SANS for Biology
A Tutorial

William T. Heller, Ph.D.

SNS-HFIR Users Meeting

October 10, 2007

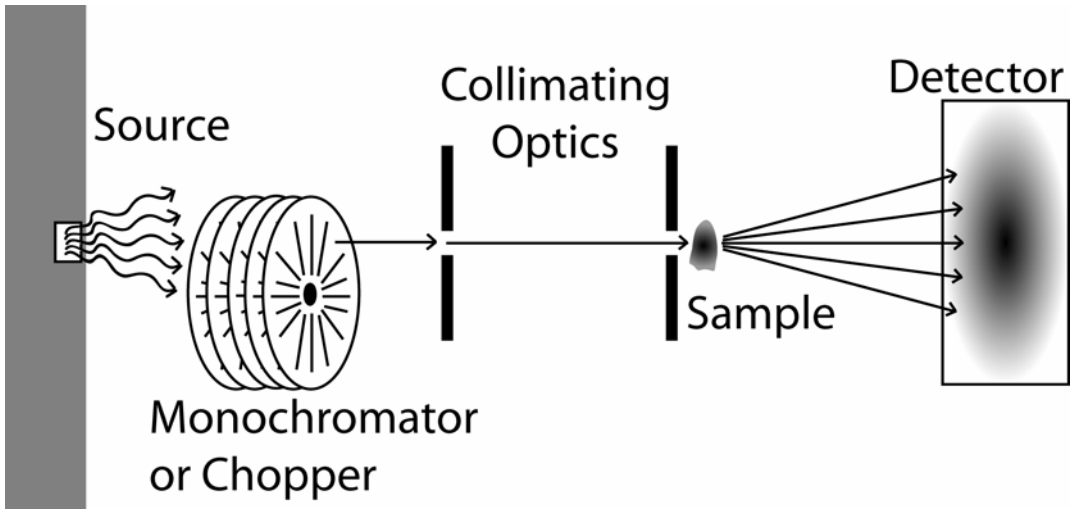
SANS is a powerful tool for studying Bio-macromolecular complexes

- Samples a large range of length scales
  - 1-100 nm
- Samples are in dilute solution
- Directly complementary to SAXS
- Complements high-resolution structural techniques
  - Crystallography
  - NMR
  - Particularly powerful for multi-component complexes and dynamic systems

OAK RIDGE National Laboratory

Center for Structural Molecular Biology · CSMB · Oak Ridge National Laboratory

SANS Instruments



Source: A spallation source or a reactor

Monochromator/Chopper: Defines wavelength(s)

Collimating Optics: Defines the angular divergence of the beam

Determines the maximum size probed

Detector: Collects the neutrons scattered by the sample

SNS and HFIR have large area detectors

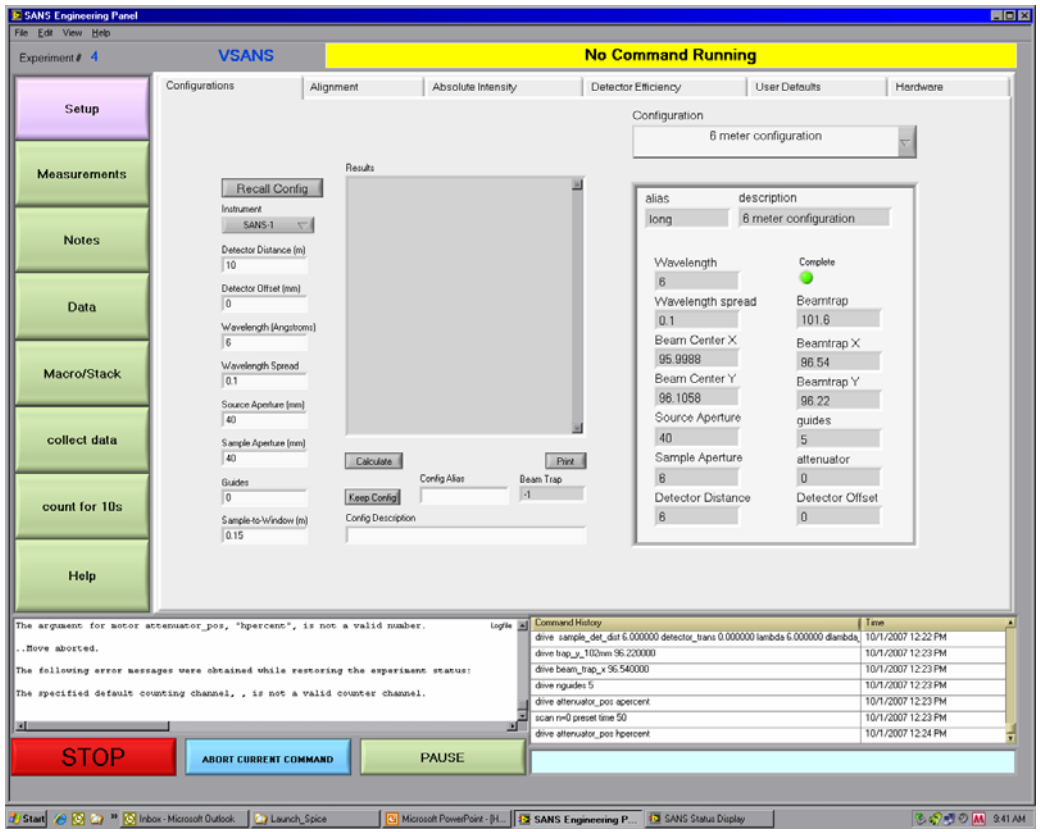
OAK RIDGE National Laboratory

Center for Structural Molecular Biology · CSMB · Oak Ridge National Laboratory

SANS Instruments

The HFIR SANS instrument control program is based on Spice, originally developed for the HFIR Triple Axis Instruments (Robertson, Lumsden and Yethiraj)

SANS data reduction will be provided as a suite of IGOR routines



OAK RIDGE National Laboratory

Center for Structural Molecular Biology · CSMB · Oak Ridge National Laboratory

SANS Basics

SANS probes differences in scattering length density within a sample

The measurement probes the time and ensemble average

For a dilute solution of proteins, the data is azimuthally isotropic

$$I(q) = \left| \left\langle \int_V (\rho(\vec{r}) - \rho_s) e^{-i\vec{q}\bullet\vec{r}} d^3 r \right\rangle \right|^2$$

$$\left|\vec{q}\right| = (4\pi \sin\theta) / \lambda$$

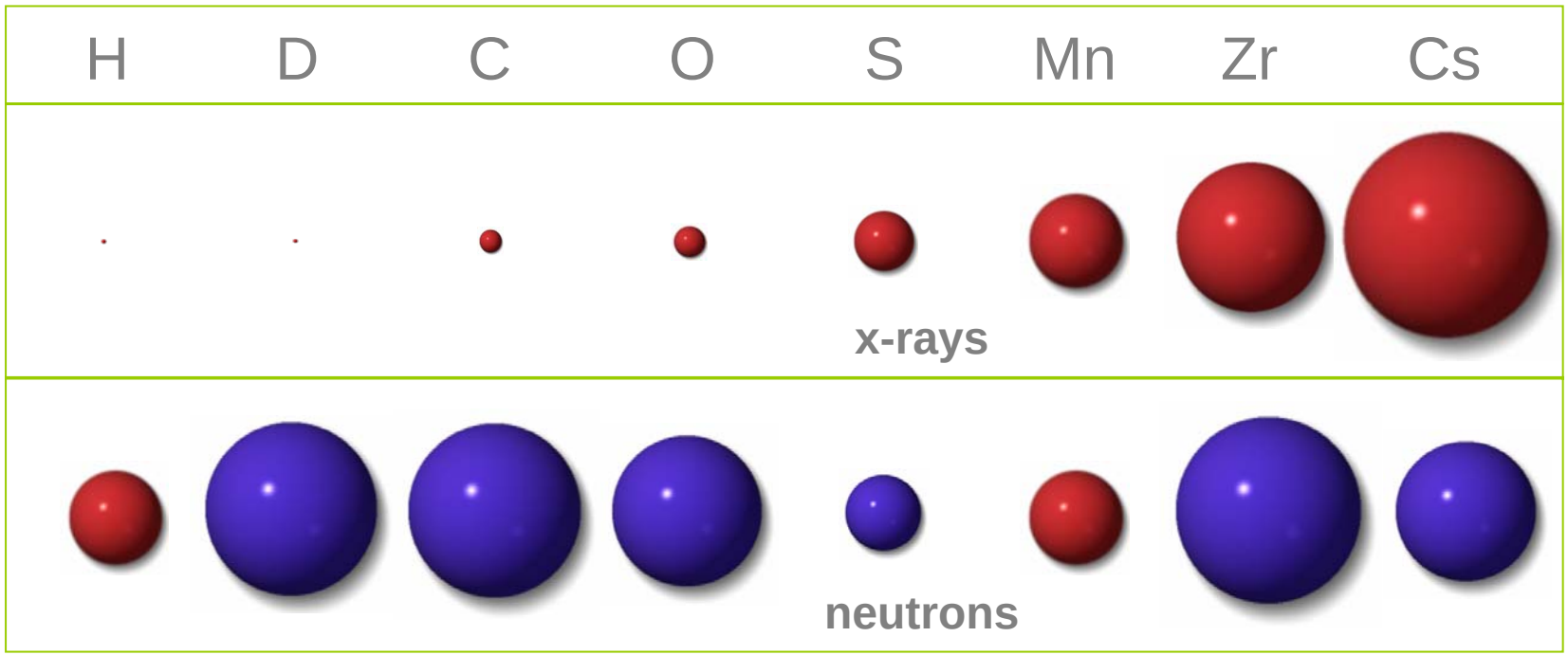
OAK RIDGE National Laboratory

Center for Structural Molecular Biology · CSMB · Oak Ridge National Laboratory

SANS Basics

X-ray and neutron scattering are essentially the same, except...

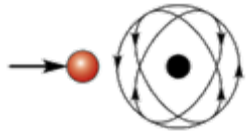
- X-rays scatter from electrons
- Neutrons scatter from nuclei



OAK RIDGE National Laboratory

Center for Structural Molecular Biology
CSMB
Oak Ridge National Laboratory

SANS Basics

Neutrons see nuclei and distinguish between isotopes 

	C	N	O	H	D
b_{coh} (fm)	+6.65	+9.36	+5.81	-3.74	+6.67
σ_{coh} (barns)	5.56	11.03	4.23	1.76	5.59
σ_{inc} (barns)	0	0.49	0	80.27	2.05

Large difference in the cross-section among isotopes that are important to Biology!

Using SANS with contrast variation (H_2O/D_2O mixtures) and selective deuterium labeling of different subunits in a complex, it is possible to determine the low-resolution structures of the subunits within the complex and the structure of the complex

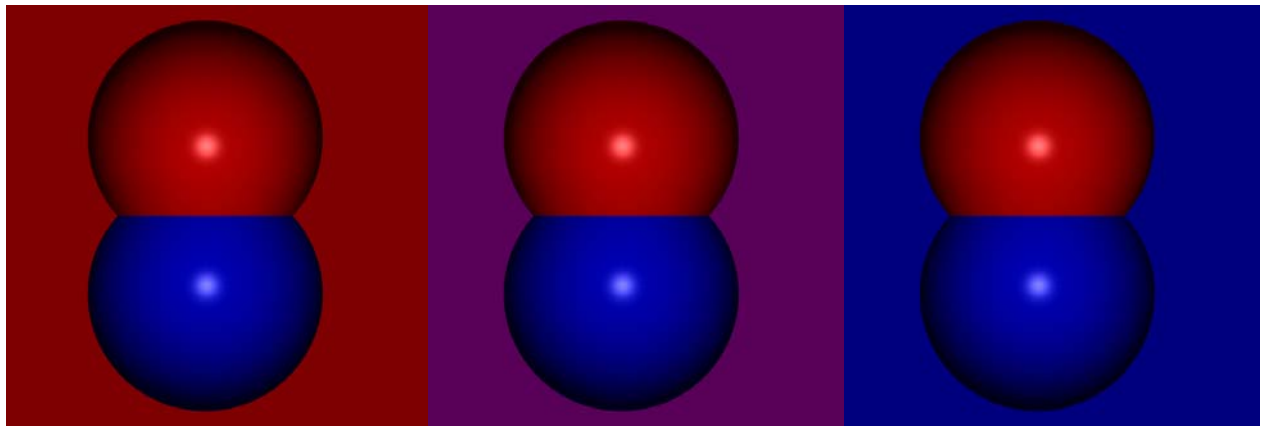
OAK RIDGE National Laboratory

Center for Structural Molecular Biology · CSMB · Oak Ridge National Laboratory

SANS Basics

One way to think about contrast variation of selectively labeled complexes is that you are changing the “color” of the background

This differentially highlights components of a multi-component system relative to the solvent and the other subunits



$$I(q,\Delta\rho_1,\Delta\rho_2) = \Delta\rho_1^2 I_1(q) + \Delta\rho_2^2 I_2(q) + \Delta\rho_1\Delta\rho_2 I_{12}(q)$$

$I_1(q)$, $I_2(q)$ and the cross-term $I_{12}(q)$ are “Basic Scattering Functions”

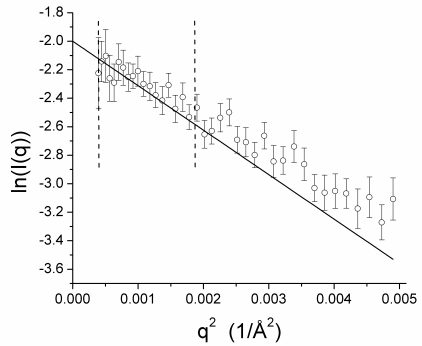
OAK RIDGE National Laboratory

CSMB Center for Structural Molecular Biology Oak Ridge National Laboratory

SANS Data Analysis

- Guinier Analysis for Radius of Gyration R_g

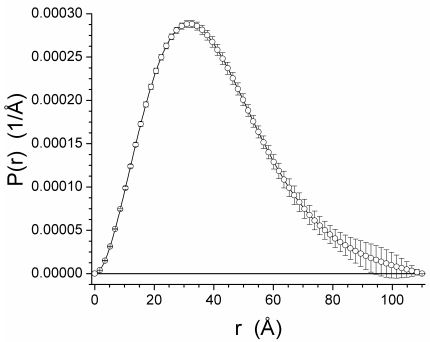
$$\ln\left(I\left(q\right)\right) = \ln\left(I\left(0\right)\right) - \frac{q^2 R_g^2}{3}$$



Fitting can also be done for radius of gyration of cross-section, R_c, and thickness, R_t

- Data Fitting for *P(r)*

$$P(r) = \frac{1}{2\pi^2}\int_0^\infty dq \cdot (qr) \cdot I(q)\sin(qr)$$



The program GNOM from D. Svergun's group at EMBL is the most popular software for *P(r)* fitting

OAK RIDGE National Laboratory

Center for Structural Molecular Biology • CSMB • Oak Ridge National Laboratory

SANS Data Modeling

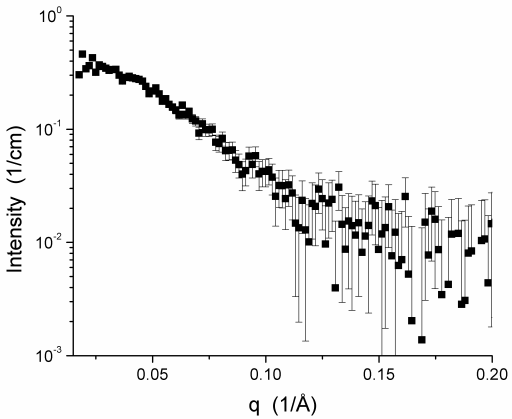
Guinier and P(r) analyses are:

sensitive to changes in conformation

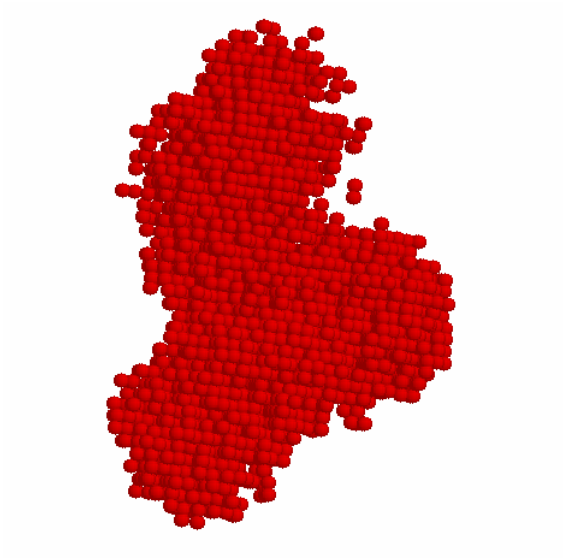
tools for determining quality of samples

Structural Biology is a very visual science

Turn Data



Into Structure



OAK RIDGE National Laboratory

Center for Structural Molecular Biology · CSMB · Oak Ridge National Laboratory

SANS Data Modeling

What kinds of modeling are available?

…and some authors…

- Spherical Harmonics
  - Svergun, Stuhrmann, Grossman, etc.
- Aggregates of Spheres
  - Svergun, Doniach, Chacón, Heller, etc.
- Sets of High-resolution Structures
  - Svergun, Heller, Grishaev, Gabel, etc.
- Simple Shapes and Custom Approaches for Specific Problems
  - Henderson, Zhao, Gregurick, Heller, etc.

The problem dictates the most appropriate modeling method

Creativity and time are the only limits

...programming ability doesn't hurt, either

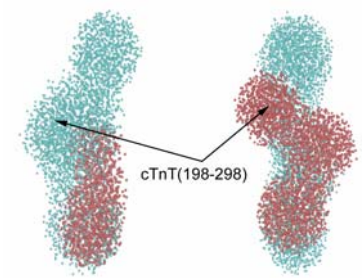
OAK RIDGE National Laboratory

Center for Structural Molecular Biology · CSMB · Oak Ridge National Laboratory

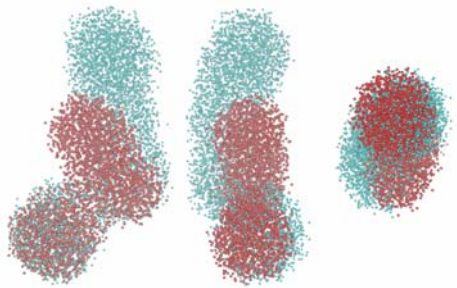
SANS Data Modeling

Spherical Harmonic, Aggregate and Simple Shape Methods

- No initial structural data exists of complete system
- Investigating conformational transitions

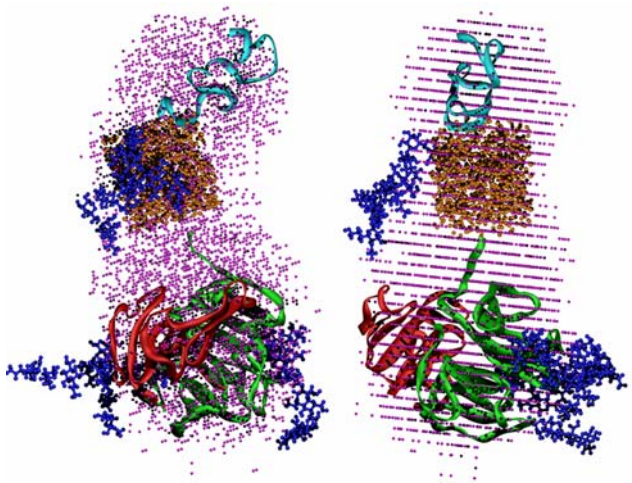


The location of Troponin T(198-298) was inferred using SANS with contrast variation



Bis-phosphorylation of Troponin I produces a bend in the structure.

Heller, W. T., et al. (2003)
Biochemistry **42**: 7790-7800.



Models of human plasma vitronectin produced from SAXS data help visualize global conformation when only fragmentary structural information exists

Lynn, G. W., et al. (2005)
Biochemistry **44**: 565-574.

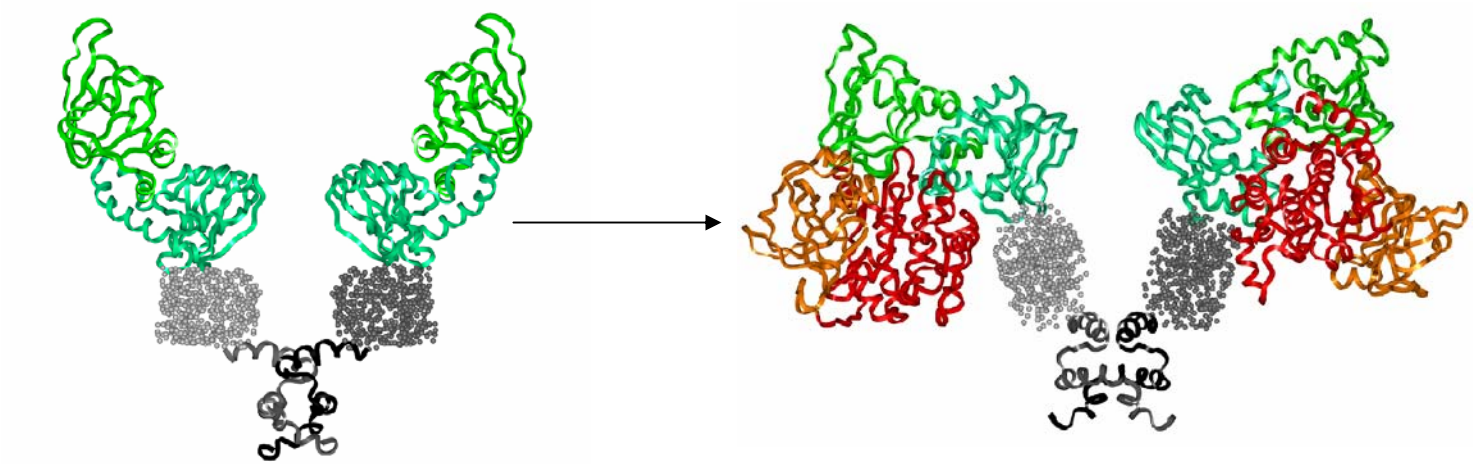
OAK RIDGE National Laboratory

Center for Structural Molecular Biology • CSMB • Oak Ridge National Laboratory

SANS Data Modeling

Modeling Building from Sets of High-resolution structures

- Structures of subunits of a complex have been determined by crystallography or NMR
- Mechanisms exist for filling in missing sequences



The conformation of the regulatory dimer of protein kinase A changes in response to binding of the catalytic subunits

Heller, W. T., et al. (2004) *J. Biol. Chem.* **279**: 19084-19090.

Vigil, D., et al. (2004) *J. Mol. Biol.* **337**: 1183-1194.

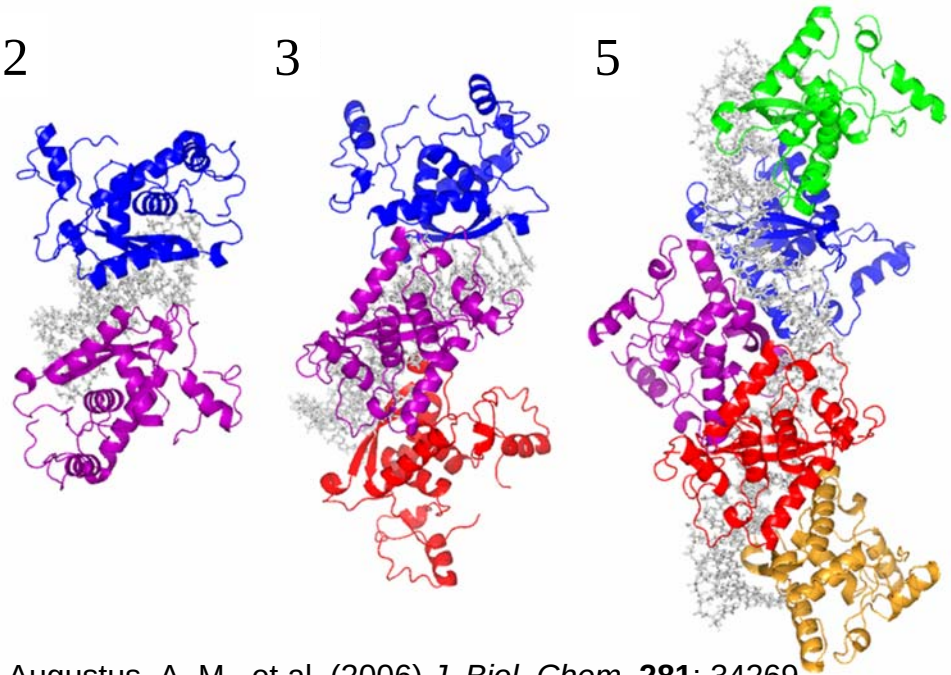
The goal is to construct biologically relevant models using scattering data

Methionine Repressor Protein (MetJ)-DNA Complexes: Influence of DNA Length on the Complex Structure

The methionine repressor protein controls the expression of serveral genes in the *met* regulon of *E. coli* responsible for the biosynthesis of methionine (Leonard D. Spicer, Duke Univ.)

SANS with contrast variation was used to construct models of MetJ in complex with different lengths of DNA

- **2 "consensus" operators (2 & 3 MetJ binding motifs)**
- **1 native operator (5 MetJ binding motifs)**



Augustus, A. M., et al. (2006) *J. Biol. Chem.* **281**: 34269

The MetJ protein packs very differently on the longest sequence of DNA used (the native operator)

The study supports DNA length-dependent differences observed in *met* regulation

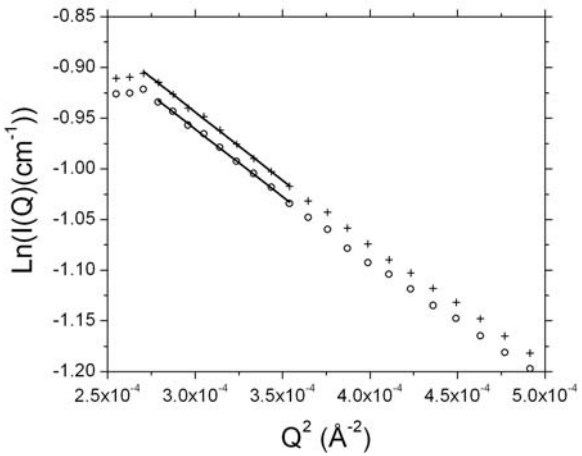
OAK RIDGE National Laboratory

Center for Structural Molecular Biology • CSMB • Oak Ridge National Laboratory

Photosystem I in Detergent Solution

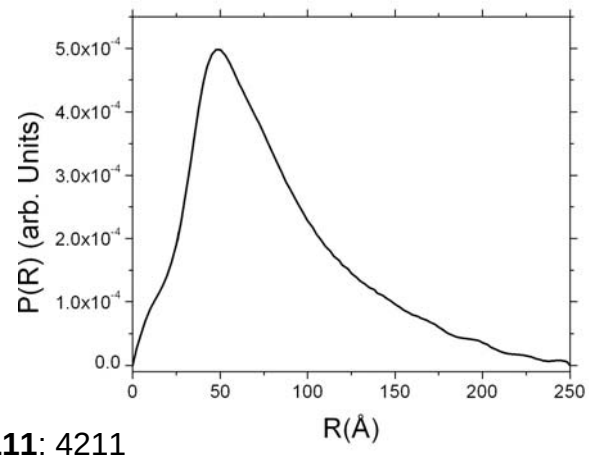
- Very large, multisubunit complex
- Responsible for trans-membrane electron transfer

Investigate the solution structure of the Triton X-100 solubilized complex



The data suggest a large, monodisperse scattering particle

- Linear Guinier region
- well-defined P(r)



Use modeling to understand the structure

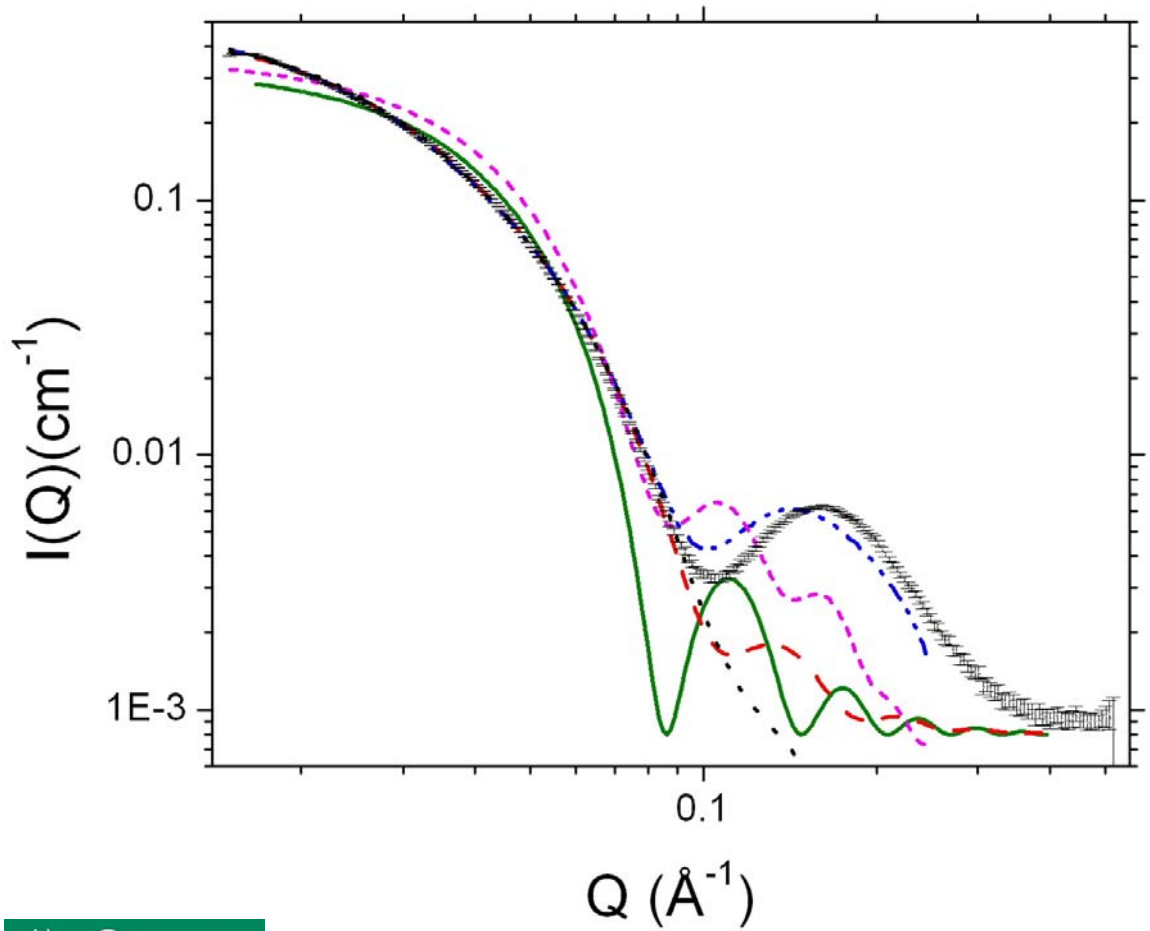
O'Neill, H. M., et al (2007) *J. Phys. Chem B* **111**: 4211

OAK RIDGE National Laboratory

CSMB Center for Structural Molecular Biology Oak Ridge National Laboratory

Photosystem I in Detergent Solution

...maybe it isn't so simple



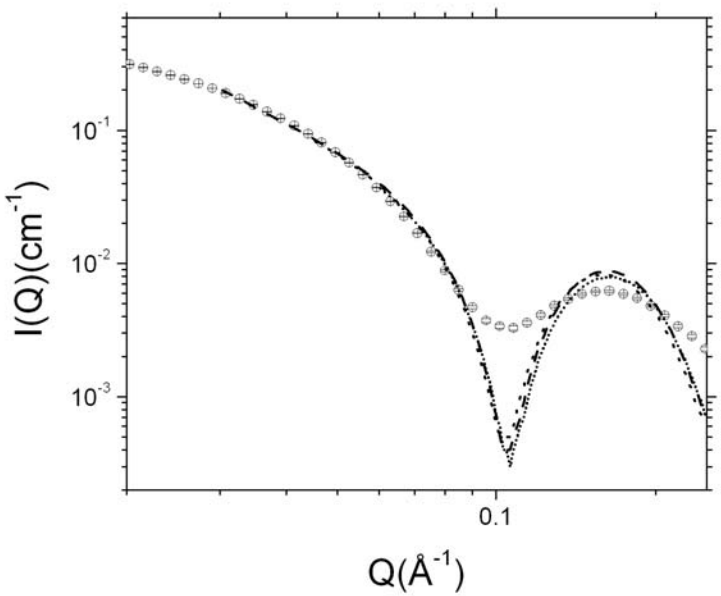
- ✓Sphere (—)
- ✓Ellipsoid of revolution (– –)
- ✓Triaxial Ellipsoid (····)
- ✓Core-Shell Triaxial Ellipsoid (-··)
- ✓Crystal structure in a detergent disk (----)

None of the models reproduce the position of the broad peak at wide angle

O'Neill, H. M., et al (2007) *J. Phys. Chem B* **111**: 4211

Photosystem I in Detergent Solution

What if the complex is denatured?



Using a “beads-on-a-string” model, the position of the peak can be reproduced

The low-q region is less well-fit

The data suggest some polydispersity in the micelle size

Some components of PS-I, which remains photoactive, are likely to be correctly folded

When modeling, one needs to remain open to the possibility that the system is doing something unexpected

O’Neill, H. M., et al (2007) *J. Phys. Chem B* **111**: 4211

Center for Structural Molecular Biology

William T. Heller: hellerwt@ornl.gov

Volker S. Urban: urbanvs@ornl.gov

Gary W. Lynn: lynngw@ornl.gov

Dean A. Myles: mylesda@ornl.gov

This work was supported by the Office of Biological and Environmental Research project KP1102010 of the U. S. Department of Energy, under contract No. DE-AC05-00OR22725 with Oak Ridge National Laboratory, managed and operated by UT-Batelle, LLC. The submitted manuscript has been authored by a contractor of the U.S. Government under Contract DE-AC05-00OR22725. Accordingly, the U.S. Government retains a nonexclusive royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

OAK RIDGE National Laboratory