Small-angle X-ray scattering from biomacromolecular solutions

Dmitri Svergun

EMBL

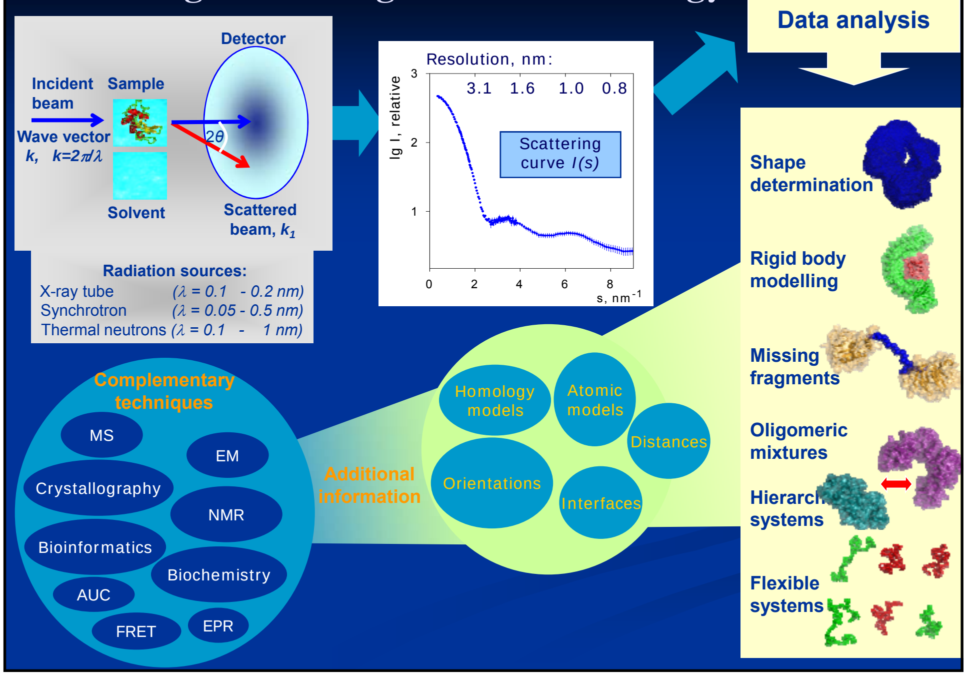
Biological SAXS @ EMBL-HH

Group leader: D. Svergun
Staff : M. Petoukhov, C.Blanchet, D.Franke, A.Kikhney, H.Mertens
Postdocs: A.Tuukkanen, M.Graewert , A.Spilotros, C.Jeffries, A.Panjkovich
Predocs: G.Tria, M.Kachala, E.Valentini, N.Hajizadeh

Major tasks:

- Development of data analysis methods
- Running and developing SAXS beamlines
- User support and collaborative projects
- Education and training

Small-angle scattering in structural biology

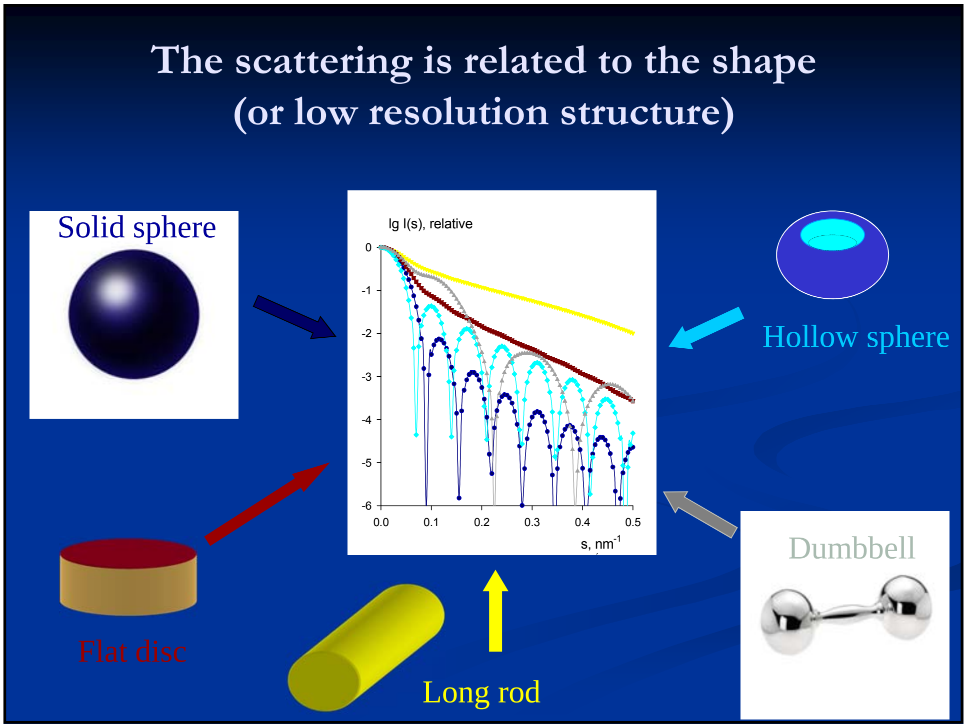


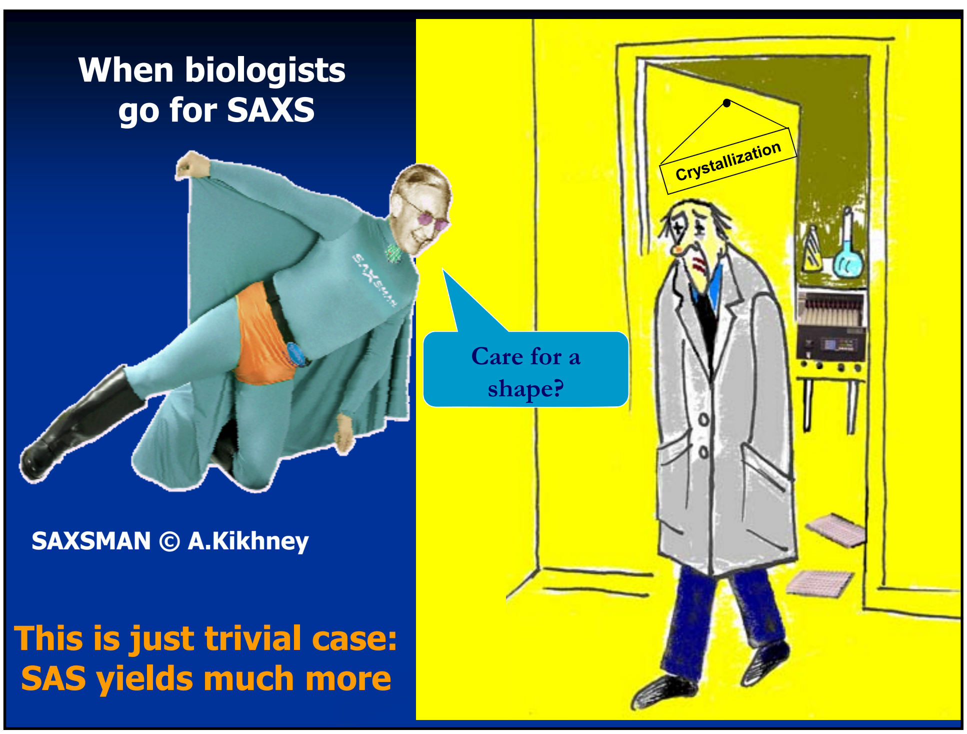
Crystal versus solution

Crystal versus solution

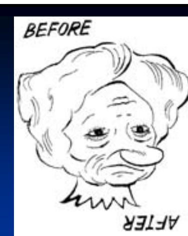
- In solution, no crystallographic packing forces are present

- For SAXS solution studies, one does not need to grow crystals
- SAXS is not limited by molecular mass and is applicable under nearly physiological conditions
- Using solution SAXS, one can more easily observe responses to changes in conditions
- SAXS permits for quantitative analysis of complex systems and processes


The scattering is related to the shape
(or low resolution structure)
Solid sphere
lg I(s), relative
0
-1
-2
-3
-4
-5
-6
0.0
0.1
0.2
0.3
0.4
0.5
s, nm-1
Hollow sphere
Flat disc
Long rod
Dumbbell

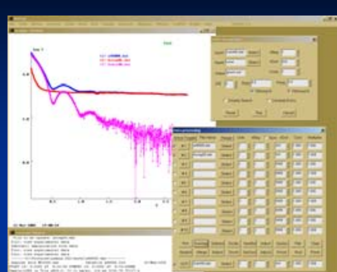

When biologists
go for SAXS
SAXSMAN © A.Kikhney
This is just trivial case:
SAS yields much more
Care for a
shape?
Crystallization

Some words of caution

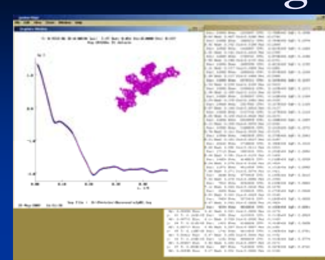


- Sample preparation
- Experiment
- Data processing
- Unambiguous interpretation
- Changing conditions
- Relation to function

Methods development at EMBL-Hamburg



Employed by over 9500 users worldwide

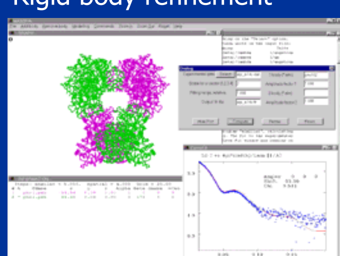


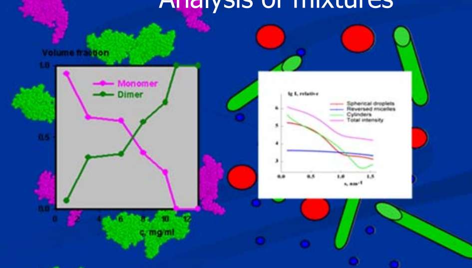
Data processing and manipulations

Ab initio modeling suite

Rigid body refinement

Analysis of mixtures



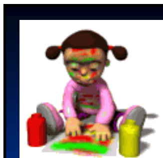


Konarev, P.V., Petoukhov, M.V., Volkov, V.V. & Svergun, D.I. (2006). *J. Appl. Cryst.* **39**, 277

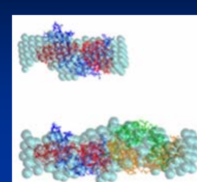
Scattering from dilute macromolecular solutions (monodisperse systems)

$$I(s) = 4\pi \int_0^D p(r) \frac{\sin sr}{sr} dr$$

The scattering is proportional to that of a single particle averaged over all orientations, which allows one to determine size, shape and internal structure of the particle at low (*1-10 nm*) resolution.

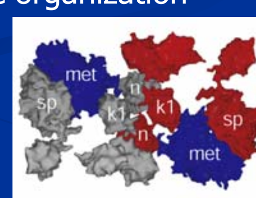


"Simple" monodisperse systems



Shape and conformational changes of macromolecules and complexes

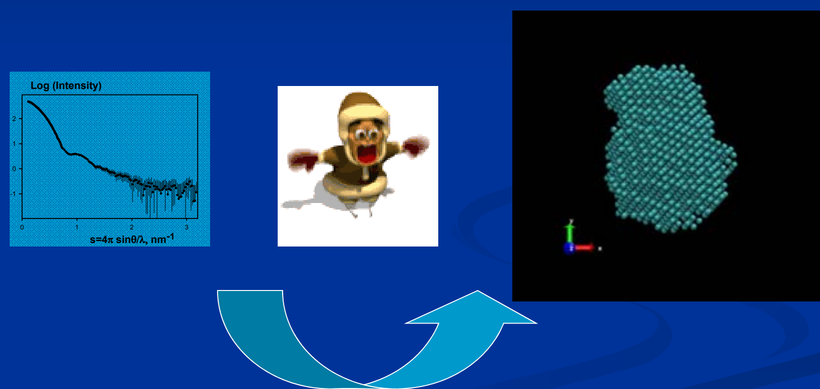
Validation of high resolution models and oligomeric organization



Rigid body models of complexes using high resolution structures

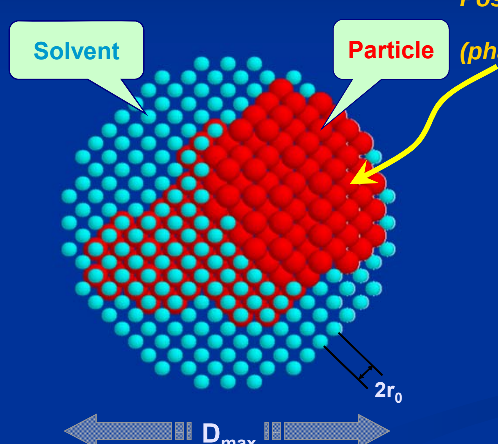
Addition of missing fragments to high resolution models

Ab initio methods



Bead (dummy atoms) model

A sphere of radius D_{max} is filled by densely packed beads of radius $r_0 << D_{max}$



Vector of model parameters:

Position (j) = X(j) = $\begin{cases} 1 & \text{if particle} \\ 0 & \text{if solvent} \end{cases}$

(phase assignments)

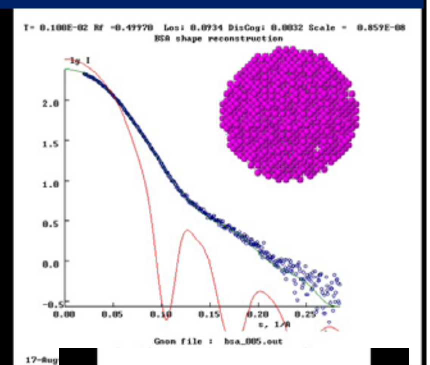
Number of model parameters $M \approx (D_{max} / r_0)^3 \approx 10^3$ is too big for conventional minimization methods – Monte-Carlo like approaches are to be used

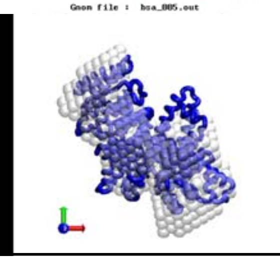
But: This model is able to describe rather complex shapes

Chacón, P. *et al.* (1998) *Biophys. J.* 74, 2760-2775.

Svergun, D.I. (1999) *Biophys. J.* 76, 2879-2886

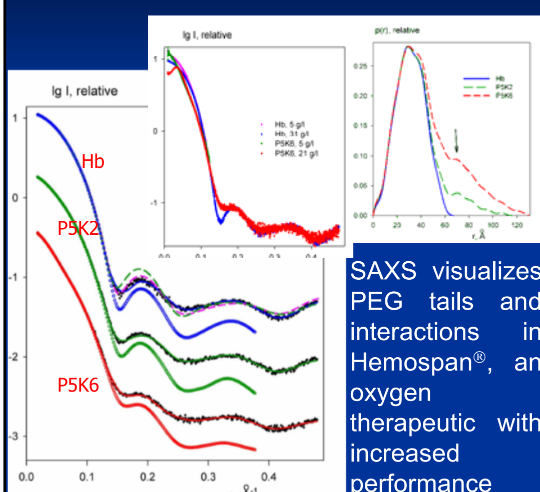
How to reconstruct 3D from 1D



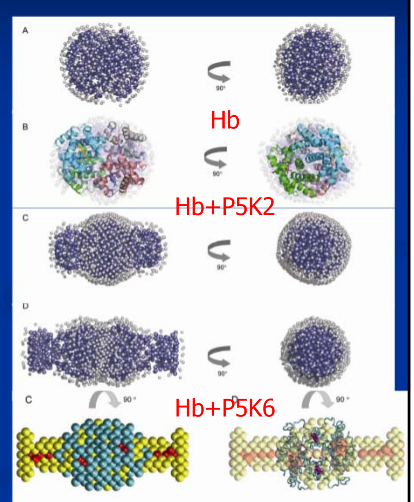


The 3D model is required not only to fit the data but also to fulfill (often stringent) physical and/or biochemical constrains

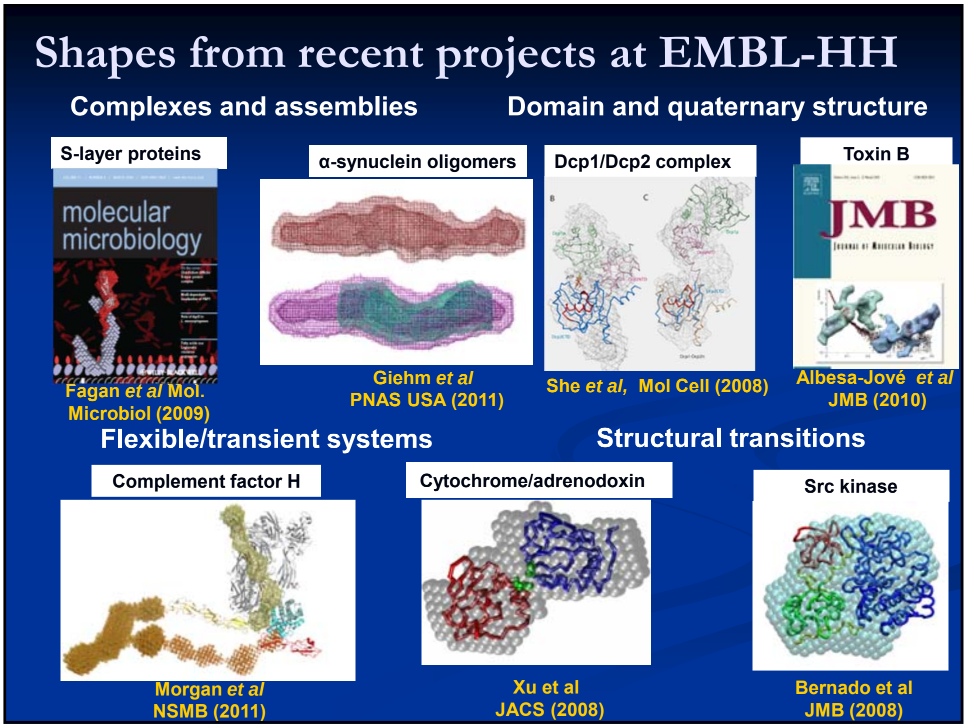
Poly(ethylene) glycol-conjugated hemoglobin: implications for a new oxygen therapeutic

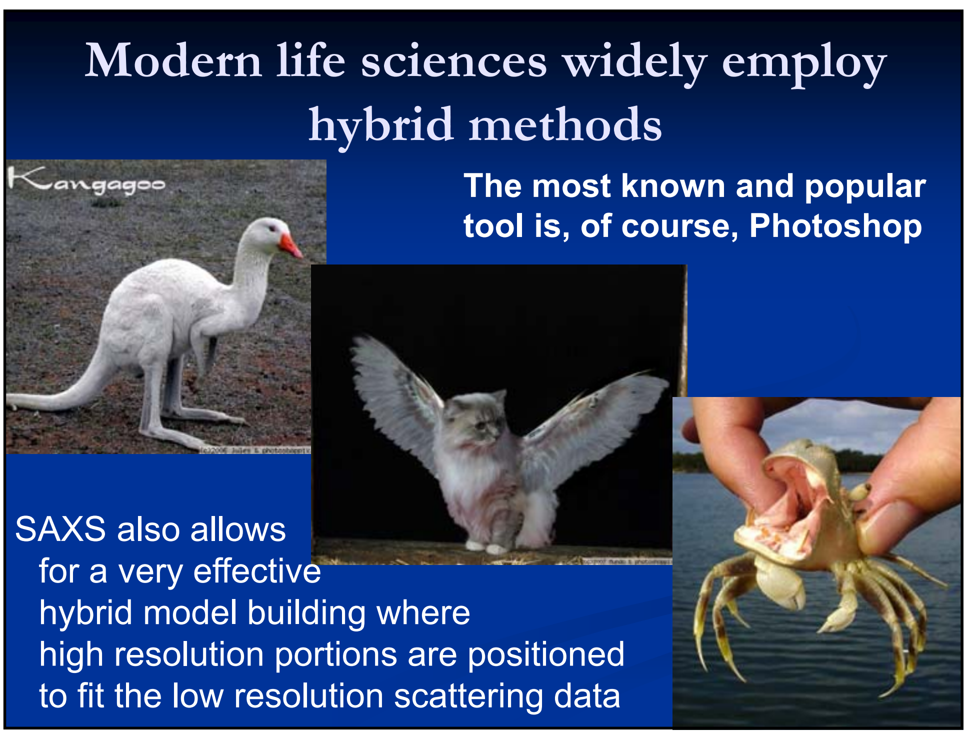


SAXS visualizes PEG tails and interactions in Hemospan®, an oxygen therapeutic with increased performance

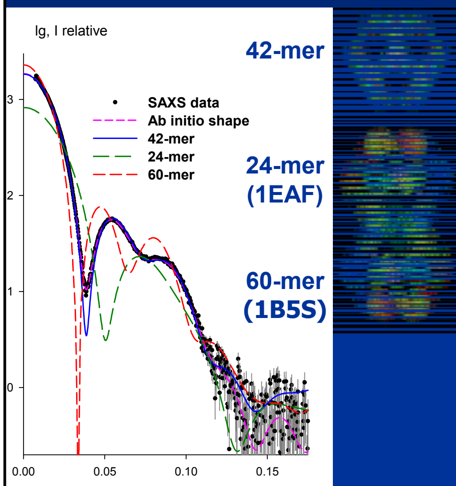


Svergun DI, Ekström F, Vandegriff KD, Malavalli A, Baker DA, Nilsson C & Winslow RM (2008) *Biophys J.* **94**, 173-181.

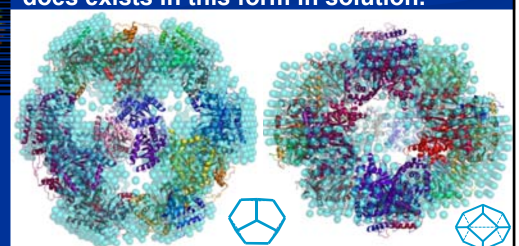

Shapes from recent projects at EMBL-HH
Complexes and assemblies
Domain and quaternary structure
S-layer proteins
molecular microbiology
Fagan et al Mol. Microbiol (2009)
α-synuclein oligomers
Giehm et al PNAS USA (2011)
Dcp1/Dcp2 complex
She et al, Mol Cell (2008)
Toxin B
JMB
Albesa-Jové et al JMB (2010)
Flexible/transient systems
Structural transitions
Complement factor H
Morgan et al NSMB (2011)
Cytochrome/adrenodoxin
Xu et al JACS (2008)
Src kinase
Bernado et al JMB (2008)


Modern life sciences widely employ hybrid methods
Kangagoo
The most known and popular tool is, of course, Photoshop
SAXS also allows for a very effective hybrid model building where high resolution portions are positioned to fit the low resolution scattering data

Catalytic core of E2 multienzyme complex is an irregular 42-mer assembly

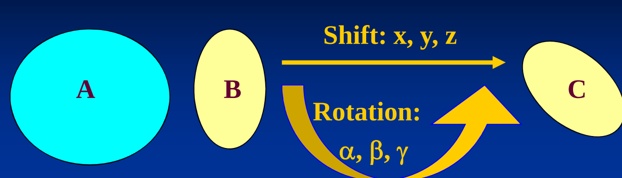


The E2 cores of the dihydrolipoyl acyl-transferase (E2) enzyme family form either octahedral (24-mer) or icosahedral (60-mer) assemblies. The E2 core from Thermoplasma acidophilum assembles into a unique 42-meric oblate spheroid. SAXS proves that this catalytically active 1.08 MDa unusually irregular protein shell does exists in this form in solution.



Marrott NL, Marshall JJ, Svergun DI, Crennell SJ, Hough DW, Danson MJ & van den Elsen JM. (2012) *FEBS J.* **279**, 713-23

Principle of rigid body modelling



Using spherical harmonics, the amplitude(s) of arbitrarily rotated and displaced subunit(s) are analytically expressed *via* the initial amplitude and the six positional parameters: $C_{lm}(s) = C_{lm}(B_{lm}, \alpha, \beta, \gamma, x, y, z)$.

The scattering from the complex is then rapidly calculated as

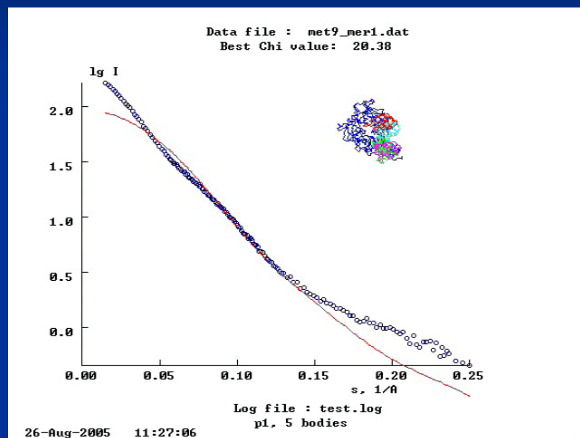
$$I(s) = I_A(s) + I_B(s) + 4\pi^2 \sum_{0}^{\infty} \sum_{-l}^{l} \mathrm{Re}\left[A_{lm}(s) C^*_{lm}(s)\right]$$

Svergun, D.I. (1991). *J. Appl. Cryst.* **24**, 485-492

A global refinement run with distance constraints

A tyrosine kinase MET (118 kDa) consisting of five domains

Program SASREF

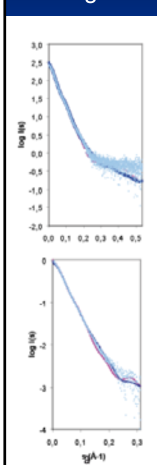


Single curve fitting with distance constraints: C to N termini contacts

Gherardi, E., Sandin, S., Petoukhov, M.V., Finch, J., Youles, M.E., Ofverstedt, L.G., Miguel, R.N., Blundell, T.L., Vande Woude, G.F., Skoglund, U. & Svergun, D.I. (2006) *PNAS USA,* **103**, 4046.

Architecture of nuclear receptor heterodimers on DNA direct repeat elements

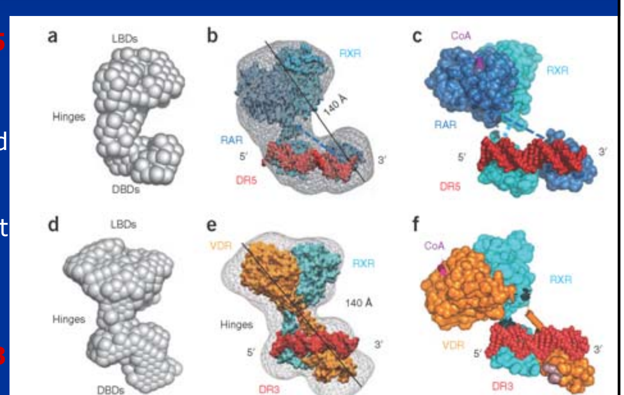
Nuclear hormone receptors (NHRs) control numerous physiological processes through the regulation of gene expression. SAXS, SANS and FRET were employed to determine the solution structures of NHR complexes, RXR–RAR, PPAR–RXR and RXR–VDR, free and in complex with the target DNA



RXR–RAR–DR5

Ab initio and rigid body models of NHRs complexed with direct repeat elements

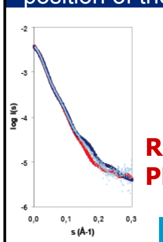
RXR–VDR–DR3

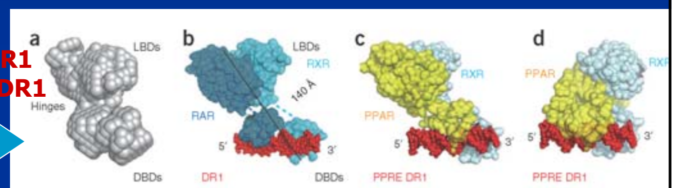


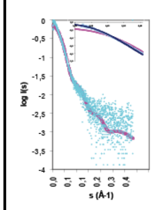
N.Rochel, F.Ciesielski, J.Godet, E.Moman, M.Roessle, C.Peluso-Iltis, M.Moulin, M. Haertlein, P.Callow, Y.Mely, D.Svergun & D.Moras (2011) *Nat Struct Mol Biol* 18, 564-70

Architecture of nuclear receptor heterodimers on DNA direct repeat elements

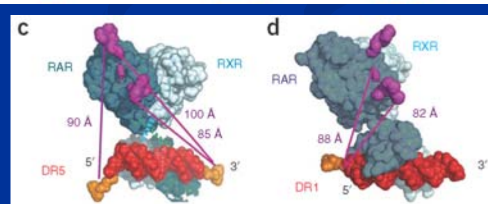
NHR-DNA complexes show extended asymmetric shape and reveal conserved position of the ligand-binding domains at the 5′ ends of the target DNAs. Further, the binding of only one coactivator molecule per heterodimer, to RXR's partner, is observed.







The models and the polarity of RXR–RAR–DR5 and RAR–RXR–DR1 were validated using neutron scattering and FRET

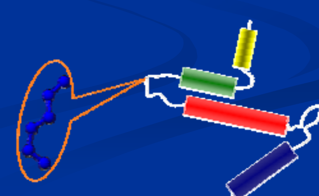


N.Rochel, F.Ciesielski, J.Godet, E.Moman, M.Roessle, C.Peluso-Iltis, M.Moulin, M. Haertlein, P.Callow, Y.Mely, D.Svergun & D.Moras (2011) *Nat Struct Mol Biol* 18, 564-70

Addition of missing fragments

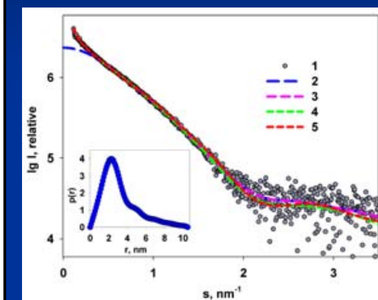
- Flexible loops or domains are often not resolved in high resolution models
- Their tentative configuration can be reconstructed by fixing the known portion and adding the missing parts to fit the scattering from the full-length macromolecule.
- Moreover, addition of missing fragments can be combined with rigid body refinement (programs BUNCH and CORAL)

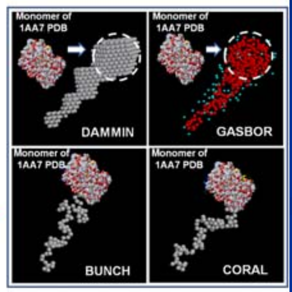
Petoukhov, M. V. & Svergun, D. I. (2005). *Biophys. J.* **89**, 1237-1250

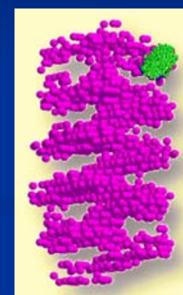


C-terminal flexibility of the influenza virus M1 protein

Influenza A virus matrix protein M1 is one of the most important and abundant proteins in the virus particles. The low resolution SAXS models reveal a polarized M1 molecule consisting of a compact NM-fragment, compatible with its crystal structure, and an extended and partially flexible C-terminal domain.

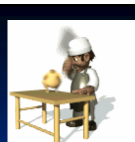




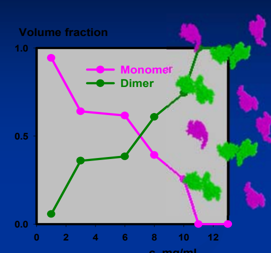


The M1 monomers co-exist in solution with a small fraction of large assemblies revealing a layered architecture similar to that observed in the authentic influenza virions. The flexibility of the C-terminus is an essential feature relevant for the multifunctionality of the entire protein.

E.V. Shtykova, L.A. Baratova, N.V. Fedorova, V.A. Radyukhin, A.L. Ksenofontov, V. V. Volkov, A.V. Shishkov, A.A. Dolgov, L.A. Shilova, O.V. Batishchev, C.M. Jeffries, D.I. Svergun (2013), PLoS One, 8(12):e82431.

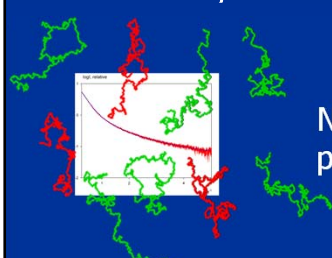


Life is more complicated: mixtures and processes

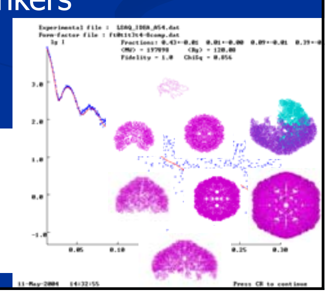


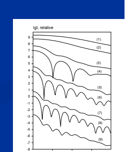
Equilibrium oligomeric mixtures

Stoichiometry and complex formation



Natively unfolded proteins and multidomain proteins with flexible linkers





Protein folding/unfolding kinetics

Assembly/disassembly processes

Scattering from mixtures

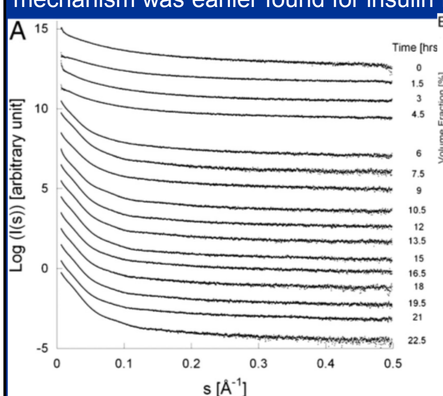
$$I(s)=\sum_k v_k I_k(s)$$

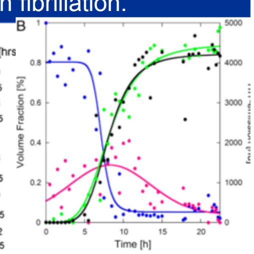
The scattering is proportional to that of a single particle averaged over all orientations, which allows one to determine size, shape and internal structure of the particle at low *(1-10 nm)* resolution.

For equilibrium and non-equilibrium mixtures, solution scattering permits to determine the number of components and, given their scattering intensities $I_k(s)$, also the volume fractions

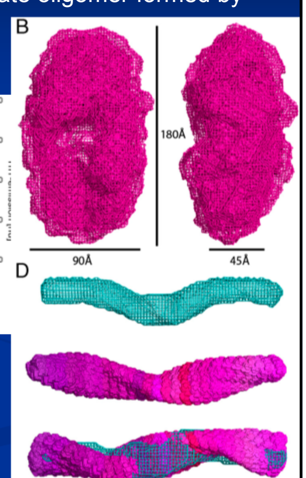
Fibrillation of α-synuclein (αSN)

Aggregation of αSN leads to Parkinson decease. The fibril formation process characterized by SAXS reveals that there exists an intermediate oligomer formed by several partially unfolded αSN molecules, and the mature fibril is formed by association of these oligomers. Such mechanism was earlier found for insulin fibrillation.



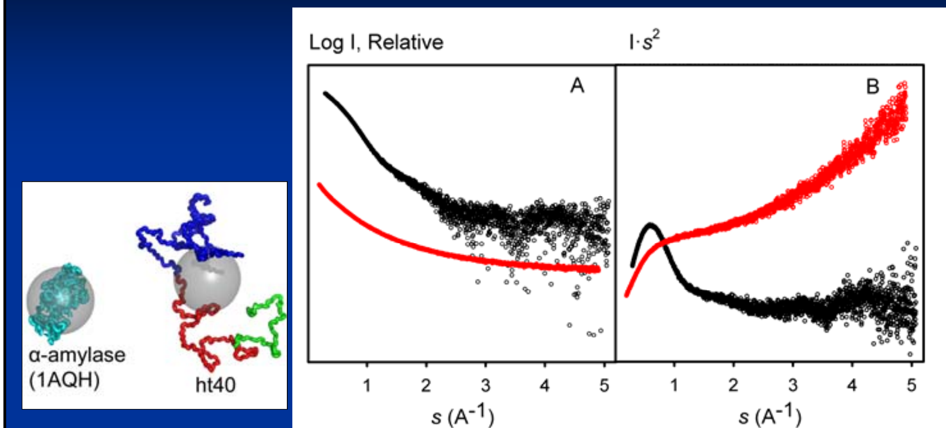


The oligomer is shown to disrupt liposomes, i.e. it is potentially cytotoxic



Giehm, L., Svergun, D.I., Otzen, D.E. & Vestergaard, B. (2011) *PNAS USA,* **108,** 3246

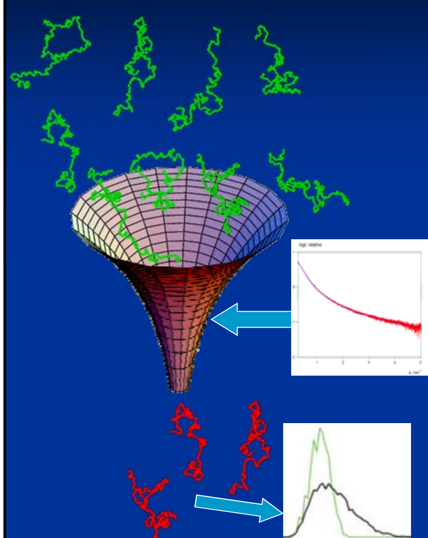
SAXS from folded *vs* disordered protein



Folded: relatively small R_g and D_{max}, bell-shaped Kratky plot (e.g. for folded α-amylase (448 AAs) R_g=2.4 nm)

Disordered: large R_g and D_{max}, increasing Kratky plot (e.g. for IUP tau (441 AAs) Rg=6.5 nm)

Quantitative assessment of flexibility



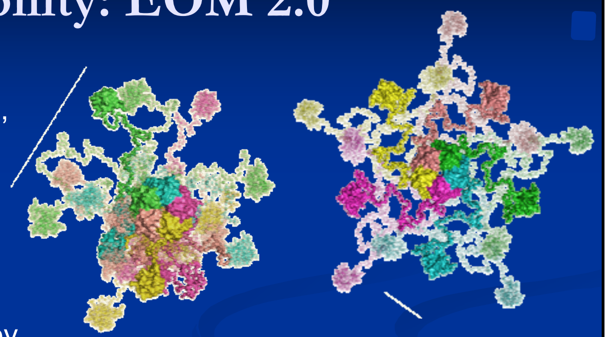
- Automated classification (folded, partially or completely unfolded) is available
  D.Franke DATCLASS
- More quantitative estimates are provided by ensemble methods
- One generates a large pool covering the conformational space and selects sub-ensemble(s) such that their mixture fits the available experimental data
- The structural properties of the selected ensemble(s) are compared to those of the pool

EOM, Bernadó et al. (2007)
J. Am. Chem. Soc. **129**, 5656.

30

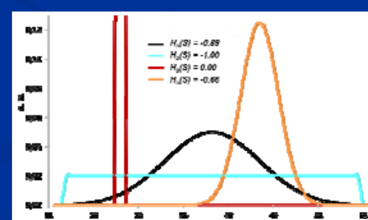
Quantitative assessment of flexibility: EOM 2.0

- Extended possibilities for the pool generation, e.g. use of (partial) point symmetry
- Broader search capabilities
- Quantification of flexibility using enthropy and variation



$$R_{flex} = -\mathrm{H_b}(\mathrm{S}) = \sum_{i=1}^{n} p(x_i)\log_b(p(x_i))$$

$$\mathrm{R}_\sigma = \sigma_{ensemble} / \sigma_{pool}$$

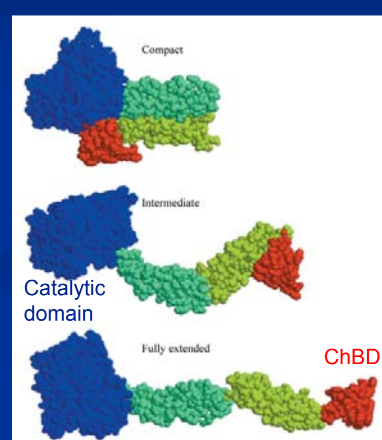


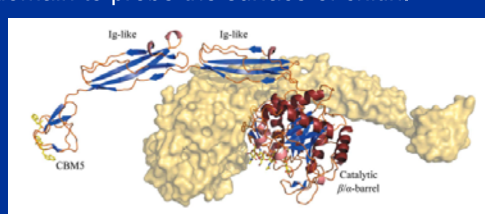
G.Tria, M.Kachala

Crystal structures of substrate-bound chitinase from *Moritella marina* and its structure in solution

Chitinases break down glycosidic bonds in chitin and only few crystal structures are reported because of the flexibility of these enzymes.

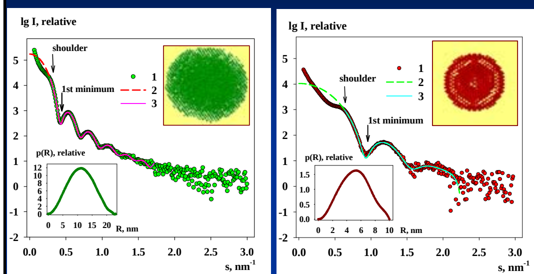
The dimeric crystal structure (at BESSY) of chitinase 60 from M. marina (MmChi60) contains four domains: catalytic, two Ig-like, and chitin-binding (ChBD). SAXS (at EMBL) demonstrates that MmChi60 is monomeric and flexible in solution. The flexibly hinged Ig-like domains may thus allow the catalytic domain to probe the surface of chitin.





P. H. Malecki, C. E. Vorgias, M. V. Petoukhov, D. I. Svergun and W. Rypniewski. *Acta Cryst.* (2014) **D70**, 676-684

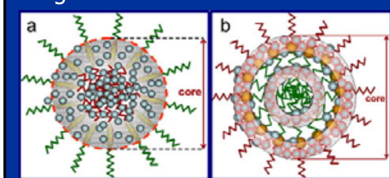
Not only proteins: magnetic nanoparticles

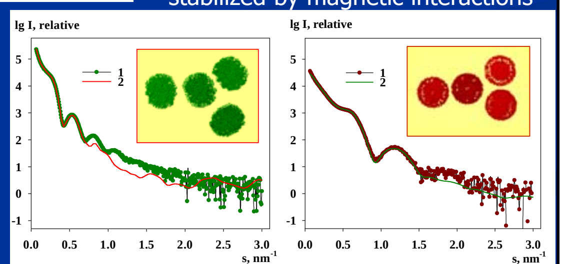


Highly monodisperse NPs are prepared by thermal decomposition of iron compounds including oxygen-containing ligands in boiling surfactants. The NPs are coated by phospholipids with PEG Tails to become soluble.

Rigid body analysis reveals equilibrium clusters of the NPs stabilized by magnetic interactions

Ab initio analysis: peculiarities of organization of different NPs



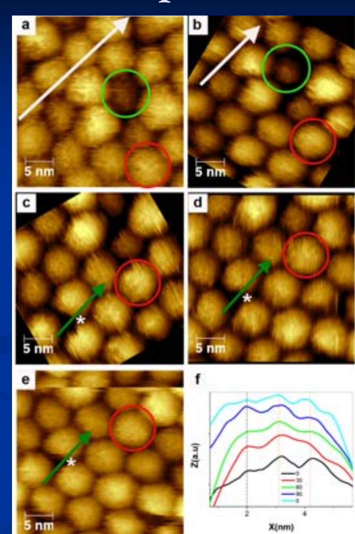


Shtykova, E.V, Huang, X., Remmes, N., Baxter, D., Dixit, S., Stein, B., Dragnea, B., Svergun, D. I. & Bronstein, L. M. (2007) J. Phys. Chem. C, **111**, 18078-18086

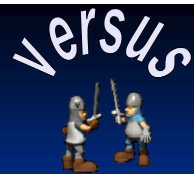
Mixed Monolayer Protected Gold Nanoparticles

Gold nanoparticles (NPs) are efficiently synthesized using self-assembled monolayers (SAMs) as stabilizing agents. SAMs of thiolated ligand molecules on gold substrates are ordered two-dimensional crystals. When mixtures of molecules are used, these may spontaneously form domains. The ligand shell provides the NP with important interfacial properties.

Scanning tunneling microscopy on dodecanethiol-hexanethiol (C12 : C6) 2:1 protected gold nanoparticles reveals stripe-like domains persistent across many images and retain their direction and overall morphology at different scan angles



M. Moglianetti, Q.K. Ong, J. Reguera, K.M. Harkness, M. Mameli, A. Radulescu, J. Kohlbrecher, C. Jud, D.I. Svergun, F. Stellacci (2014) Chem. Sci., **5**, 1232-1240

X-rays  neutrons

- **X-rays:** scattering factor increases with atomic number, no difference between H and D
- **Neutrons:** scattering factor is irregular, may be negative, huge difference between H and D

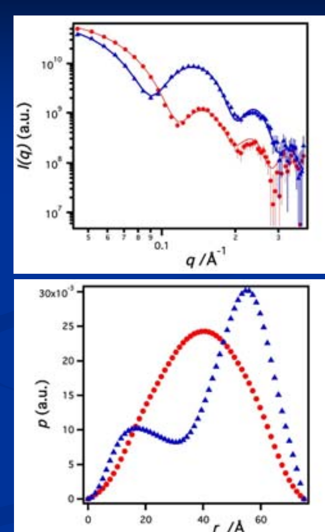
Element	H	D	C	N	O	P	S	Au
At. Weight	1	2	12	14	16	30	32	197
N electrons	1	1	6	7	8	15	16	79
b_X,10^{-12} cm	0.282	0.282	1.69	1.97	2.16	3.23	4.51	22.3
b_N,10^{-12} cm	-0.374	0.667	0.665	0.940	0.580	0.510	0.280	0.760

Neutron contrast variation

Mixed Monolayer Protected Gold Nanoparticles

SANS was performed on the samples with specifically deuterated dodecanethiol (C6:d-C12; red circles) and hexanethiol (d-C6:C12; blue triangles).

Contrast, 10^{-6} Å^{-2}	C6: d-C12	d-C6: C12
Phase 1: gold	1.4	1.4
Phase 2: C12	2.5	-3.4
Phase 3: C6	-3.4	1.9

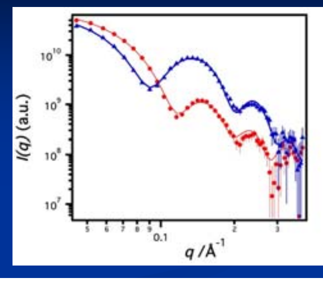


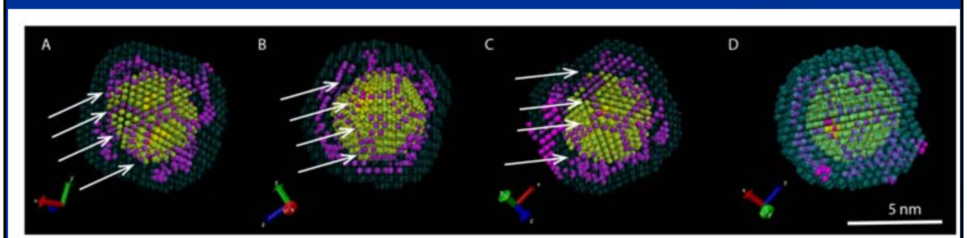
The scattering patterns and distance distribution functions from the two hybrid particles are distinctly different indicating that SANS sees the internal structure

M. Moglianetti, Q.K. Ong, J. Reguera, K.M. Harkness, M. Mameli, A. Radulescu, J. Kohlbrecher, C. Jud, D.I. Svergun, F. Stellacci (2014) Chem. Sci., **5**, 1232-1240

Mixed Monolayer Protected Gold Nanoparticles

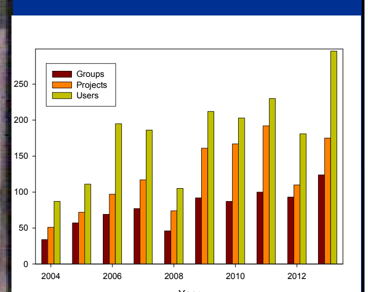
The shape and internal structure of the particles was reconstructed using a multi-phase model depicting the three components (yellow: gold; magenta: C6; cyan: C12 and fitting both data sets from specifically deuterated thiols.





M. Moglianetti, Q.K. Ong, J. Reguera, K.M. Harkness, M. Mameli, A. Radulescu, J. Kohlbrecher, C. Jud, D.I. Svergun, F. Stellacci (2014) Chem. Sci., **5**, 1232-1240

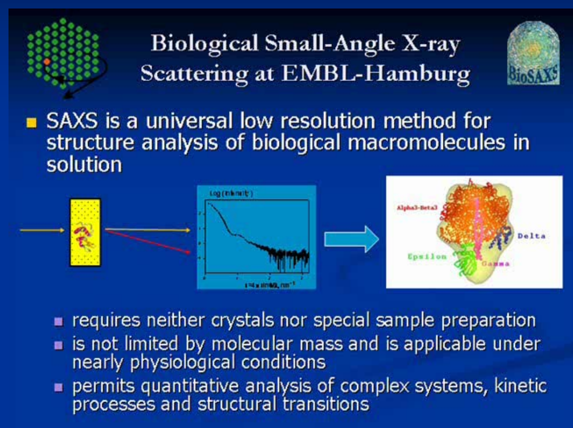
Small-angle scatterers and biologists



Beamyear 2013: 335 user-visits, 291 unique users, 159 projects, 119 unique groups (of those, 10 mail-in/remote and 3 proprietary)

CRG access (HZG): 39 user-visits, 33 unique users, 23 projects, 19 unique groups (NOT counted in the above graph)

EMBL SAXS X33 beamline, 1979-2012



This is how it worked in 2005

X33 SAXS beamline of the EMBL

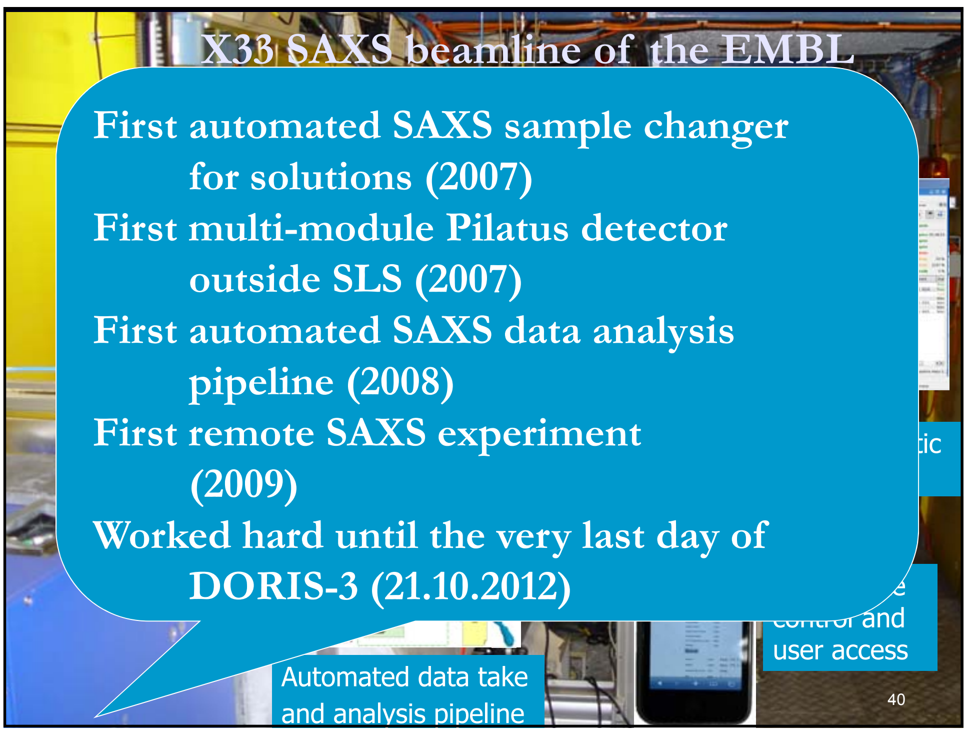
First automated SAXS sample changer for solutions (2007)

First multi-module Pilatus detector outside SLS (2007)

First automated SAXS data analysis pipeline (2008)

First remote SAXS experiment (2009)

Worked hard until the very last day of DORIS-3 (21.10.2012)

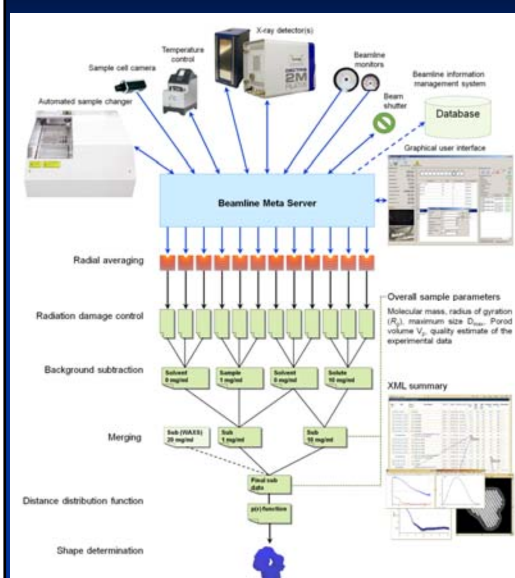


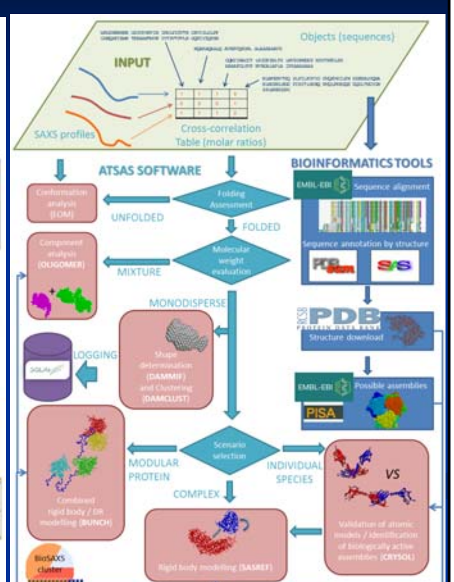
High brilliance beamline P12 (Petra-3)

(P)reincarnation:

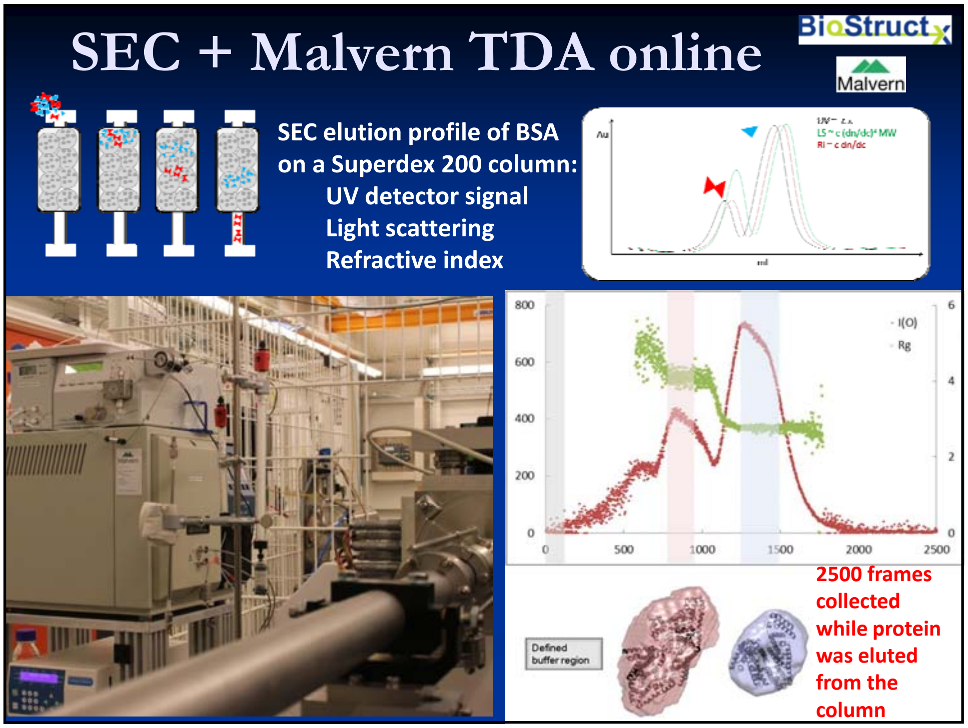
BioSAXS P12 beamline inherits not only the soul but also fully utilizes the automation developed for X33

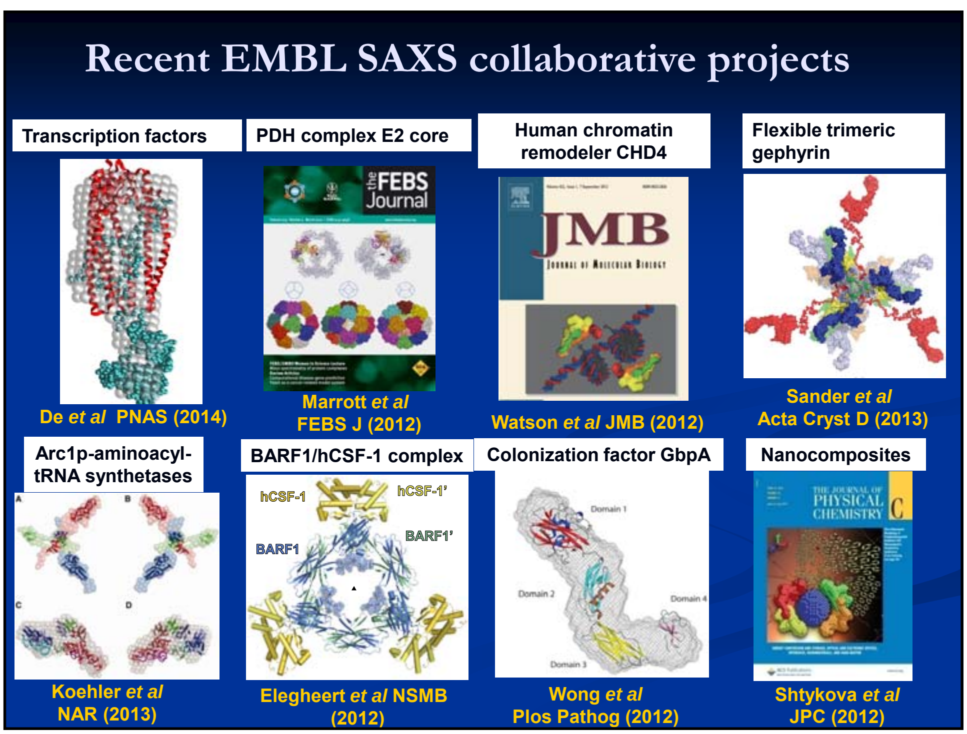
Automation of SAS experiment and analysis





Processing with an automated pipeline and interpretation with an expert system

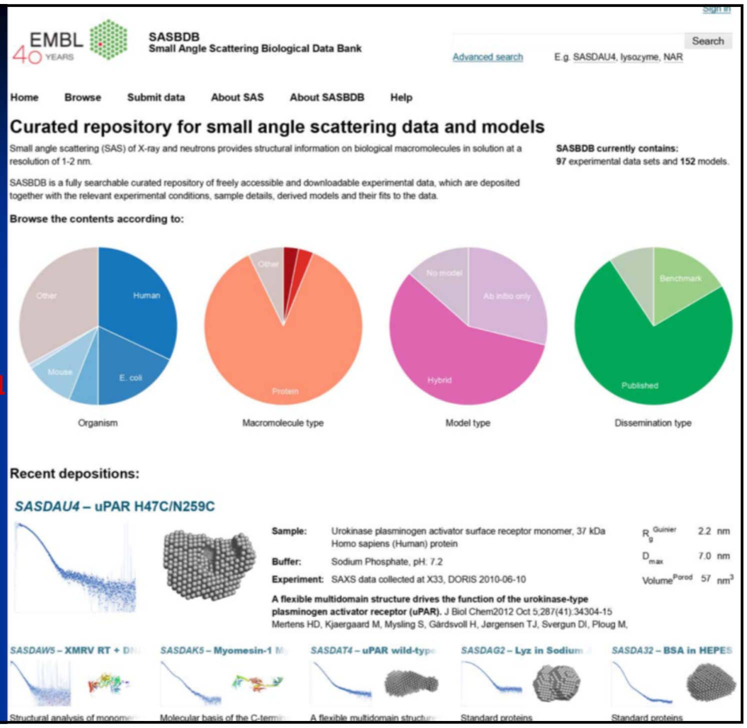

SEC + Malvern TDA online
BioStructX
Malvern
SEC elution profile of BSA
on a Superdex 200 column:
UV detector signal
Light scattering
Refractive index
I(0)
Rg
2500 frames collected while protein was eluted from the column
Defined buffer region


Recent EMBL SAXS collaborative projects
Transcription factors
De et al PNAS (2014)
PDH complex E2 core
Marrott et al
FEBS J (2012)
Human chromatin remodeler CHD4
JMB
Watson et al JMB (2012)
Flexible trimeric gephyrin
Sander et al
Acta Cryst D (2013)
Arc1p-aminoacyl-tRNA synthetases
Koehler et al
NAR (2013)
BARF1/hCSF-1 complex
hCSF-1
hCSF-1'
BARF1
BARF1'
Elegheert et al NSMB (2012)
Colonization factor GbpA
Domain 1
Domain 2
Domain 3
Domain 4
Wong et al
Plos Pathog (2012)
Nanocomposites
THE JOURNAL OF PHYSICAL CHEMISTRY C
Shtykova et al
JPC (2012)

SAS models and data deposition

SASBDB

www.sasbdb.org



E.Valentini
A.Kikhney

Recent reviews on solution SAS

Blanchet CE, Svergun DI (2013) Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. Annual Review of Physical Chemistry 64(1): 37–54.

Schneidman-Duhovny D, Kim S, Sali A. (2012) Integrative structural modeling with small angle X-ray scattering profiles. BMC Structural Biology 12(1):17.

Graewert MA, Svergun DI (2013) Impact and progress in small and wide angle X-ray scattering (SAXS and WAXS). Curr Opin Struct Biol 23: 748-754.

Rambo RP and Tainer JA (2013) Super-resolution in solution X-ray scattering and its applications to structural systems biology., Annu Rev Biophys. 42, 415-441

Books on SAXS

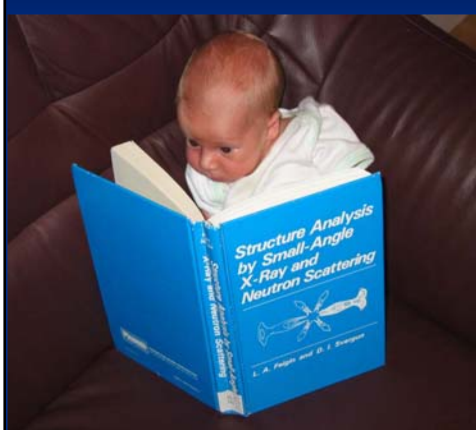
" The origins" (no recent edition) : Small Angle Scattering of X-rays. A. Guinier and A. Fournet, (1955), in English, ed. Wiley, NY

Small-Angle X-ray Scattering: O. Glatter and O. Kratky (1982), Academic Press. PDF available on the Internet at http://physchem.kfunigraz.ac.at/sm/Software.htm

Structure Analysis by Small Angle X-ray and Neutron Scattering. L.A. Feigin and D.I. Svergun (1987), Plenum Press.
PDF available on the Internet at http://www.embl-hamburg.de/ExternalInfo/Research/Sax/reprints/feigin_svergun_1987.pdf

Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules. D.I,Svergun, M.H.J. Koch, P.A.Timmins, R.P. May (2013) Oxford University Press, London.

Young scientists were always very interested in SAXS



… in preparation for EMBL PhD program …
(courtesy of F.Gabel, IBS, Grenoble)

The brand new SAS book is not yet available for everyone (courtesy of M.Graewert, EMBL-HH)

Typical tasks addressed by biological SAS (often in combination with other methods)

- Oligomeric struct
  (e.g. equilibri
- Quaterna
  and bir
- Conf
  (e.g.
- Kine
- Qua
- Analy
  concen

Synchro

Me
A ount of purified ml range)
Concentration range: from sub-mg/ml to hundreds mg/ml

Not just high throughput: important applications of SAXS/SANS are functional complexes and processes (flexible, dynamic, transient, evolving), where SAS is among the few methods providing quantitative structural information. And not only proteins: SAS is also applicable to various nanostructural systems

Acknowledgments

EMBL-Hamburg: D.Franke, A. Kikhney, P. Konarev, M. Petoukhov, M. Roessle, C.Blanchet, H.Mertens, A.Spilotros, G.Tria, M. Graewert, C.Jeffries, A.Tuukkanen, M.Kachala, E.Valentini
http://www.embl-hamburg.de/biosaxs

Collaborative projects at EMBL beamlines:

- **Calmodulin, MMP12:** I.Bertini, C.Luccinat (CERM, Florence)
- **BARF1, Flt3:** S. Savvides (Gent University)
- **Insulin, α-synuclein:** B.Vestergaard (Pharmaceutical Uni Copenhagen)
- **PEG-hemoglobin:** F.Ekström (FOM, Sweden), K.D.Vandegriff (Sangart, USA)
- **Myomesin:** N.Pinotsis (University Vienna), M.Wilmanns (EMBL, Hamburg)
- **Nuclear receptors:** D.Moras (CNRS Strasbourg)
- **S-layer proteins:** N. Fairweather, K. Brown (Imperial College, London)
- **Frataxin:** A.Pastore (NIMR MRC, London)
- **E2 core:** J. van den Elsen (University of Bath)
- **eRF1/3:** H.Song (Institute of Molecular and Cell Biology, Singapore)
- **Magnetic NPs:** L. Bronstein (Indiana University), E.Shtykova (ICRAS Moscow)
- **Chitinase:** W. Rypniewski (Poznan University),
- **Quantum dots:** L. Bronstein (Indiana University)
- **M1 protein:** E.Shtykova (ICRAS Moscow) L.Baratova (Moscow University)
- **Arcp1:** C.Koehler, D.Suck (EMBL, Heidelberg)
- **Gold NPs:** F.Stellacci (EPFL, Lausanne)