attention ? », la motivation est ensuite de leur ressort.

Découvrir certains travaux des sciences de l'éducation m'a d'ailleurs fait prendre conscience que beaucoup de choses ont déjà été testées et étudiées en détail, et que cette vision transmissive de l'enseignement devrait être derrière nous depuis longtemps. Il importe maintenant de poursuivre ce travail de littérature, d'expérimenter, de disséminer ces pédagogies, sans quoi nous ne dépasserons jamais les vieux schémas, qui continuent à se perpétuer, même parmi les nouvelles générations.

Annexes

Annexe A1 2018-2019 : TPs Photographie (7 x 4h). Licence 2 DLST.

Sous la responsabilité de Martial Balland, j'ai encadré des TPs en deuxième année de licence (Département licence sciences et technologies). Le parcours de deux ans au DLST, commun entre plusieurs disciplines de sciences physiques, comporte au quatrième semestre des TPs sur un thème précis, que les étudiants peuvent choisir. Parmi ces choix, se trouvent les TPs Photographie, qui incluent des expériences d'optique visant à comprendre le fonctionnement d'un appareil photo, mais aussi un projet à mener en parallèle et à présenter lors d'une séance supplémentaire. Les sept séances de TPs couvrent des sujets distincts, des lentilles au capteur CCD, et mettent en œuvre plusieurs types de travaux : constructions sur banc optique, utilisation d'un appareil Réflex prêté par l'université, ou prise en main de logiciels de traitements d'images (Gimp). Les séances de quatre heures sont listées ci-dessous :

1. Lentilles minces
2. Construction d'un téléobjectif
3. Aberrations et profondeur de champ
4. Utilisation d'un appareil photo Réflex
5. Traitement d'images
6. Eclairage et photographie en couleur
7. Capteur CCD
8. Oraux de projet

L'évaluation a porté donc d'une part sur la rédaction d'un compte-rendu sur une des séances (tirage au sort à la fin de la dernière séance), et d'autre part sur l'oral de projet. Le choix de ne rendre qu'un seul compte-rendu a été discuté avec Martial Balland et l'équipe pédagogique qui ont assuré ces mêmes TPs avec d'autres groupes. Lors des précédentes années, ils avaient constaté que le rendu une semaine à l'autre ne permettait pas aux étudiants de prendre le recul nécessaire pour fournir un compte-rendu scientifique de qualité. Il a donc été décidé de réduire le nombre de compte-rendus à rendre, d'étendre le temps de rédaction, mais d'augmenter le niveau d'exigence. D'un autre côté, nous avons voulu pousser les étudiants à adopter une démarche scientifique au travers d'un projet. Les étudiants ont dû choisir une problématique, en rapport avec la photographie, par exemple comment mesurer la vitesse d'une voiture avec un appareil photo, et mettre en œuvre une expérience pour y répondre. L'université a prêté des appareils Réflex en-dehors des séances, mais les étudiants étaient aussi libres d'utiliser leur propre matériel. Les incertitudes et limites de la méthode choisie devaient par contre être mises en avant lors de l'oral.

TP Photo notations (draft)

Compte-rendu TP (/20) :

- Forme compte-rendu (/5) : plan, articulation des parties, façon d'écrire…
- Contenu scientifique (/5) : théorie, réponses aux questions, calculs, équations
- Partie expérimentale (/5) : la(les) problématique(s) du TP a(ont) été compris, le protocole mis en lien avec la théorie et les résultats
- Résultats et discussion (/5) : présentation des résultats (rigueur, c'est-à-dire incertitudes, chiffres significatifs, graphiques…), savoir discuter les résultats en lien avec la problématique du TP, ce qui a été obtenu comme résultats, le protocole etc.

Oral de projet (/20) :

- Forme présentation (/5) : clarté de l'oral, organisation des diapos, façon de présenter les photos et les résultats (si chiffrés : incertitudes, chiffres significatifs, présentation des graphes etc)
- Contenu scientifique (/5) : théorie, compréhension des concepts utilisés et/ou équations si présentes
- Mise en œuvre du projet / démarche scientifique (/5) : présentation du sujet, problématique, protocole utilisé, comment les résultats sont exposés et discutés en lien avec le sujet
- Réponses aux questions (/5)

Figure 12 – Grille de notations du CR et de l'oral de projet.

Annexe A2 2018-2020 : CM/TD Séries et transformées de Fourier (2 x [13,5h]). Licence 3 Physique-Chimie.

Pendant deux années consécutives, 2018-2019 et 2019-2020, j'ai assuré au premier semestre la seconde partie du cours/TD de Mathématiques pour la physique, en Licence 3 Physique-Chimie. L'UE se divise en une première partie introduisant des concepts généraux d'algèbre linéaire (espaces vectoriels, matrices...) dont Elisabeth Charlaix, la responsable de l'UE, a la charge, et une seconde partie sur les séries et transformées de Fourier, sur laquelle j'ai remplacé ma directrice de thèse, Judith Peters, pour deux ans. Cette partie s'est généralement étalée sur une première séance d'une heure et demie, puis quatre séances de trois heures, coupée par la semaine des vacances de la Toussaint, que j'ai mise à profit pour faire travailler les étudiants sur des devoirs-maison. Suivant les conseils de Judith Peters, je me suis appuyée sur un support de cours de mathématiques pour la physique de Bahram Houchmandzadeh [1], qui comporte également des exercices. Néanmoins, j'ai eu toute la liberté possible pour concevoir et organiser les enseignements sur ma partie, tant que je respectais les conditions auxquelles Elisabeth Charlaix et moi nous étions mises d'accord :

- Inclure les séries de Fourier, mais aussi les transformées de Fourier, et une introduction en fin de cours sur la distribution de Dirac.
- Mettre l'accent sur les applications, ce qu'elle ferait aussi dans sa partie.
- Le premier contrôle continu avant les vacances de la Toussaint porterait sur sa partie, et nous nous partagerions l'examen terminal (3h, dont 1,5h chacune).

Etant donné le format cours/TD, j'ai alterné entre des présentations frontales, et des séances d'application directement en lien avec le cours. Je demandais un minimum de travail à la maison, notamment dans la préparation des TDs. De plus, sur certains exercices redondants, mais essentiels pour s'entraîner à calculer pour l'examen, je n'ai pas donné de correction formelle en classe, mais je restais disponible par mail si les étudiants avaient des questions. Le contenu du cours est resté globalement le même sur les deux années. Cependant suite à des ateliers de formation sur l'enseignement, et l'obligation de convertir une partie de mes enseignements en distanciel, j'ai changé le format la deuxième année. Ainsi, j'ai organisé les séances par séquences de 20 à 30 minutes, alternant rapidement entre éléments de cours et applications. J'ai mis l'accent sur les jeux-cadres (via des QCM Kahoot par exemple) et le travail en groupe. De plus, dés le début du cours, les étudiants avaient une grille concernant les objectifs à atteindre pour ma partie, et chaque exercice indiquait la compétence travaillée, pour leur permettre plus d'autonomie dans le travail personnel. Concernant l'examen terminal, j'ai rédigé un sujet sur table chaque année, participé aux rattrapages (mise en place d'un QCM Moodle avec Elisabeth Charlaix, ou oraux), et corrigé ma partie. J'ai également eu l'opportunité d'assister aux jurys de premier semestre des L3 Physique-Chimie, et de discuter avec l'équipe pédagogique des obstacles liés aux enseignements de mathématiques en physique, et des solutions possibles pour les surmonter.

1. `https://www-liphy.univ-grenoble-alpes.fr/pagesperso/bahram/`

Grille objectifs - Page 1 de 3

Compétences générales	**Non atteint**	**Insuffisant**	**Suffisant**	**Master**
C1 – Savoir calculer sans erreurs (dérivation, intégration, IPP, trigo…)	Règles de calcul non connues, jamais appliquées, résultat faux	Règles de calcul connues, mais mal appliquées / mal comprises	Règles de calcul connues et appliquées, mais avec quelques erreurs (pré-facteur, inattention…)	Règles de calcul connues et appliquées, sans aucune erreur
C2 – Comprendre les termes physiques d'une équation	Aucun terme n'est interprété ou compris	Compréhension globale, assez vague	Compréhension globale de chaque terme de l'équation (variables, dérivées…)	Compréhension de chaque terme, des liens qui existent en terme de physique
C3 – Savoir interpréter un résultat par rapport à la physique	Aucun commentaire sur le résultat, aucun lien fait avec le contexte	Commentaire sur l'ordre de grandeur	Commentaire sur l'ordre de grandeur, lien avec le contexte du problème	Commentaire sur l'ordre de grandeur, lien avec le contexte du problème, et avec d'autres notions de physique d'autres cours

Séries de Fourier				
S1 – Connaître la décomposition d'une fonction dans la base de Fourier	Définition non connue	Forme générale connue, mais ni les coefficients, ni les facteurs de la base	Expression connue en totalité, coefficients et facteurs compris	Expression connue en totalité, démonstration du cours comprise (origines de la base…)
S2 – Savoir décomposer une fonction dans la base de Fourier	Décomposition non appliquée	Décomposition connue, mais application incorrecte	Décomposition appliquée, mais avec quelques erreurs (pré-facteur…) qui ne dénaturent pas le résultat	Décomposition appliquée, sans aucune erreur

Grille objectifs - Page 2 de 3

	Non atteint	**Insuffisant**	**Suffisant**	**Master**
S3 – Savoir décomposer une fonction sur un intervalle quelconque [a;b]	Règle non connue	Règle connue, mais application incorrecte	Règle appliquée, mais avec quelques erreurs (pré-facteur…) qui ne dénaturent pas le résultat	Règle appliquée, sans aucune erreur
S4 – Utiliser une série de Fourier pour calculer une somme	Méthode ni connue, ni appliquée	Méthode partiellement comprise, résultat faux	Méthode comprise, appliquée avec quelques erreurs (pré-facteur…)	Méthode comprise et parfaitement appliquée
S5 – Connaître l'égalité de Parceval et savoir l'appliquer	Egalité ni connue, ni appliquée	Egalité connue, mais application incorrecte	Egalité connue et appliquée, avec quelques erreurs minimes	Egalité connue et appliquée, sans aucune erreur
S6 – Connaître les cas physiques intéressants où utiliser la base de Fourier	Aucun cas connu	Quelques cas connus	Cas connus, ainsi que les raisons	Cas et raison connus, lien fait avec d'autres cours
S7 – Savoir calculer et tracer le spectre d'une fonction	Définition non connue, aucun calcul, aucun tracé	Erreurs dans calculs et tracés	Erreurs minimes dans calculs et tracés	Calculs et tracés corrects
S8 – Connaître et utiliser la base des sin() ou cos() purs	Définition non connue, non appliquée	Définition connue, mais application incorrecte	Définition appliquée, avec quelques erreurs minimes (pré-facteur…)	Définition appliquée, sans aucune erreur
S9 – Connaître et utiliser la base des exp() purs	Définition non connue, non appliquée	Définition connue, mais application incorrecte	Définition appliquée, avec quelques erreurs minimes (pré-facteur…)	Définition appliquée, sans aucune erreur
S10 – Avoir compris les applications et les interprétations physiques. Savoir les refaire.	Applications non comprises (ni début, ni aucune étape)	Applications partiellement comprises (début, moins de la moitié des étapes)	Applications globalement comprises. Peut refaire une partie en autonomie.	Applications totalement comprises, implications physiques incluses. Peut refaire la majorité en autonomie.

Grille objectifs - Page 3 de 3

	Non atteint	**Insuffisant**	**Suffisant**	**Master**
Transformées de Fourier				
T1 – Connaître l'expression d'une TF et d'une TF inverse d'une fonction, et savoir les calculer	Expression ni connue, ni appliquée	Expression connue, mais application incorrecte	Expression connue et appliquée, mais avec quelques erreurs minimes (pré-facteur…)	Expression connue et appliquée, sans aucune erreur
T2 – Connaître et savoir appliquer les propriétés et opérations sur les TF	Propriétés ni connues, ni appliquées	Propriétés connues, mais non mises en lien avec le problème, ou mal appliquées	Propriétés connues et appliquées, mais avec quelques erreurs minimes (pré-facteur…)	Propriétés connues et appliquées, sans aucune erreur
T3 – Connaître et savoir appliquer l'égalité de Parseval – Plancheral	Egalité ni connue, ni appliquée	Egalité connue, mais application incorrecte	Egalité connue et appliquée, avec quelques erreurs minimes	Egalité connue et appliquée, sans aucune erreur
T4 – Connaître la définition du produit de convolution et ses propriétés pour la TF. Savoir les appliquer.	Définition et propriétés ni connues, ni appliquées	Définition et propriétés connues, mais non ou mal appliquées	Définition et propriétés connues et appliquées, mais avec quelques erreurs minimes (pré-facteur…)	Définition et propriétés connues et appliquées, sans aucune erreur
T5 – Avoir compris les applications et les interprétations physique liées. Savoir les refaire.	Applications non comprises (ni début, ni aucune étape)	Applications partiellement comprises (début, moins de la moitié des étapes)	Applications globalement comprises. Peut refaire une partie en autonomie.	Applications totalement comprises, implications physiques incluses. Peut refaire la majorité en autonomie.

Figure 13 – Grille critériée distribuée en début de cours (année 2020-2021). (Retour Vers un meilleur alignement pédagogique en CM/TD).

TD 3 : Transformées de Fourier.

Exercice 1 : décroissance exponentielle — C1/T1

1. Calculez la TF de $f(\ ) = \mathrm{e}^{-\gamma|t|}$. (M)
2. En déduire la TF inverse de $\ (\omega) = \frac{336}{1764+\omega^2}$. NB : rapportez-vous au résultat précédent, et notez bien que la TF et la TF inverse sont des opérations linéaires. (M)

Exercice 2 : l'incontournable fonction porte — C1/C3/T1

1. Calculez la TF de la fonction porte : (M)

$$\begin{cases} \ (x) = 0 & \text{si } |x| > 1 \\ \ (x) = 1 & \text{si } |x| \quad 1 \end{cases}$$

2. Calculez la TF de $a^{-1} \ \ \left(\frac{x}{a}\right)$. (M)
3. Que devient cette TF quand $a \to 0$?

Exercice 3 : fonction d'Heaviside — C1/T1

Calculez la TF de $f(\ ) = \ \ (\ )\mathrm{e}^{-\gamma t}$, avec $\ \ (\ )$ la fonction d'Heaviside, définie comme :

$$\begin{cases} \ (\ ) = 0 & \text{si } \ < 0 \\ \ (\ ) = 1 & \text{si } \ > 0 \end{cases}$$

On écrira le résultat sous la forme complexe $\ = a + \ \ .$ (M)

Exercice 4 : train d'onde — C1/C3/T1

1. Calculez la TF d'un train d'onde : $f(x) = \ \ \left(\frac{x}{a}\right)\mathrm{e}^{ik_0 x}$. (M)
2. k_0^{-1} est la période de l'onde et a son extension spatiale. Discutez les diverses limites et leurs effets sur la TF.

Exercice 5 : Gaussienne normale — C1/T1

Sachant que $\int_{-\infty}^{+\infty} \mathrm{e}^{-x^2}\mathrm{d}x = \sqrt{\ }$, calculez la TF de $f(x) = \mathrm{e}^{-x^2/2}$ et montrez que $f(\ ) = \sqrt{2\ }\,\mathrm{e}^{-q^2/2}$.
Pour cela :
- Notez que $x^2 + 2 \ \ x = (x + \ \ )^2 + \ ^2$.
- On fera aussi un changement de variable sur x pendant les calculs pour se ramener à une intégrale de type $\int_{-\infty}^{+\infty} \mathrm{e}^{-X^2}\mathrm{d} \ \ = \sqrt{\ }$.

Exercice 6 : Gaussienne de largeur a — C1/C3/T2

1. Calculez la TF de $f(x) = \frac{1}{a}\mathrm{e}^{-\frac{x^2}{2a}}$. On utilisera le résultat précédent et la propriété de la TF pour le changement d'échelle vue en cours. (M)
2. Que devient cette transformée quand $a \to 0$?
3. En utilisant la TF calculée au 1., déduisez la TF des fonctions suivantes : $f_a(x) = \mathrm{e}^{-\frac{x^2}{2a}}$, $f_b(x) = \mathrm{e}^{-\frac{2x^2}{a}}$. (M)
4. Toujours en utilisant la TF calculée au 1., déduisez la TF inverse des fonctions suivantes : $f_c(\ ) = \mathrm{e}^{-\frac{aq^2}{2}}$, $f_d(\ ) = \mathrm{e}^{-aq^2}$, $f_e(\ ) = \sqrt{2\ }\,\mathrm{e}^{-aq^2}$. (M)

Exercice 7 : et de deux portes convoluées naquit le triangle — C1/T1/T4

1. Calculez la TF de la fonction porte de longueur 1, définie par : (M)

$$\begin{cases} \ (x) = 0 & \text{si } |x| > \frac{1}{2} \quad \left(x < -\frac{1}{2} \text{ ou } x > \frac{1}{2}\right) \\ \ (x) = 1 & \text{si } |x| \quad \frac{1}{2} \quad \left(-\frac{1}{2} \quad x \quad \frac{1}{2}\right) \end{cases}$$

2. On considère la fonction triangle définie par :

$$\begin{cases} \ (x) = 1 + x & \text{si } x < 0 \\ \ (x) = 1 - x & \text{si } x \geq 0 \end{cases}$$

   En réalité, cette fonction peut s'écrire comme la convolution d'une fonction porte sur elle-même : $\ (x) = \ \ (x) \circledast \ \ (x)$.
   En utilisant les propriétés du produit de convolution dans les TF, calculez alors la TF de la fonction triangle. (M)
3. Vous pouvez vous amuser à retrouver ce résultat en calculant directement la TF de $\ (x)$.
   Rappel : $1 - \cos(2a) = 2\sin^2(a)$.

Figure 14 – TD 3 sur les transformées de Fourier (année 2020-2021). Chaque exercice indique les compétences en lien avec la grille critériée, et les (M) indique si l'exercice est en ligne sur Moodle (possibilité de vérifier ses résultats). (Retour Vers un meilleur alignement pédagogique en CM/TD).

Code compétence	Non atteint	Insuffisant	Suffisant	Master
C1 - calculs	0	1	2	3
C2 - termes équations	0	0.25	0.75	1
C3 - interprétation physique	0	1	2	3
S1 - définition base de Fourier	0	0.25	0.75	1
S3 - décomposition sur [a ;b]	0	0.5	1	2
S7 - spectre d'une fonction	0	0.25	0.75	1
T1 - calcul TF	0	0.5	1	2
T2 - propriétés et opérations TF	0	0.5	1	2
T4 - produit de convolution	0	0.5	1	2
T5 - applications TF	0	1	2	3
Total	0	5.75	12.25	20

Figure 15 – Barème de l'examen terminal 2020-2021. (Retour Vers un meilleur alignement pédagogique en CM/TD).

Sheet1

DMs 1 et 2	Non atteint	Insuffisant	Suffisant	Au-delà
Bonus	0	0.5	1.5	2

Critère	Non atteint	Insuffisant	Suffisant	Au-delà
EXERCICE 1 (/8)				
Décomposer une fonction sur la base de Fourier	0	1	2	3
Utiliser une série pour obtenir une somme	0	1	2	3
Connaître le spectre d'une fonction	0	0.5	1.5	2
EXERCICE 2 (/12)				
Traduire et manipuler une équation dans l'espace réciproque	0	1	2	3
Connaître et appliquer les propriétés du produit de convolution	0	0.5	1.5	2
Connaître et appliquer les propriétés du pic de Dirac	0	0.5	1.5	2
Obtenir un résultat correct pour u(x,t)	0	0.5	1.5	2
Savoir interpréter un résultat	0	1	2	3

Figure 16 – Barème de l'examen terminal 2019-2020. (Retour Vers un meilleur alignement pédagogique en CM/TD).

Questionnaire Moodle de fin de semestre (16 réponses) :

La partie Séries et Transformées de Fourier vous a-t-elle paru accessible ?

— Très accessible (69%)

— Assez accessible (31%)

— Peu accessible (0%)

— Pas du tout accessible (0%)

Quels sont les éléments qui vous ont le plus aidé à suivre ce cours ? A contrario, quels sont ceux qui vous ont manqué pour que ce cours soit plus accessible ? (texte libre)

Figure 17 – Nuage de mots correspondant aux réponses (généré sur `nuagedemots.co`).

Comment avez-vous ressenti la gestion du temps pendant les cours ?

— Trop rapide pour suivre (0%)

— Rapide, mais on peut suivre (87.5%)

— Lent, mais on ne décroche pas (12.5%)

— Trop lent, et on décroche (0%)

Concernant le nombre d'exercices et d'applications, il en aurait fallu :

— Beaucoup plus (0%)

— Un peu plus (100%)

— Un peu moins (0%)

— Beaucoup moins (0%)

Est-ce que les différentes activités proposées (DM, travaux de groupe, quizz Kahoot...) vous ont aidé à mieux intégrer les éléments du cours et des TD ?

— Tout à fait (94%)

— Assez (6%)

— Peu (0%)

— Pas du tout (0%)

Votre degré d'implication personnelle (participation en cours, travail à réaliser à la maison, etc.) était :

— Très important (6%)

— Assez important (69%)

— Peu important (25%)

— Pas important du tout (0%)

Est-ce que vous avez le sentiment de sortir de ces six séances de cours en ayant appris quelque chose ? Cela englobe autant des éléments de théorie, que des savoir-faire (calculs...)

— Oui, totalement (87.5%)
— Un peu (12.5%)
— Pas vraiment (0%)
— Non, pas du tout (0%)

Qu'est-ce que vous avez préféré dans cet enseignement ? (texte libre)

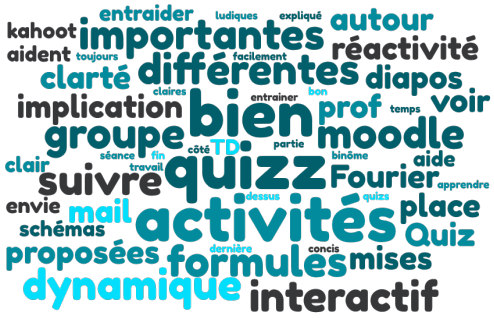


Figure 18 – Nuage de mots correspondant aux réponses (généré sur `nuagedemots.co`).

Quels sont les points qu'il faudrait améliorer selon vous ? (texte libre)

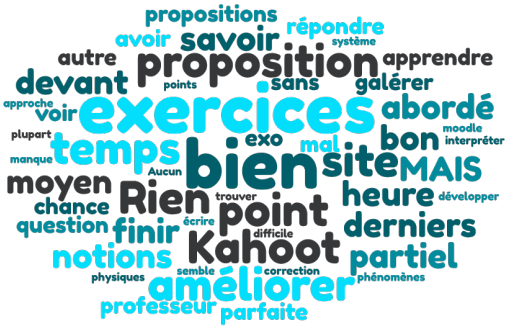


Figure 19 – Nuage de mots correspondant aux réponses (généré sur `nuagedemots.co`).

Avez-vous des suggestions d'activités, de format, que vous auriez aimé avoir pour ce cours ? Par exemple, des activités que vous avez rencontrées dans d'autres cours, et qui aurait été applicable dans cette partie ? (texte libre)

Figure 20 – Nuage de mots correspondant aux réponses (généré sur `nuagedemots.co`).

Annexe A3 2019-2020 : CMs (4 x 2h) et TDs (5 x 2h) d'opto-électronique. 2ème année IUT 1 Mesures Physiques.

Dans le cadre d'une année de décharge d'enseignements, Guillermo Martin m'a recrutée pour assurer le cours magistral d'opto-électronique, consistant en quatre cours de deux heures, ainsi qu'un groupe de TD (cinq séances de deux heures), en deuxième année de DUT, à l'IUT 1 de Mesures Physiques.

Guillermo Martin m'a directement envoyé et présenté les supports de cours à utiliser, diapositives et TDs. Ceux-ci sont reliés à des TPs, dans lesquels les étudiants appliquent directement les notions introduites. J'ai donc eu une contrainte concernant le contenu des quatre séances de cours :

1. Sources : lampes à décharge, DEL, laser.
2. Transport et codage de l'information : fibres optiques, modulateurs électro-optiques.
3. Détecteurs : détecteurs photoélectriques, détection synchrone, grandeurs.

Cependant, j'ai tout de même conservé une certaine marge de manœuvre sur la forme.

Dans l'équipe enseignante, le principal défi auquel nous avons été confrontés, notamment en cours et en TD, est la complexité de certaines notions ; biréfringence, ellipsoïde des indices, utilisation de matrices ainsi que de tenseurs, à présenter à des étudiants de deuxième année.

Dans ce cadre-là, j'ai modifié les slides, de façon à présenter les notions une à une, maintenir l'attention des étudiants, et prendre le temps d'expliquer en parallèle sur le tableau plus de détails. J'ai aussi décomposé certains raisonnements, et mis en exergue la signification physique de chaque équation qui apparaissait. D'un autre côté, que ce soit en amphithéâtre ou en TD, j'ai tenté de pousser à la discussion, d'interagir avec les étudiants, mais avec du recul je manquais de méthodes plus formelles et plus cadrées pour les pousser à participer, notamment avec des plus grands groupes. Enfin, j'ai dû rédiger l'examen de fin de semestre, et je devais également m'occuper de la correction. Néanmoins le confinement de Mars 2020 a contraint l'IUT à annuler l'examen au vu des nombreuses contraintes qu'il y avait alors pour le convertir en QCM en ligne, et seuls les compte-rendus ont compté dans la note.

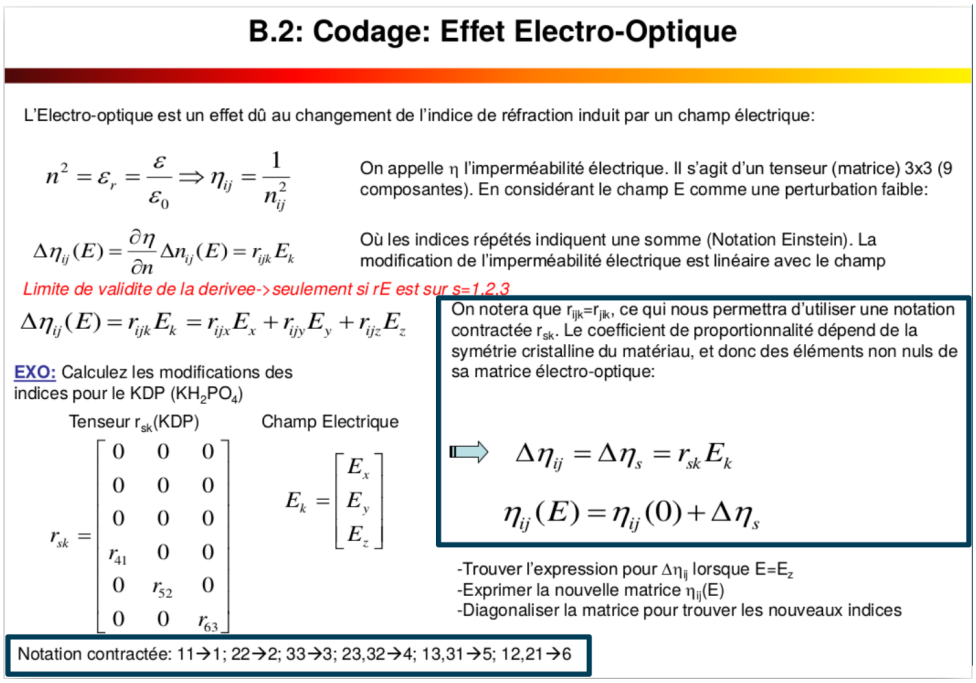


Figure 21 – Adaptation d'une diapositive en opto-électronique. En haut, le support initial, dont les parties encadrées en bleu sont montrées dans leur nouvelle version en bas. Figure liée à la partie Travailler le support de cours.

Annexe A4 2019-2020 : TPs d'optique géométrique (5 x 4h). 1ère année IUT 1 Mesures Physiques.

Au second semestre de première année de DUT, Irène Ventrillard m'a intégrée à son équipe en charge des TPs d'optique géométrique, à l'IUT 1 Mesures Physiques. Ces TPs sont constitués de quatre séances de quatre heures couvrant plusieurs sujets introduits auparavant en cours et TDs :

1. Des objets et des images
2. Focométrie
3. Appareil photographique
4. Réfractométrie avec un goniomètre

Les deux premiers TPs, introduisant des notions vues en cours, incluent également des oraux avec transparents, en réponse à une problématique, par exemple la rédaction d'un protocole pour créer un objet virtuel. Chaque groupe passe au moins une fois à l'oral, doit expliquer la démarche suivie en cinq minutes, et répondre aux éventuelles questions de leurs collègues.

En parallèle, des compte-rendus de TP doivent être rédigés pendant les séances mêmes, et corrigés par les enseignants pour la semaine d'après. Ils sont notés suivant une grille critériée précise, établie par l'équipe enseignante de l'IUT, connue des étudiants dés la première séance, et en lien avec un cours de méthodologie du premier semestre. Chaque TP met en œuvre différents types de compétences pour la rédaction, par exemple le second TP met en avant le calcul d'incertitudes, tandis que les deux derniers pondèrent en plus de ça la présentation des résultats (graphiques, comparaison avec la théorie).

Les deux premiers compte-rendus ne sont pas formellement notés, mais un retour est donné, afin de montrer aux étudiants ce qui est attendu d'eux, tandis que les deux derniers sont notés et constituent un tiers de la note finale. Il est à noter que les compte-rendus n'évaluent que la forme, la rédaction, et non le contenu scientifique.

Les connaissances pratiques sont évaluées à part, lors d'une cinquième séance, et cet examen pratique compte pour deux tiers.

Néanmoins, en raison du confinement de Mars 2020, je n'ai pu encadrer que les deux premiers TPs, et les autres ont été annulés. L'évaluation n'a porté que sur la rédaction d'un compte-rendu pour le TP1, le seul pour lequel nous avions pu faire un retour aux étudiants.

Toute l'équipe enseignante est restée en contact le long du confinement, notamment pour répondre aux questions des étudiants (mise en place d'un salon Discord). Nous avons aussi longuement discuté de la mise ou place ou non d'une alternative à l'examen pratique, mais avons finalement considéré que l'examen concernant les CMs et les TDs était suffisant.

Grille critériée utilisée pour évaluer le compte rendu de TP à rendre tapé

Thème	Compétences attendues	Niveau 0	Niveau 1	Niveau 2	Niveau 3
Organisation générale	**Savoir rédiger et organiser un CR**	Phrases sans sens ou orthographe et/ou grammaire déplorable	Structure claire + Moins de 5 fautes par page ne nuisant pas à la compréhension	Niv 1 + Utilisation d'un vocabulaire pertinent + Numéro et titre aux annexes éventuelles	Niv 2 + Au maximum une faute par page + Mise en forme correcte du document
Introduction	**Savoir introduire le sujet d'un TP et formuler sa problématique**	Pas d'introduction	Enoncé recopié	Reformulation avec ses propres mots des objectifs et des méthodes	Niv 2 + Présentation du contexte de l'étude ou illustration avec un exemple non mentionné dans l'énoncé
Description du dispositif expérimental	**Savoir réaliser et utiliser des schémas ou figures pour illustrer un CR**	Présence trop faible ou inexistante de schémas ou schémas incompréhensibles	Présence de schémas lisibles	Niv 1 + Schémas clairs avec un titre faisant apparaître des paramètres et informations utiles	Niv 2 + Présence de l'ensemble des paramètres et informations utiles + Référence explicite aux schémas dans le texte[1]
Protocole	**Savoir décrire un protocole de mesure**	Pas de protocole	Rédaction des tâches effectuées	Niv 1 + Niveau de détail adapté (description des réglages à effectuer, des consignes de sécurité, des précautions particulières…)	Niv 2 + Utilisation d'un vocabulaire pertinent + Discussion des limites du protocole
Tableaux / Unités	**Savoir présenter des mesures et des grandeurs calculées**	Tableau non fourni	Présence des mesures brutes + Structure et mise en forme correctes du tableau	Niv 1 + Unités pour toutes les grandeurs + Constantes à part + Numéro et titre au tableau	Niv 2 + Affichage d'un nombre correct de chiffres significatifs[2]
Graphiques	**Savoir tracer des graphiques et utiliser des courbes de tendance**	Graphique non fourni	Représentation correcte des points de mesures + Présence des noms des grandeurs sur les axes + Légende éventuelle	Niv 1 + Unités sur les axes + Echelles adaptées + Numéro et titre au graphique	Niv 2 + Affichage correct de la modélisation adaptée[3] + Points aberrants éventuels détectés et exclus de la série de mesures exploitables

[1] Par exemple : "voir tableau n°2", "voir schéma n°1" ou "d'après la courbe présentée sur le graphique n°5, ..." ou "comme le montrent les résultats en annexe 1, ..."

[2] Le nombre correct de chiffres significatifs est lié :
• pour une grandeur mesurée : à la précision de l'appareil de mesure utilisé
• pour une grandeur calculée : au nombre de chiffres significatifs des grandeurs utilisées dans le calcul

Grille critériée utilisée pour évaluer le compte rendu de TP à rendre tapé

Thème	Compétences attendues	Niveau 0	Niveau 1	Niveau 2	Niveau 3
Commentaires des résultats	**Savoir décrire les résultats obtenus**	Pas de description des résultats obtenus	Ordres de grandeurs et/ou sens de variation décrits	Niv 1 + Référence explicite à la figure correspondante[1]	Niv 2 + Qualité des résultats commentée + Justification des points aberrants éventuels
	Savoir analyser les résultats obtenus	Pas d'analyse des résultats	Explication succincte de la théorie + Présence des valeurs attendues si pertinent	Niv 1 + Méthode correcte permettant de comparer les résultats avec la théorie[4]	Niv 2 + Commentaires + Formulation d'hypothèses pour expliquer les écarts
Conclusion	**Savoir formuler une synthèse globale**	Pas de conclusion	Simple énumération des tâches réalisées pendant la séance	Bilan synthétique des différentes mesures réalisées et des méthodes mises en œuvre	Niv 2 + Commentaires et critiques des différentes mesures et/ou méthodes et/ou thèmes abordés
Incertitudes	**Savoir calculer une incertitude et présenter un résultat avec son incertitude**	Pas de calcul d'incertitude	Evaluation ou estimation de l'incertitude sur une grandeur mesurée	Niv 1 + Calcul de l'incertitude sur une grandeur calculée avec la formule de propagation des incertitudes	Niv 2 + Présentation du résultat et de son incertitude avec le nombre de chiffres significatifs adapté

[3] Selon les cas, la modélisation adaptée peut être par exemple :
• la courbe de tendance adéquate + indication de son équation et du facteur de corrélation si pertinent
• une linéarisation, un calcul de dérivée ou autre + le tracé associé et sa modélisation éventuelle

[4] Selon les cas, un ou plusieurs points à développer :
• Justification du choix de la modélisation[3]
• Comparaison d'une valeur mesurée/calculée ou d'un coefficient d'une courbe de tendance avec une valeur théorique + calcul d'écart relatif si pertinent

Remarque : La pondération entre les compétences n'est pas uniforme et peut varier en fonction des spécificités de chaque TP.

Figure 22 – Grille critériée utilisée pour l'évaluation des compte-rendus de TP. (Retour à la section L'évaluation des travaux pratiques).

Bibliographie et formations

[1] Mieux APPRENDRE. Les jeux de Thiagi. 2021. URL : https://ressources.mieux-apprendre.com/presentation-des-jeux/.

[2] Jean-Pierre ASTOLFI. L'erreur, un outil pour enseigner. 13ème édition. ESF sciences. Paris, 2020. ISBN : 978-2-7101-4181-5.

[3] John BIGGS. « Enhancing teaching through constructive alignment ». In : Higher Education 32 (1996), p. 347-364. DOI : 10.1007/BF00138871.

[4] Markus BRAUER. Enseigner à l'université - Conseils pratiques, astuces, méthodes pédagogiques. Sous la dir. d'Armand COLIN. Paris, 2011. ISBN : 978-2-200-25458-2.

[5] Anne DÉCORET-AHIHA. Formation ADEA - Se développer dans son métier d'enseignant. Grenoble, 2020.

[6] Julien DOUADY et Christian HOFFMAN. Formation Adum - Evaluer les apprentissages des étudiants. Grenoble, 2020.

[7] Julien DOUADY et al. Stage d'Autrans. Stage Label RES et Marche d'Approche. Autrans, 2019.

[8] Dieudonné LECLERCQ. Conception d'Interventions et Construction de Produits pour la formation. Liège, 1995.

[9] Rolland VIAU. « Chapitre 9 : Savoir motiver les étudiants ». In : Se former à la pédagogie de l'enseignement supérieur. Cheneliere. 2014, p. 236-254. ISBN : 9782921793490.