Next generation
X-ray Solution Scattering

Lee Makowski
Northeastern University
**WAXS Data Set from Hb – 150 mg/ml**

SAXS is more intense than WAXS

SAXS can be used to determine Rg; shape; oligomeric state… (see Dmitri talk!)

WAXS cannot be used to extend the resolution of shape determinations nor improve the estimate of Rg…
Wide-angle x-ray solution scattering

most future applications will require measurement of small (and in some cases - very small) differences in intensity

critical features for advances:

very high stability
strong signal (but not so much intensity as to FRY the samples)
minimal background
accounting for all features of the data/background
Wide-angle x-ray solution scattering

test molecular models
detect/characterize structural changes
characterize ensembles
MADMAX - anomalous WAXS (atomic positions)
characterize structural intermediates

membrane proteins
TR-WAXS (see Philip's talk)
greatly enhanced by computational methods
Solution scattering pattern from a protein is:
accurately predicted by atomic coordinates of protein
closely related to the distribution of interatomic distances in the protein
inadequate in information content to completely specify the structure of the protein

scattering pattern

spherically averaged autocorrelation function = histogram of lengths of interatomic vectors

atomic coordinates

pair distribution function
Success of this approach also provides strong evidence that MD approaches are getting water of hydration correct.

discrepancies are suggestive of small structural fluctuations of protein in solution.
Ensembles…

sets of accessible conformations accessible from one another through structural fluctuations

modulated by ligands, environment…

increasingly implicated in protein function

give rise to predictable changes in solution scattering
WAXS studies of protein ensemble – how does polymorphism affect the scattering pattern?

scattering from spheres of 14; 15 and 16 Å radius minima at ~ 1/(radius)

so…the broader the ensemble; the greater the effect

WAXS is highly sensitive to this effect…
... can characterize an ensemble by exhaustive enumeration...
Representatives of families of conformations abundant under different conditions

Catalytic domain – blue
SH2 domain green
SH3 domain yellow

peptides modulate the relative abundances of conformations in hck Kinase - shifting equilibrium from largely inactive (states 1-4) to largely active conformations (5-9)

Hck-YEEI – high-affinity mutant
Multidomain assembled states of Hck tyrosine kinase in solution
Sichun Yang*, Lydia Blachowicz*, Lee Makowskise, and Benoit Roux*••

Representatives of families of conformations abundant under different conditions

Catalytic domain – blue
SH2 domain green
SH3 domain yellow

Hck-YEEI – high-affinity mutant
… but for many applications an exhaustive enumeration is not necessary, possible, or appropriate to the question being asked
Fluctuations are expressed in the scattering pattern through the variation of interatomic vector lengths.

Interatomic vectors of a certain length are replaced by a distribution of vector lengths...
WAXS pattern always gives you a pair-distribution function $p(r)$; 
A simple model for increased fluctuations is to replace every 
interatomic distance in the protein by a distribution of distances

$$F(r) = \exp\left(-\frac{(r-r_\circ)^2}{2\sigma}\right)$$

where $\sigma$ is a function of vector length:

$$\sigma = ar^n$$
Ubiquitin

Is very rigid in aqueous solution

Calculated ~ observed

Discrepancy between calculated and observed due largely to fluctuations in the structure in solution

With Tobin Sosnick…
Effect of single amino acid substitution:

Change L50 to E…

What is effect?

all features of the pattern are severely muted

$$\sigma \sim 0.7-0.8^*r^{0.5}$$

For $$r = 10$$ … $$\sigma \sim 2.1$$ Å – 20% - so denaturing…

4.7 Å peak decreases more rapidly in model than in data… as might be expected for a global model

With Tobin Sosnick…
**MADMAX** - multi-wavelength anomalous diffraction using medium angle x-ray solution scattering

\[ I(k) = \sum_i \sum_j f_i f_j^* \left( \frac{\sin(\mathbf{k} \cdot \mathbf{r}_{ij})}{\mathbf{k} \cdot \mathbf{r}_{ij}} \right) \quad \text{[Debye formula]} \]

\[ f = f^o + f' + if'' \]

\[ I(k) = \sum_i \sum_j \left[ f_i^o f_j^o + 2 f_i^o f_j' + f_i' f_j' + f_i'' f_j'' \right] \left( \frac{\sin(\mathbf{k} \cdot \mathbf{r}_{ij})}{\mathbf{k} \cdot \mathbf{r}_{ij}} \right) \]

\[ \Delta I(k) = I_e(k) - I_o(k) \]

\[ = 2 \sum_i \sum_j f_i^o f_j' \left( \frac{\sin(\mathbf{k} \cdot \mathbf{r}_{ij})}{\mathbf{k} \cdot \mathbf{r}_{ij}} \right) + \sum_i \sum_j \left[ f_i' f_j' + f_i'' f_j'' \right] \left( \frac{\sin(\mathbf{k} \cdot \mathbf{r}_{ij})}{\mathbf{k} \cdot \mathbf{r}_{ij}} \right) . \]
**MADMAX** - multi-wavelength anomalous diffraction using medium angle x-ray solution scattering

\[ I(k) = \sum_i \sum_j f_i f_j^* \left( \frac{\sin(kr_{ij})}{kr_{ij}} \right) \]

\( f = f^o + f^' + if^" \)

\[ I(k) = \sum_i \sum_j \left[ f_i^o f_j^o + 2 f_i^o f_j^' + f_i^' f_j^' + f_i^" f_j^" \right] \left( \frac{\sin(kr_{ij})}{kr_{ij}} \right). \]

\[ \Delta I(k) = I_e(k) - I_o(k) \]

\[ = 2 \sum_i \sum_j f_i^o f_j^' \left( \frac{\sin(kr_{ij})}{kr_{ij}} \right) + \sum_i \sum_j \left[ f_i^' f_j^' + f_i^" f_j^" \right] \left( \frac{\sin(kr_{ij})}{kr_{ij}} \right). \]

*cross terms* \hspace{2cm} *pure anomalous terms*
MADMAX - multi-wavelength anomalous diffraction using medium angle x-ray solution scattering

can determine distance between anomalous scatterer and center of mass of the protein

similar to not-so-heavy atom derivative

~ 6 electrons

collect data at absorption edge and remote from edge
Myoglobin - calculated vs observed
Hemoglobin - calculated vs observed
what about fluctuations that occur during function?

can intermediate states be observed?
Catalyzes the interconversion of AMP; ADP and ATP

\[ 2 \text{ADP} \rightleftharpoons \text{AMP} + \text{ATP} \]

Flaps cover the reagents once in the active site

Form of scattering suggests flaps very flexible in unliganded form
Adenylate kinase +/- inhibitor … and during catalysis

During catalysis closely resembles inhibited form

but cannot be constructed as linear combination of the two endpoints

Reduced chi-squares
apo – inhib 42.75
apo-adp 6.58
Inhib-adp 5.23
Cycling Adenylate Kinase

Can we see what some of these states look like?

Aden and Wolf-Watz, 2007
Singular Value Decomposition

data from \( \sim 100 \) conditions

SVD suggests 4 basis vectors can be used to reconstruct all data sets

suggests that end points plus at least 2 intermediate states are sufficiently abundant to give rise to observable scattering…
the future...

> small differences critical
> high signal
> low noise
> background must be completely accounted for
> high stability - need to be able to scale hundreds of patterns taken sequentially
Thanks

Bob Fischetti
Dave Gore
Diane Rodi
Suneeta Mandava

Sanghyun Park (ANL)  Chien Ho (CMU)
Jaydeep Bardhan (Rush)  George Phillips (UW)
David Minh (ANL)  Mike Daily (UW)
Jyotsana Lal (ANL)  Quiang Cui (UW)
Sichun Yang (Case)  
Benoit Roux (UCh)  Steve Kent (U Ch)
Tobin Sosnick (UCh)  Vladimir Torbeev (U Ch)
Karl Freed (UCh)  Celia Schiffer (UMass)
Juoko Virtanen (UCh)  