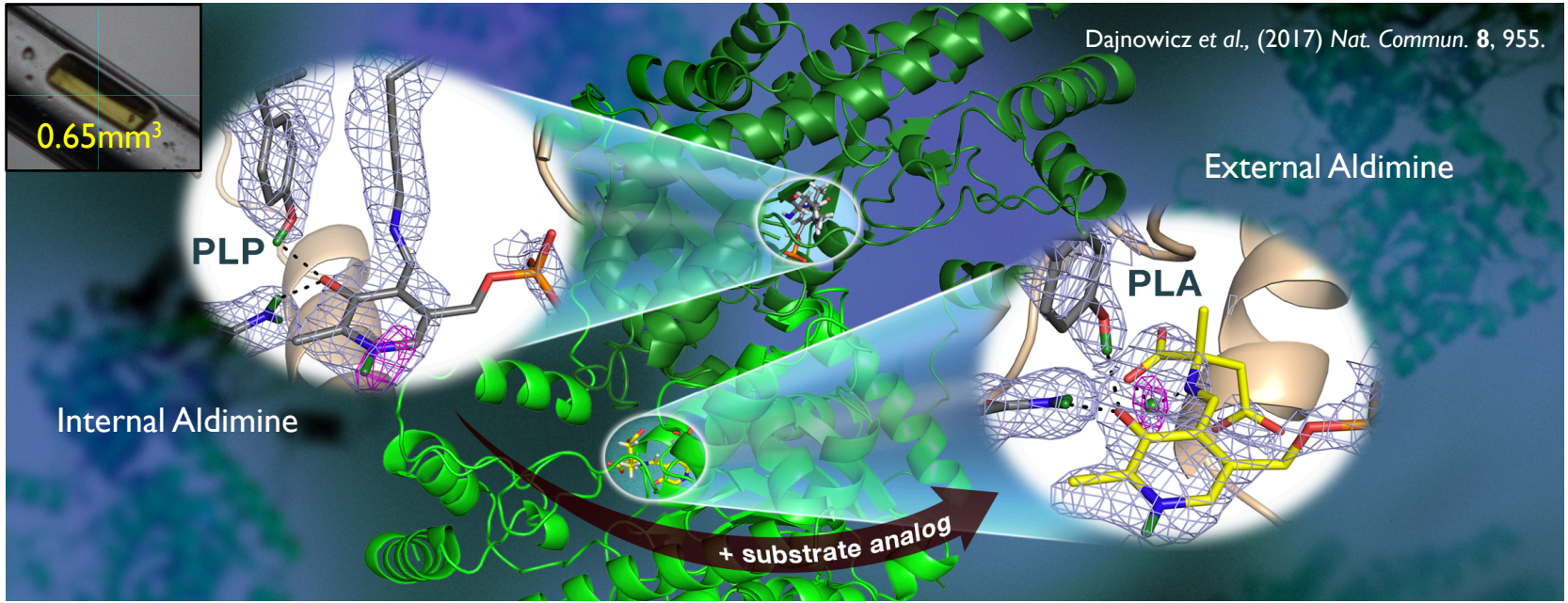
NEUTRONS FOR SOCIETY

Seeing the chemistry in biology using neutron crystallography

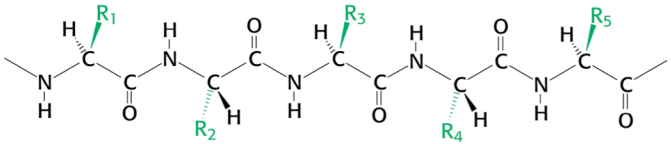


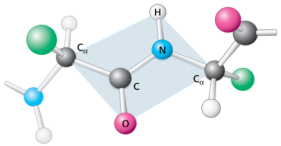
Matthew Blakeley, Large-Scale Structures Group, Institut Laue-Langevin

NEUTRONS FOR SOCIETY

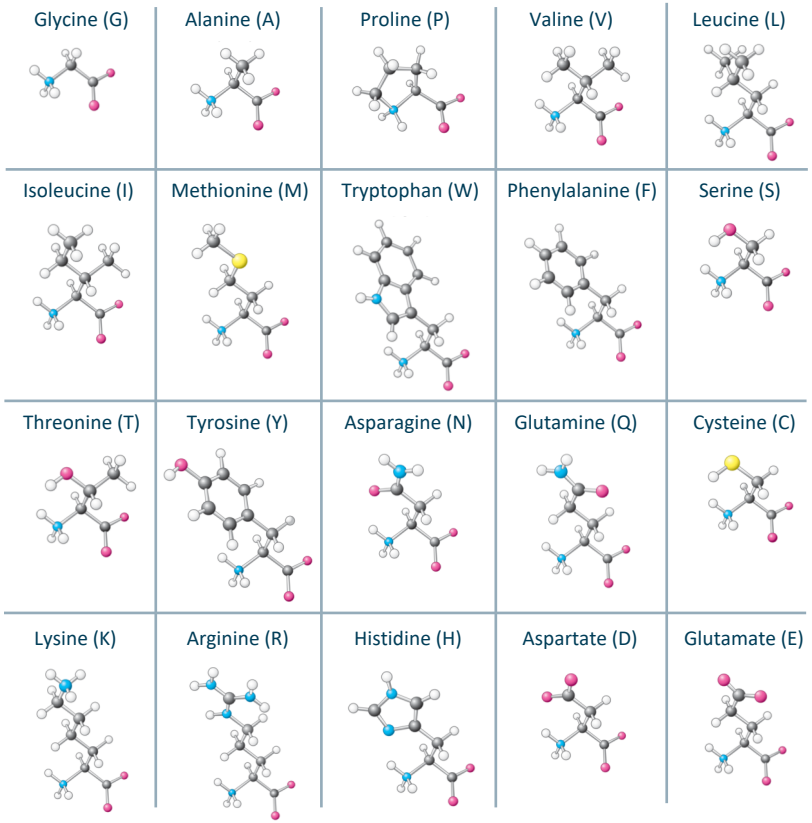
Hydrogen atoms, hydrogen-bonding and proton transfer play critical roles in biological structure and function

- Proteins are the workhorses of the cell, performing a myriad of essential functions.
- Composed of chains of amino-acids (from 20 naturally occurring) linked together via peptide bonds.



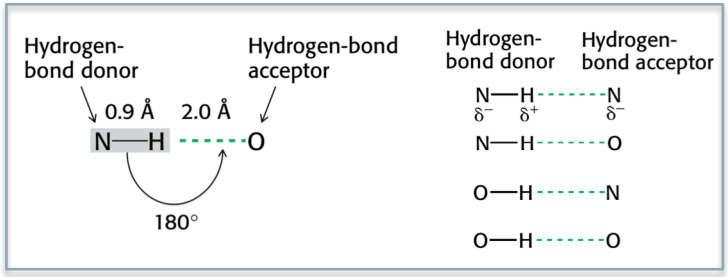


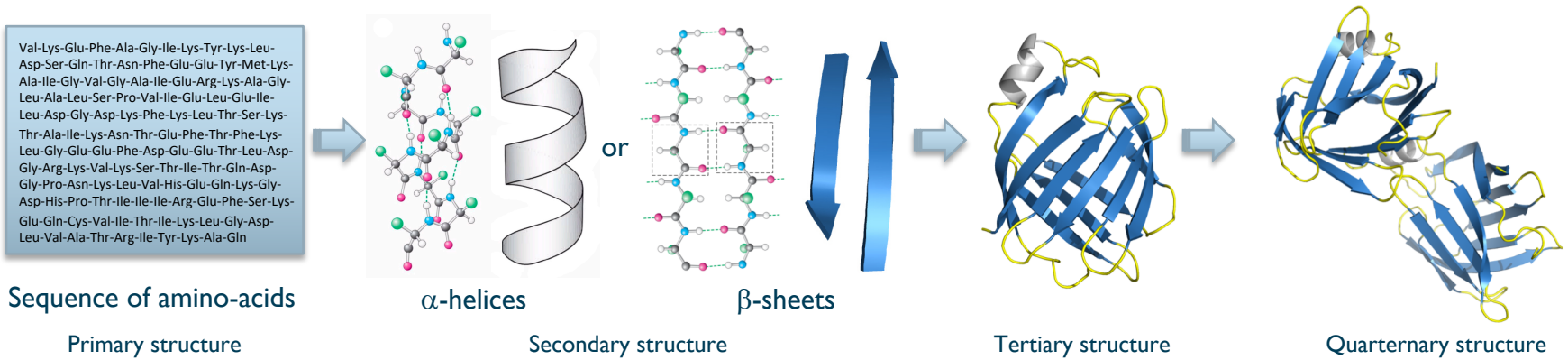
- *H-atoms account for ~half of the atoms of a protein, and play key roles in protein structure and function.*



Protein-folding and structure stabilization

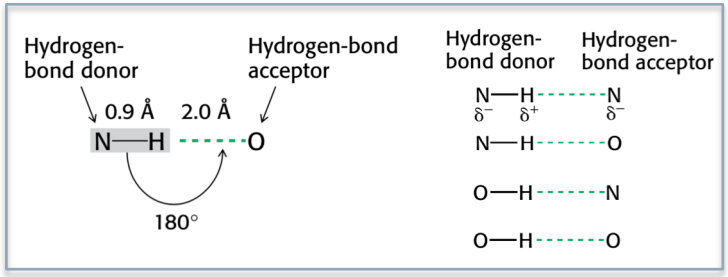
- 3D structures of proteins are defined by their linear sequence of amino-acid residues (primary structure).
- *Networks of H-bonds are essential for (i) the correct folding and (ii) stabilization of protein structures and their macromolecular complexes (e.g. protein-protein, protein-DNA etc.).*

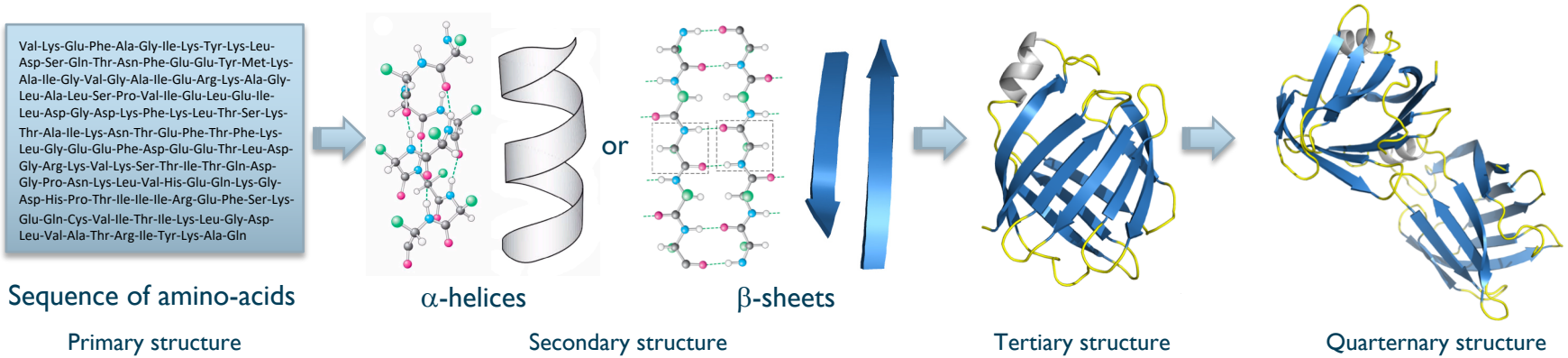




Protein-folding and structure stabilization

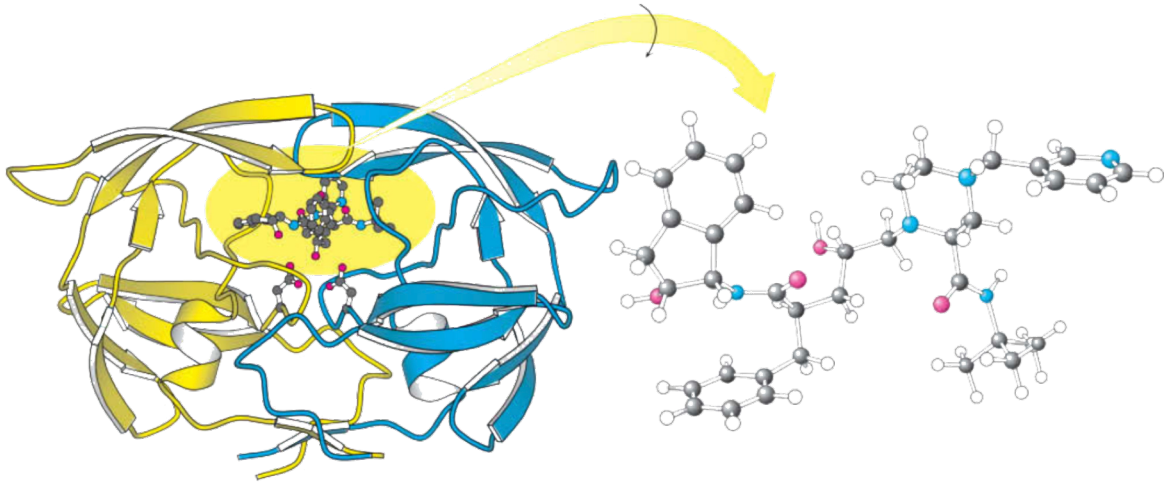
- 3D structures of proteins are defined by their linear sequence of amino-acid residues (primary structure).
- *Mutations in amino-acid sequence alter the H-bond networks, which can reduce or enhance the protein's propensity to mis-fold or dissociate (e.g. amyloidogenic diseases).*





Small-molecule ligand binding

- Small molecule ligands (*e.g.* substrates, inhibitors, activators, carbohydrates etc.) dock in the binding sites of their target biological macromolecules *via* direct and water-mediated H-bonds, and through hydrophobic interactions.

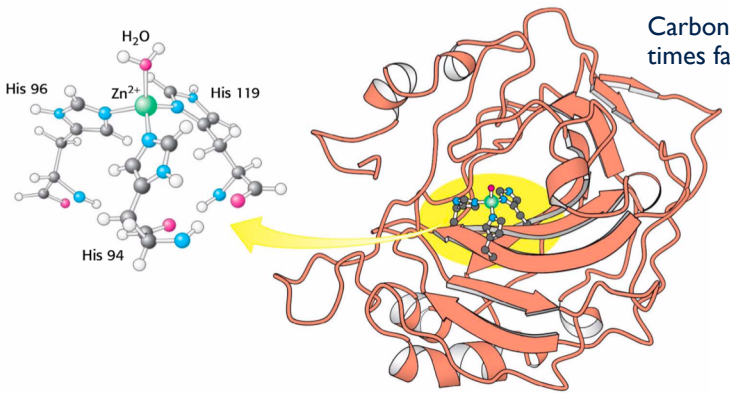


Anti-retroviral inhibitor (indinavir) binding to HIV-1 protease

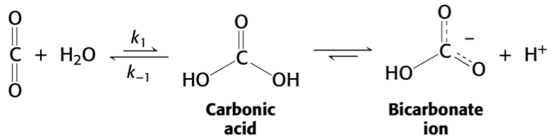
- *Details of the H-bonding and hydrophobic interactions allows us to visualize how the ligands bind.*
- *For structure-based drug design these details are crucial towards identifying ways to enhance drug-binding and reduce drug-resistance.*

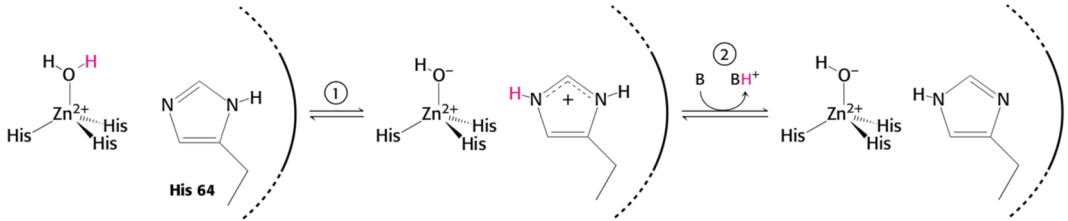
Enzyme mechanisms

- Enzymes act as biological catalysts, increasing the reaction rate by lowering the activation energy. Almost all metabolic processes in the cell need enzyme catalysis in order to occur at rates fast enough to sustain life.
- *H-atoms/protons of amino-acids and waters located in the active site are involved (directly/in-directly) in the reaction.*
- Details of amino-acid protonation states & water positions/orientations (and discrimination between H_2O, OH^-, H_3O^+) required for determining the correct catalytic pathway. NB* *Amino-acid side-chain pKa's can be shifted in enzyme active-sites!*



Carbonic anhydrase 10^7 times faster!





- *Understanding enzyme mechanisms can be hugely beneficial for the design of effective enzyme inhibitors or activators.*

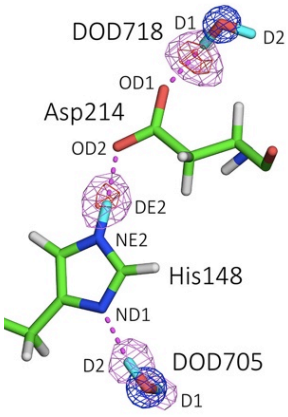
Neutron diffraction and the hydrogen isotopes

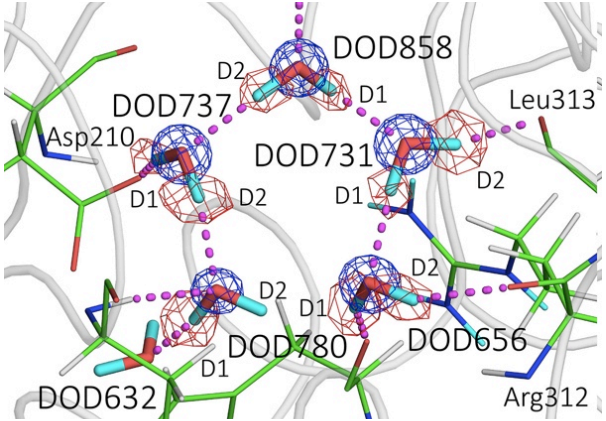
- Neutrons are scattered by atomic nuclei (*cf.* X-rays are scattered by electrons).
- The coherent scattering length of ^{1}H and ^{2}H (D) are of similar magnitude to the other common elements of biological macromolecules (C, N, O, S, etc).
- Advantageous to exchange ^{1}H for ^{2}H (D) to enhance visibility (S/N) in nuclear scattering length density maps.

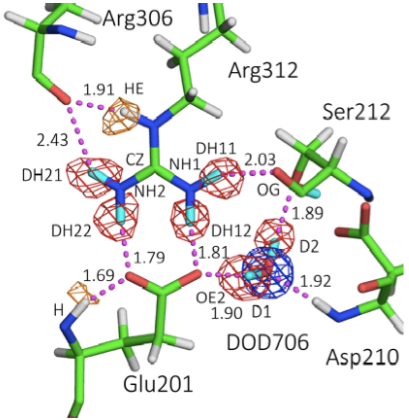
Isotope	Atomic number	Neutron incoherent cross-section (barns, 1 barn = 10^{-24} cm^2	Neutron coherent scattering length, *b* (10^{-12} cm)	X-ray scattering length (sin θ = 0)	X-ray scattering length (sin θ = 0.5 Å^{-1})
^{1}H	1	80.27	-0.374	0.28	0.02
^{2}H (D)	1	2.05	0.667	0.28	0.02
^{12}C	6	0.00	0.665	1.69	0.48
^{14}N	7	0.50	0.937	1.97	0.53
^{16}O	8	0.00	0.580	2.26	0.62
^{32}S	16	0.00	0.280	4.51	1.90

Neutron macromolecular crystallography (nMX)

- Neutron diffraction data from single-crystals of biological macromolecules allows H-atom/H^+ positions and D-atom/D^+ positions to be revealed at resolutions (d_{min}) of ~1.5Å and 2.5Å, respectively.
- Provides key details of H-bonding, protonation and hydration required for understanding many biological processes *e.g.* protein-folding/stabilization, small-molecule ligand-binding and enzyme mechanisms.
- Radiation damage-free structures can be determined at room-temperature (*cf.* X-rays and electrons) or low-T.



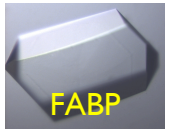


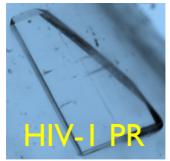


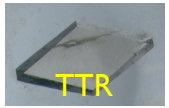
Shibazaki *et al.*, (2018) *J. Mol. Biol.* **430**, 5094.

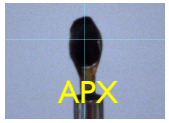
Sample preparation – level of deuteration

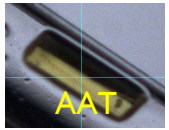
- H/D-exchanged (~15% H to D): Soaking in, or vapour diffusion of deuterated buffers.
- Perdeuterated (near to complete exchange): Expression of bacteria fed deuterated media (D-Lab, LSG).





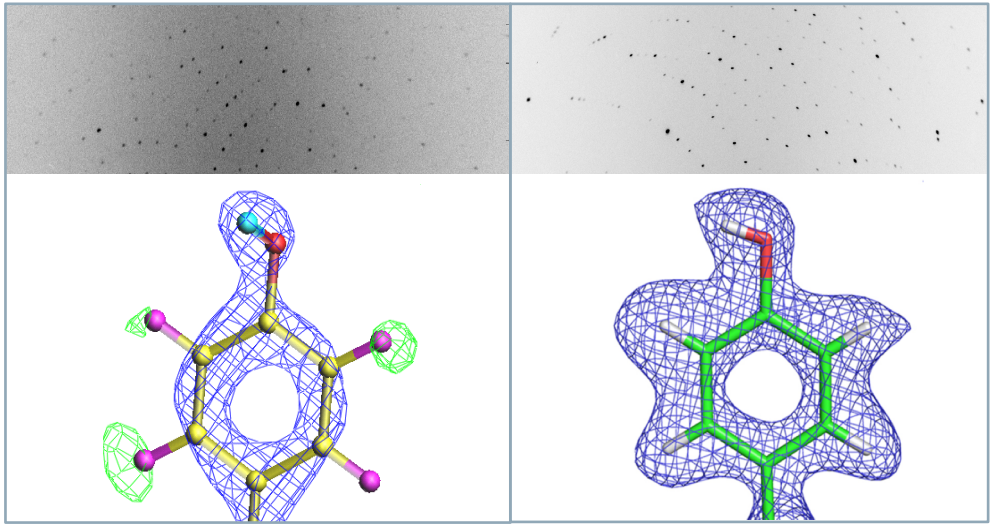






	H	D (^{2}H)
Coherent	-3.74	+6.67
Incoherent	80.27	2.05

Protein	Crystal volume (mm^3)	M.Wt (kDa)
FABP	0.05	15
HIV-1 PR	0.10	22
TTR	0.11	28
APX	0.14	29
PKAc	0.10	44
AAT	0.65	93

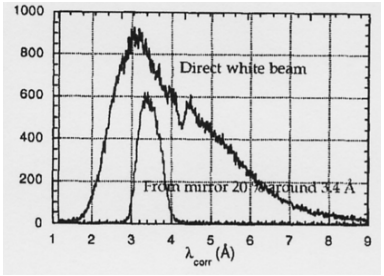


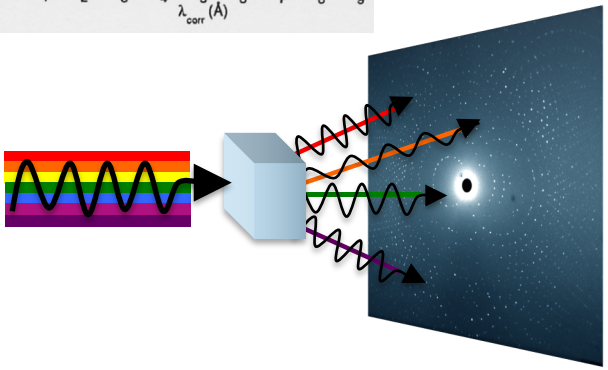
Perdeuteration allows (i) smaller crystals *cf.* H/D-exchanged and (ii) avoids map cancellation issues at CH_n groups.

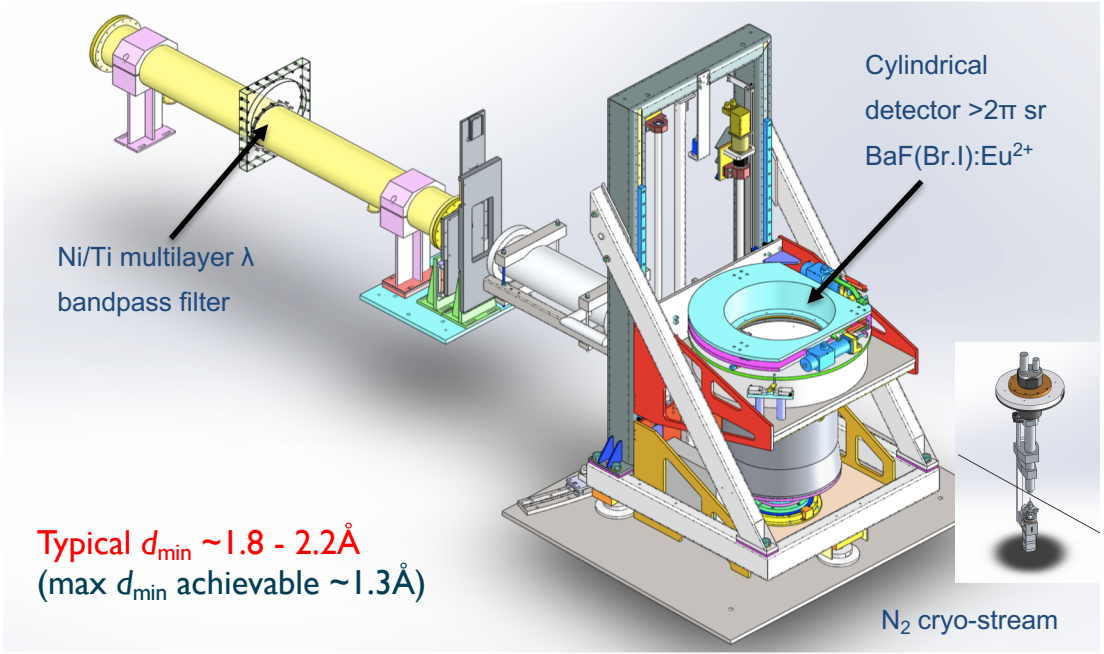
Haertlein *et al.*, (2016) *Methods Enzymol.* **566**, 113.

Data collection – Laue Diffractometer LADI

- Neutron guide H143 transports 'cold' neutrons (2.5Å < λ < 10Å) from the reactor to the primary spectrometer.
- Ni/Ti multilayer filter (δλ/λ~30%) selects wavelengths required for data collection (typically 2.8Å < λ < 3.8Å).
- Quasi-Laue methods in combination with a large (>2π sr) cylindrical detector enhances data collection efficiency.



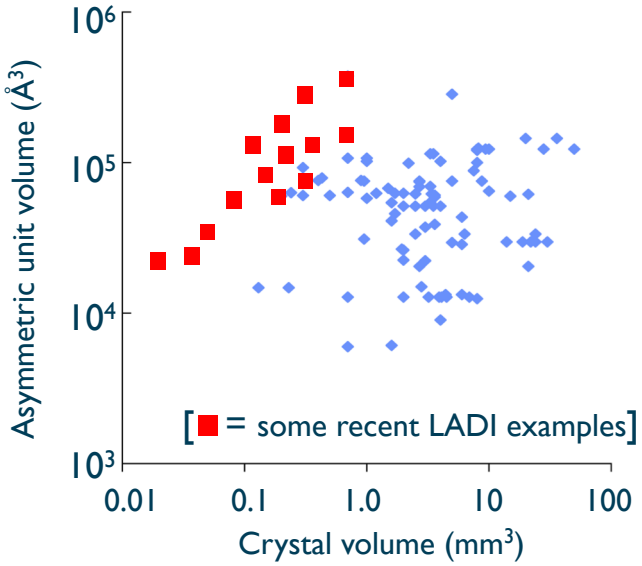


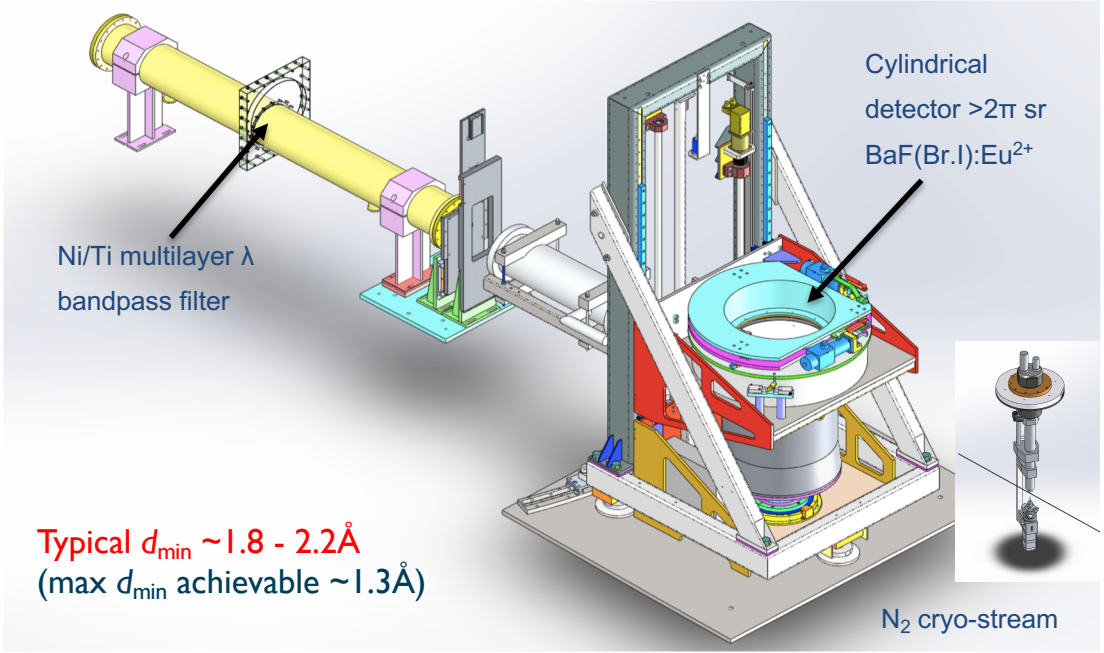


Blakeley & Podjarny (2018) *Emerg. Top. Life. Sci.* **2**, 39; Blakeley *et al.*, (2015) *IUCrJ* **30**, 464; Blakeley *et al.*, (2010) *Acta Cryst. D***66**, 1198; Blakeley *et al.*, (2009) *Cryst. Rev.* **15**, 157.

Data collection – Laue Diffractometer LADI

- Collection at room- or cryo-temperatures from H/D-exchanged or perdeuterated crystals
- Crystal volumes typically from ~0.05 to 1mm^3 (*unit-cell vol., space group, level of deuteration etc*).
- Data collection from a few hours up to ~2 weeks (*crystal vol., unit-cell vol., space group, level of deuteration etc*)

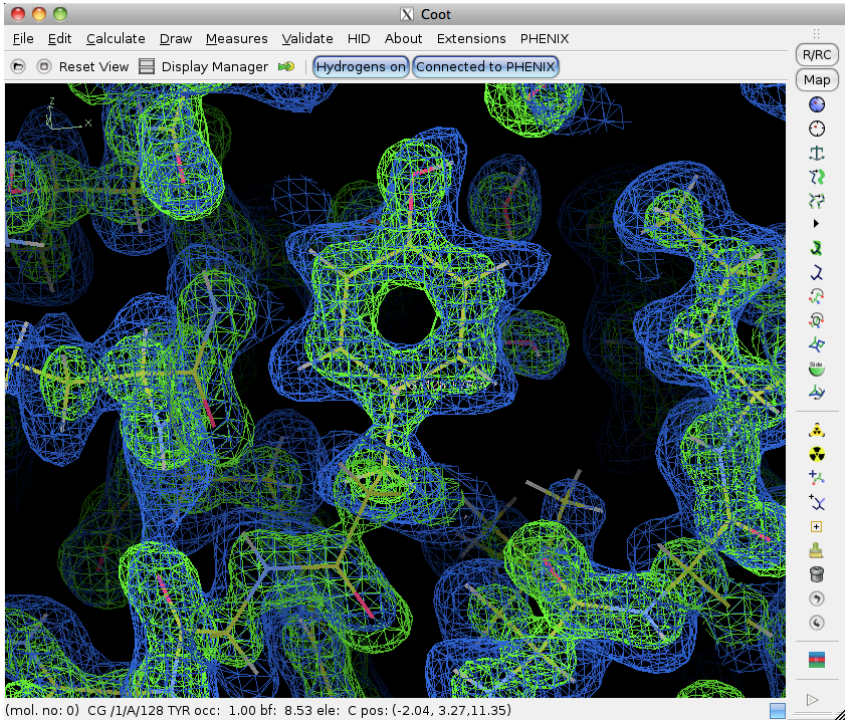




Blakeley & Podjarny (2018) *Emerg. Top. Life. Sci.* **2**, 39; Blakeley *et al.*, (2015) *IUCrJ* **30**, 464; Blakeley *et al.*, (2010) *Acta Cryst. D***66**, 1198; Blakeley *et al.*, (2009) *Cryst. Rev.* **15**, 157.

Laue diffraction data processing / Structural refinement

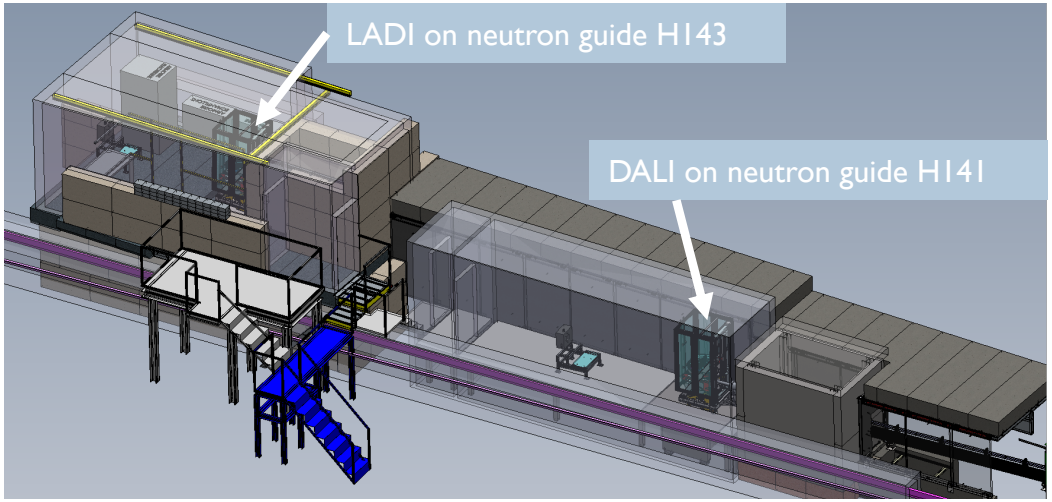
- Neutron Laue diffraction data indexed (*h, k, l*) and integrated (I, sig(I)) using *LAUEGEN*.
- Intensities are λ-normalized using *LSCALE*.
- Data then processed with standard X-ray software from *CCP4* (https://www.ccp4.ac.uk/).
- Structural refinement (*x,y,z,* B, occ) against neutron data, X-ray data or both in a joint X-ray/neutron strategy using *PHENIX* software suite and the molecular visualization program *COOT*.
- Electron density shown in **green.**
- Nuclear scattering length density in **blue.**



Campbell *et al.,* (1998) *J. Appl. Cryst* **31**, 496; Arzt *et al.,* (1999) *J. Appl. Cryst* **32**, 554; Afonine *et al.,* (2010) *Acta Cryst. D***66**, 1153; Emsley *et al.,* (2010) *Acta Cryst. D***66**, 486.

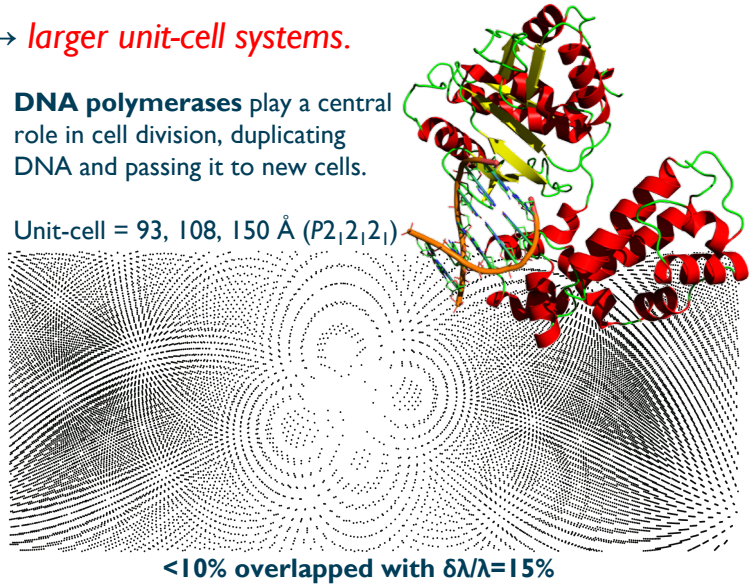
Extending the limits for nMX - DALI (Coquelle/Fuard/Clancy/Ruiz/Ollivier)

- nMX is an expanding field and for several years over-demand for LADI has been high (>2.5).
- New instrument DALI installed at ILL as part of Endurance Programme to extend capacity and capability.
- Neutron velocity selector (NVS) provides ~2.7x higher transmission of neutrons (*cf.* multilayer) → *smaller crystals.*
- NVS can deliver a reduced bandwidth (δλ/λ) → less spatial overlaps → *larger unit-cell systems.*



DNA polymerases play a central role in cell division, duplicating DNA and passing it to new cells.

Unit-cell = 93, 108, 150 Å ($P2_12_12_1$)

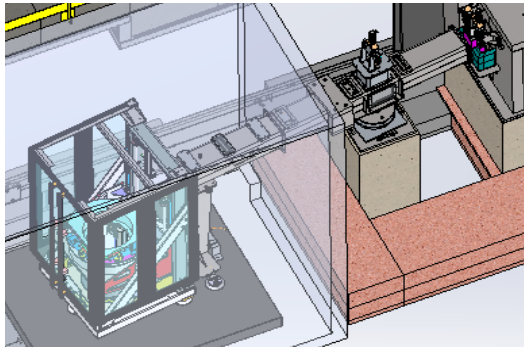


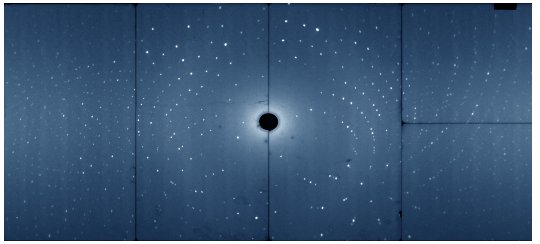
<10% overlapped with δλ/λ=15%

- Commissioning during the pandemic a challenge but ILL staff have made huge efforts, allowing DALI to be in the user programme for the 3rd cycle - *a huge thank you to everyone involved!*

Current status - DALI

- Commissioning close to complete - software, background issues etc. resolved. Installation and testing of the cryo-stream system to be done in April.
- Flux at sample position **2.9x higher than LADI** (using λ_{range} 3.2 - 4.0 Å, $\delta\lambda/\lambda$ ~22%).
- Comparative tests have been made with 'standard' crystals → data reduction and refinement statistics per resolution shell correlate well with the flux gain.





	Starting	Final
R-work	0.2274	0.1789
R-free	0.2598	0.2327
Bonds	0.093	0.002
Angles	0.571	0.544

	R-work	R-free	%complete
30.21 - 3.27	0.1347	0.2014	93.9%
3.27 - 2.60	0.2133	0.2628	80.6%
2.60 - 2.27	0.2102	0.2340	71.1%
2.27 - 2.06	0.2143	0.2605	62.0%
2.06 - 1.91	0.2553	0.2880	50.5%
1.91 - 1.80	0.3412	0.4094	33.0%

- Currently using the spare SANS NVS (tilted at -5 degrees), to go to the lowest wavelength range possible (3.2 - 4.0 Å) → restricts max. d_{min} achievable to ~1.9 Å
- Spare SANS NVS with $\delta\lambda/\lambda$ ~22% too wide for the larger unit-cells we wish to study.
  - *Both of these issues will be resolved with the arrival of the optimised NVS!*

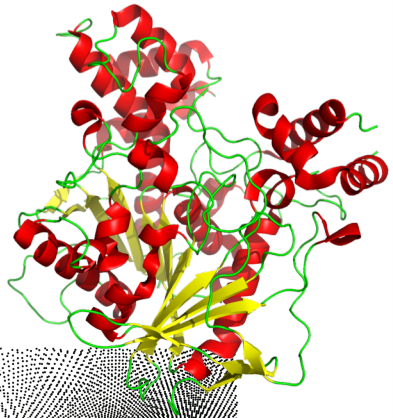
Central λ	3.13 Å	3.57 Å	3.80 Å
Tilt (deg)	-3	-1	0
Transmission (%)	61	78	79
δλ/λ FWHM (%)	11	9.3	8.7
Useful δλ/λ (%)	18	15	14
Flux gain (cf. LADI)	2.1	2.5	2.7

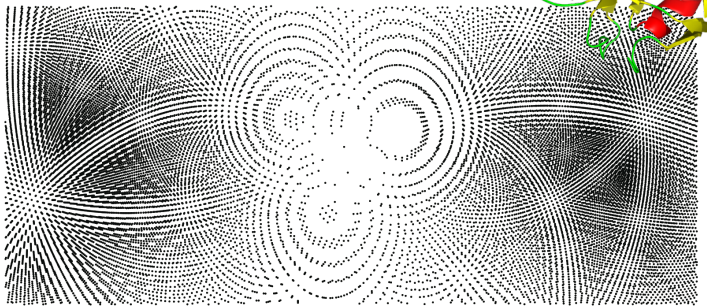
Future Science - DALI

Human acetylcholinesterases catalyze the breakdown of the neurotransmitter acetylcholine.

They are the target of inhibition by nerve agents and pesticides.

Unit-cell = 126, 126, 134 Å ($P3_112$)





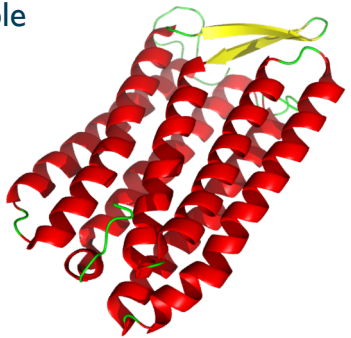
<15% overlapped with δλ/λ=15%

Data collected on LADI for hAcChE crystals (~0.1mm^3) diffracted to a d_{min} of 3.5 Å

Transmembrane receptor proteins enable cells to sense and respond to their environment by undergoing conformational changes on ligand binding or light absorption.

e.g. Archaerhodopsin-3

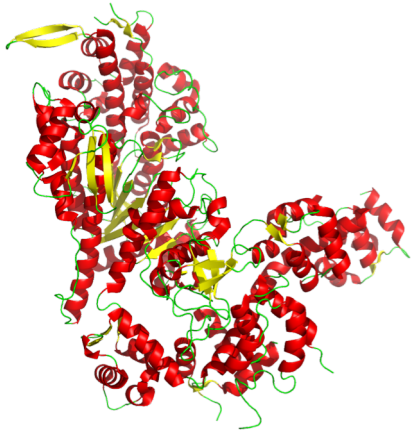
Unit-cell = 45, 47, 104 Å ($P2_12_12_1$)



Myosin proteins are a superfamily of proteins which bind actin, hydrolyze ATP and transduce force. Thus most are located in muscle cells.

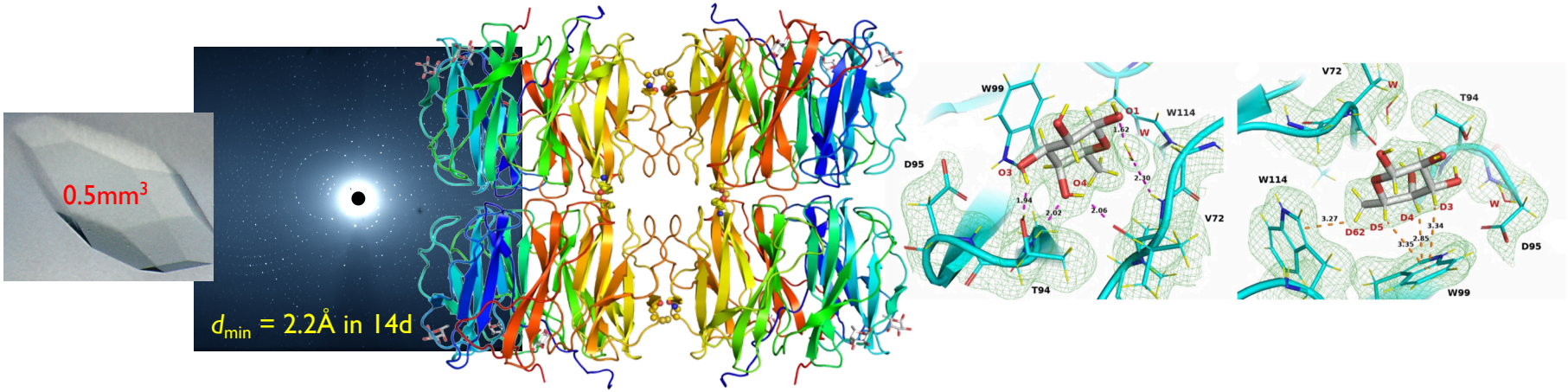
e.g. Myosin A full-length

Unit-cell = 90, 115, 170 Å ($P2_12_12_1$)



PLL lectin carbohydrate-binding studies (Imberty/Gajdos/Devos/Forsyth, France)

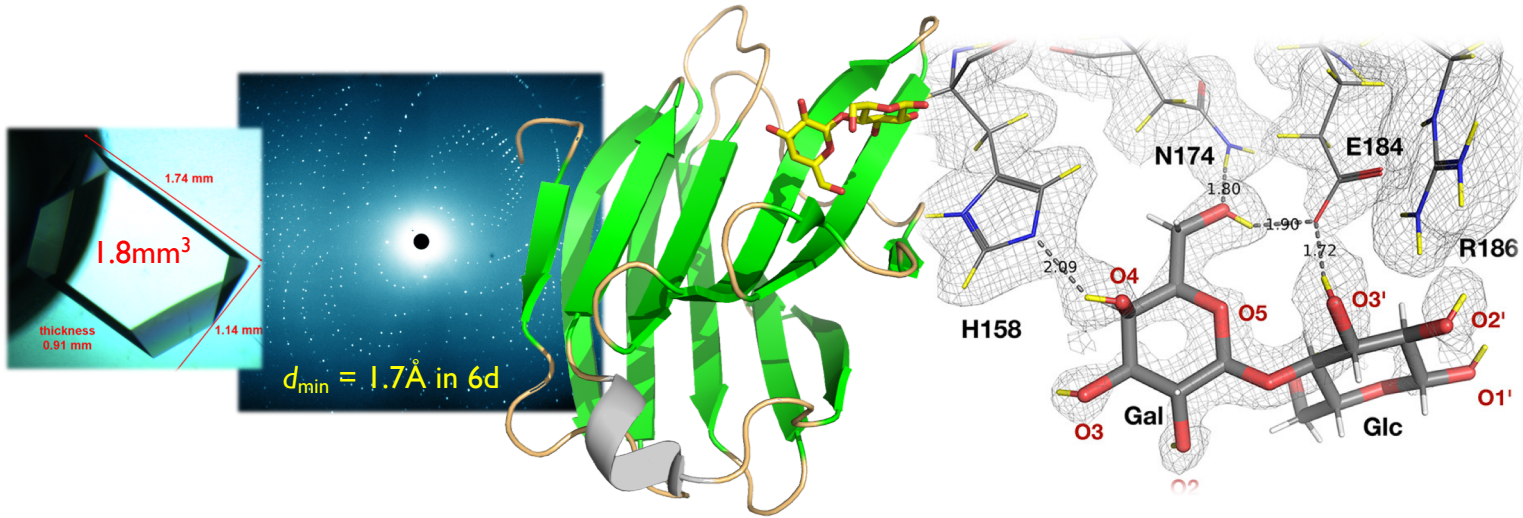
- Lectins bind to carbohydrates reversibly and specifically. Carbohydrate-mediated recognition & adhesion are key events in the interaction of bacteria with a host.
- *PLL* is a L-fucose-specific lectin from bacteria (*P. luminescens*) that lives in a symbiotic relationship with nematodes.
- Neutron data from crystals of *apo* PLL (H/D-exchanged) and PLL/fucose complex (*both fully deuterated!*) revealed details of H-bonding, hydration, and CH-π stacking interactions between fucose and the aromatic rings of tryptophan.



Gajdos *et al.*, (2021) *Structure.* S0969-2126(21)00076-9 (*online ahead of print*); Gajdos *et al.*, (2021) *Glycobiology* **31**, 151; Kumar *et al.*, (2016) *J. Biol. Chem.* **291**, 25032

Human Galectin-3 inhibitor design studies (Logan/Manzoni, France)

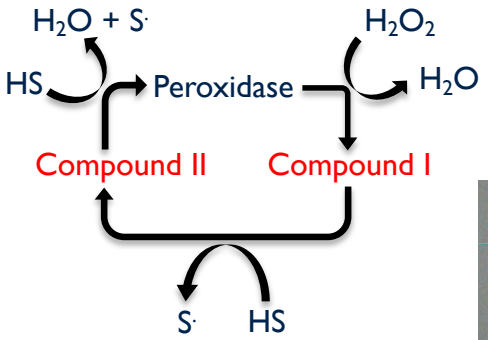
- *Galectin-3* binds galactose-containing moieties on glycoproteins, and is involved in cell-to-cell adhesion, cell growth, and cell differentiation. Important drug target since it is implicated in breast cancer and heart disease.
- Details of H-bonding, protonation and hydration for Gal3C complexes with (i) lactose and (ii) glycerol and in the *apo* form confirmed that the design of inhibitors should be based on the disaccharide core with the addition of non-sugar groups to increase the binding affinity (*e.g.* TD139, phase II clinical trials for pulmonary fibrosis).

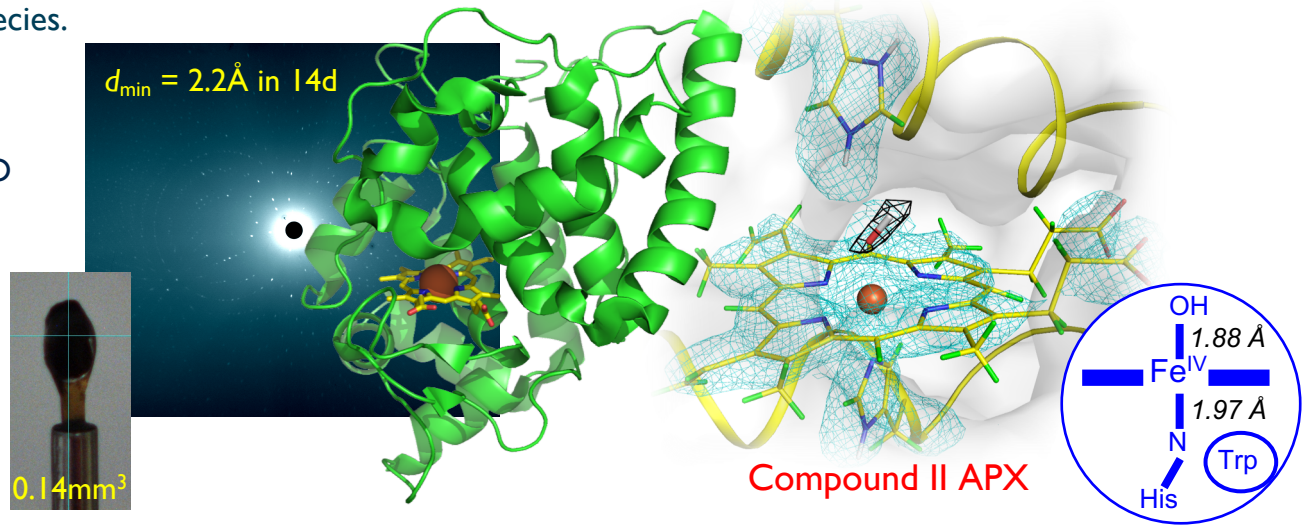


Logan (2020) *Methods Enzymol.* **634**, 201; Manzoni *et al.,* (2018) *J. Med. Chem.* **61**, 4412; Manzoni *et al.,* (2016) *Acta Cryst. D* **72**, 1194

Heme peroxidases enzyme mechanism studies (Moody/Raven/Kwon, UK)

- *Heme peroxidases* carry out a wide range of oxidations using highly reactive **ferryl intermediates.**
- Cryo-trapping (at 100K) the transient reaction intermediates Compound I in cytochrome *c* peroxidase & Compound II in ascorbate peroxidase (APX) allowed us to reveal their chemical species as Fe(IV)=O and Fe(IV)-OH, respectively.
- Neutron data for APX/ascorbate complex allowed us to visualize the protons in a PCET pathway, which included an unexpected neutral Arg species.

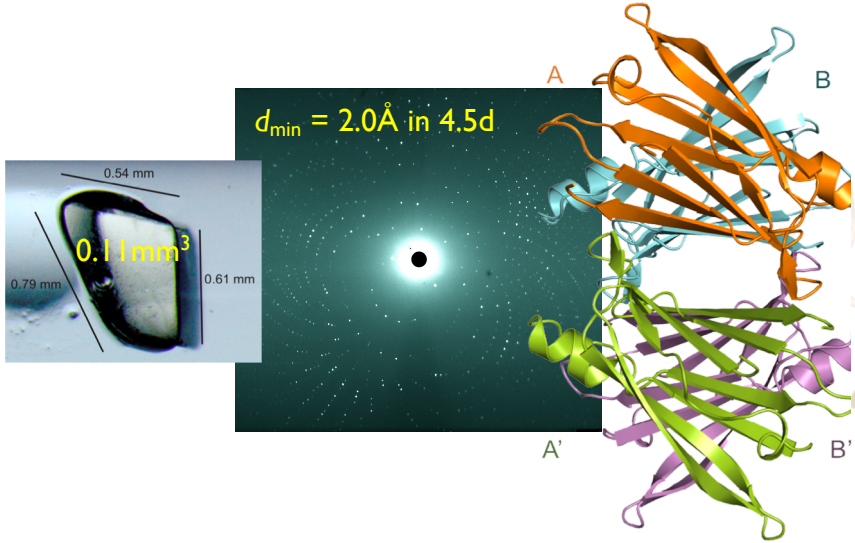


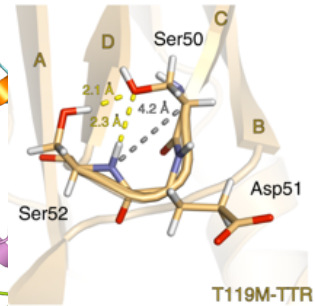


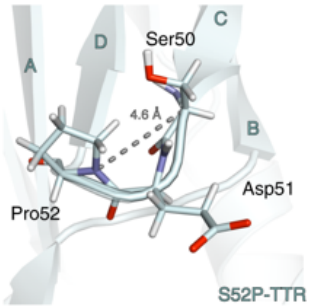
Kwon *et al.*, (2020) *PNAS* **117**, 6484; Moody & Raven (2018) *Acc. Chem. Res.* **51**, 427; Kwon *et al.*, (2016) *Nat. Commun.* **7**, 13445; Casadei *et al.*, (2014) *Science* **345**, 193

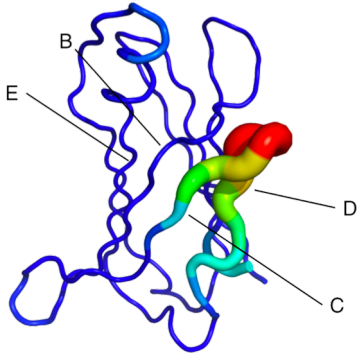
Transthyretin amyloidosis (Yee/Forsyth/Cooper, France/UK)

- *Transthyretin* (TTR) is a transport protein that carries thyroxine and retinol-binding protein bound to retinol.
- Wild-type TTR and certain mutants are unstable (*e.g.* S52P), while other mutants are stable (*e.g.* T119M). Unstable forms have a propensity to dissociate to form amyloid fibrils leading to familial amyloid polyneuropathy (FAP).
- In T119M and WT TTR, Ser52 forms two H-bonds with Ser50, while in S52P the absence of these H-bonds creates a looser CD loop. In T119M and the S52P/tafamidis complex, additional H-bonds are formed between chains B and B'





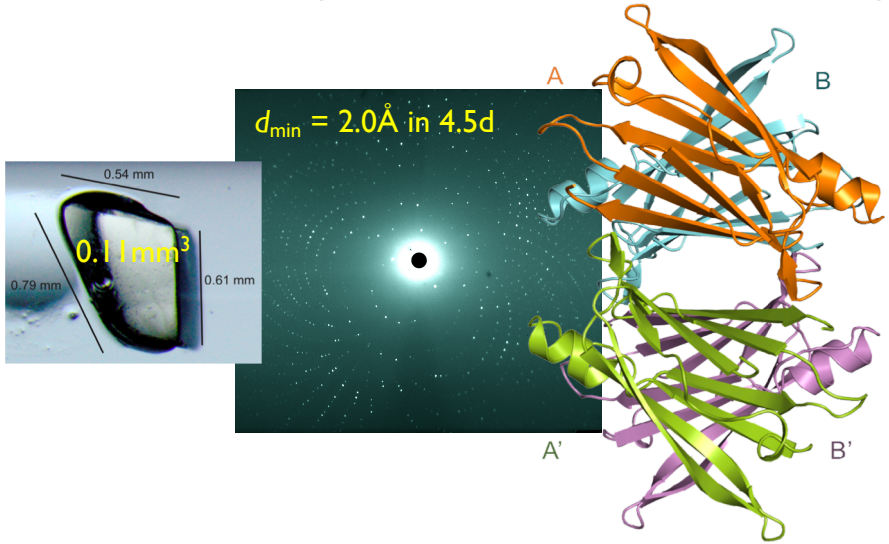


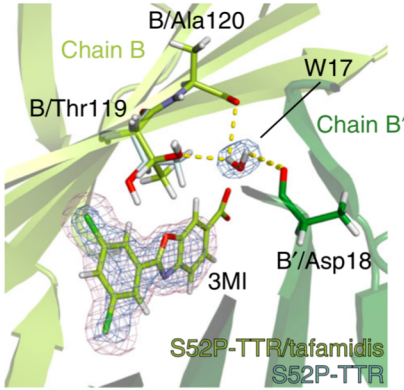


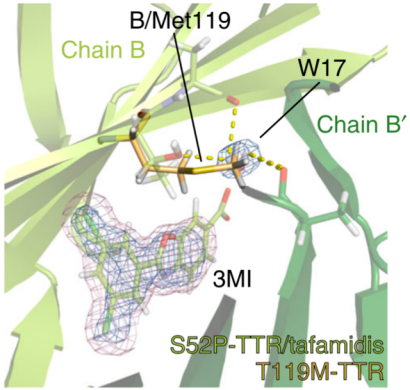
Yee *et al.*, (2019) *Nat. Commun.* **10**, 925; Yee *et al.*, (2017) *J. Appl. Cryst.* **50**, 660; Yee *et al.*, (2016) *Angew. Chem. Int. Ed. Engl.* **55**, 9292; Haupt *et al.*, (2014) *IUCrJ* **1**, 429.

Transthyretin amyloidosis (Yee/Forsyth/Cooper, France/UK)

- *Transthyretin* (TTR) is a transport protein that carries thyroxine and retinol-binding protein bound to retinol.
- Wild-type TTR and certain mutants are unstable (*e.g.* S52P), while other mutants are stable (*e.g.* T119M). Unstable forms have a propensity to dissociate to form amyloid fibrils leading to familial amyloid polyneuropathy (FAP).
- In T119M and WT TTR, Ser52 forms two H-bonds with Ser50, while in S52P the absence of these H-bonds creates a looser CD loop. In T119M and the S52P/tafamidis complex, additional H-bonds are formed between chains B and B'



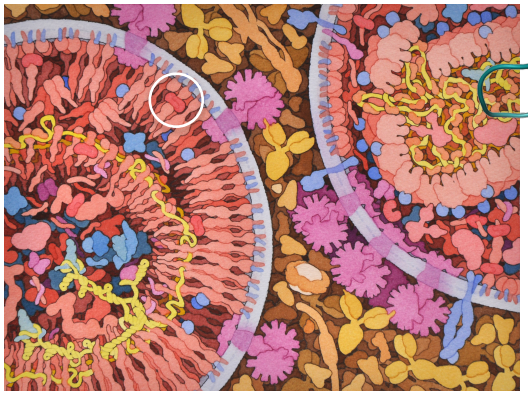


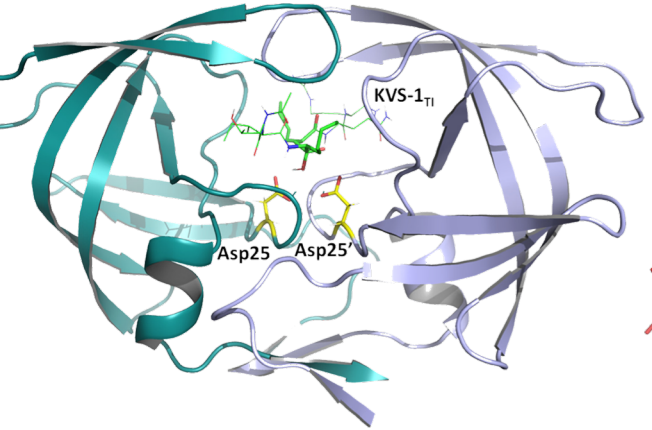


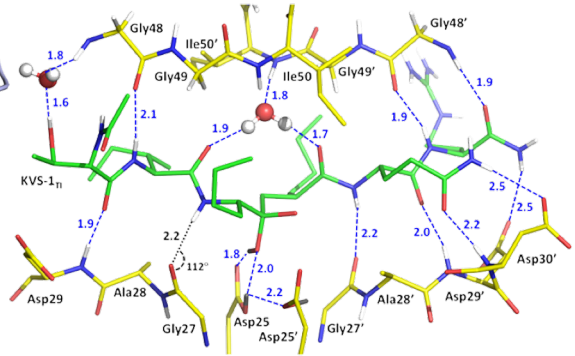
Yee *et al.*, (2019) *Nat. Commun.* **10**, 925; Yee *et al.*, (2017) *J. Appl. Cryst.* **50**, 660; Yee *et al.*, (2016) *Angew. Chem. Int. Ed. Engl.* **55**, 9292; Haupt *et al.*, (2014) *IUCrJ* **1**, 429.

HIV-1 protease drug-binding/enzyme mechanism studies (Kovalevsky, USA)

- *HIV-1 protease* is an essential enzyme in the life-cycle of HIV and is a clinical drug target. RT neutron studies of wild-type enzyme drug complexes and triple mutant enzyme drug complexes at different pHs have revealed how the different drugs bind and the key interactions, which allowed us to suggest ways to enhance the binding and limit drug-resistance.





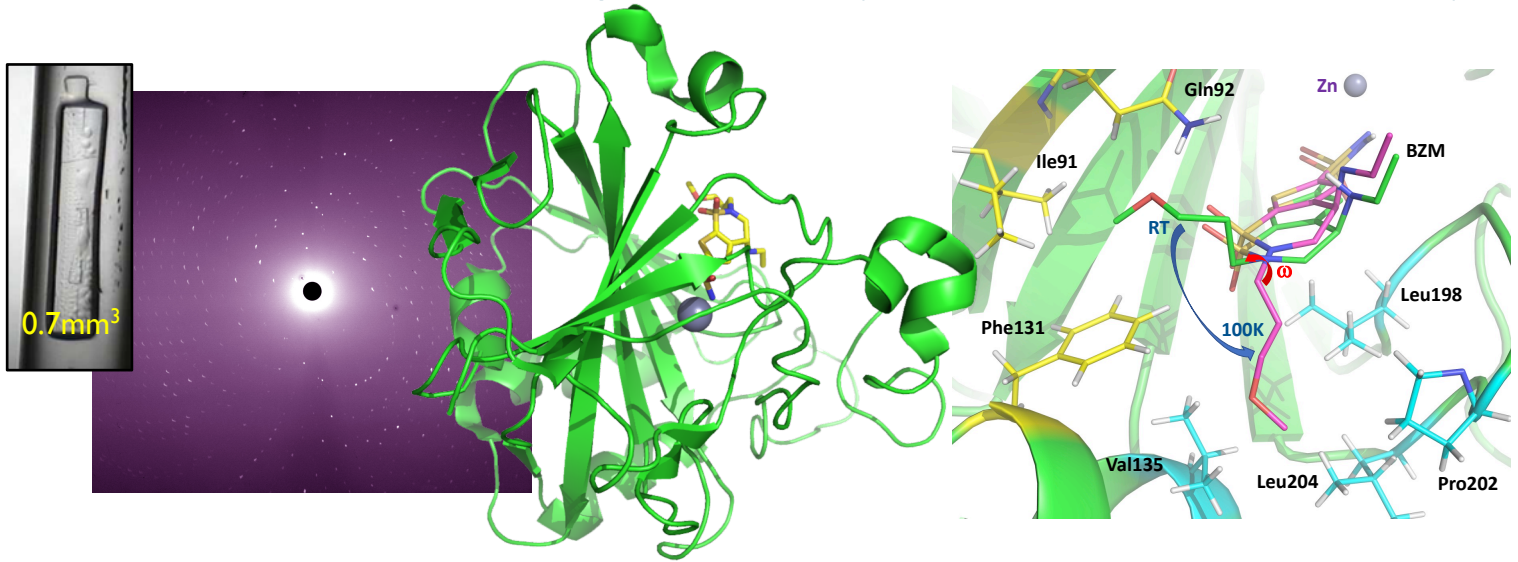


- By using a peptidomimetic inhibitor KVS-1 containing a reactive nonhydrolyzable ketomethylene isostere, we showed that the tetrahedral intermediate is an oxyanion, rather than a gem-diol (Kumar *et al.,* (2020) *ACS Omega* **5**, 11605).

Kovalevsky *et al.,* (2020) *Methods Enzymol.* **634**, 257; Gerlits *et al.,* (2017) *J. Med. Chem.* **60**, 2018.
Gerlits *et al.,* (2016) *Angew. Chem.* **55**, 4924; Weber *et al.,* (2013) *J. Med. Chem.* **56**, 5631.

Human carbonic anhydrase-II drug-binding studies (McKenna/Fisher/Kovalevsky, USA/Sweden)

- *Carbonic anhydrase-II* is a target of sulfonamide drugs used against glaucoma and breast cancer.
- H-bonding, protonation and hydration patterns for 3 different clinical inhibitors provided insights for the design of isoform-specific drugs.
- *Differences observed between 100K X-ray and RT neutron (*cf.* HIV-1 PR and SARS-CoV-2 M^{pro} studies).



Kovalevsky *et al.,* (2018) *Structure* **26**, 383; Aggarwal *et al.,* (2016) *IUCrJ* **3**, 319; Michalczyk *et al.,* (2015) *PNAS* **112**, 5673; Fisher *et al.,* (2012) *J. Am. Chem. Soc.* **134**, 14726

NEUTRONS FOR SOCIETY

Seeing the chemistry in biology using neutron crystallography

- Neutron crystallography allows the positions of all the H-atoms and protons in biological macromolecules to be directly visualized at room temperature.

- Given the important roles H-atoms, protons and H-bonding play in biological systems and processes, neutron crystallography is a very useful technique for structural biology, answering questions that are unattainable using other techniques.

- The new DALI instrument further extends the capabilities and capacity for nMX at ILL, allowing the use of smaller crystals and the study of larger macromolecules and their complexes.