What neutrons can do for you: From diffraction on protein crystals to small angle neutron scattering with contrast matching

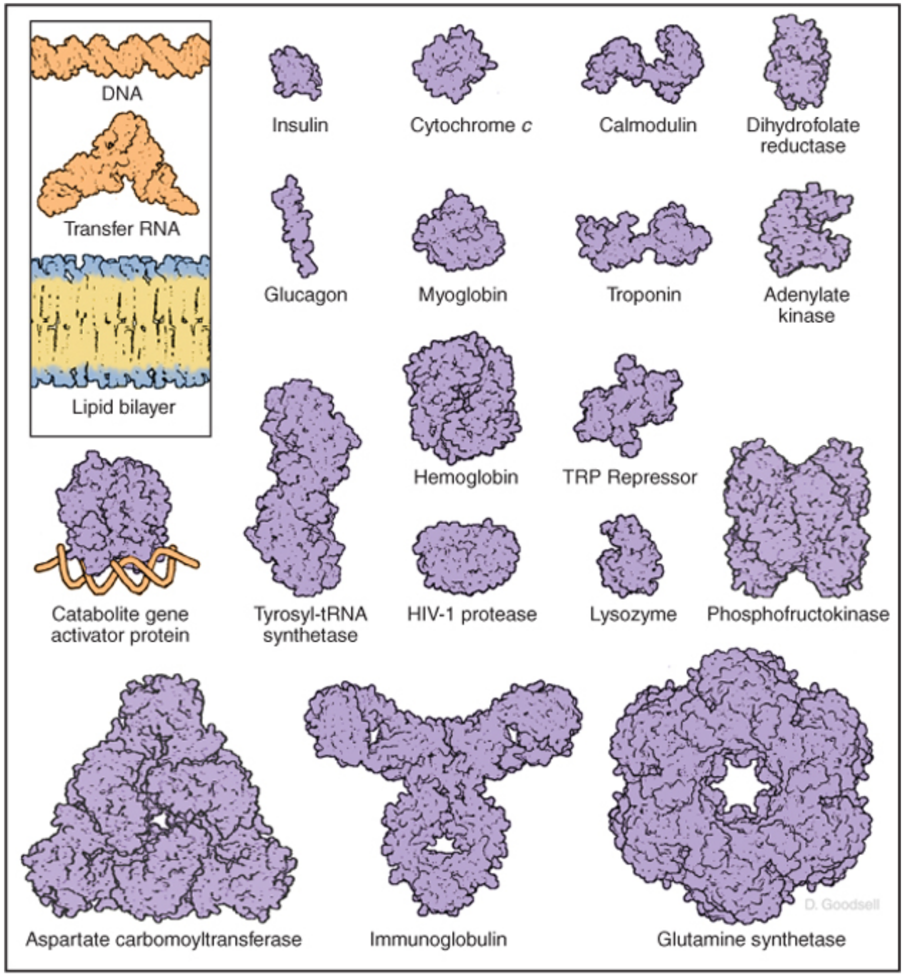
Seminar, National Academy of Science, Mongolia

May 23rd 2022 | Tobias E. Schrader

Outline

- Motivation: Why do we need protein structures at atomic resolution?
- neutron protein crystallography
- one application example: PL^{pro} of SARS-CoV2
- Small angle neutron scattering
- Contrast matching with small angle neutron scattering
- one application example: matched bicelles for membrane proteins
- Summary

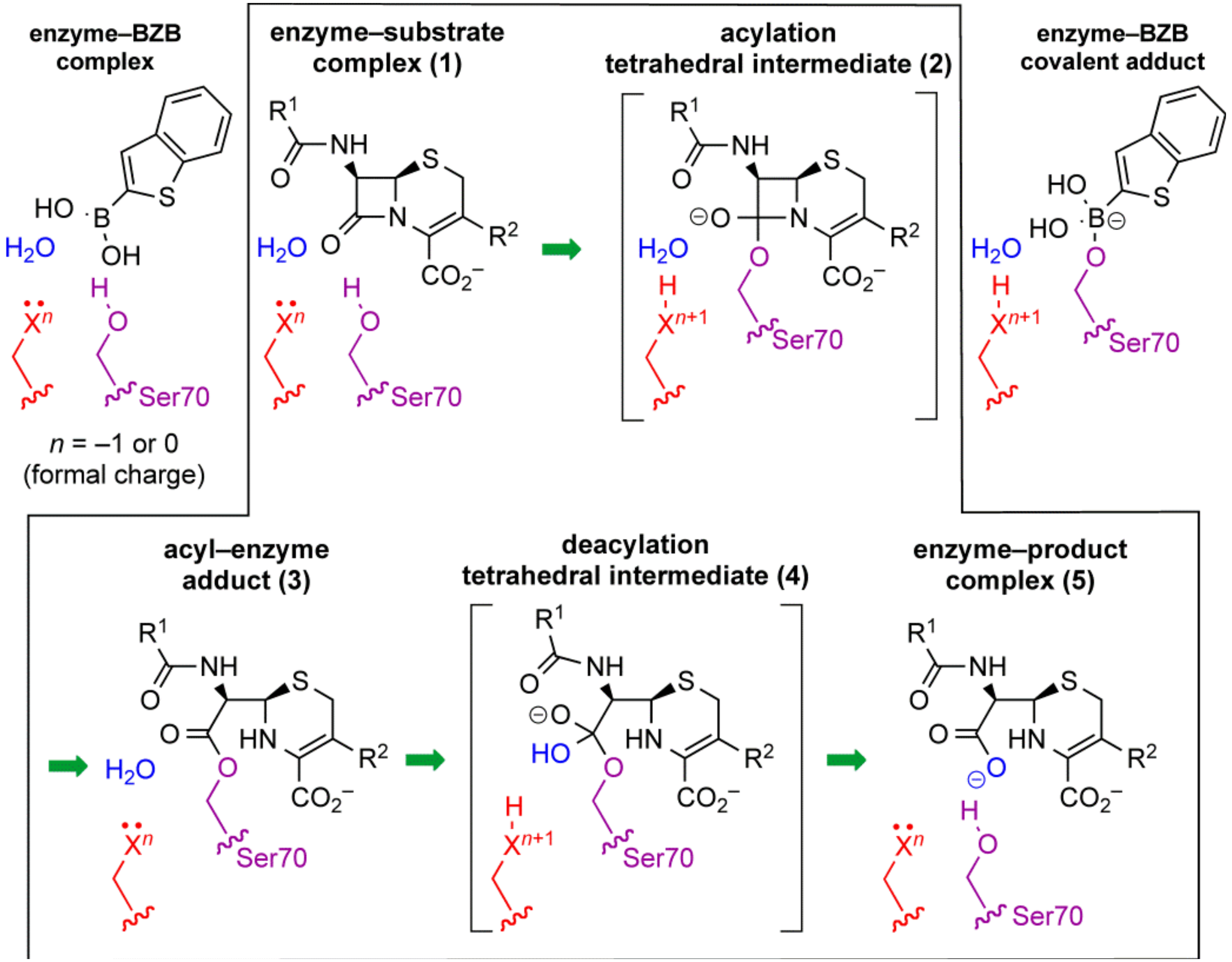
Proteins are structured macromolecules and come in different sizes and shapes



© Elsevier. Pollard et al: Cell Biology 2e - www.studentconsult.com

The structure is crucial for the protein‘s function

β-lactamase: An enzyme that destroys β-lactam antibiotics



The catalytic cycle of a class A β-lactamase illustrated for a cephalosporin substrate (inside box) and the mode of inhibition by BZB (outside box). The general base employed is not necessarily the same for acylation and deacylation. The overall reaction pathway for β-lactam hydrolysis of a cephalosporin-like substrate by the class A β-lactamase enzymes.

Figures from: Tomanicek et al., doi: 10.1074/jbc.M112.436238

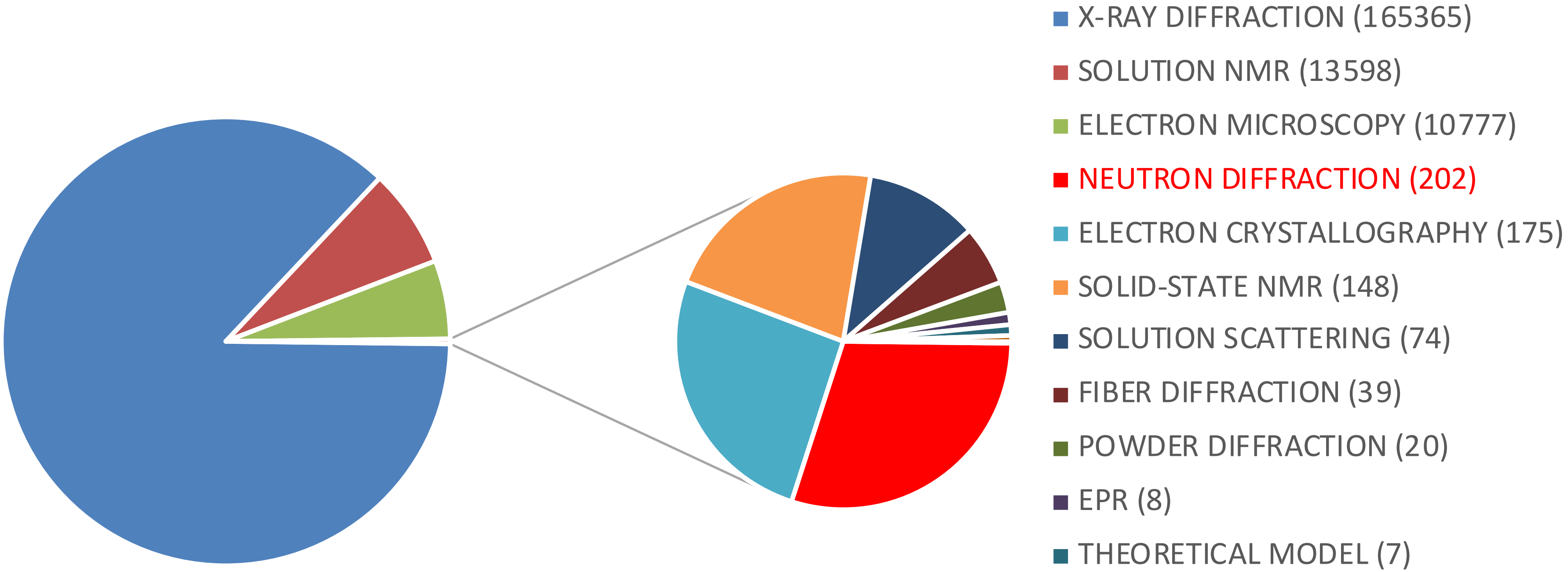
How do we find out about protein structures?

Why do we need experimental studies on proteins?

- **MD-Simulations** suffer from non-perfect force fields: Especially the long range electrostatics is not reproduced very well. But proteins use defined and structure related electrostatics to move the acidity constants of side chains in order to make them fullfill their tasks. MD-simulations cannot model bond breaking and forming very well since the quantum chemistry nature of this process is not included in the theoretical foundation of MD.
- **Ab initio quantum chemical calculations** are still too demanding to model the complete active centre of a protein (including its substrate)

Most structures are obtained by x-ray crystallography, available neutron structures in protein data bank: ca. 200

Entries in the protein data base as of 9th of May 2022

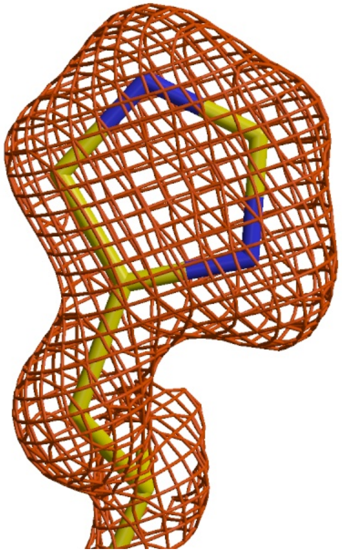


Total number of structures: 190418

http://www.rcsb.org/

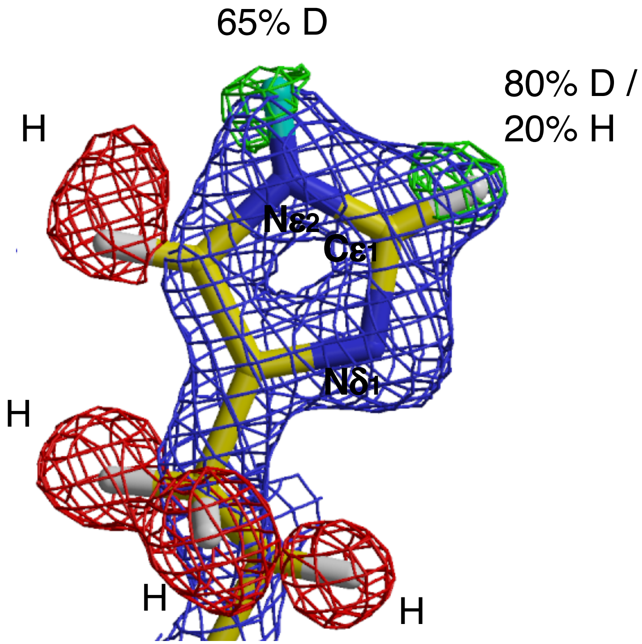
X-ray data versus neutron data on the same protein:

X-ray d_{min}= 1.5Å:



🟧 2Fo-Fc map; +1.5σ

neutrons d_{min}= 1.5Å:



🟦 2Fo-Fc map; +1.5σ

🟥 Fo-Fc omit-map; -3.0σ

🟩 Fo-Fc omit-map; +3.0σ

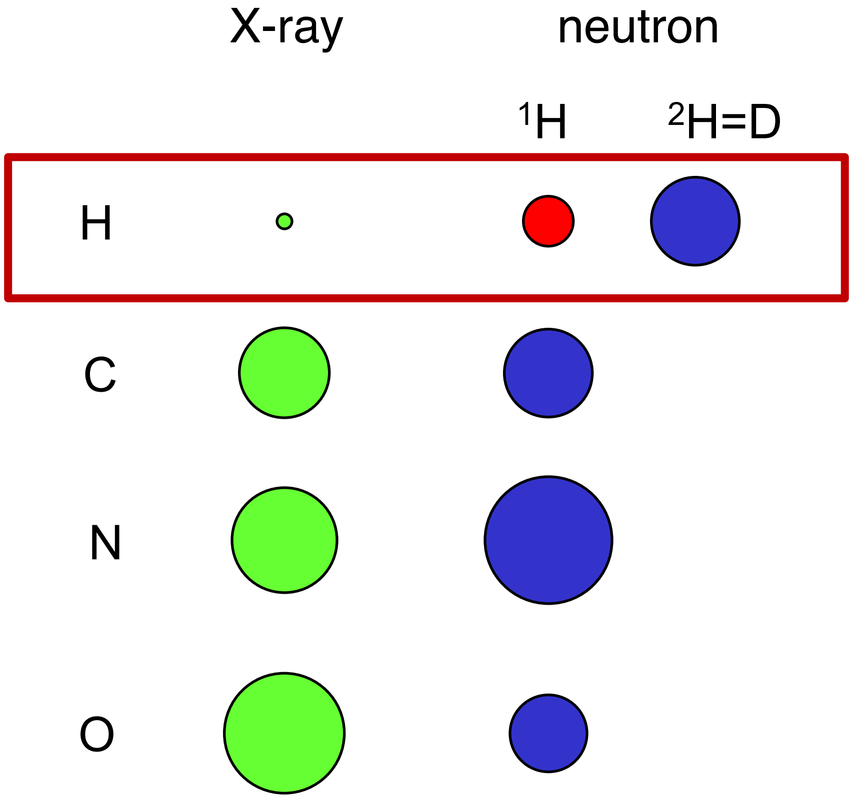
Niimura N, Chatake T, Ostermann A, Kurihara K, Tanaka T. (2003) Z. Kristallogr. 218:96

Advantages of structure determination with neutrons:

Comparison of form factors (X-ray) and scattering lengths (neutrons):

Nucleus	atomic number	scattering length [10^{-12} cm]
^{1}H	1	**-0.378**
^{2}H	1	**0.667**
^{12}C	6	**0.665**
^{15}N	7	**0.921**
^{16}O	8	**0.581**

σ_{coh} of ^{1}H is 1.8×10^{-28} m^{2} but
σ_{incoh} of ^{1}H is 80.2×10^{-28} m^{2}
Large background from hydrogen atoms!



diameters correspond to:
form factor / scattering length
(scaled for C-atom)

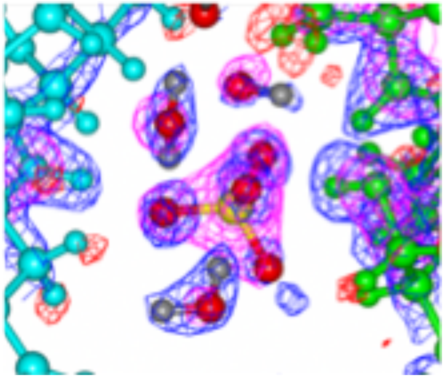
Advantages of Structure Determination with Neutrons

Hydrogen/deuterium atoms can be resolved even at a resolution of $d_{min} \approx 2.5$ Å (for ^2H). Therefore one can determine:

→ protonation states of amino acid side chains and ligands

→ deuterium exchange as a measure of flexibility and accessibility (discrimination between **H** / **D**)

→ solvent structure including hydrogen atoms



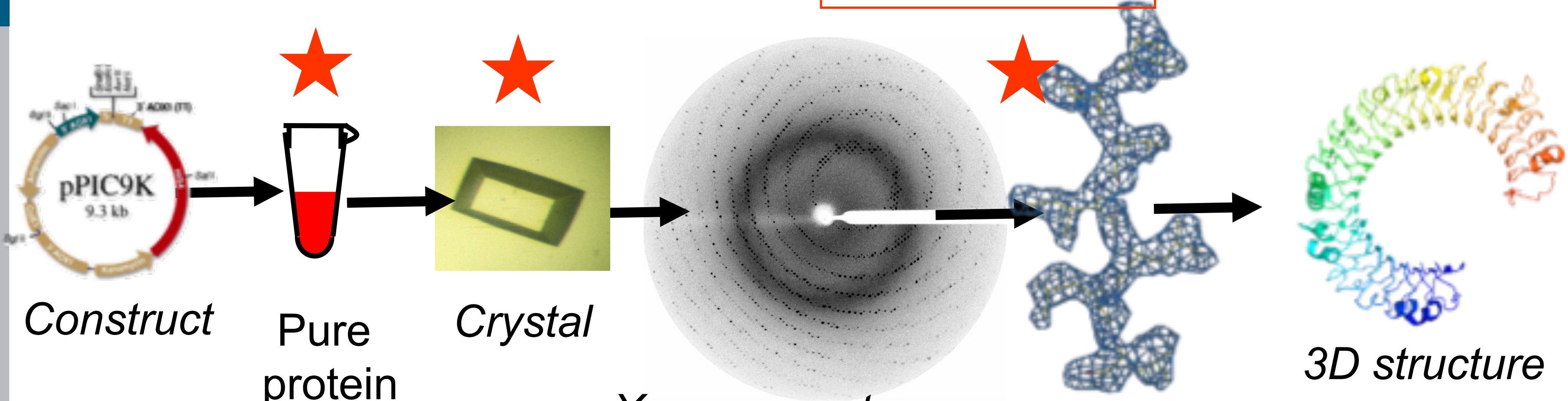
Water network in the contact region between two myoglobin molecules in the crystal.

x-ray map (magenta): contour level of $+2.7\sigma$
nuclear map (red): contour level of -1.75σ
nuclear map (blue): contour level of $+2.3\sigma$

Much less radiation damage as compared to x-rays: Metallo-proteins can be measured without reducing the metal centres

Protein crystallography in general, valid for both x-rays and neutrons as probes

Crystallography: Overiew over the process



Phase problem

Construct

Pure protein

Crystal

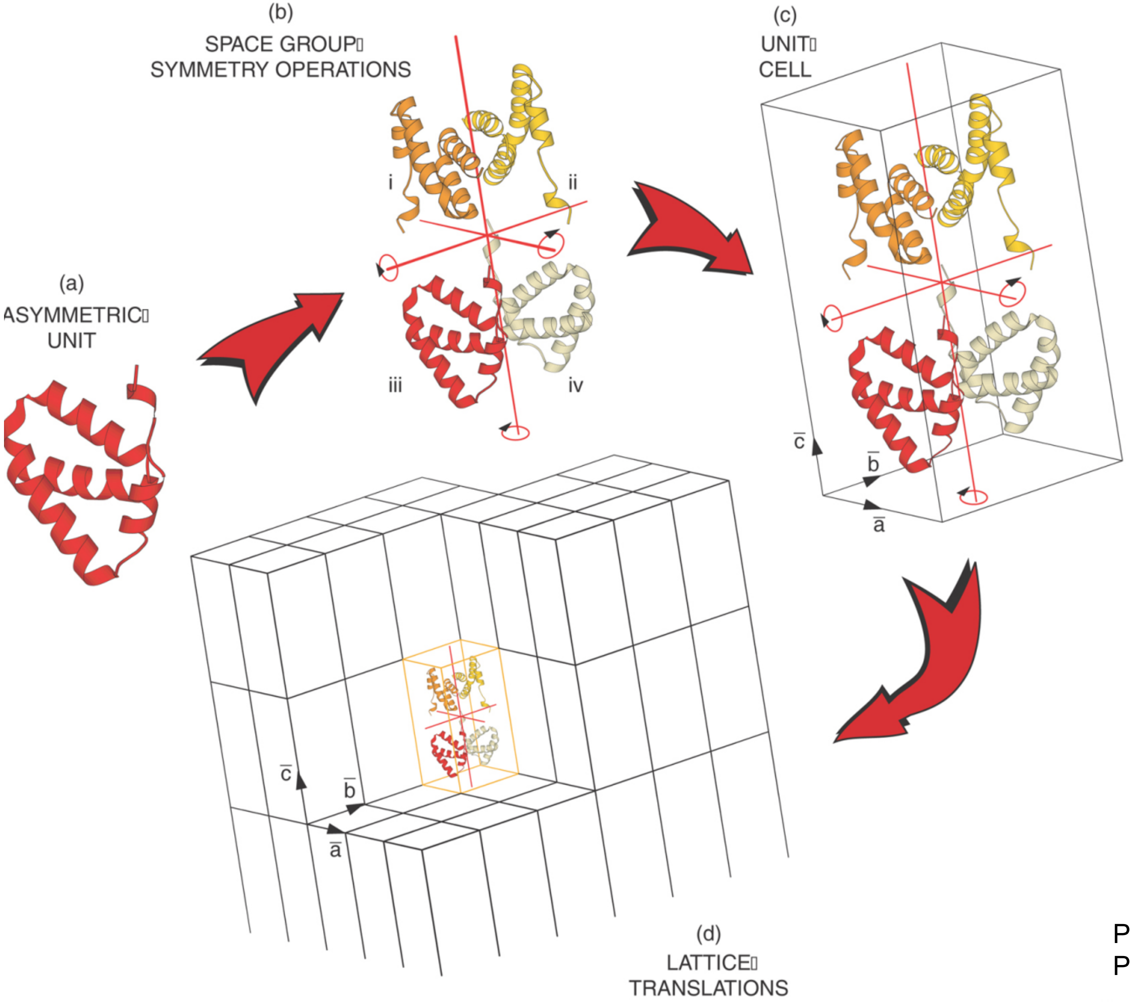
X-ray or neutron Diffraction

Electron density (x-rays) or nuclear scattering length density

3D structure

Harma Brondijk, Crystal and Structural chemistry, Utrecht University

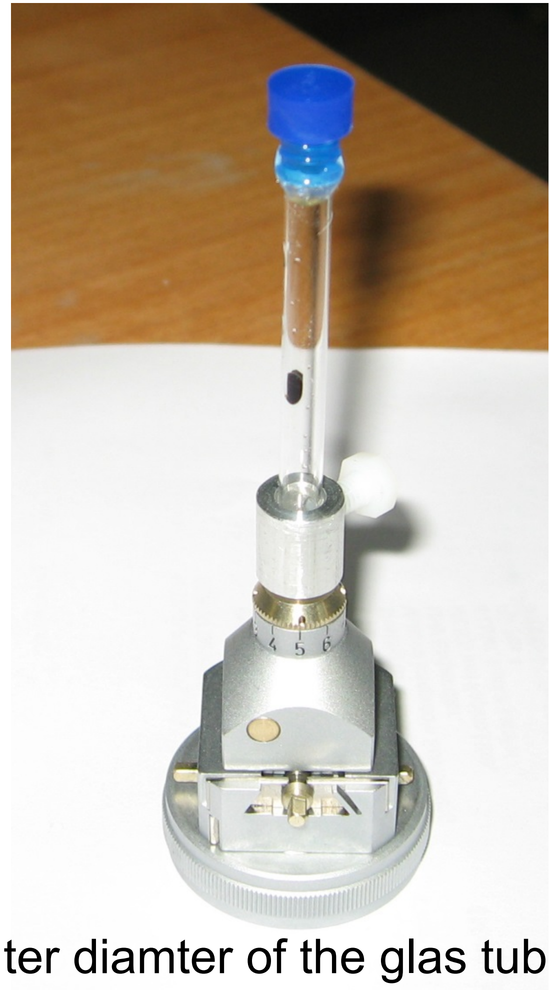
How a typical protein crystal looks like…



Picture taken from Lecture of
Prof. Locher at ETH Zürich

fig 2.2

Size considerations of protein crystals



size:

x-ray-crystallography:

ca. 10 µm x 10 µm x 10 µm

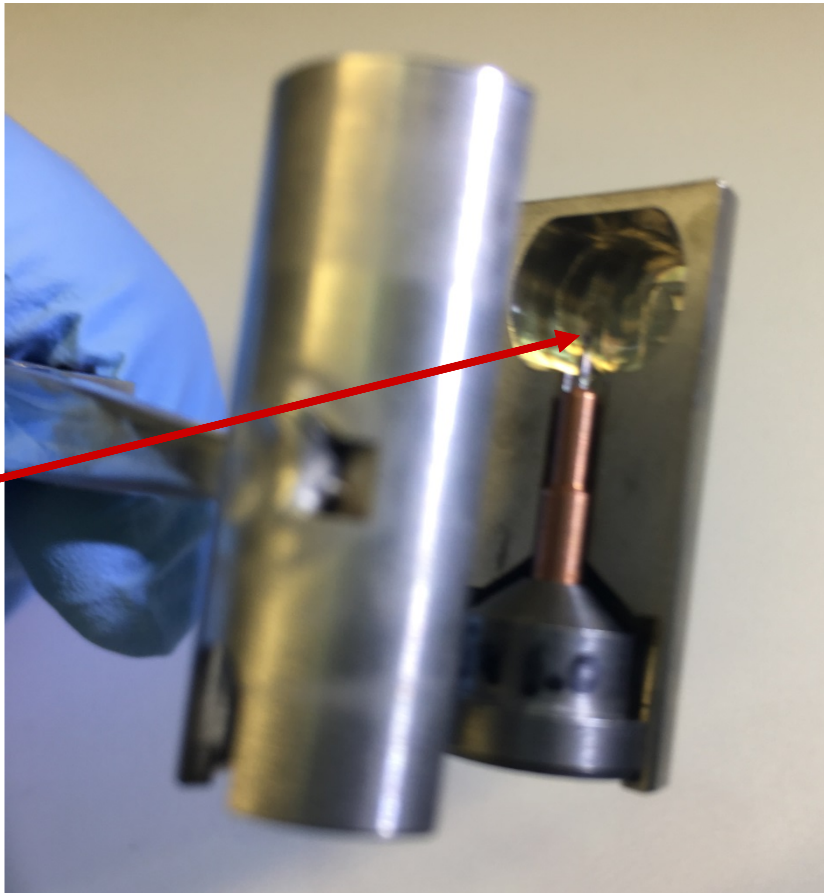
typically cryoprotectants needed to facilitate measurements at low (80 K) temperatures

neutron protein crystalography:

The desirable size should be around 1 mm x 0.5 mm x 0.5 mm (depending on the protein/space group)

Outer diamter of the glas tube: 5 mm

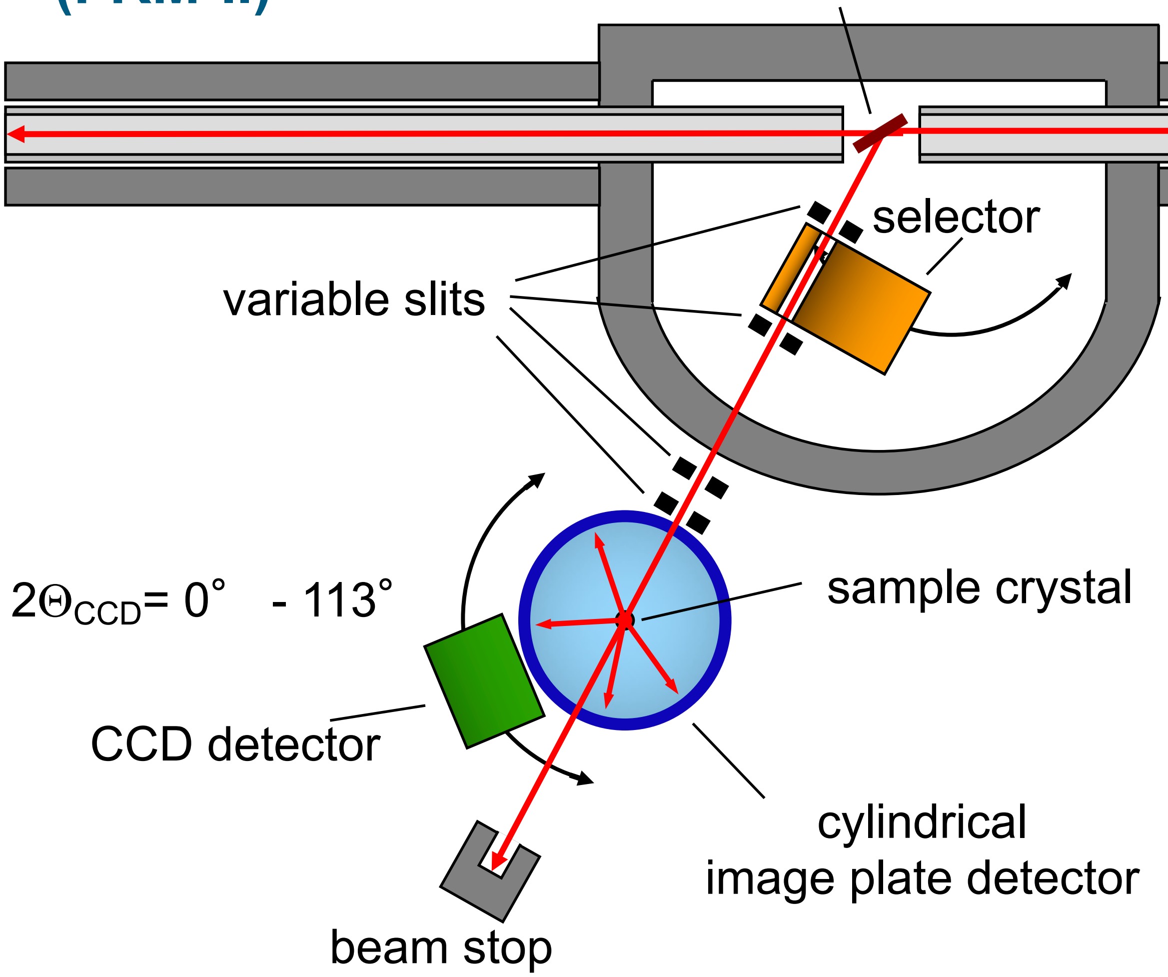
Cryo-mounting of large crystals



sample crystal

- Avoid hydrogenated polymers in the loop, use capton (Mitigen) or carbon meshes instead (especially when you have a fully deuterated protein)
- Make sure that your crystal fits into the cryoTong: We prefer the 18 mm one.

Schematic overview over BioDiff: A neutron protein diffractometer: collaboration between JCNS and TUM (FRM-II)



selector

variable slits

sample crystal

$2\Theta_{CCD}$= 0° - 113°

CCD detector

cylindrical
image plate detector

beam stop

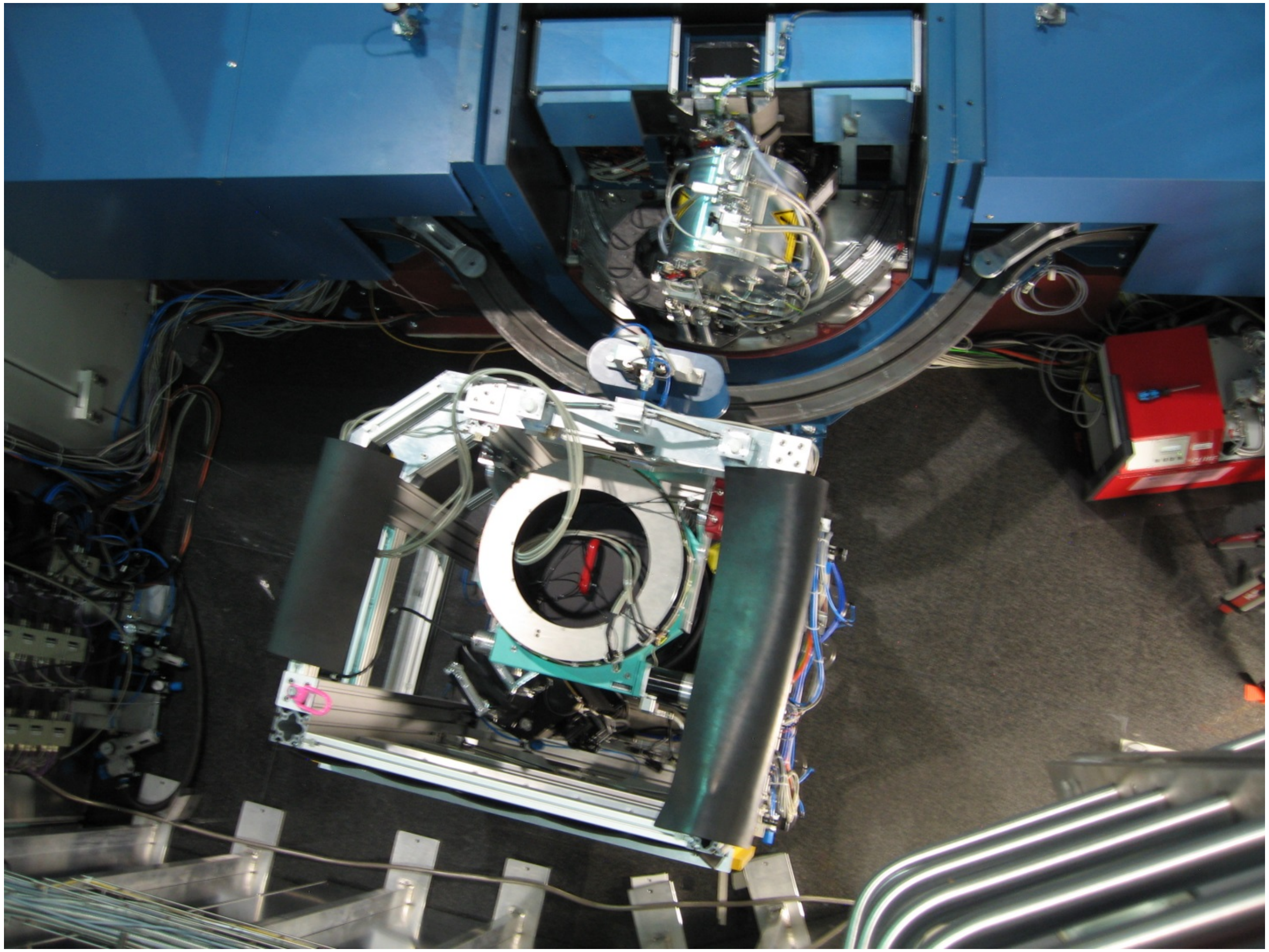
$2\Theta_{mono}$= 42° - 112°

Wavelength range:
2.4 – 5.6 Å with selector
2.4 – 6.1 Å without selector

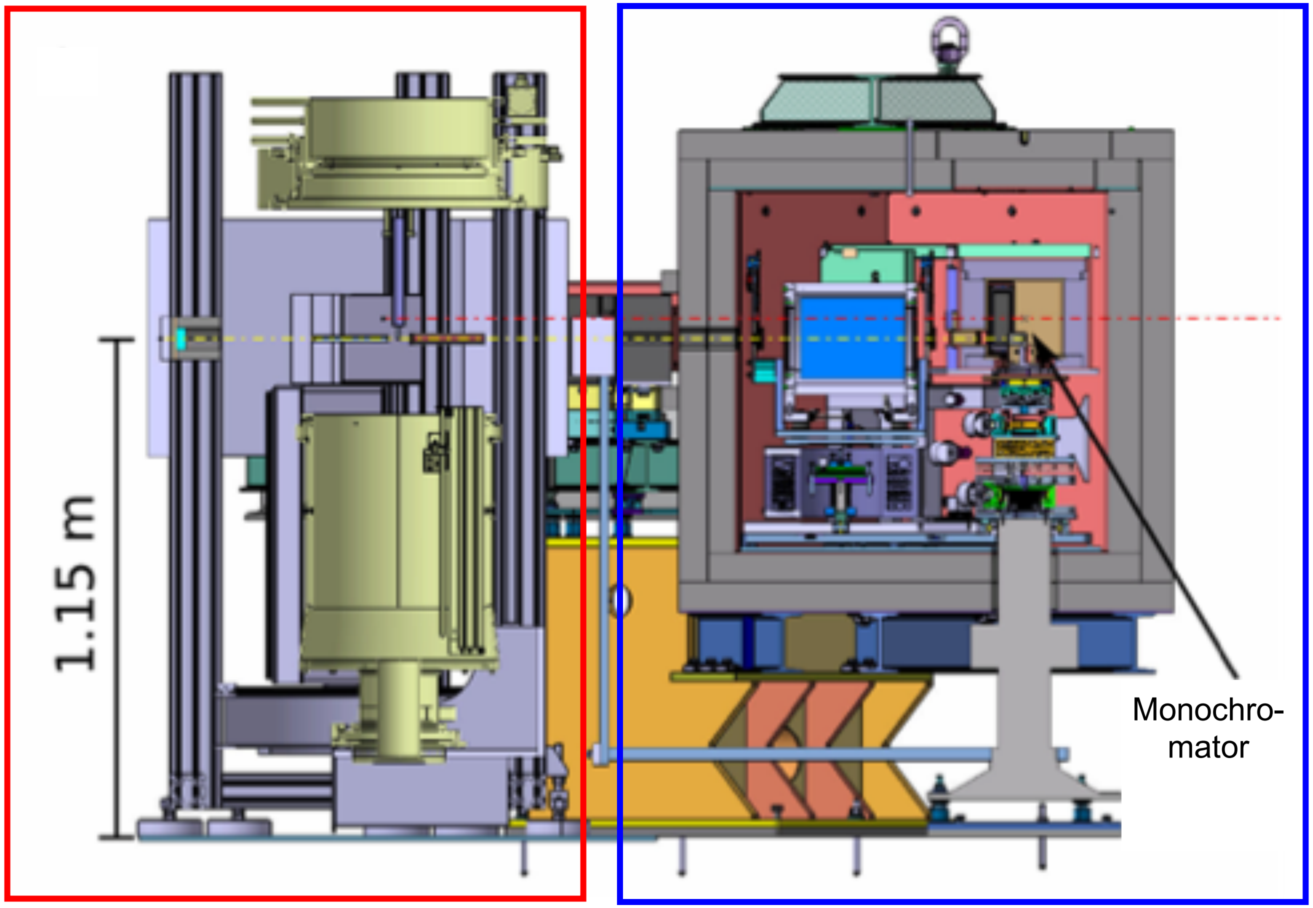
Bragg's law

$$n\lambda = 2d \sin\theta$$

BioDiff, the corresponding view in reality:



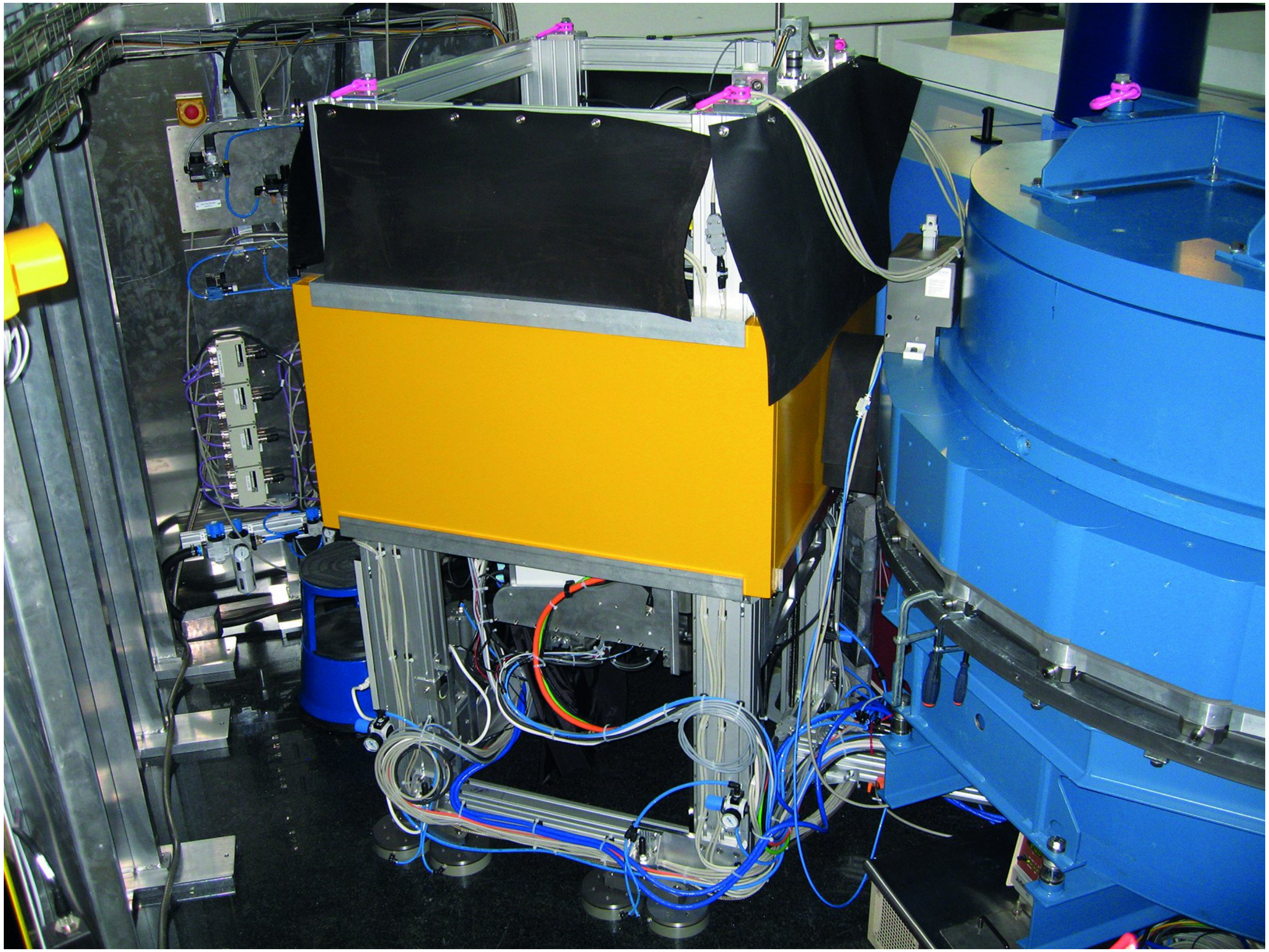
The Simultaneous Construction-phase in Garching and Jülich



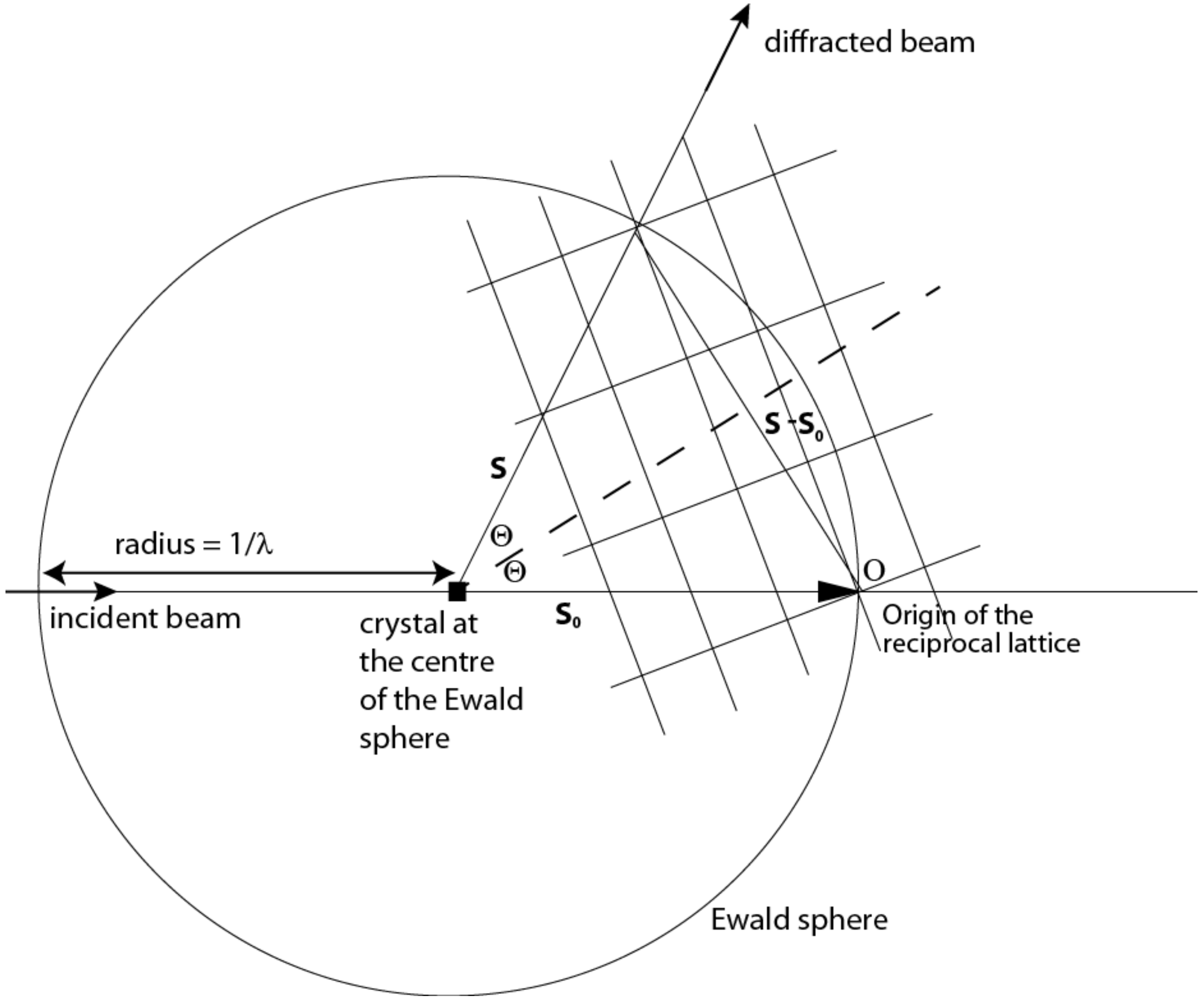
Detector unit, constructed and built in Garching (Ph. Jüttner, MLZ)

Monochromator-shielding, constructed and built in Juelich (B. Laatsch, ZEA-1 Engineering)

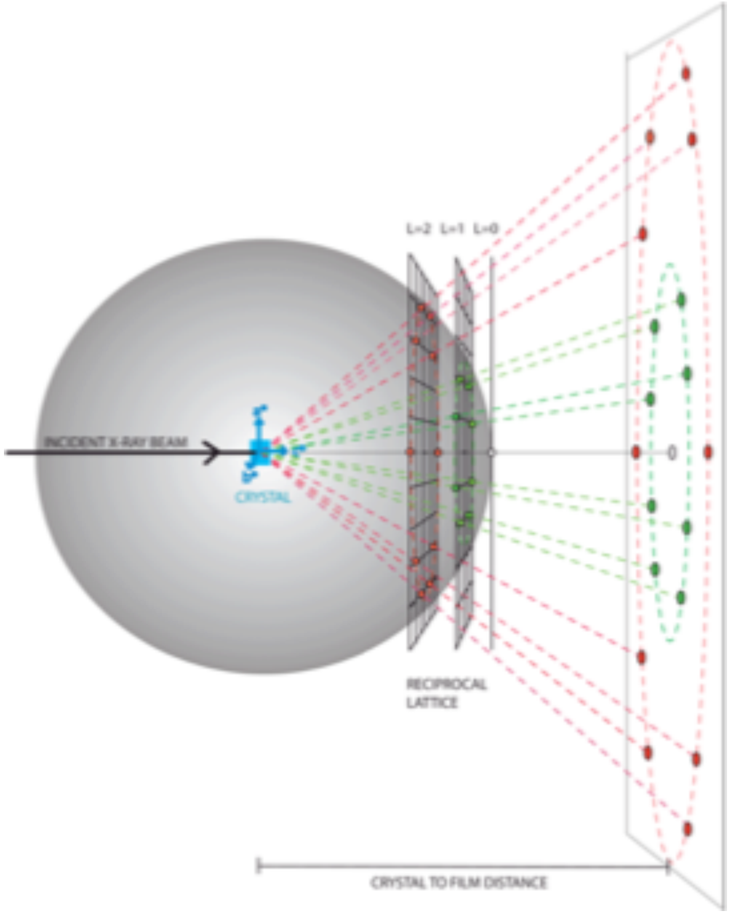
A Most Recent View of the Instrument BioDiff

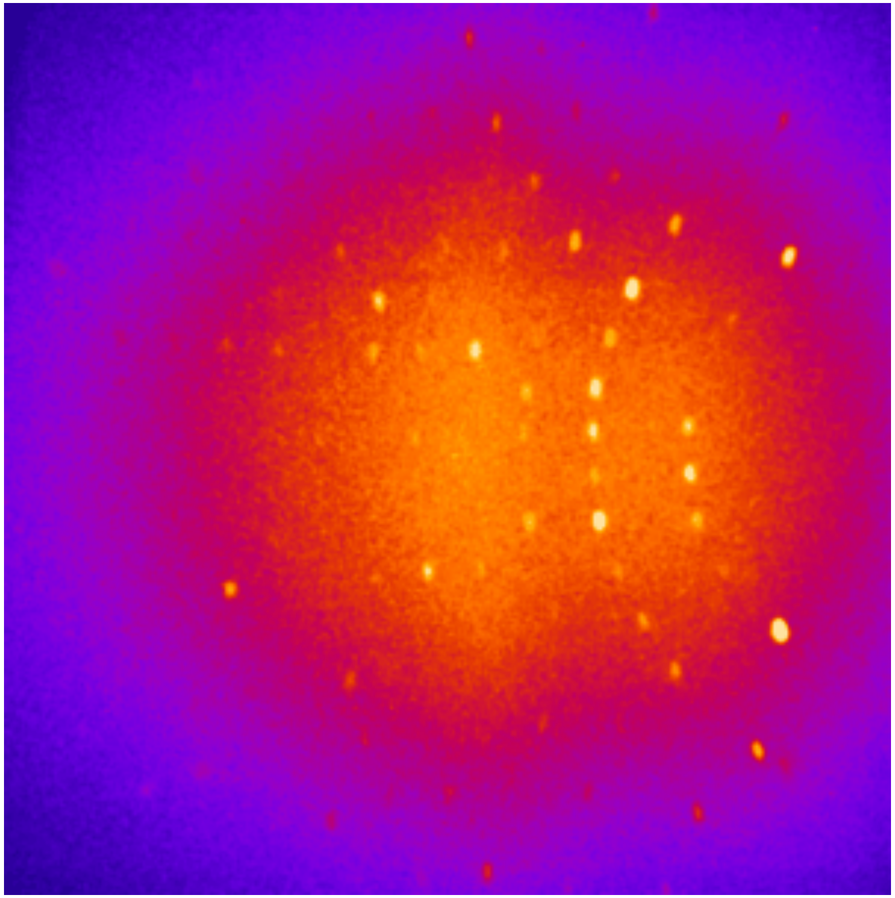


Ewald construction and Bragg‘s Law



Myoglobin protein crystal (deuterated mother liquor) full data set recorded with CCD-camera

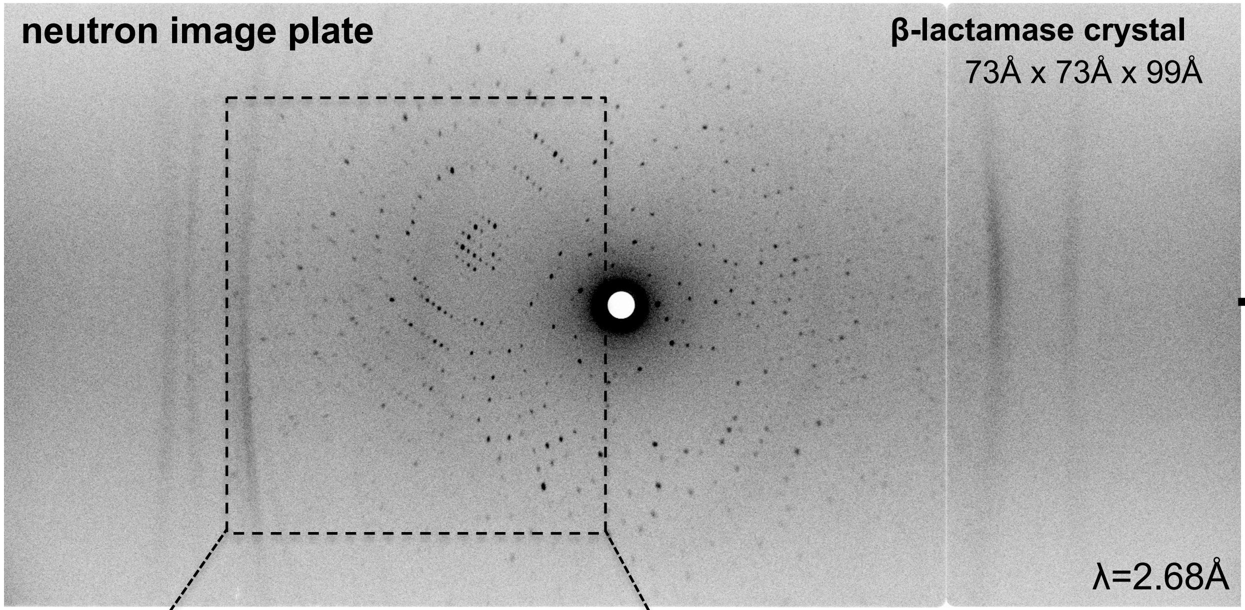


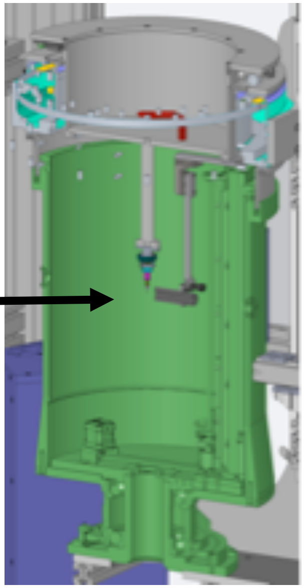


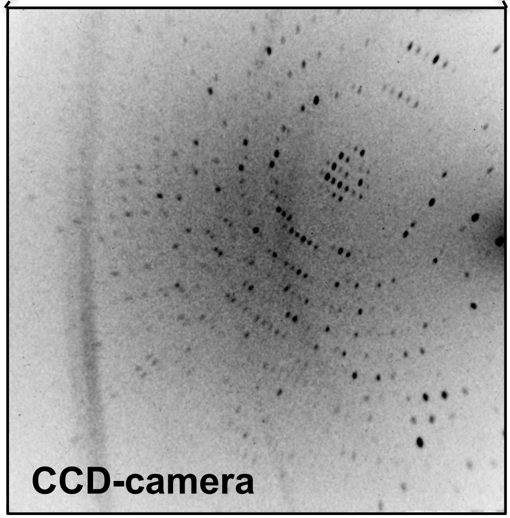
● prim. beam

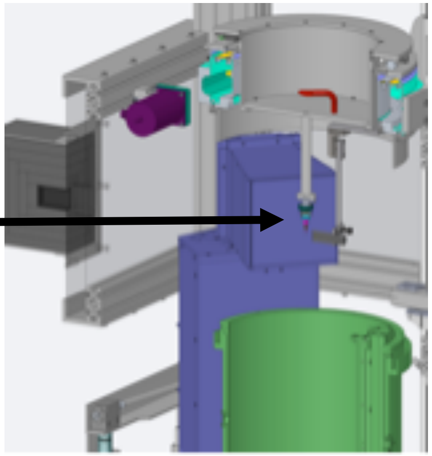
BioDiff: exposure time per frame: 20 minutes, sample: Myoglobin in deuterated mother liquor

The Two Detectors of BIODIFF









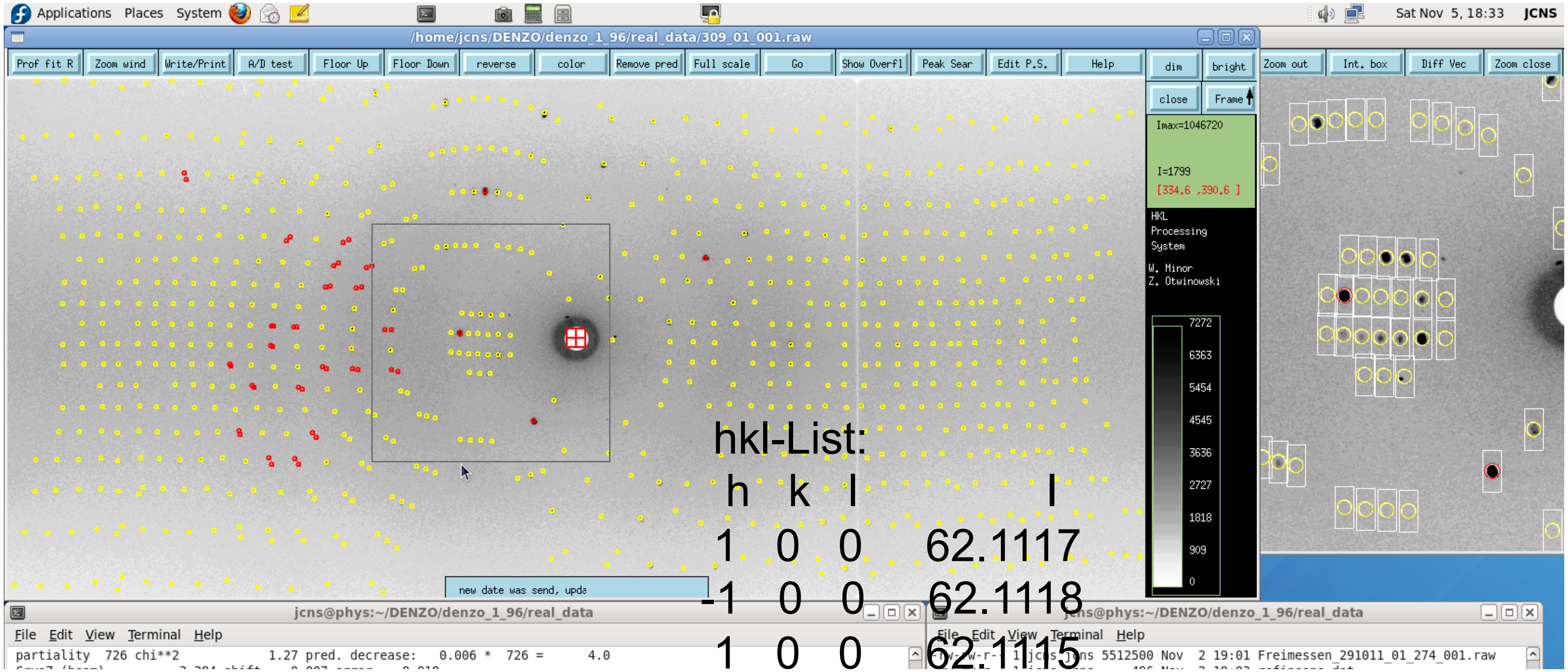
NIP-scanner

- larger solid angle
- readout time ≥ 4 min

CCD-camera

- smaller solid angle
- readout time ≥ 1 sec

Integration of partial Bragg peaks with the commercial software hkl-denzo up to d_{min}=1.5 Å



Imax=1046720

I=1799
[334.6 ,390.6]

HKL
Processing
System

W. Minor
Z. Otwinowski

new date was send, upda

```
partiality  726 chi**2         1.27 pred. decrease:   0.006 *  726 =       4.0
```

```
-rw-r--r-- 1 jcns jcns 5512500 Nov  2 19:01 Freimessen_291011_01_274_001.raw
```

hkl-List:

h	k	l	I
1	0	0	62.1117
-1	0	0	62.1118
1	0	0	62.1115
-1	0	0	62.1120
0	0	-1	33.5555
1	0	-1	33.5589
0	0	1	33.5533
-1	0	1	33.5511

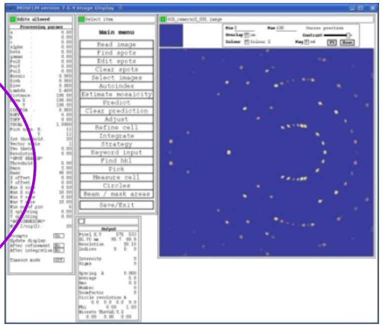
ca. 300 images

Flow chart of data treatment and model building

Scans at varying crystal orientation
Scan := Series of detector images

Data reduction
- determination of crystal orientation, unit cell dimensions etc.
- Calculating integral of reflection intensities

hkl-list for each scan:
h k l Intensity Intensity error



-MOSFLM
-HKL-denzo
- HKL2000
(comercial)

Scaling of each hkl list to match each other

-SCALA (CCP4-program package)

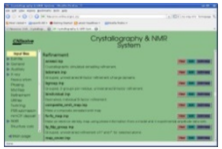
Unified hkl-list of measurement := complete data set

Calculation of a first map ⟷ Additional information from the solution of the phase problem

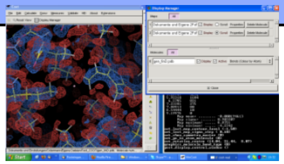
Structure refinement
-Refinement of atom coordinates displacements
- Calculation of scattering density maps (netrons) or electron density maps (x-rays)

⟷

Map-plotting
- inspection of model to fit the map)
- real space changes and refinement to the model

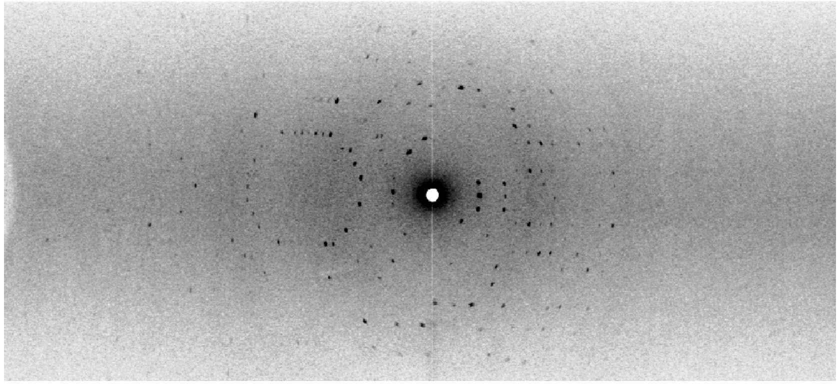


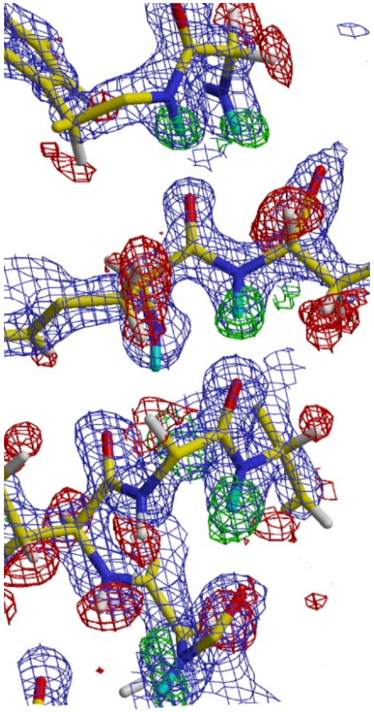
-nCNS
-PHENIX



-XtalView
-Coot

The phase problem:





$$\rho(x,y,z)=\frac{1}{V_E}\sum_{h,k,l}F_{hkl}\cdot e^{-2\pi i(h\cdot x+k\cdot y+l\cdot z)}$$

Structure factors are complex numbers: $F_{hkl} = \|F_{hkl}\| e^{-2\pi i \alpha_{hkl}}$

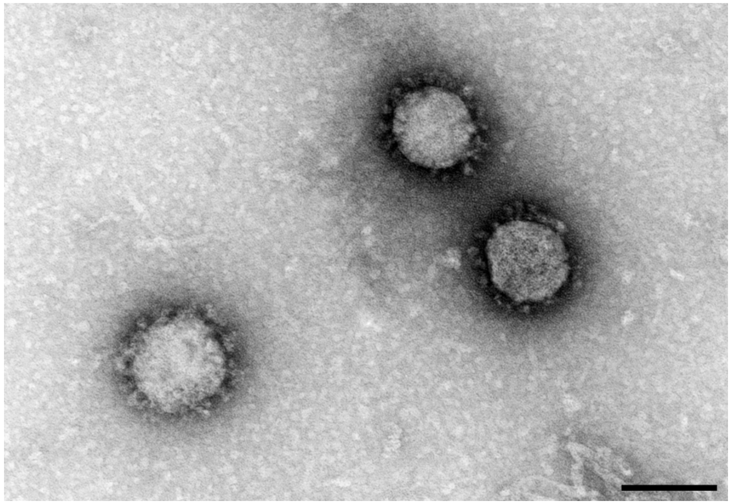
with amplitudes $\|F_{hkl}\|$ and phases α_{hkl}

→ Phase Problem, because we only record intensities: $I = \|F_{hkl}\|^2$

Application Example II: Research on Corona Virus proteins from SARS-COV-2 with x-ray and neutron crystallography

How does the virus look like?

The size is 60-160 nm



Scale bar = 100 nm,
Source: RKI web-site,
Source: *Hans R. Gelderblom, Freya Kaulbars/RKI*

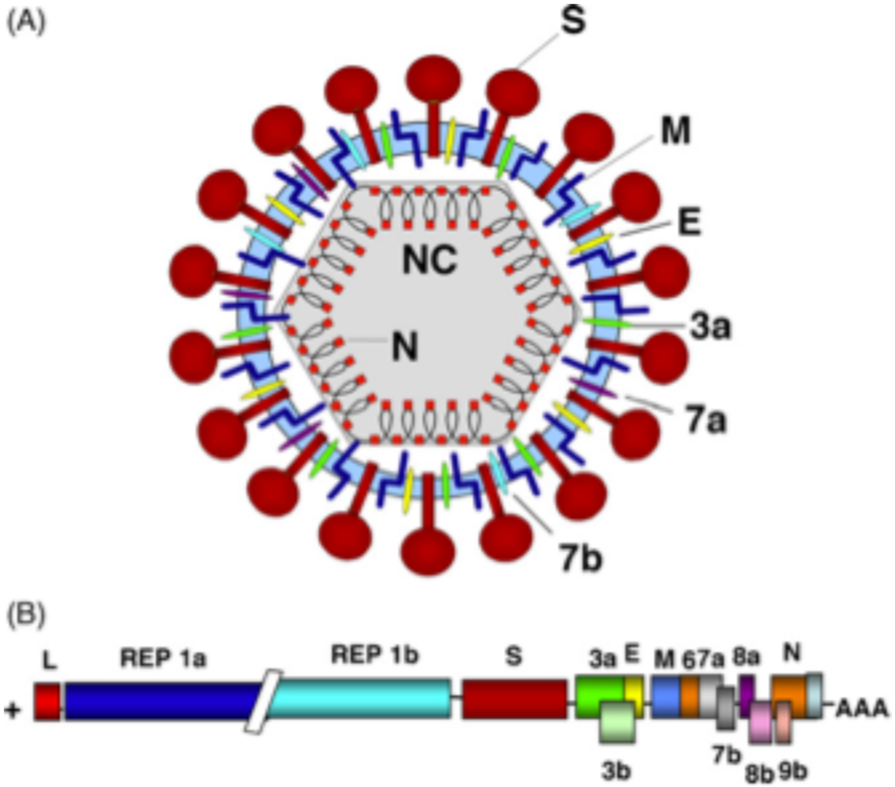
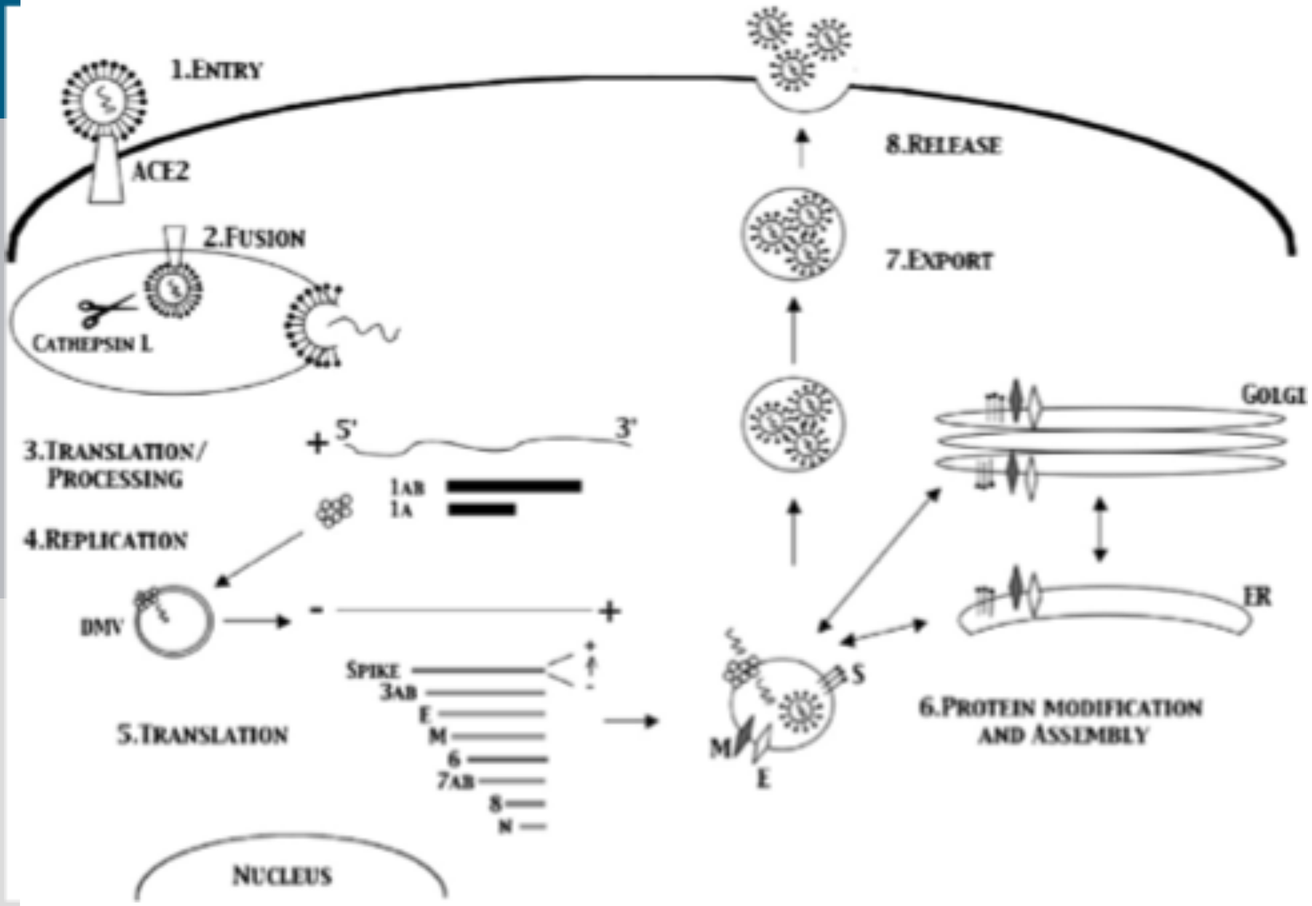


Fig. 2. Structure and genome organization of SARS-CoV. (A). Schematic diagram of SARS-CoV structure. S, spike protein; M, membrane protein; E, envelope protein; N, nucleoprotein; 3a, 7a, and 7b, structural proteins of SARSCoV. (B). Representation of a prototype SARS-CoV genome. Poly(A) tail is indicated by AAA. Numbers and letters indicate viral genes.: Virus Research 133 (2008) 45–62

How does it work?

It programms a human cell to produce replicas of it:



Virus Research 133 (2008) 101–112

Fig. 6. The coronavirus life cycle. Coronavirus entry is mediated by binding of S glycoprotein to the ACE2 receptor, cleavage by cathepsin L and activation of a fusion peptide in S2 that mediates entry via fusion through endocytic compartments [1]. Following fusion with the endosomal compartment the viral genome release into the cytosol where it is translated into the viral replicase proteins ORF1a and 1b [2]. These polyproteins are then cleaved by 2 proteases, Main Protease (Mpro) and Papain like protease, PLP, into the individual proteins necessary for replication [3]. Subgenomic RNA synthesis occurs from discontinuous transcription which joins leader RNA sequences encoded at the 5 end of the genome to the body sequences of each subgenomic RNA. The eight different subgenomic negative strands serve as template for the synthesis of like sized subgenomic mRNA [4]. Subgenomic RNAs are then translated into viral proteins which localize to their relevant compartments [5]. Assembly of virions occurs in an ERGIC like compartment in the cell. Here S, E,Mand N bound to genomic viral RNA are assembled into virions in vesicles [6]. The vesicles are then exported to the cell surface where fusion occurs with release of virions into the exterior environment [7,8].

The genome and the encoded proteins of SARS-CoV-2

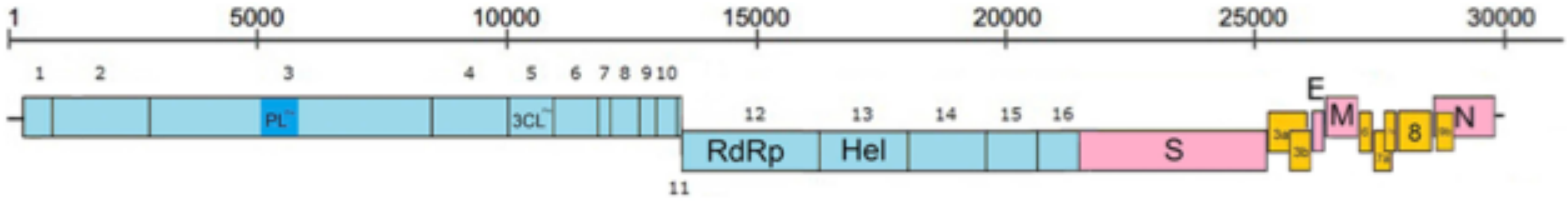


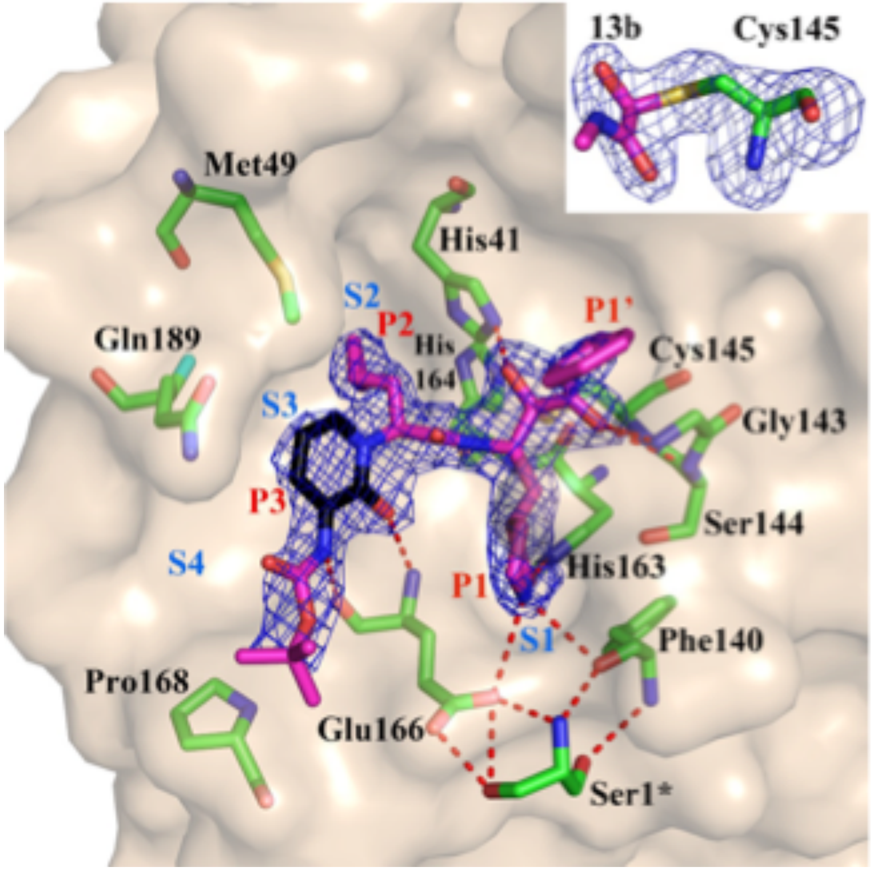
Fig. 1. Genome organization of SARS-CoV. ORF1ab with nsp1–16 are colored in blue. Structural proteins including S, E, M and N are in pink. Accessory proteins were numbered and in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Main protease cleaves viral proteins to the right length: This is why it is an important drug target.

From: Infection, Genetics and Evolution 71 (2019) 21–30

The main protease of the corona virus

Here, an inhibitor is bound to the protein



L. Zhang *et al*., *Science* 10.1126/science.abb3405 (2020).

Fig. 3. Compound 13b in the substrate-binding cleft located between domains I and II of the Mpro, in the monoclinic crystal form (space group C2). Fo-Fc density is shown for the inhibitor (contouring level: 3σ). Carbon atoms of the inhibitor are magenta, except in the pyridone ring, which is black; oxygen atoms are red, nitrogens blue, and sulfur yellow. Light-blue symbols S1, S2, S3, S4 indicate the canonical
binding pockets for moieties P1, P2, P3, P4 (red symbols) of the peptidomimetic inhibitor. Hydrogen bonds are indicated by dashed red lines. Note the interaction between the N-terminal residue of chain B, Ser1*, and Glu166 of chain A, which is essential for keeping the S1 pocket in the right shape and the enzyme in the active conformation. Inset: Thiohemiketal formed by the nucleophilic attack of the catalytic cysteine onto the α-carbon of the inhibitor in its Fo-Fc density (contoured at 3σ). The stereochemistry of the α-carbon is *S*. See fig. S8 for more details.

The first neutron structure solved for the main protease in its apo state (empty binding pocket)

A: X-ray data

B: neutron data

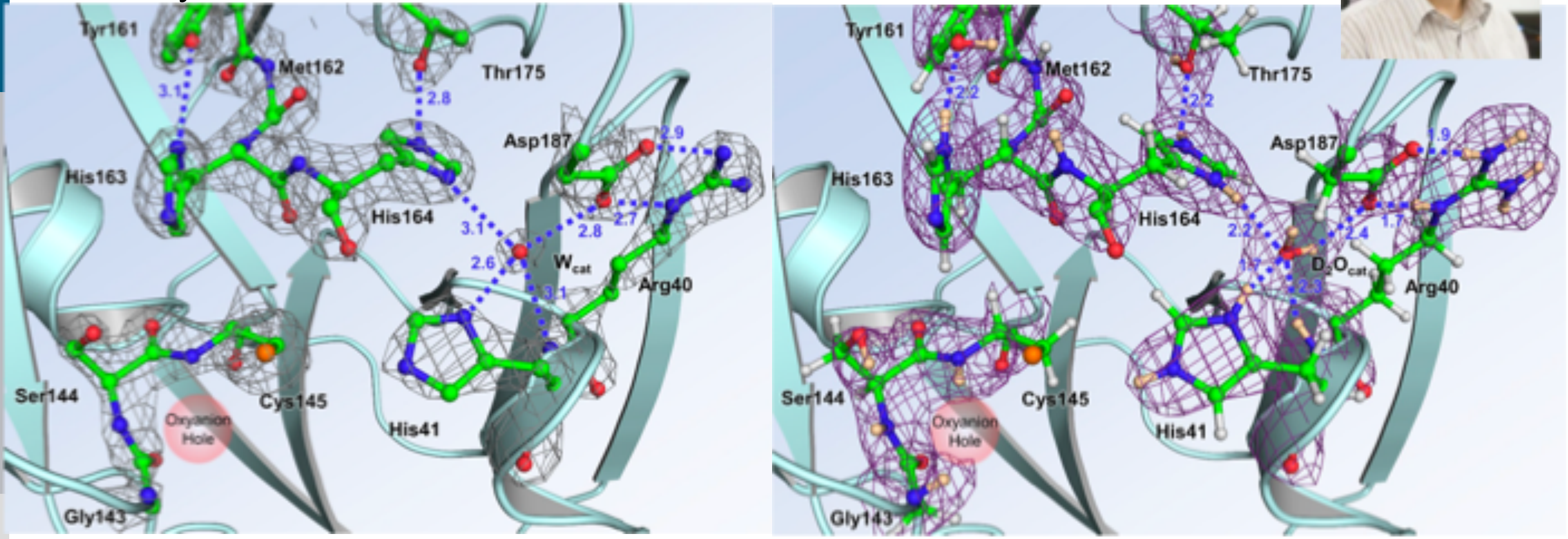
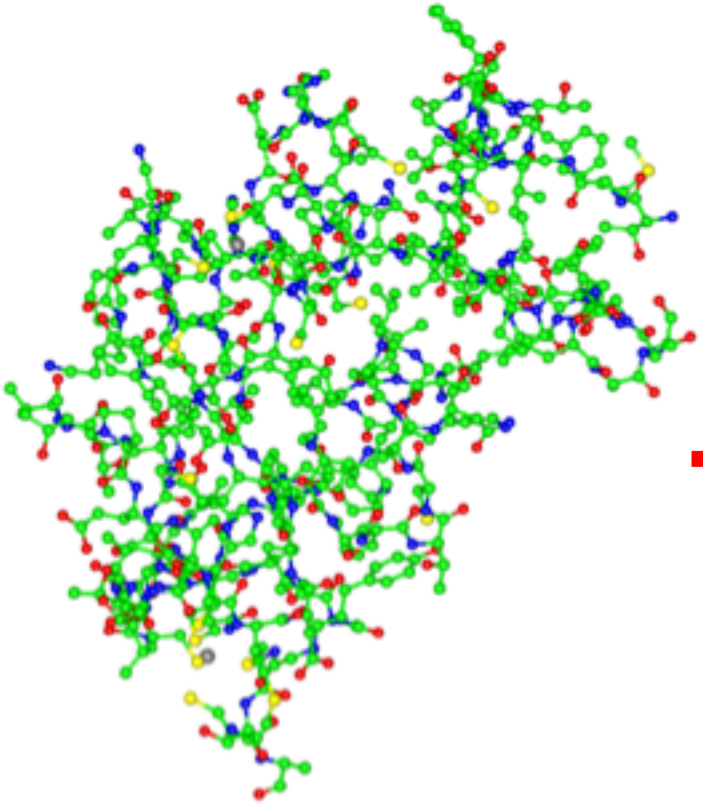


Figure 3. The catalytic site of SARS-CoV-2 3CL Mpro. (**A**) The 2FO-FC electron density map contoured at 2.0 σ level (grey mesh) with no hydrogen atoms visible. Distances between the heavy atoms in Ångstroms illustrate possible hydrogen bonds. (**B**) The 2FO-FC nuclear density map contoured at 2.0 σ level (violet mesh), allowing visualization of the actual protonation states and hydrogen bonding interactions (D…O distances are shown in Ångstroms).

From: https://www.biorxiv.org/content/10.1101/2020.09.22.308668v1, data measured at MANDI and IMAGINE at Oak Ridge National Lab, Oak Ridge, Tennessee, USA

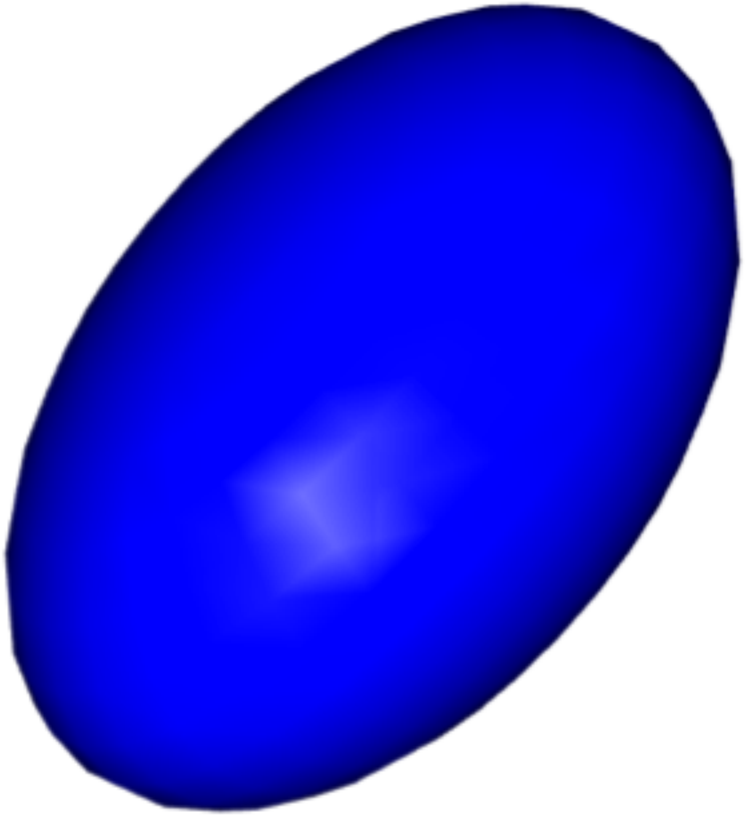
His41: doubly protonated
Cys145: deprotonated

And now the low resolution...



nsp10 of SARS-CoV2, 6zct.pdb

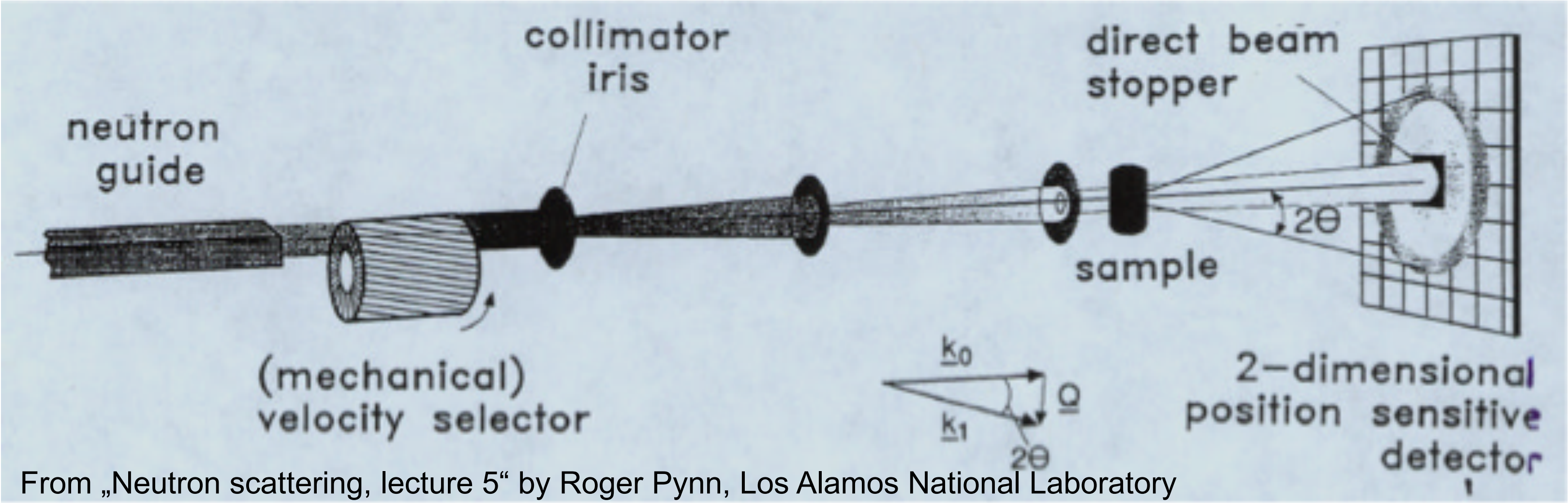
Single crystal diffraction



Small angle neutron scattering (SANS)

Small angle neutron scattering

„Pinhole SANS“



From „Neutron scattering, lecture 5“ by Roger Pynn, Los Alamos National Laboratory

„radial averaging“

$$\frac{d\Sigma}{d\Omega}(\mathbf{q}) = \frac{N}{V}\frac{d\sigma}{d\Omega}(\mathbf{q}) = \frac{1}{V}\left|\int_V \rho(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}\right|^2$$

„Rayleigh-Gans Equation“

Form factor of a sphere

For a sphere of radius r

$$P(q) = \left[\frac{3(\sin(qr) - qr\cos(qr))}{(qr)^3}\right]^2$$

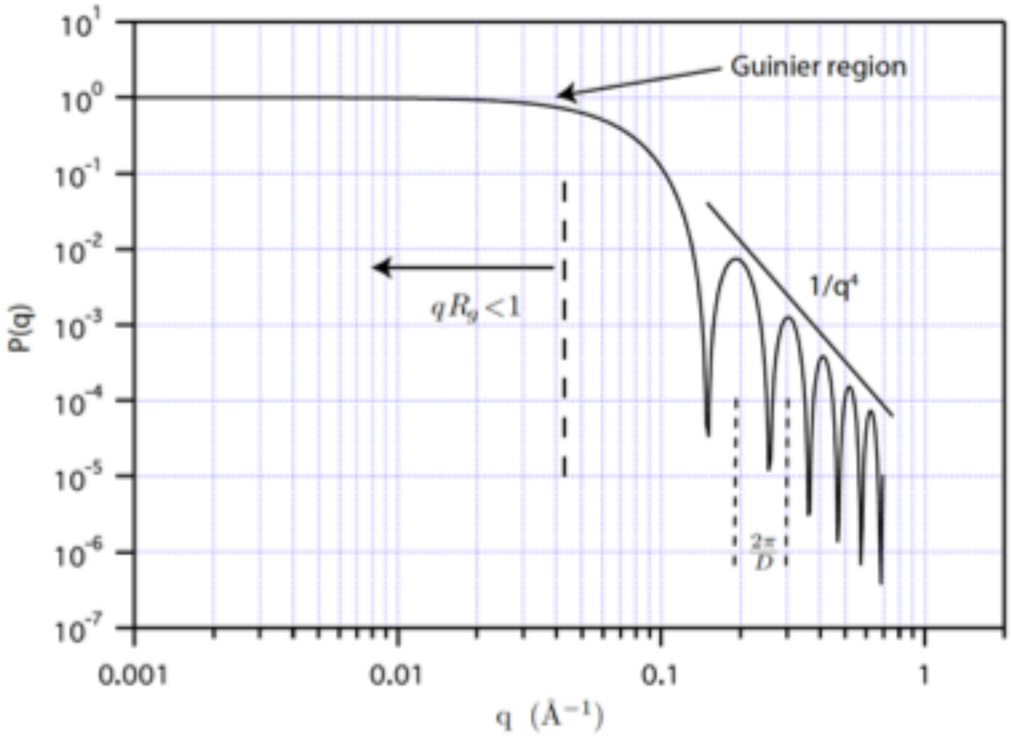


Figure 7: Form Factor spheres of radius 3Å. $R_g =$ 2.3 Å

$$R_g^2 = \frac{3}{5}R^2$$

Figure taken from „Introduction to Small-Angle Neutron Scattering and Neutron Reflectometry“, Andrew J Jackson, NIST Center for Neutron Research, May 2008

Form factor of a cylinder

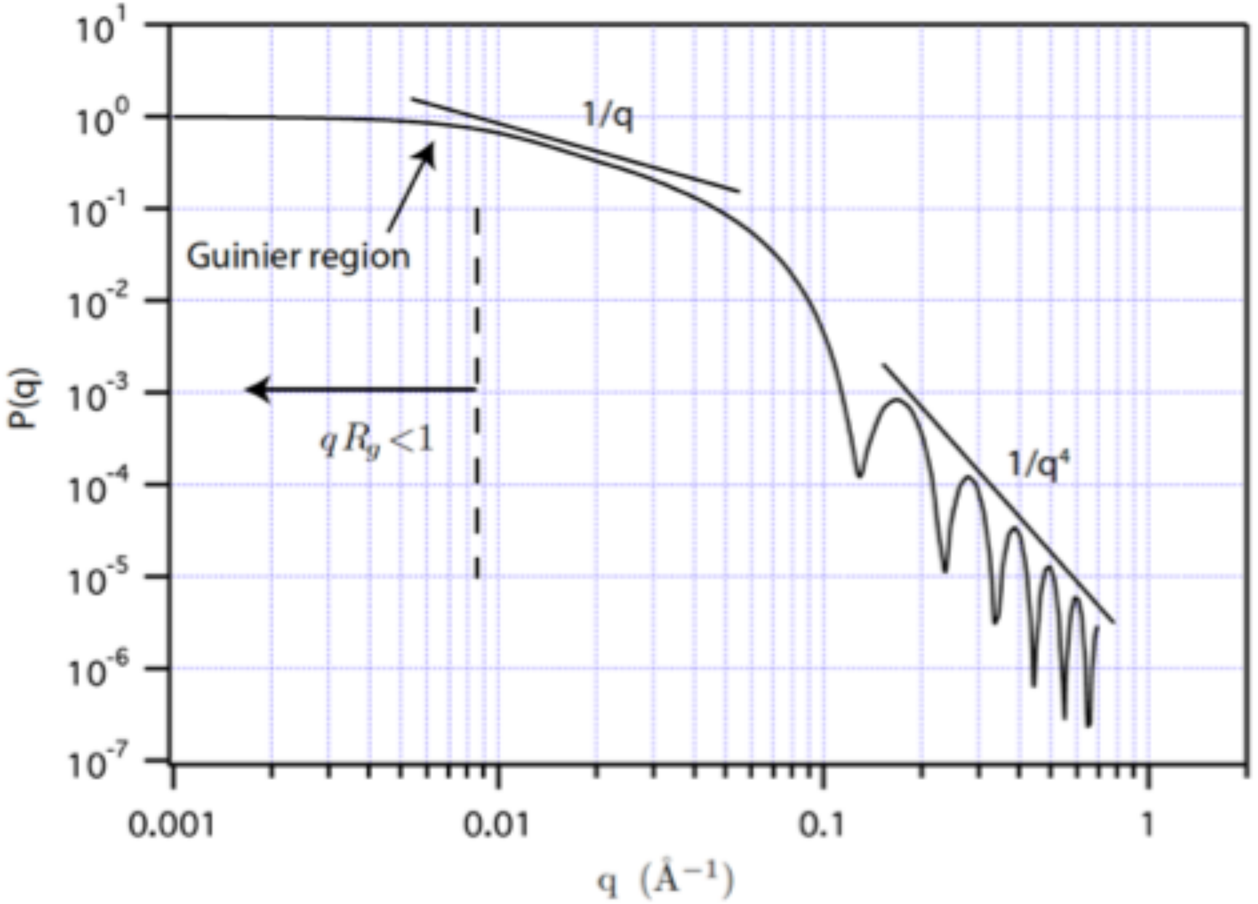
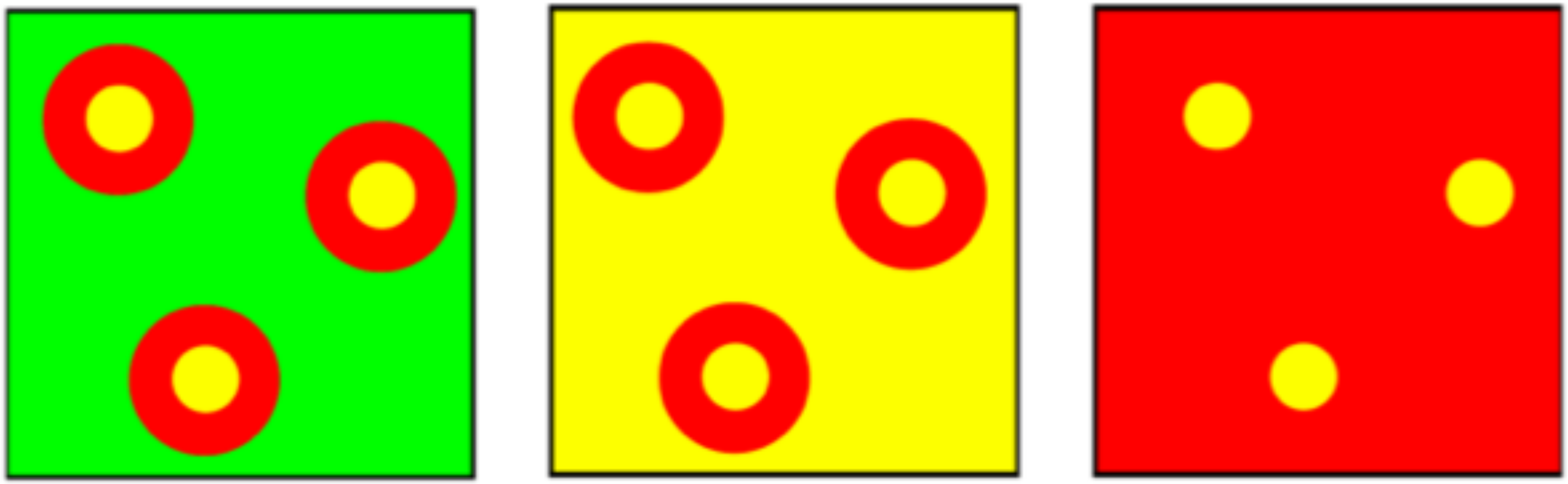


Figure 8: Form Factor for cylinders of radius 30Å and length 400Å. $R_g = 117$Å

Figure taken from „Introduction to Small-Angle Neutron Scattering and Neutron Reflectometry“, Andrew J Jackson, NIST Center for Neutron Research, May 2008

Contrast matching



Natural contrast

r solvent = r core
(shell visible)

r solvent = r shell
(core visible)

Figure 9: The effect of contrast variation on the measurable structure of a core-shell particle

Figure taken from „Introduction to Small-Angle Neutron Scattering and Neutron Reflectometry“, Andrew J Jackson, NIST Center for Neutron Research, May 2008

Contrast matching in water solutions

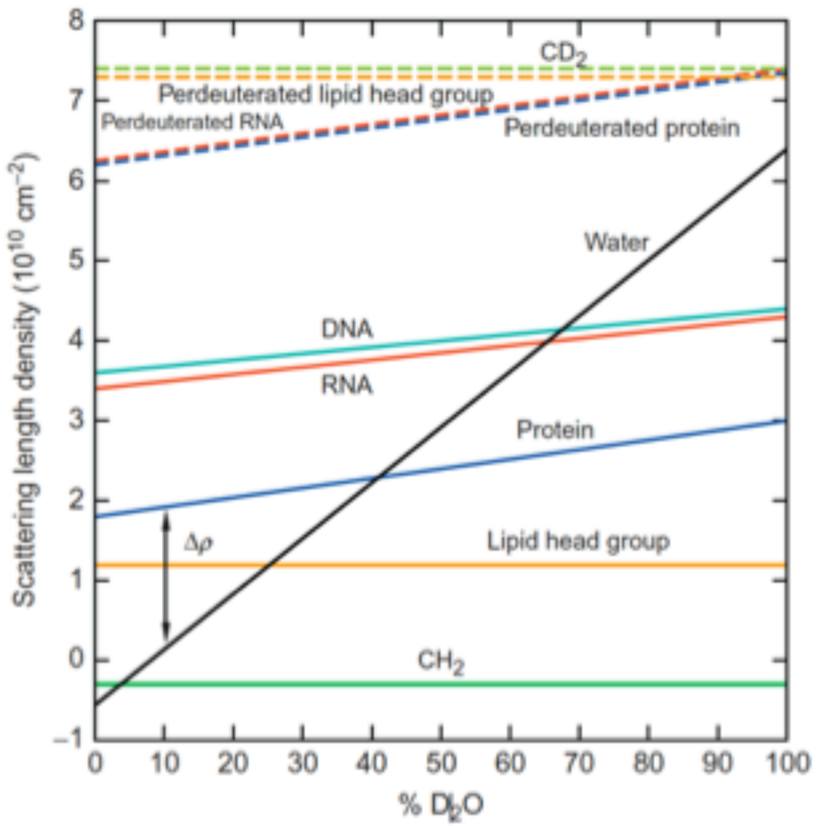
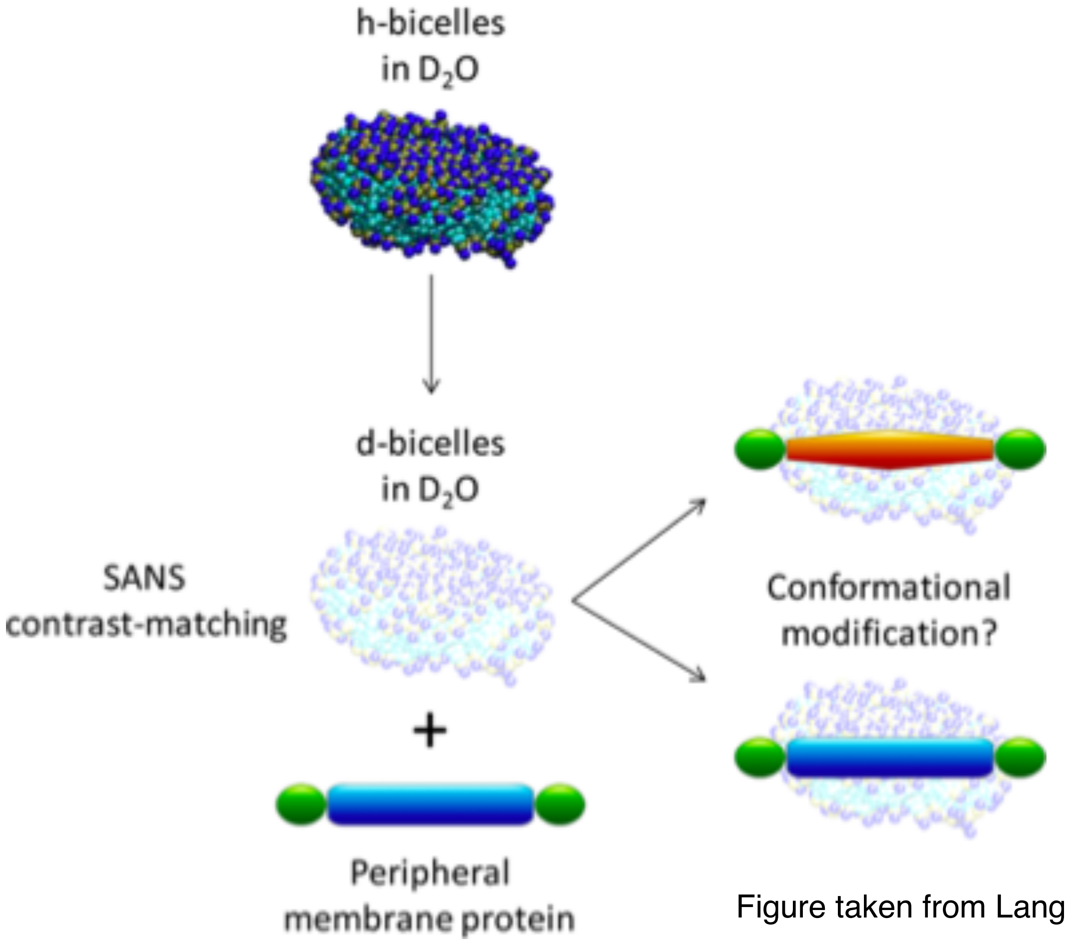


Figure 1 Plot of the neutron scattering length density versus % D_2O in the solvent for water compared to those for protein, RNA, DNA, and the components of lipids (lipid head group and CH_2), along with their perdeuterated counterparts.

From: Methods in Enzymology, Volume 566, p. 159, 2016

One application example of contrast matching SANS

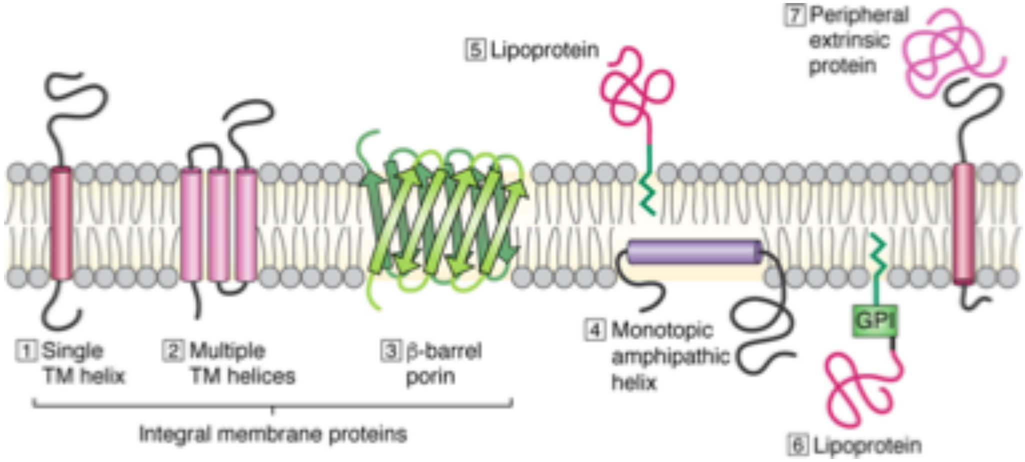
Problem: How to measure the structure of a membrane protein in its most natural environment?



Sophie Combet

Figure taken from Langmuir 2017, 33, 6572

Idea: Put a bicelle around it and match it with a mixture of deuterated and hydrogenated lipids to 100% D_2O



Classification of membrane proteins

Figure taken from J. Biol. Chem. (2019) 294(44) 15914

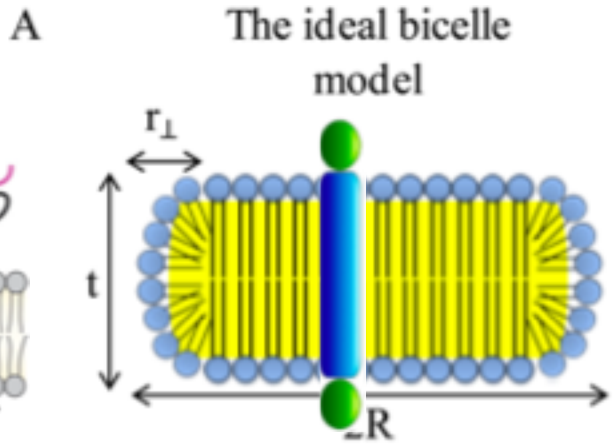
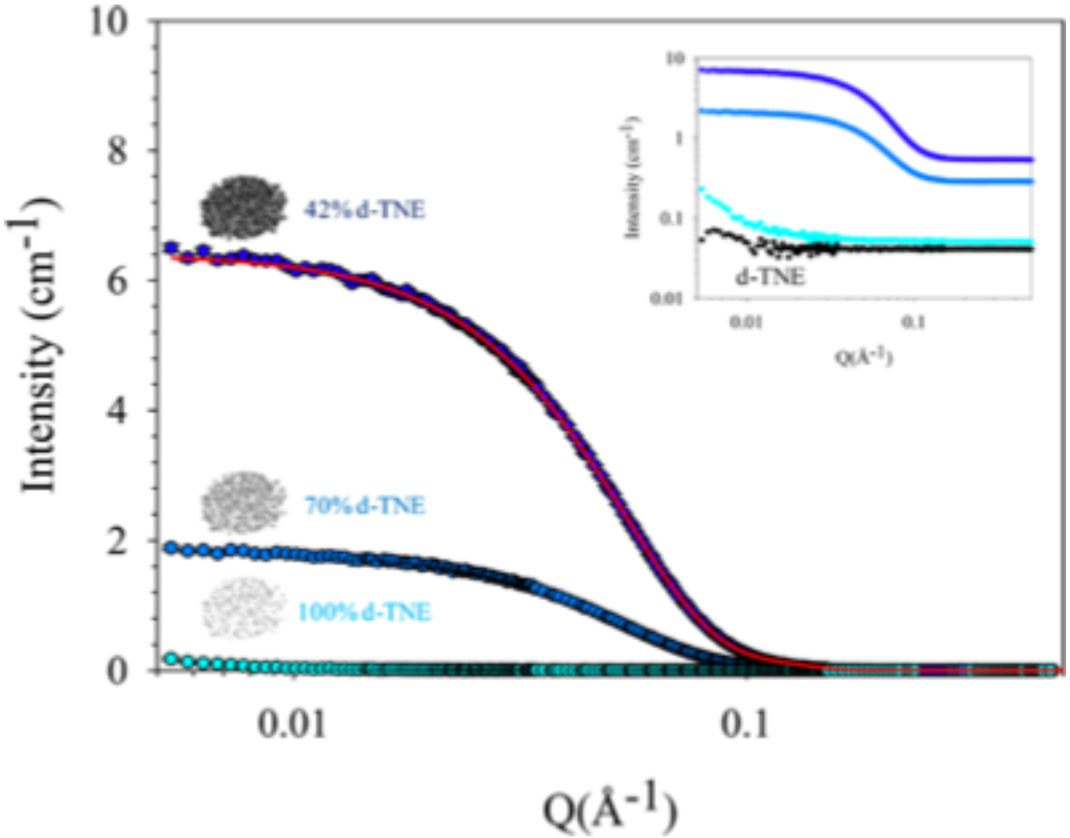


Figure 1. (A) Schematic representation of the ideal bicelle model. For DMPC/DHPC bicelles, R is the radius of the bicelle, $r\perp$ is the thickness of the rim (11 Å, the length of a DHPC molecule), and t is the thickness of the DMPC bilayer (assumed to be 40 Å).

Figure taken from Langmuir 2017, 33, 6572

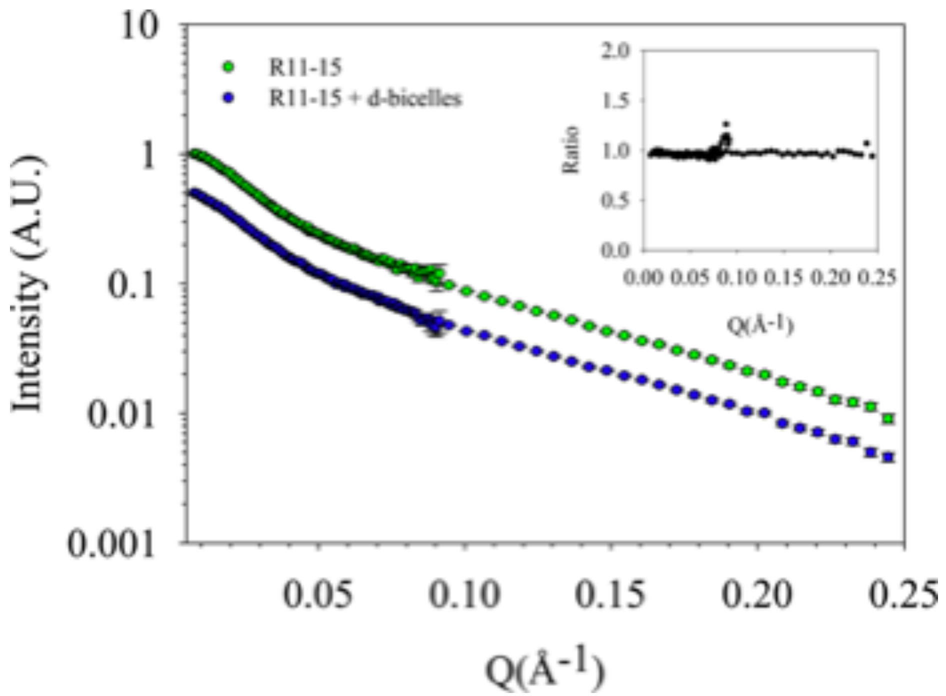
Bicelles are virtually invisible at 100%. D_2O buffer (d-TNE)



d-TNE buffer solution (20 mM Tris-d11, 150 mM NaCl, and 0.1 mM EDTA-d16, pD 7.5)

Figure taken from Langmuir 2017, 33, 6572

Only scattering from the protein is visible



protein fragment belonging to the dystrophin central domain DYS R11−15 (from the 11th to the 15th spectrin-like repeats)
= a well known very rigid protein

Figure taken from Langmuir 2017, 33, 6572

Summary

- Neutron protein crystallography is a complementary technique as compared to x-ray crystallography. Here one can determine:
  1. protonation states of amino acid side chains (important for the function of the protein)
  2. deuterium exchange as a measure of flexibility and accessibility (discrimination between **H** / **D**)
  3. solvent structure including hydrogen atoms

- Small angle neutron scattering: Here, the low resolution structure of sample in solution can be determined
- Contrast matching allows to ”hide” certain parts of the sample which then allows to measure the structure of the other part

Thanks to...

- Andreas Ostermann
- Marialucia Longo
- Livia Balacescu
- Zamaan Raza
- Tobias Weber
- Jonathan Fisher
- Leighton Coates
- Andrey Kovalevsky
- Stephan Förster

and you for your attention!

New Home source X-ray diffractometer at the MLZ in Garching

The XtaLAB Synergy-S gives great data, fast.Whether dataquality or highthroughput is your focus, theXtaLAB Synergy-S is designed to meet your needs.

Benefits include:
- Extremely high performance PhotonJet-S sources
- A robust reliable hardware platform and goniometer that just keeps on going
- Support for a wide range of accessories

Features:
Cu and Mo Source, HiPix Arc detector covering 150° of 2 θ range

The end