PAUL SCHERRER INSTITUT

PSI

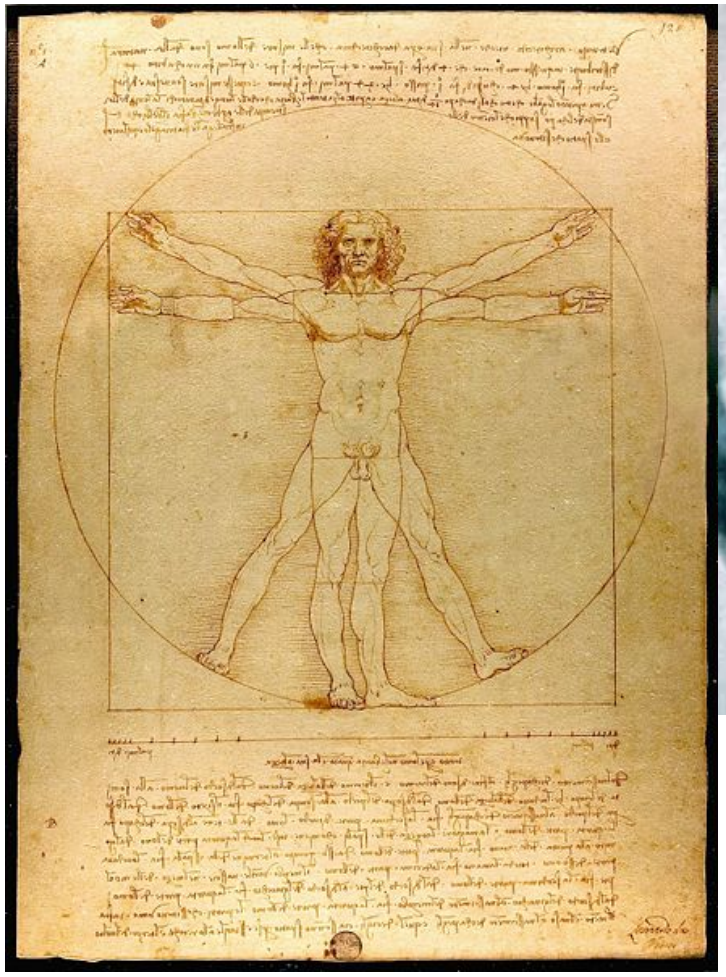
Wir schaffen Wissen – heute für morgen

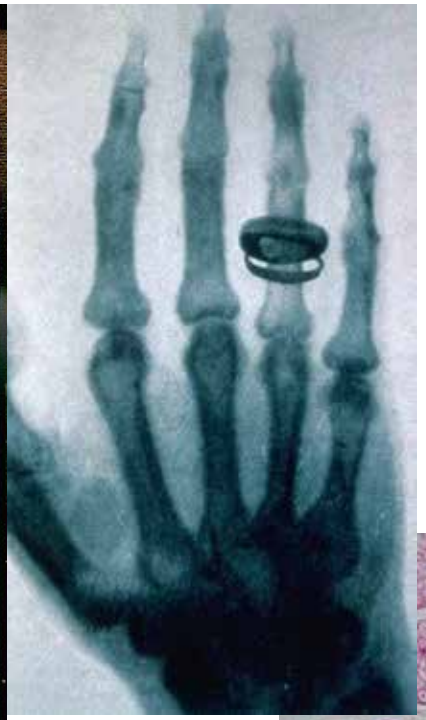
Paul Scherrer Institut
Oliver Bunk

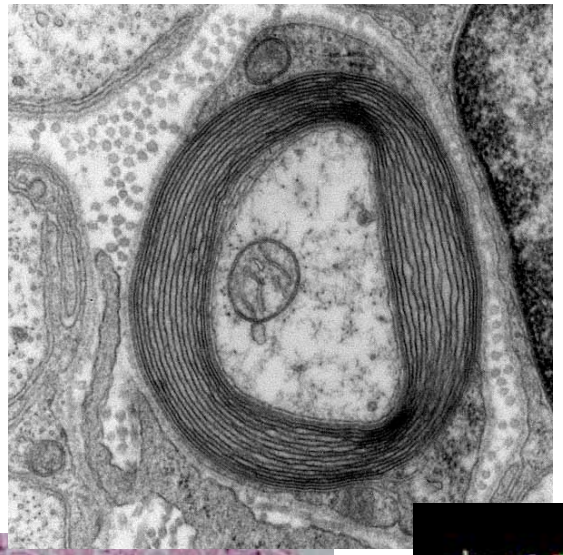
Efficient use of HDF5 with high data rate x-ray detectors: Welcome and Introduction

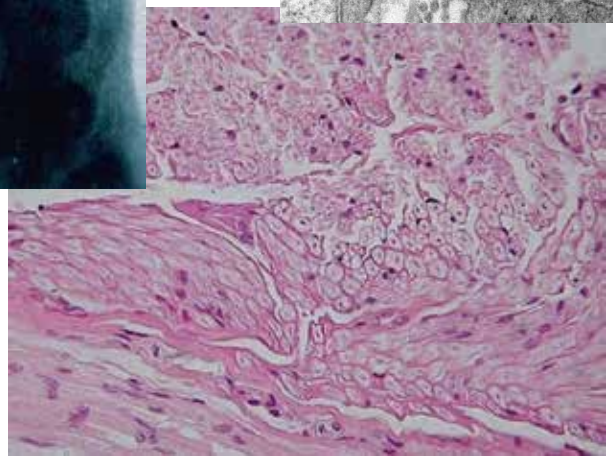
PAUL SCHERRER INSTITUT

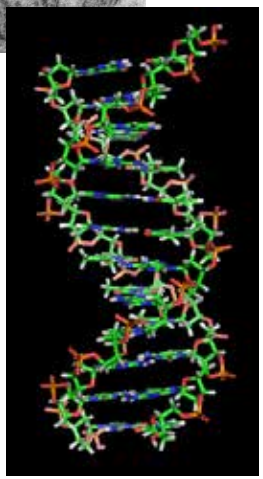
Length scales in bio imaging











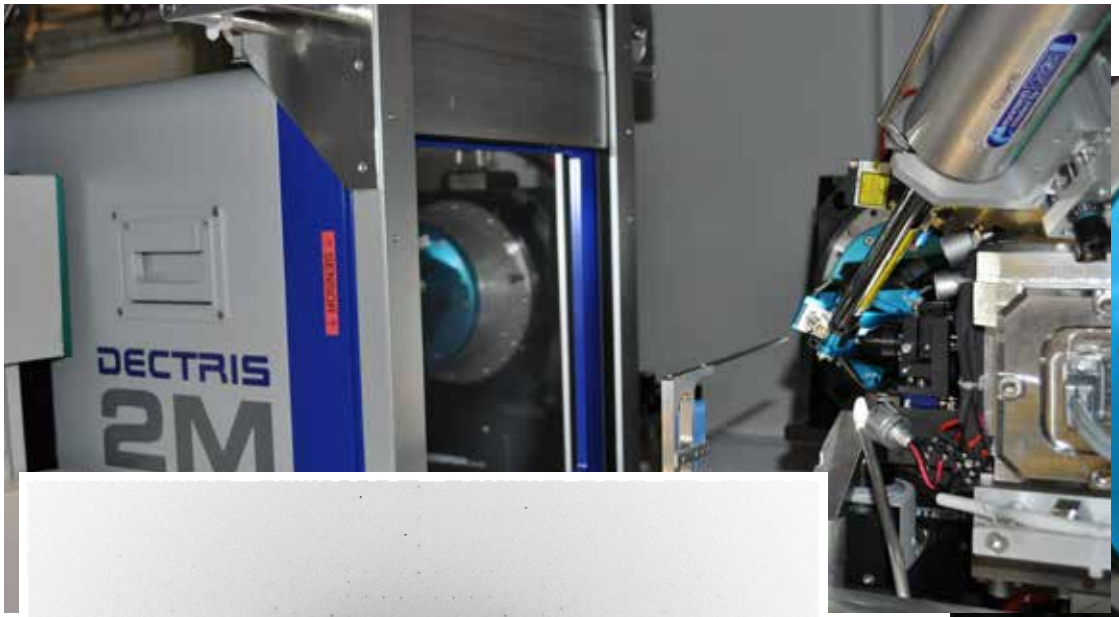
1 m

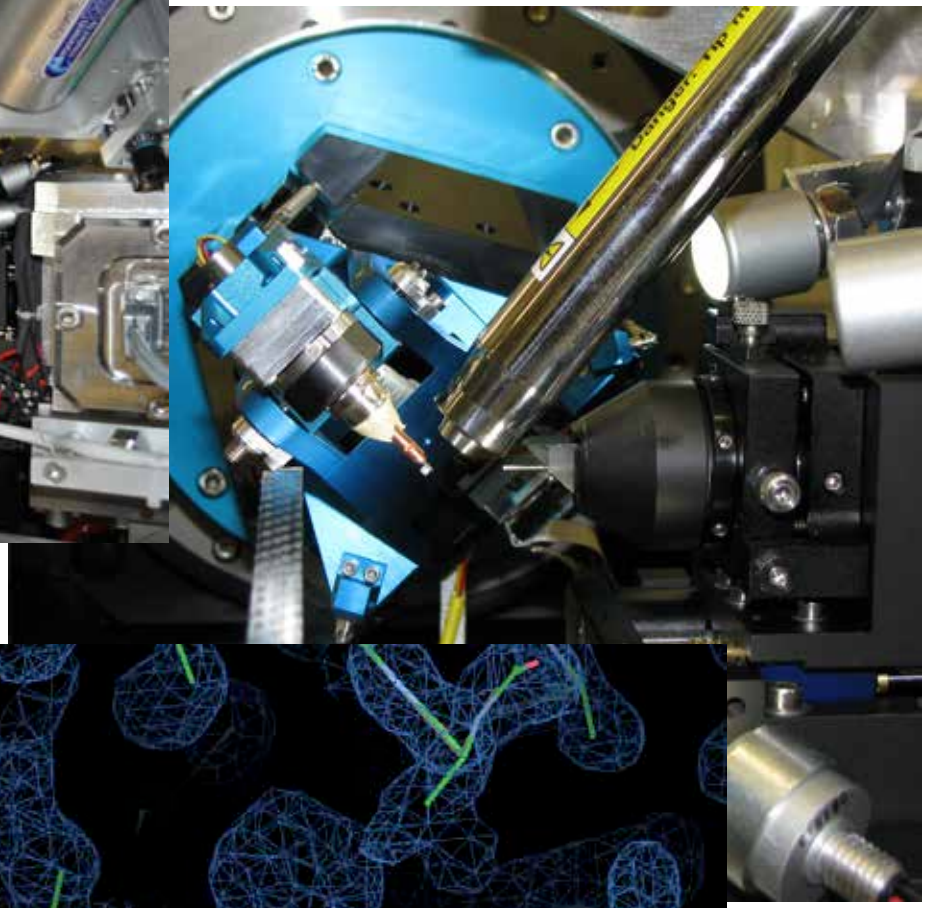
1 Å

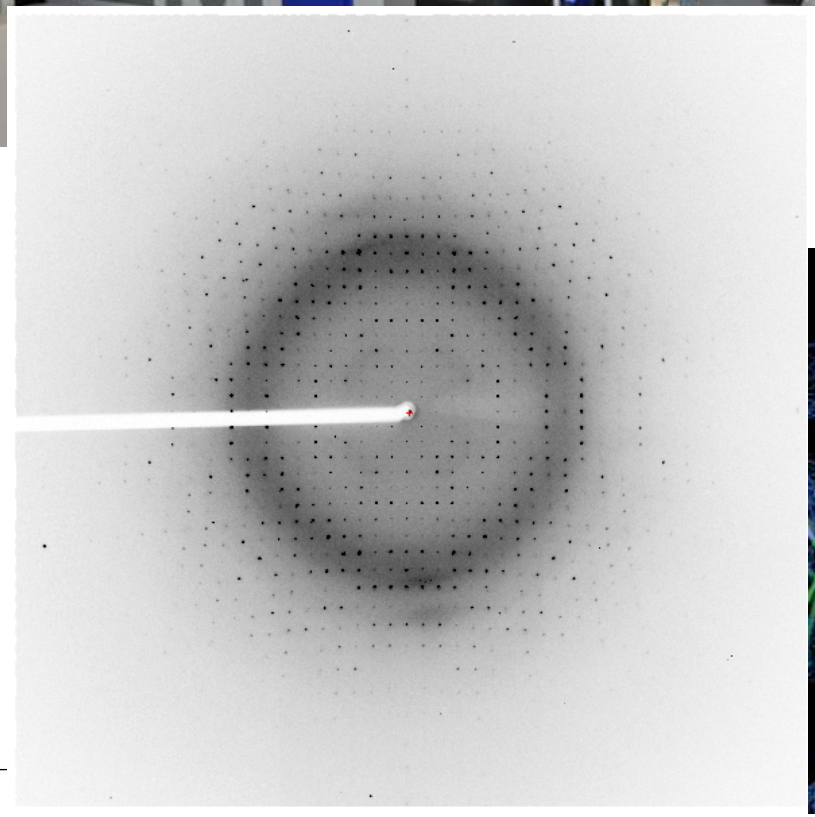
(from Wikipedia)

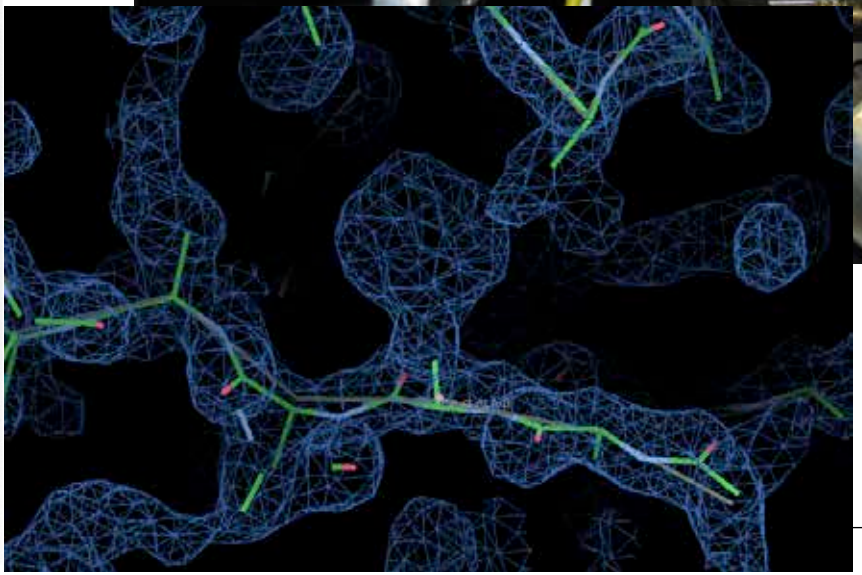
PAUL SCHERRER INSTITUT

Macromolecular crystallography (MX)

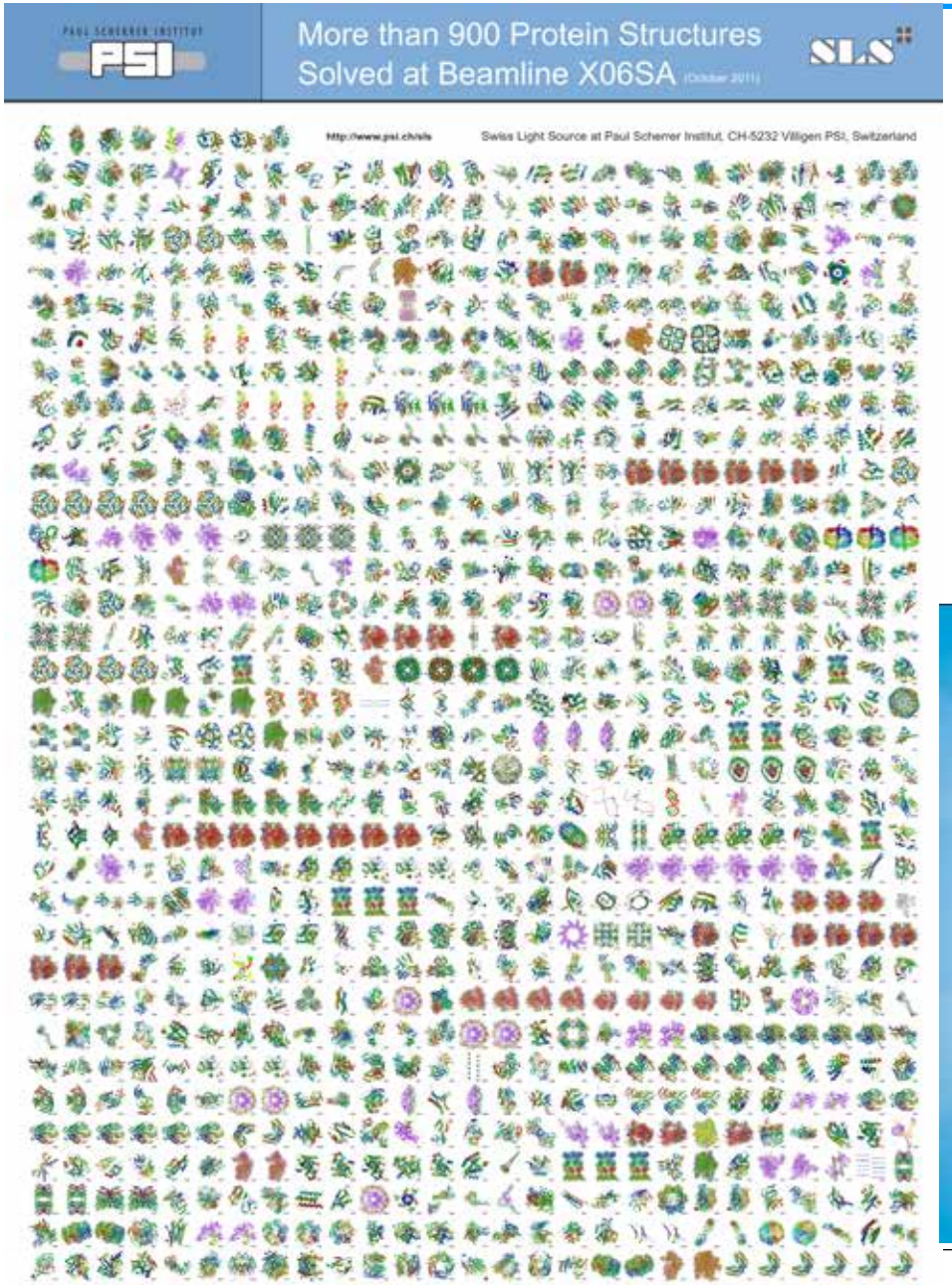




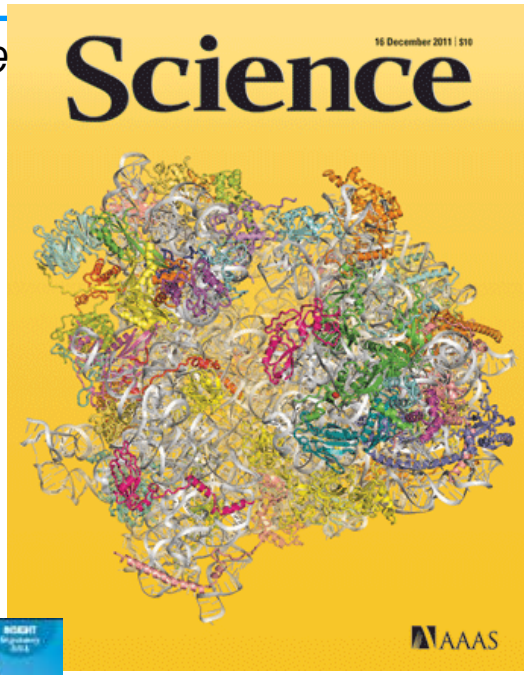




Macromolecular crystallography (MX)



Ben-Shem *et al.*, *Science* ***334****, 1524 (2011).*
80*S* ribosome from the yeast *Saccharomyces cerevisiae*,
3.3 megadaltons ,
data recorded at SLS beamline PX I / X06SA



Kato *et* al., Nature **482**, 369 (2012).
channelrhodopsin, 2.3 Å resolution,
data recorded at X06SA and BL32XU at SPring-8

MX: The tremendous effect of detector developments

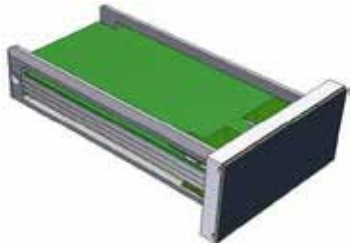
Quality and quantity

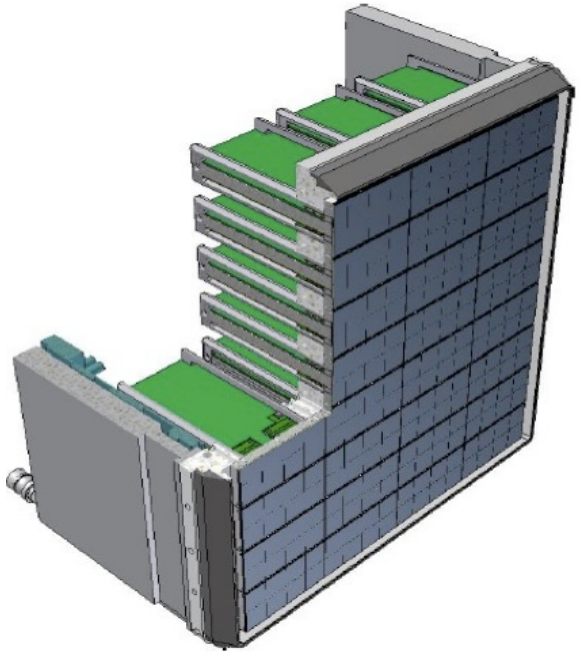
- from CCD to single photon counting
- from stop-and-go to fine-phi slicing

DECTRIS
Next Generation X-Ray Detectors

MX: From PILATUS 6M to EIGER 16M

... and from up to 12/25 Hz to 100 Hz
... and from 2/4 Gbit/s to 40 Gbit/s



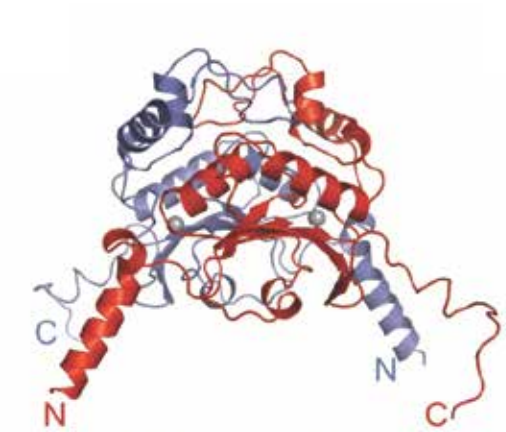


Solution scattering (SAXS)

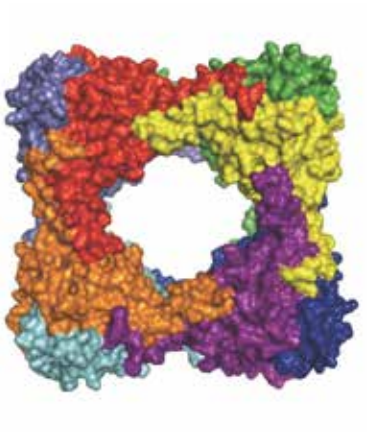
can investigate the shape and size of proteins in solution.

- sample: CS_2 hydrolase of an acidothermophilic archaeon living in mudpots of the Solfatara volcano, near Naples, Italy.
- SAXS helped to elucidate how evolution could adapt this organism's metabolism to its habitat: Instead of changing the active center of the hydrolase enzyme, the quartenary structure acts as specificity filter.
- M.J. Smeulders, T.R.M. Barends *et al*., *Nature*, **478** (2011) 412

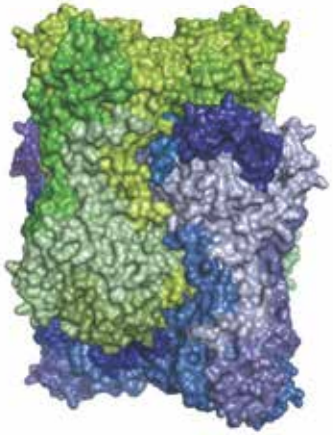
image credit:
http://en.wikipedia.org/wiki/Solfatara_%28volcano%29



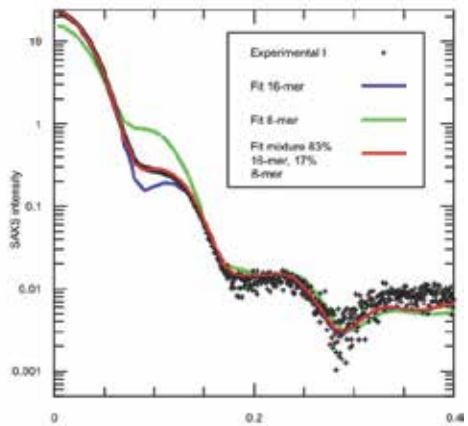
Crystal structure of the hydrolase dimer. Each monomer is colored individually, the long terminal extensions are indicated.



The long termini help to stabilize a ring-shaped octamer.



Two such rings interlock each other. One ring is shown with the monomers in various shades of blue, the other in shades of green.



SAXS results of solution (black crosses) and fits of the data (lines) assuming a pure hexadecamer, a pure octamer, and a mixture.

Time-resolved SAXS

For time-resolved measurements detector technology enabled changing from

one pump-probe repeat per data point

to

one pump-probe repeat per series of data points ('movie')

This is a factor of a few hundred!

The PILATUS 2M (2 Gbit/s) detector runs at maximum 30 Hz full frame or 300 Hz with two modules as 'region of interest'.
The duty cycle is for fast measurements awful due to the readout time of ~3ms.

The next generation EIGER detector with PSI readout system will generate 96 kbit/s/pixel, i.e., 6 GB/s/module.
Each module has two 10 GbE links.
The SAXS detector will have 18 modules, i.e., 360 Gb/s bandwidth .

Relevant time-scales in biology are seconds, milliseconds, microseconds.
→ **Some of our bio users would like to measure as fast as possible, right now.**

Scanning SAXS: Scattering meets imaging

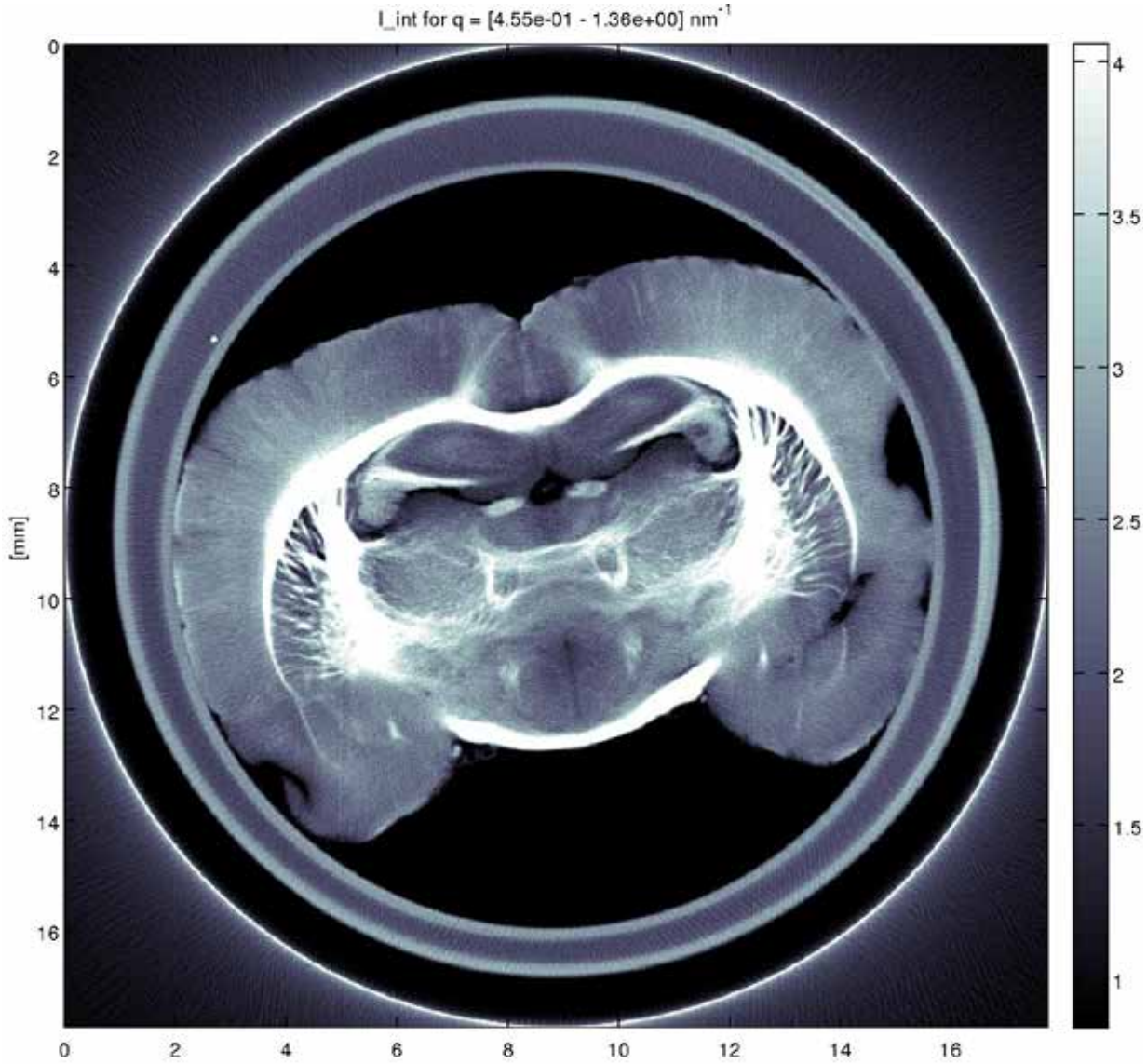
Scanning SAXS images nano-scale properties spatially resolved over extended samples.

It bridges the gap between high resolution low field of view and low resolution large field of view techniques.

Applications are for example biomedical imaging and materials science.

Each pixel in the image to the right is based on a 2.4M pixel PILATUS SAXS data frame.

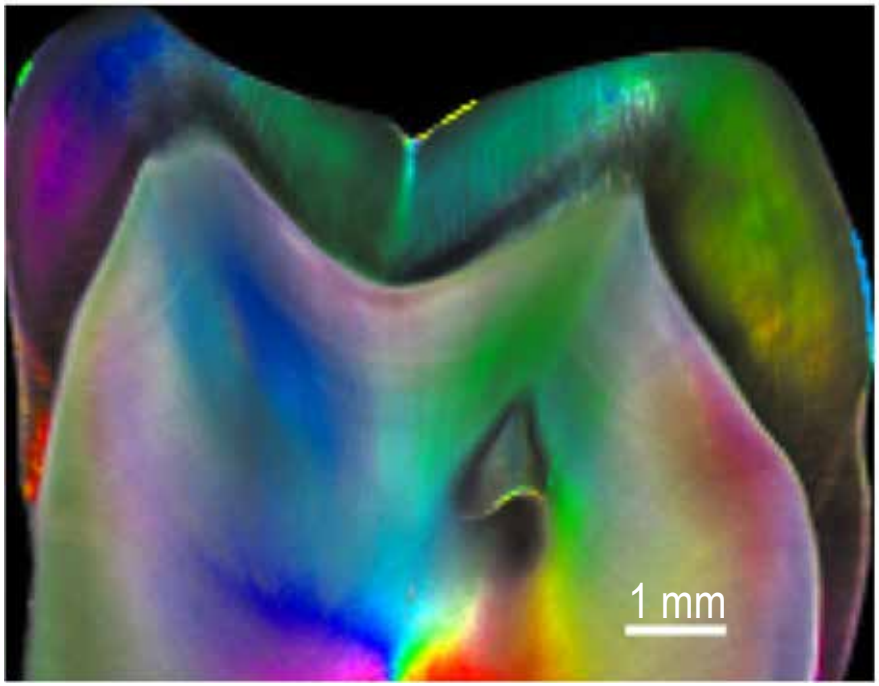
The information content of the data requires new ways of analyzing and visualizing large data sets.

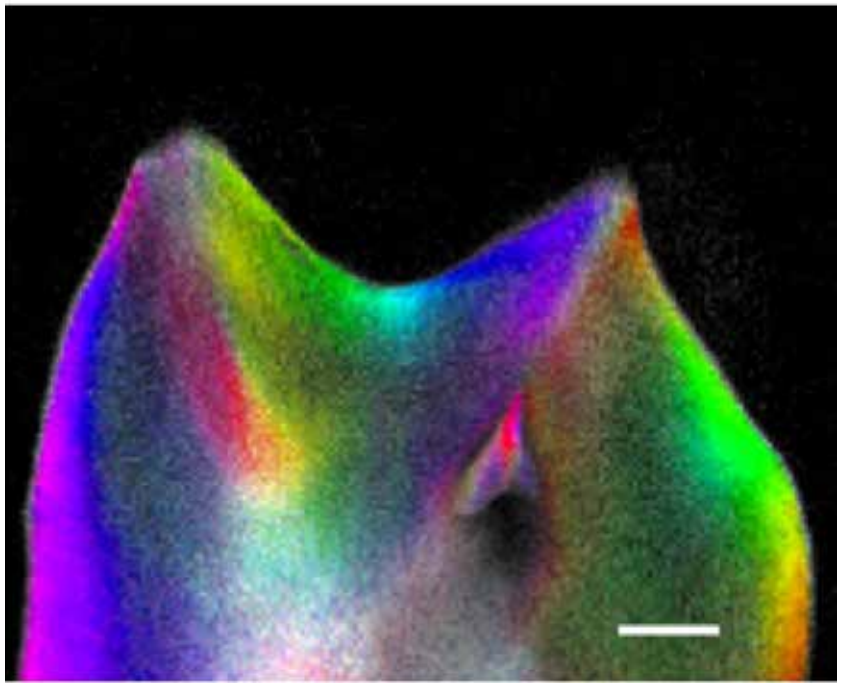


cSAXS beamline, on-the-fly scanning SAXS in tomography mode, 18.6 keV

Torben H. Jensen *et al.,* Phys. Med. Biol. **56**, 1717-1726 (2011). NeuroImage **57**, 124-129 (2011).

Scanning SAXS: Section through a human molar





185-231nm

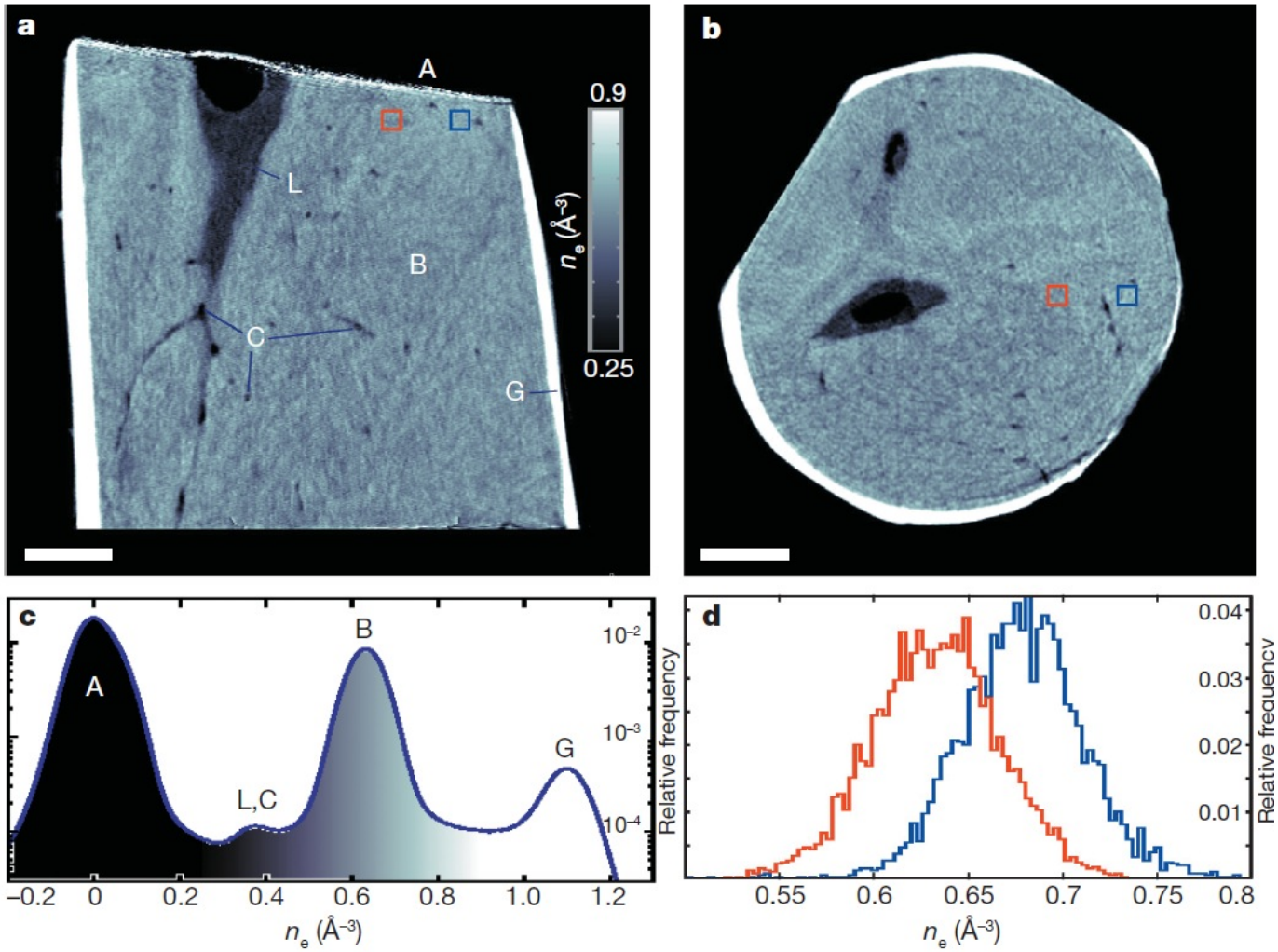
53-71nm, collagen peak extracted

- cSAXS beamline, 18.6 keV
- 20 ms exposure time per frame
- 181 x 141 points
- scan time 20 min

Bert Müller *et al.*, European Journal of Nanomedicine **3**, 30-33 (2010).

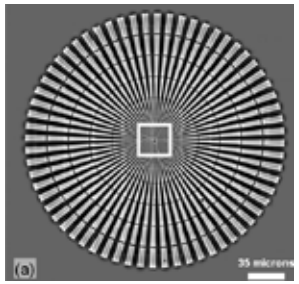
Scanning X-ray diffraction microscopy (ptychography)

CDI tomography:
highly resolving
voxel size $(65nm)^3$
resolution in 3D ~100nm
resolution in 2D ~ 10nm
quantitative results
uncertainty within voxel is 0.04 $e^-/Å^{-3}$
significantly higher sensitivity for larger volumes, e.g.,
<0.002 $e^-/Å^{-3}$ for $1\mu m^3$



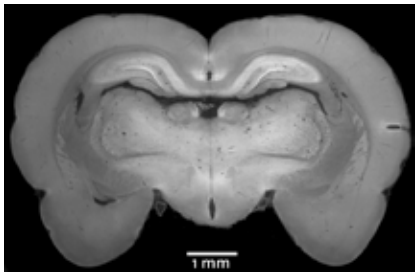
Martin Dierolf *et al.*, Nature **467**, 436-439 (2010).

X-ray tomography: towards several tomograms per second



Spatial Resolution
10 microns – 0.1 microns

Automation
Large scale studies



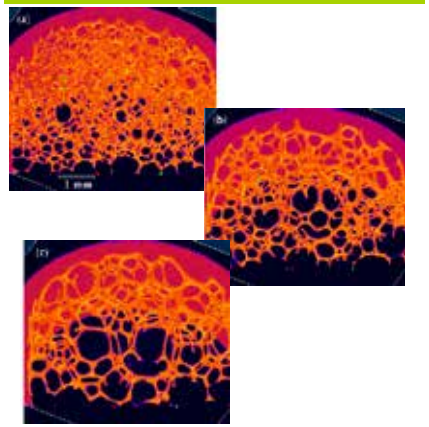
Density Resolution
Phase contrast imaging

INDUSTRIAL APPLICATIONS

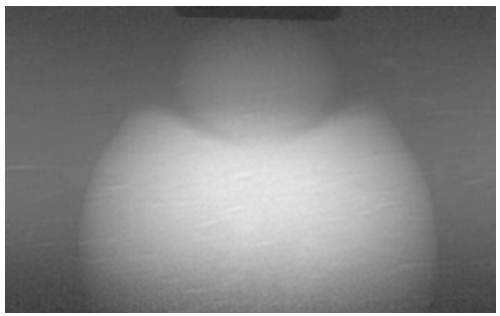
TOMCAT

Phase contrast imaging on low brilliance sources (cDPC,mammoDPC)

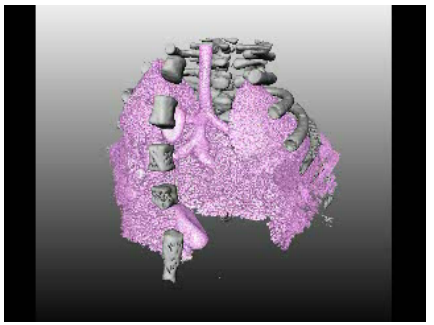
Ultrafast tomography



High-speed micro radiology



In-vivo 3D microscopy



Common requirements

Minimum requirements:

- **Data must be stored as fast as they are generated.**
- **An online feedback, i.e., on-the-fly data analysis is essential.**
- Data must be transferred in compressed form to the user institutes.
- Users should be able to read the data.
- The IT infrastructure should be affordable for all sides.
- **All changes need a strong motivation.**

Nice to have and less common requirements:

- Additional information is needed to analyze the data and should ideally be stored together with the data.
- Standards can be helpful and could be followed or created (e.g., NeXus).

For example the MX community is conservative and introducing a new data format involves convincing several programmers of scientific software and the users.

Summary

- There is a clear scientific motivation for going towards higher data rates and larger data sets.
- New detector technology like EIGER has the potential to establish a new data format – provided there is a clear scientific motivation.
- Users and facilities worldwide are potentially affected by the outcome of this project.