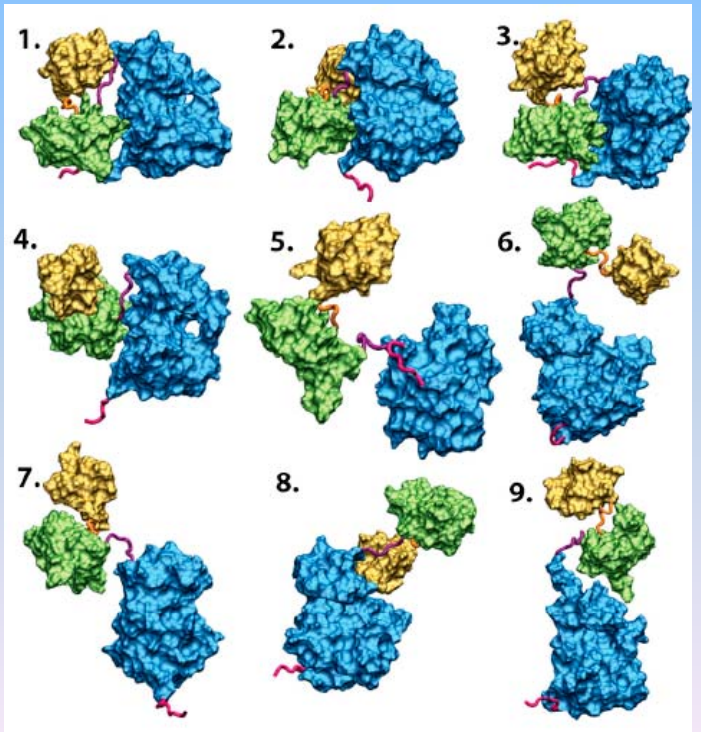
Probing Protein Ensembles and Functional Intermediates with X-ray Solution Scattering



Lee Makowski

Northeastern University

Structure based drug design…

Given the progress in

genomics (we know all the sequences)

structural biology (we know more and more of the structures)

combinatorial/synthetic chemistry (we can make most anything)

docking algorithms (we can predict binding)

Great progress in structure based drug design should be occurring

but the pipeline is not experiencing the rapid growth expected

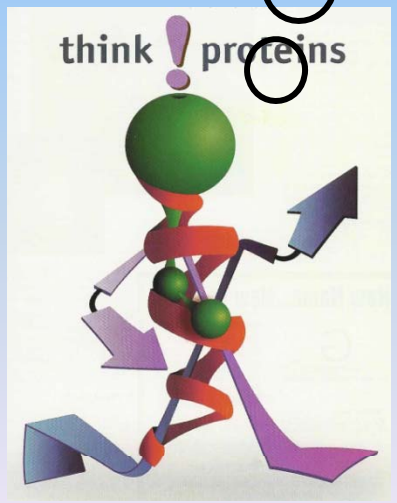
What's wrong? …perhaps…

proteins are dynamic – they occupy structural ensembles

they undergo conformational changes critical to function

Progress in characterization of intermediates and ensembles may prove a critical element in moving drug design forward…

We know that proteins don't work remotely



they don't use cell phones

they have to reach out and touch something

They move…

In response to changes in environment they:

Change shape

Change conformation

Bind a target

Release a target

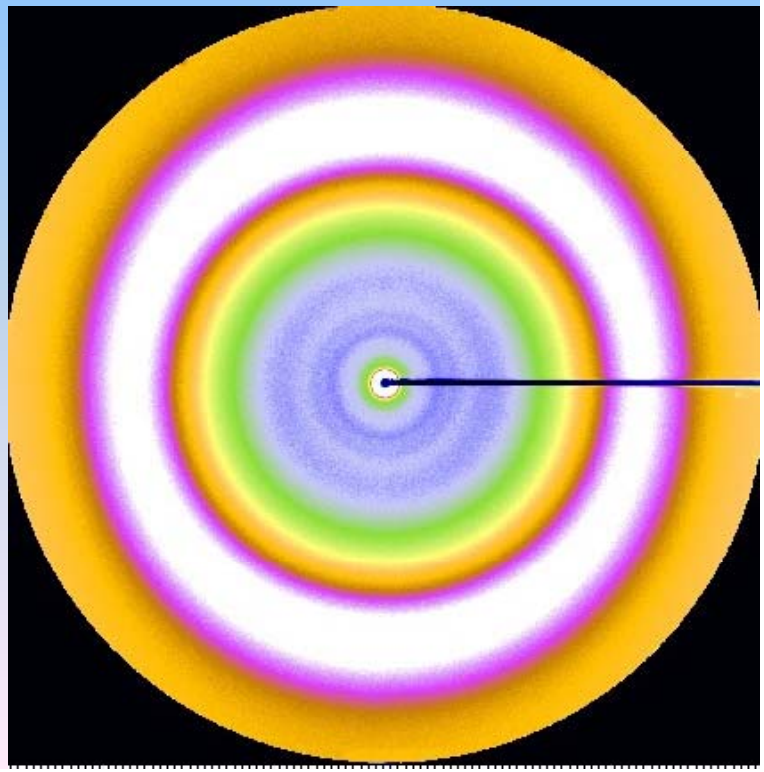
And they always…

Undergo thermal fluctuations

The fluctuations are related to the functional motions…

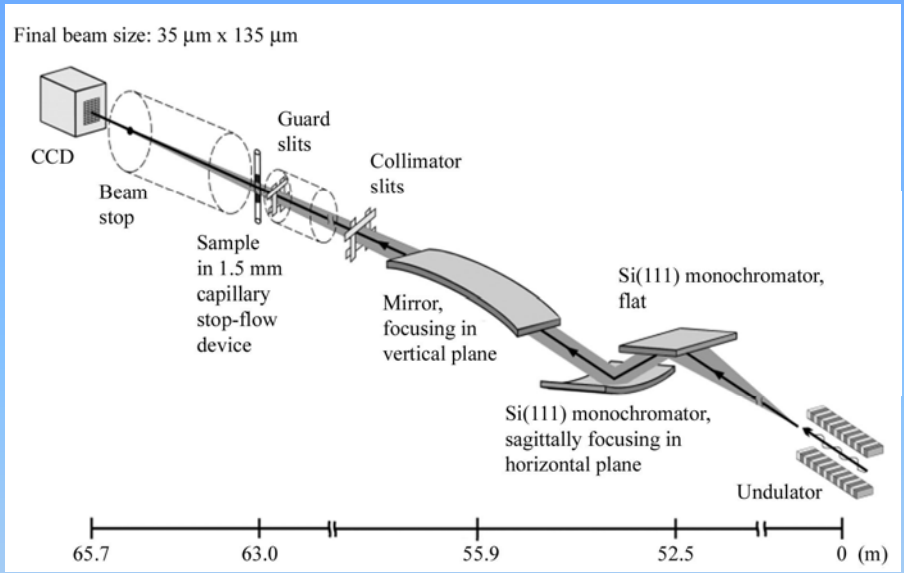
In a crystal they often don't move freely or functionally

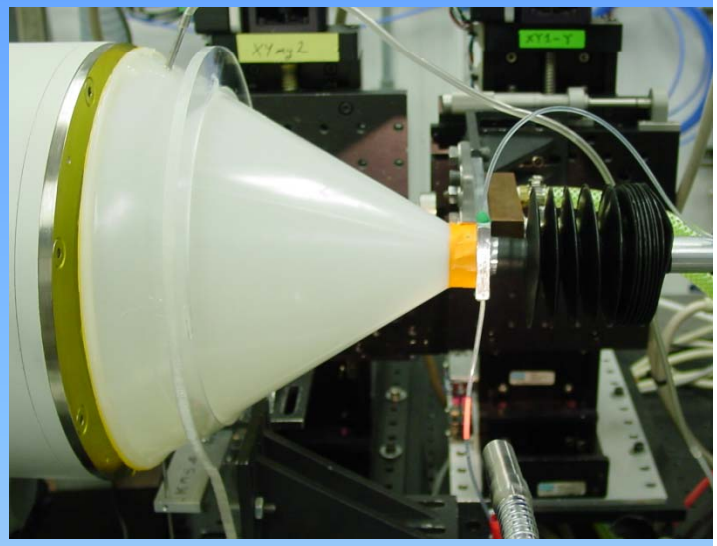
- ***So we study them in solution:***
- ***WAXS (wide angle x-ray solution scattering)***
- *To maximize information collected*
- *just like SAXS (small-angle x-ray solution scattering)*
  - *except we collect data to much higher scattering angles*
  - *the data is much **weaker** (~2 orders of magnitude)*
  - *intense scattering from buffer makes accurate measurement more difficult*
- *WAXS has much less info than crystal patterns, but…*
- *Proteins in solution are much less constrained in motions available to them*
- *WAXS is highly sensitive to structural changes*
- *Can detect changes in breadth and nature of ensembles*

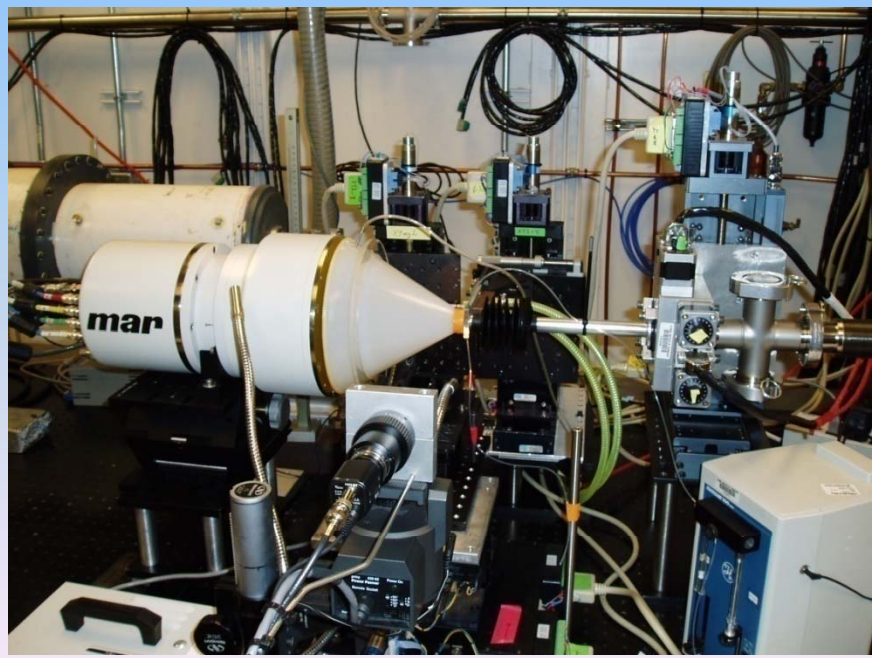


Sector 18 – Advanced Photon Source

SAXS/WAXS Experiment







Flow cell (100 ms x-ray exposure)

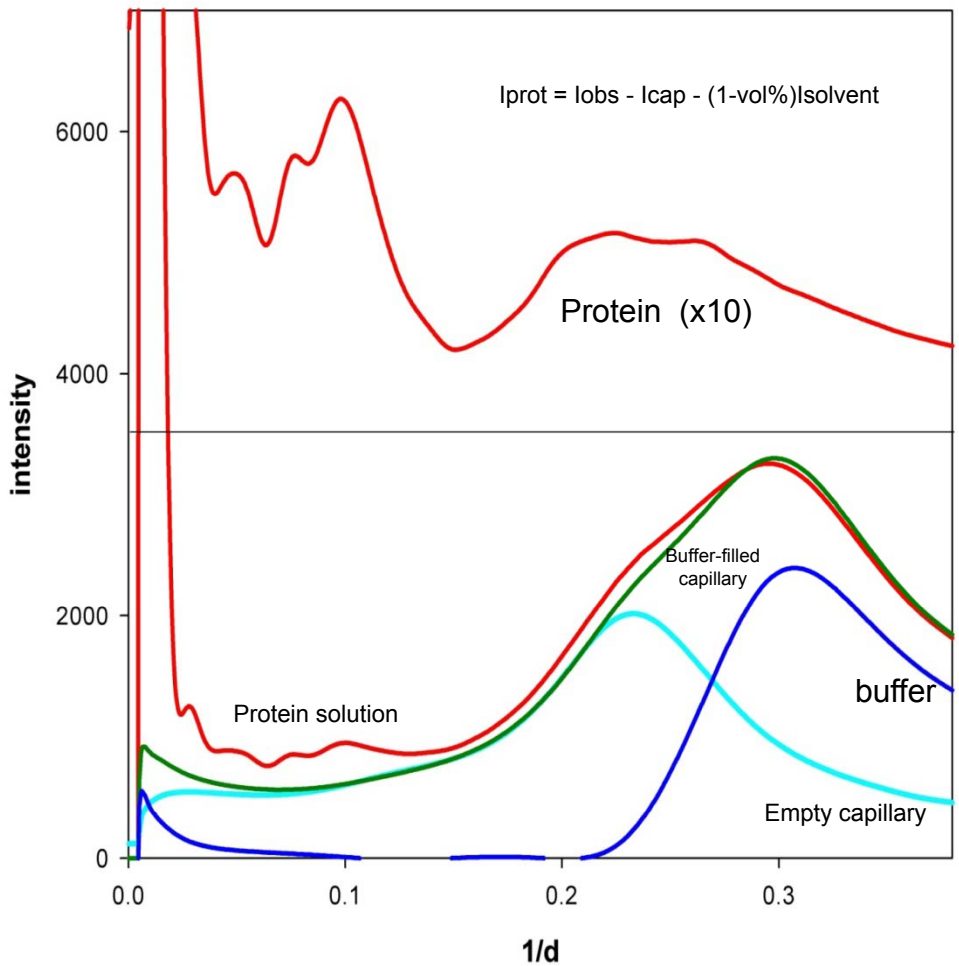
Temperature controlled

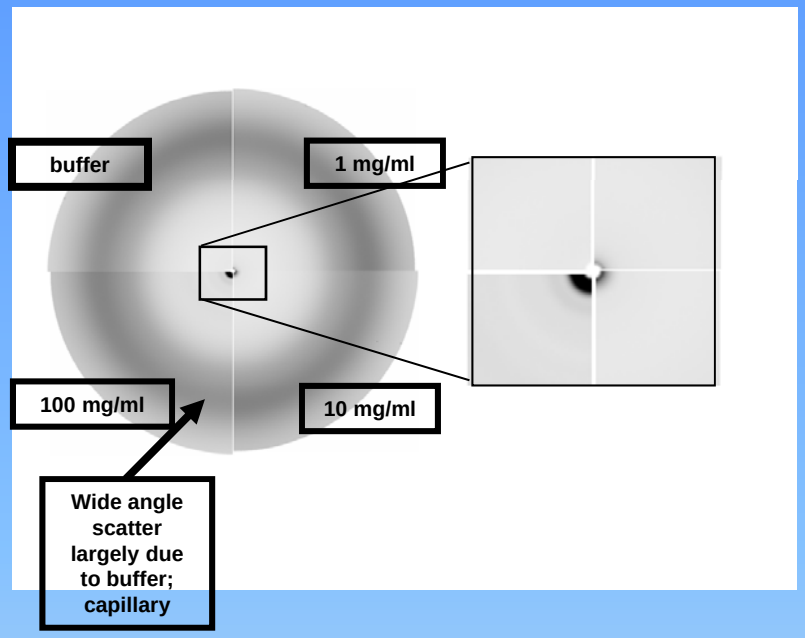
1.5 mm path length (typical)

10 microliter sample volumes possible

≥10 mg/ml concentration preferred

WAXS Data Set from Hb – 150 mg/ml





Each data set is composed of circularly averaged scattering from

(i) Empty capillary

(ii) Buffer-filled capillary

(iii) Protein solution-filled capillary

Nature of the information in WAXS

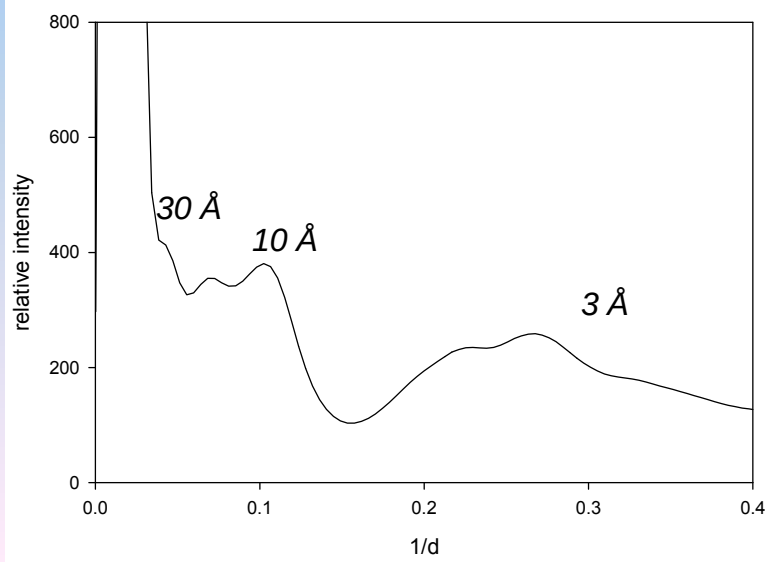
- *What sets WAXS apart from other biophysical approaches is a quantitative relationship to the protein atomic coordinates*
- *Debye Formula – a WAXS pattern can be calculated from the atomic coordinates of the protein:*

$$I(q) = \sum I_i(q) + 2 \sum\sum F_i(q)\, F_j(q)\, (\sin(qr_{ij})/(qr_{ij}))$$

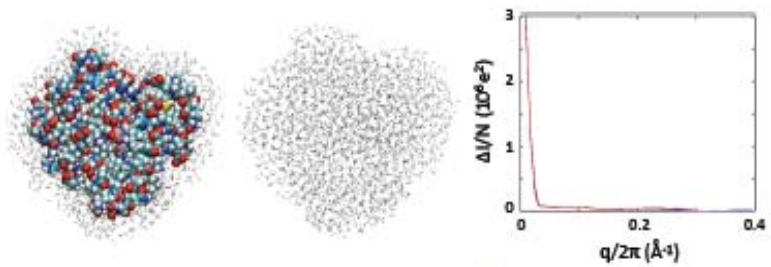
Where sum is over all atoms and r_{ij} are the interatomic vectors

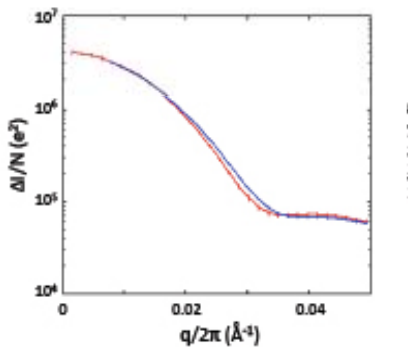
- *Can be used to calculate a <u>pair-correlation function -</u> literally a histogram of the lengths of all interatomic vectors in the protein*

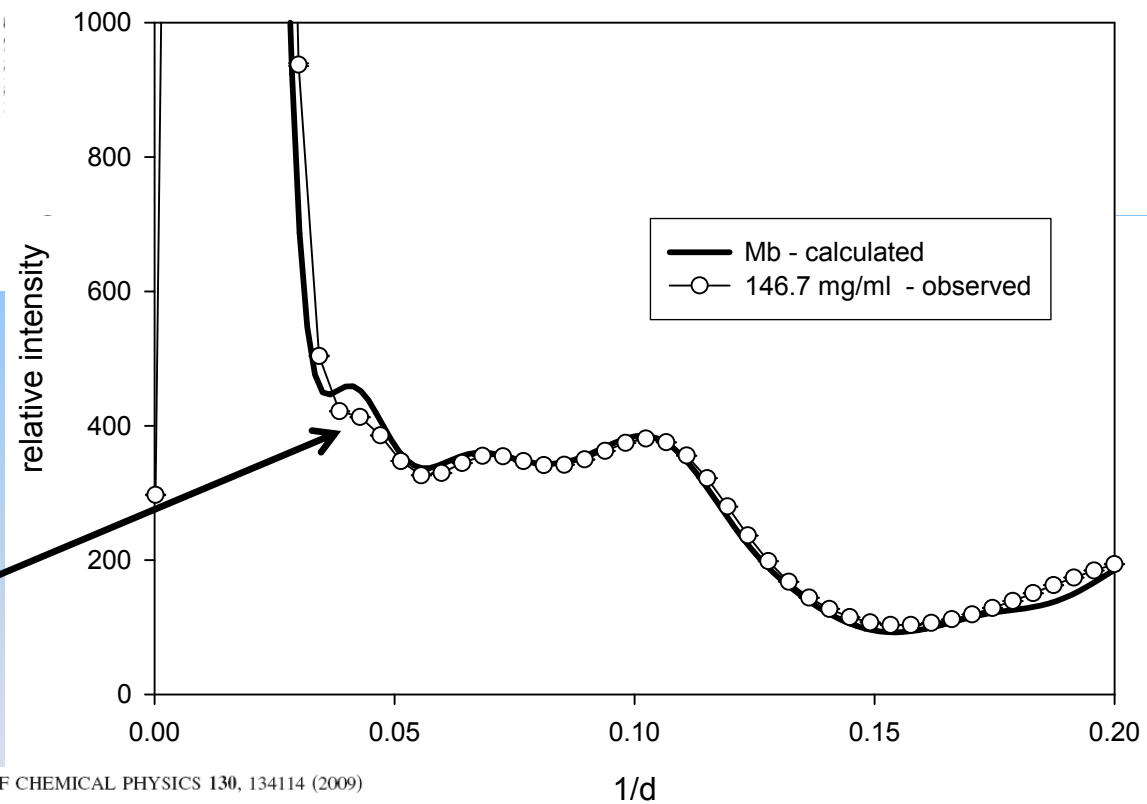
- Size and shape (radius of gyration)
- Tertiary/quaternary structure
- Secondary structure
  - Alpha helices
    - 1.5 Å axial separation of amino acids along helix
    - 5.4 Å pitch
    - 10 Å diameter (center-to-center distance)
  - Beta sheets
    - 4.7 Å strand-to-strand distance
    - 7.0 Å pleat distance (2 residue separation)



Calculate WAXS patterns from atomic coordinate sets is not trivial – water of hydration contributes significantly







Discrepancies; where they exist, often involve experimental data with weakened peaks or filled in troughs.

THE JOURNAL OF CHEMICAL PHYSICS **130**, 134114 (2009)

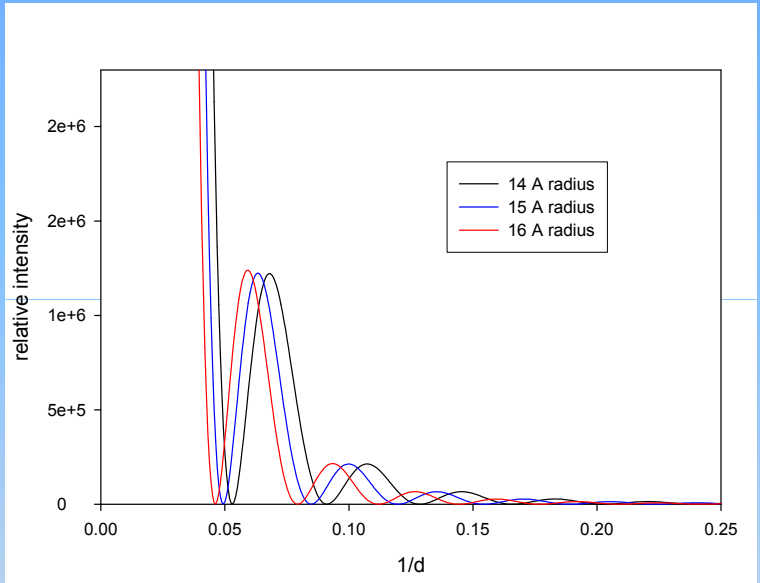
Simulated x-ray scattering of protein solutions using explicit-solvent models

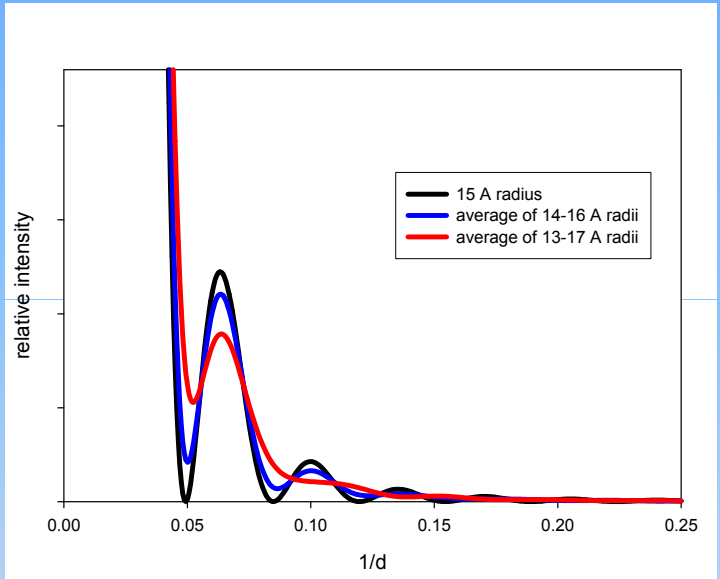
Sanghyun Park,[1,a)] Jaydeep P. Bardhan,[2] Benoît Roux,[2,3] and Lee Makowski[2]
[1]*Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA*
[2]*Biosciences Division, Argonne National Laboratory, Argonne, Illinois 60439, USA*
[3]*Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA*

(Received 29 December 2008; accepted 25 February 2009; published online 7 April 2009)

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

WAXS studies of protein ensemble – how does polymorphism effect the scattering pattern?





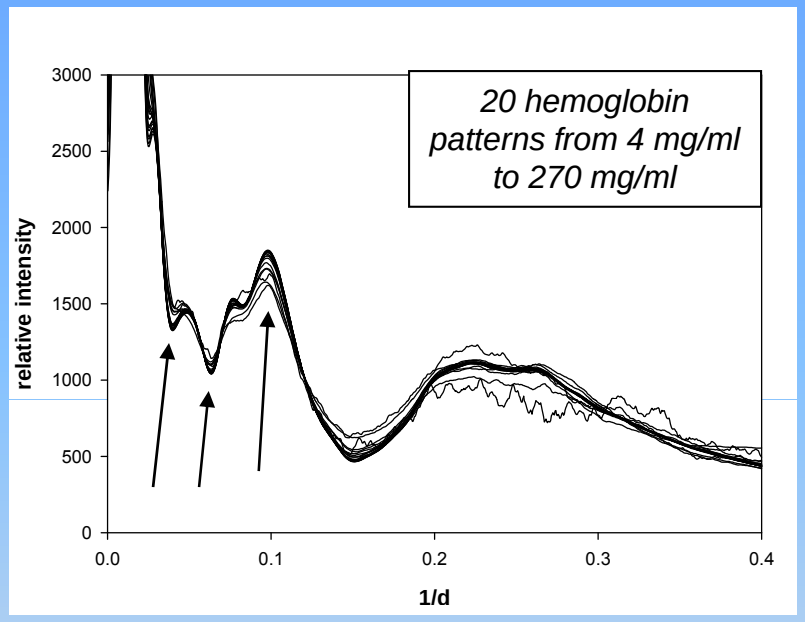
scattering from spheres of 14; 15 and 16 Å radius minima at ~ 1/(radius)

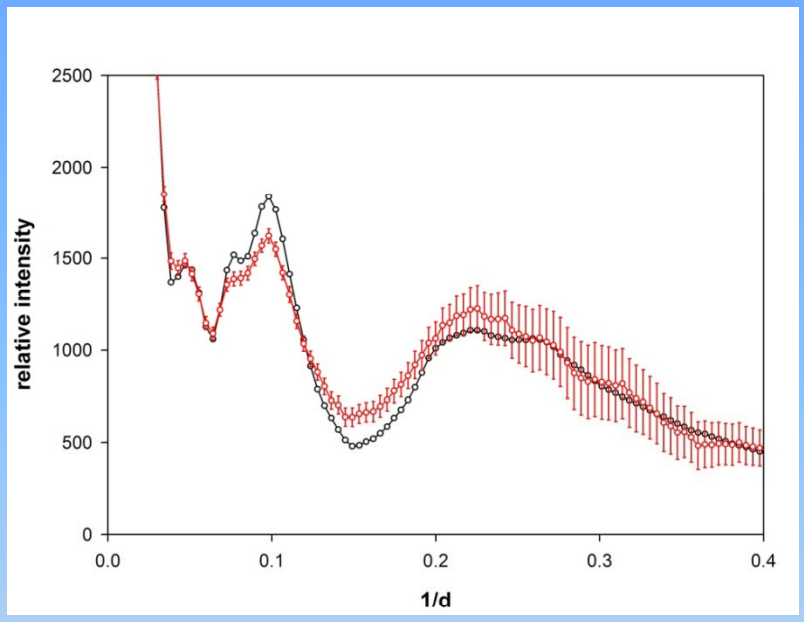
scattering from a solution of all three spheres looks like the average sphere but with minima filled in and maxima muted

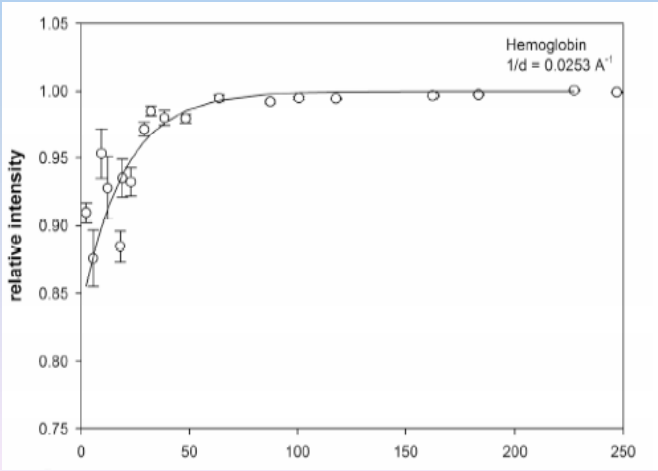
so…the broader the ensemble; the greater the effect

WAXS is highly sensitive to this effect…

Hemoglobin - Ensemble changes as function of protein concentration

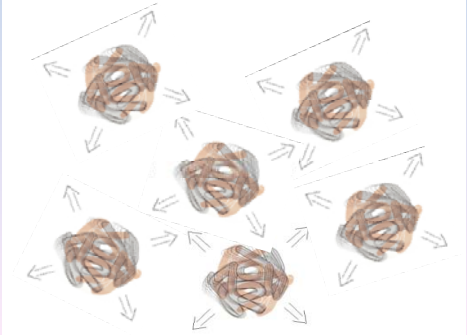


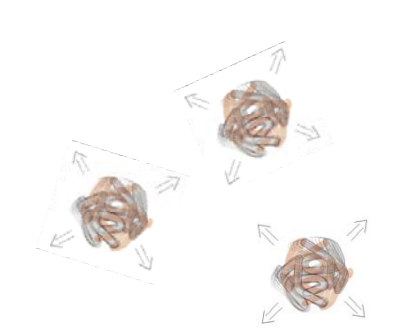




At concentrations below about 50 mg/ml, the scattering from met-Hb indicates a progressive increase in polydispersity - broadening of the structural ensemble.

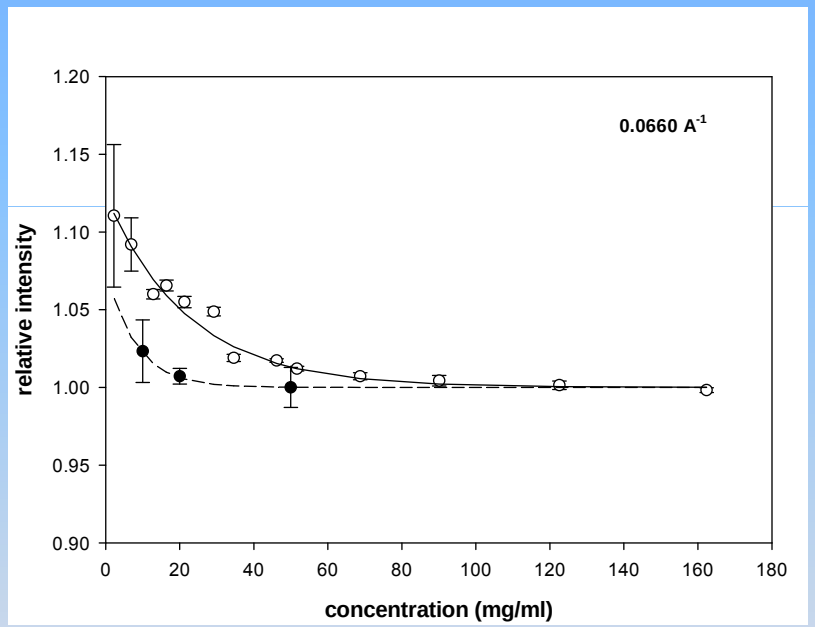
Lack of change in higher angle scatter suggests this is due to rigid body motions

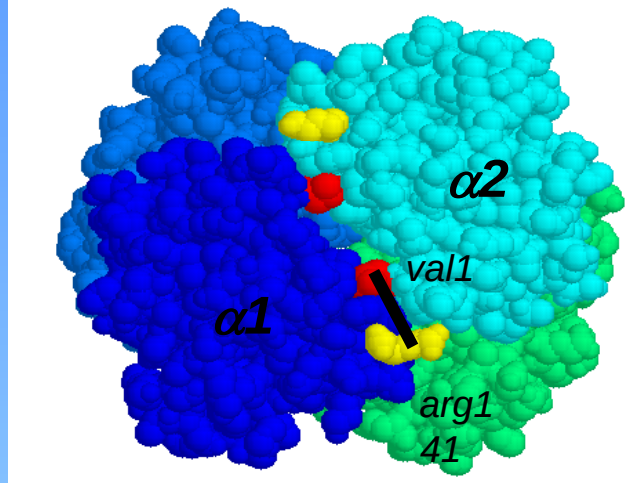


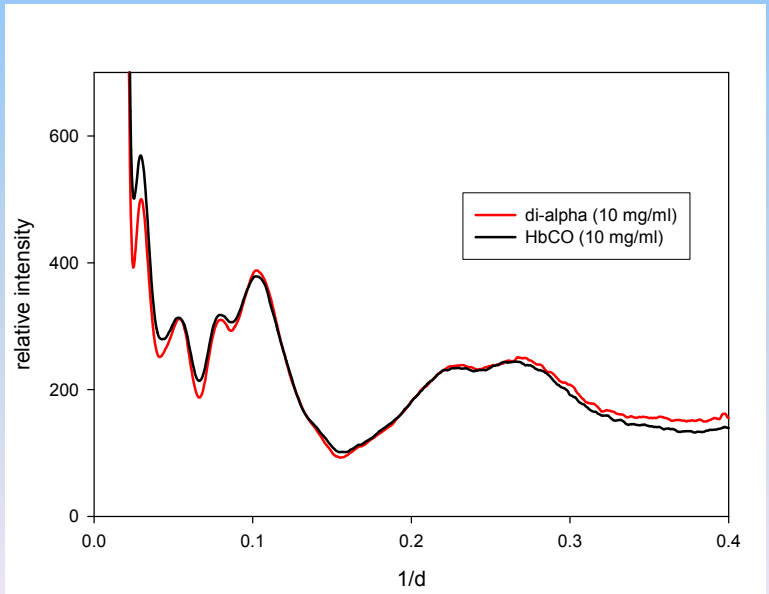


Can decrease motion of subunits -

di-α Hb is a variant in which the two α-chains are covalent linked at their termini







di-α-Hb appears far more rigid than CO-Hb

An engineered variant is far more rigid than the native protein

HIV Protease

Active site protected by two 'flaps'

When inhibitors (or substrate) bind to active site flaps fold down over them

Their flexibility is required for access to active site

Extensive information available on mutants including MDR

All studies use Q7K to prevent self-digestion

Consider two mutants:

T80N - invariant in both treated and untreated populations

G48V L63P - MDR

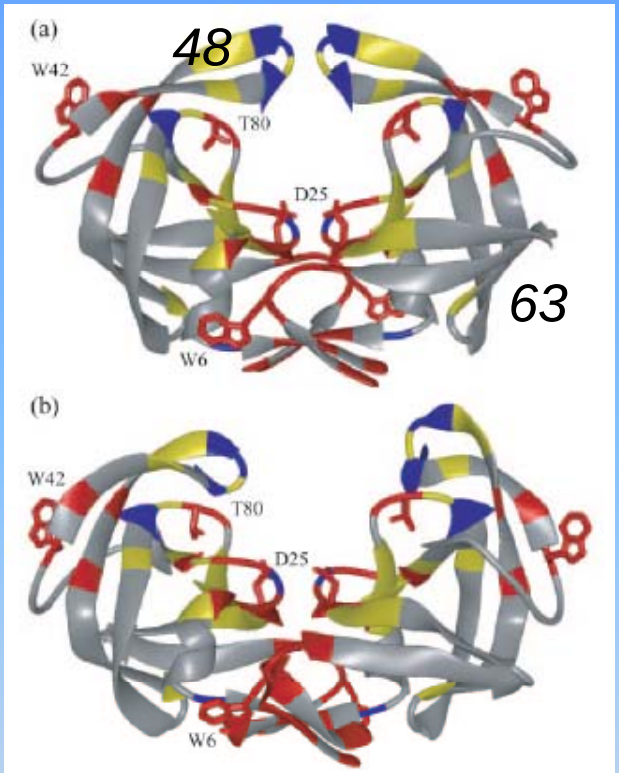
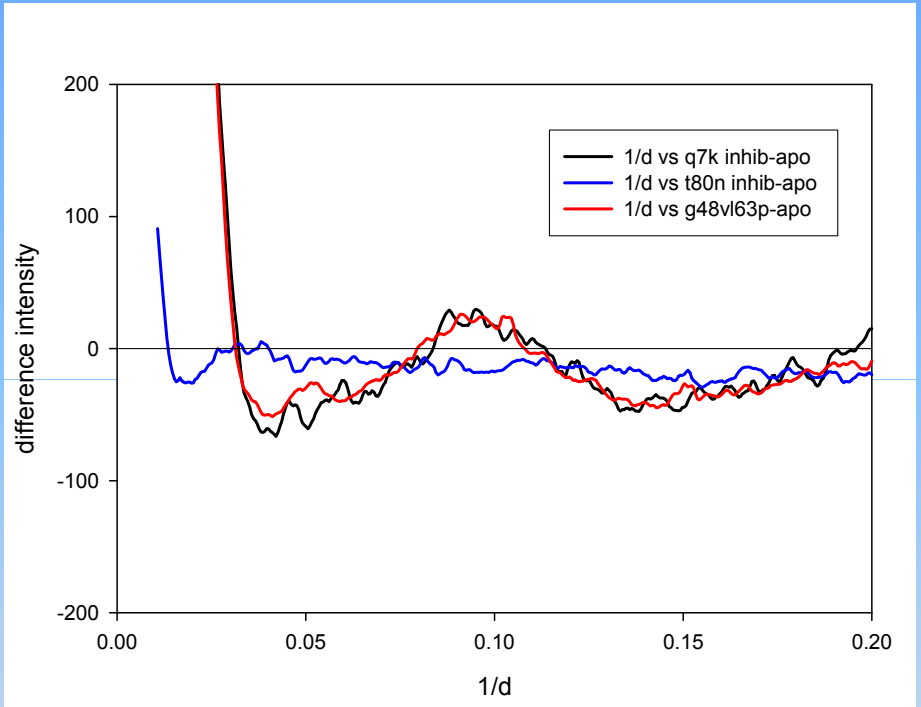


FIG. 1. HIV-1 protease colored by sequence conservation (58). Residues colored red are invariant in both untreated and treated populations of patients. Those colored yellow are invariant only in the untreated population. Invariant glycine residues are colored blue. The side chains of the catalytic aspartic acids (Asp25), Trp6, Trp42, and Thr80 are displayed. (a) Unliganded protease structure (PDB code 1HHP). (b) Protease structure following 9-ns MD simulation.

JOURNAL OF VIROLOGY, July 2006, p. 6906–6916 Vol. 80, No. 14
0022-538X/06/$08.00+0 doi:10.1128/JVI.01900-05
Copyright © 2006, American Society for Microbiology. All Rights Reserved.

Role of Invariant Thr80 in Human Immunodeficiency Virus Type 1 Protease Structure, Function, and Viral Infectivity

Jennifer E. Foulkes,[1] Moses Prabu-Jeyabalan,[1] Deyna Cooper,[1] Gavin J. Henderson,[2,3] Janera Harris,[2] Ronald Swanstrom,[2] and Celia A. Schiffer[1]*



Differences between WAXS patterns +/- inhibitor

Indicate

Wt binds inhibitor

MDR mutant binds inhibitor

t80n does not bind inhibitor

WHY ?

Detection of Functional Ligand-Binding Events Using Synchrotron X-Ray Scattering
Diane J. Rodi, Suneeta Mandava, David B. Gore, Lee Makowski and Robert F. Fischetti
J Biomol Screen 2007; 12; 994

T80N abolishes both protease activity and viral infectivity

But thr80 is not in the active site (contacts active site residues)

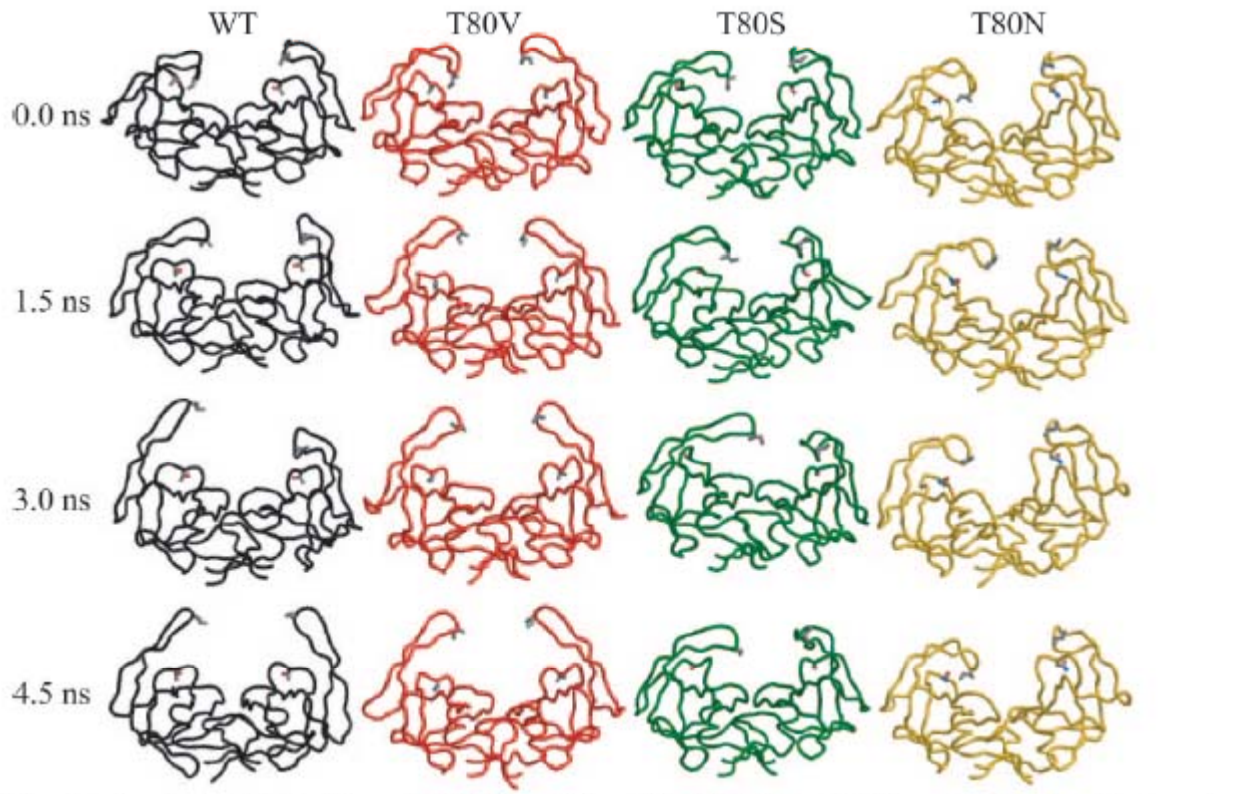


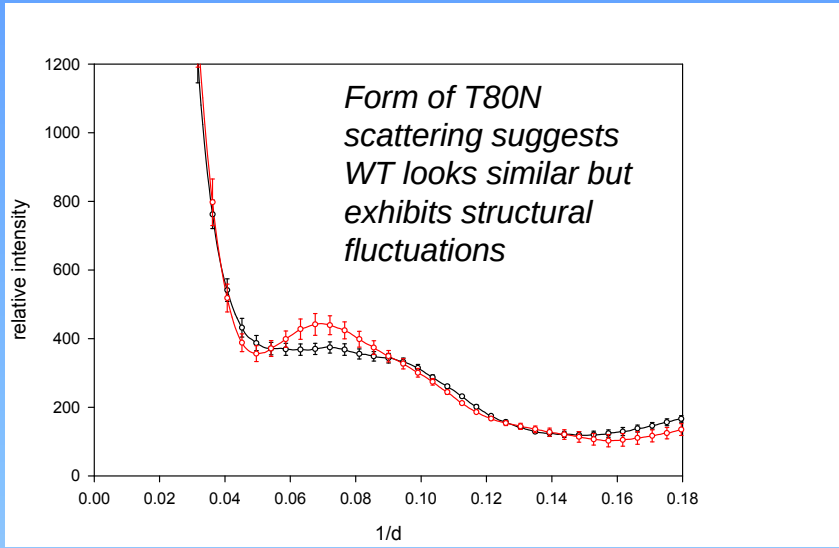
FIG. 7. Snapshots from the four MD simulations at 0.0, 1.5, 3.0, and 4.5 ns (WT, black; T80V, red; T80S, green; T80N, yellow). The flap region in the WT, T80V, and T80S simulations unfolded into the solvent within 4.5 ns. In the T80N simulations, the flap region was less mobile.

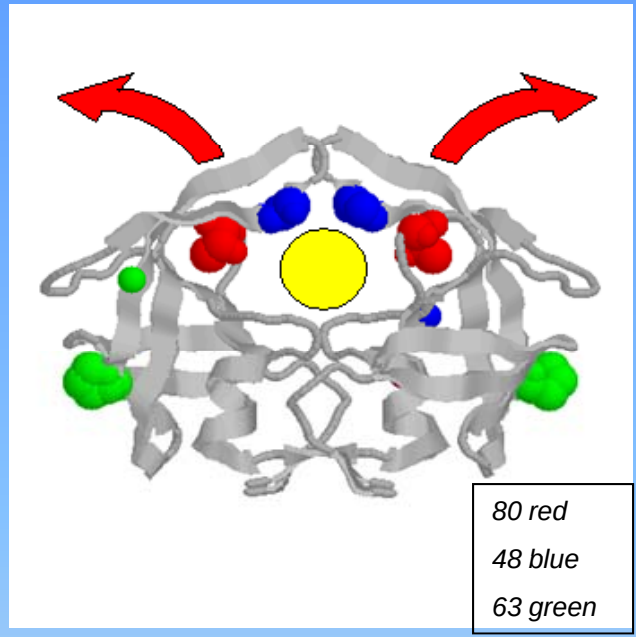
MD simulations suggest T80N inhibits motion of the flap regions

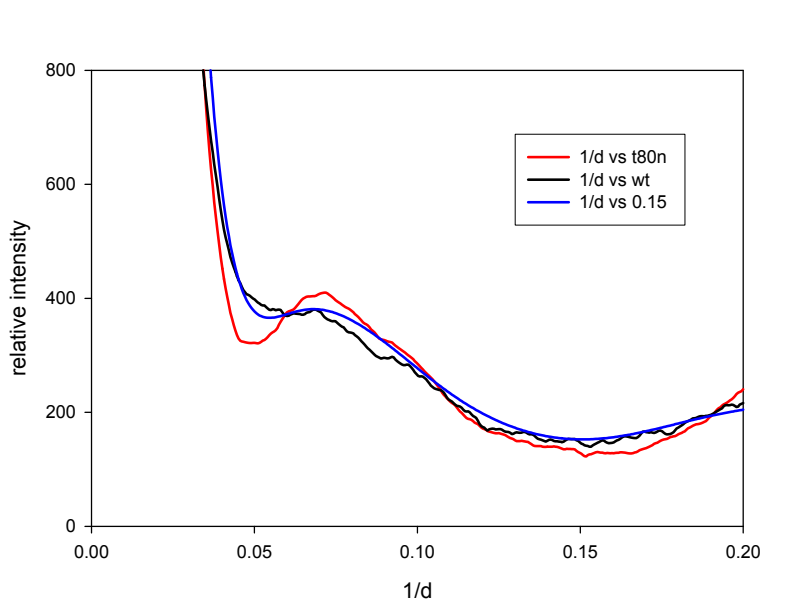
JOURNAL OF VIROLOGY, July 2006, p. 6906–6916 Vol. 80, No. 14
0022-538X/06/$08.00+0 doi:10.1128/JVI.01900-05
Copyright © 2006, American Society for Microbiology. All Rights Reserved.

Role of Invariant Thr80 in Human Immunodeficiency Virus Type 1 Protease Structure, Function, and Viral Infectivity

Jennifer E. Foulkes,[1] Moses Prabu-Jeyabalan,[1] Deyna Cooper,[1] Gavin J. Henderson,[2,3] Janera Harris,[2] Ronald Swanstrom,[2] and Celia A. Schiffer[1]*



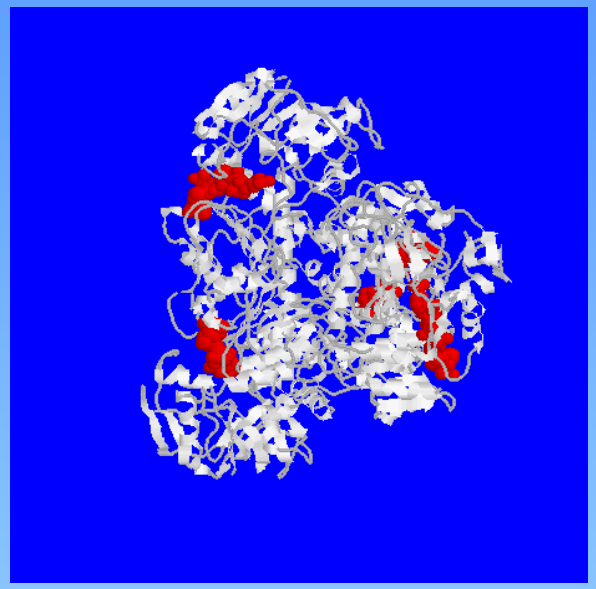


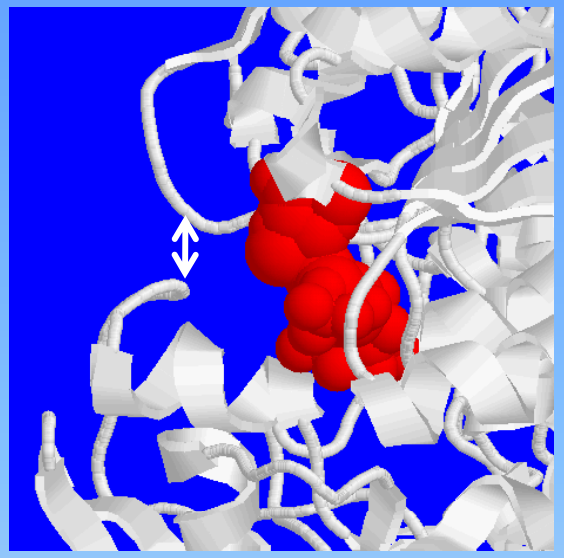


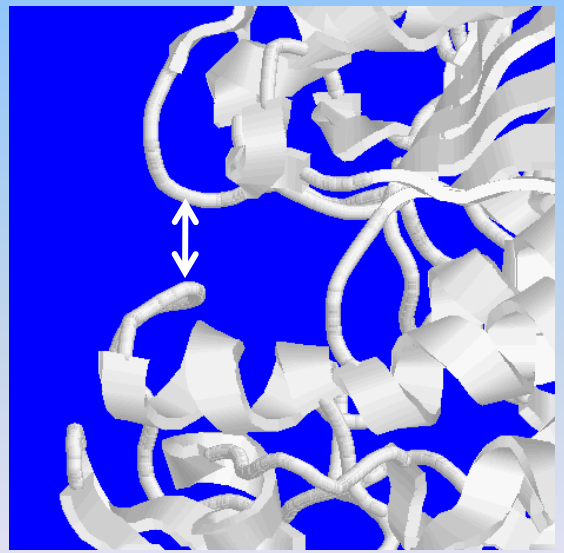
Using t80n as a model for a rigid protease - adding fluctuations results in a predicted WAXS pattern very similar to that of WT

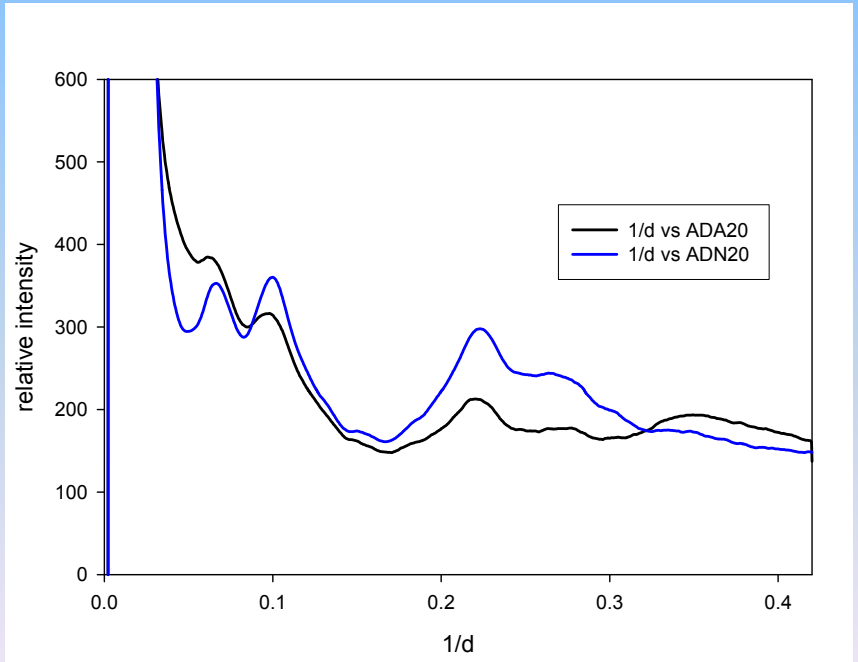
A single site mutation (t80n) causes a functionally significant suppression of structural fluctuations

Binding of NAD to ADH is accompanied by narrowing of binding site in crystallographic structures

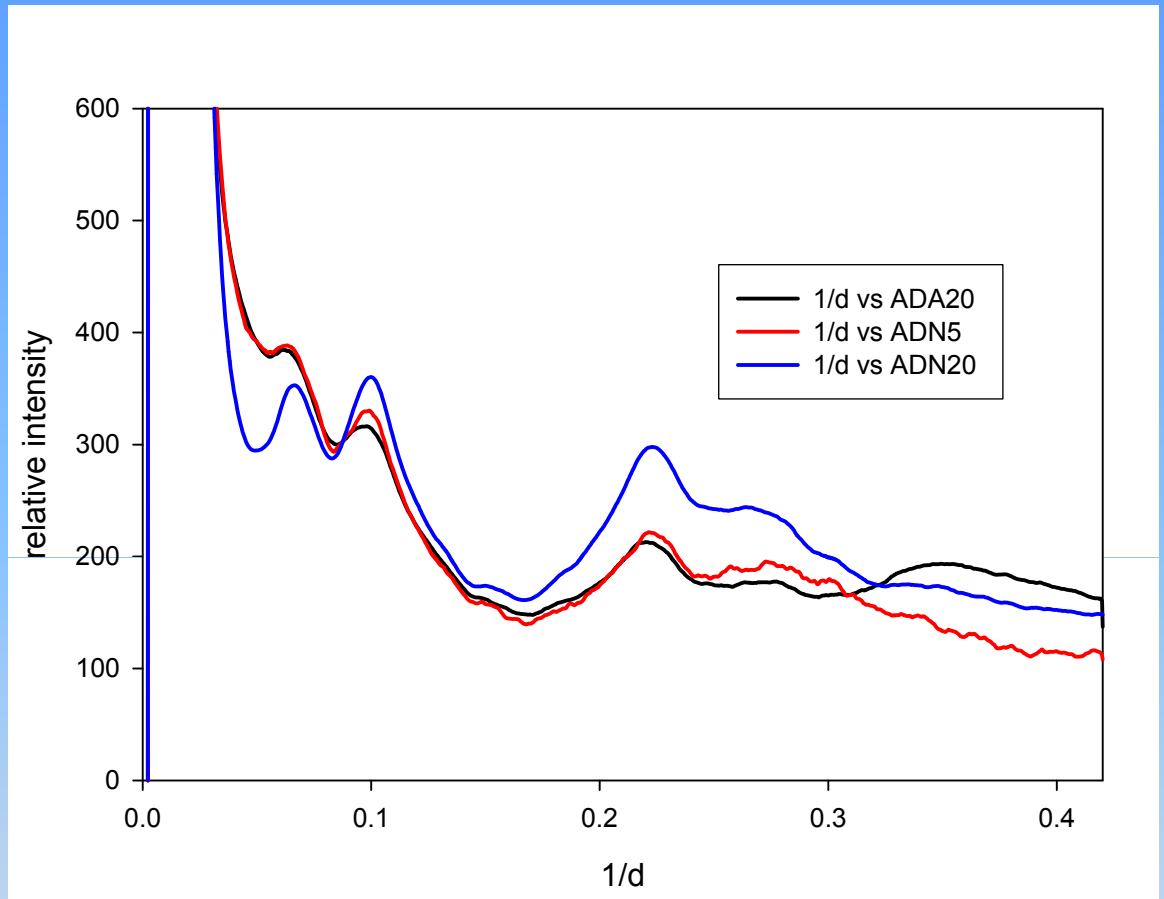








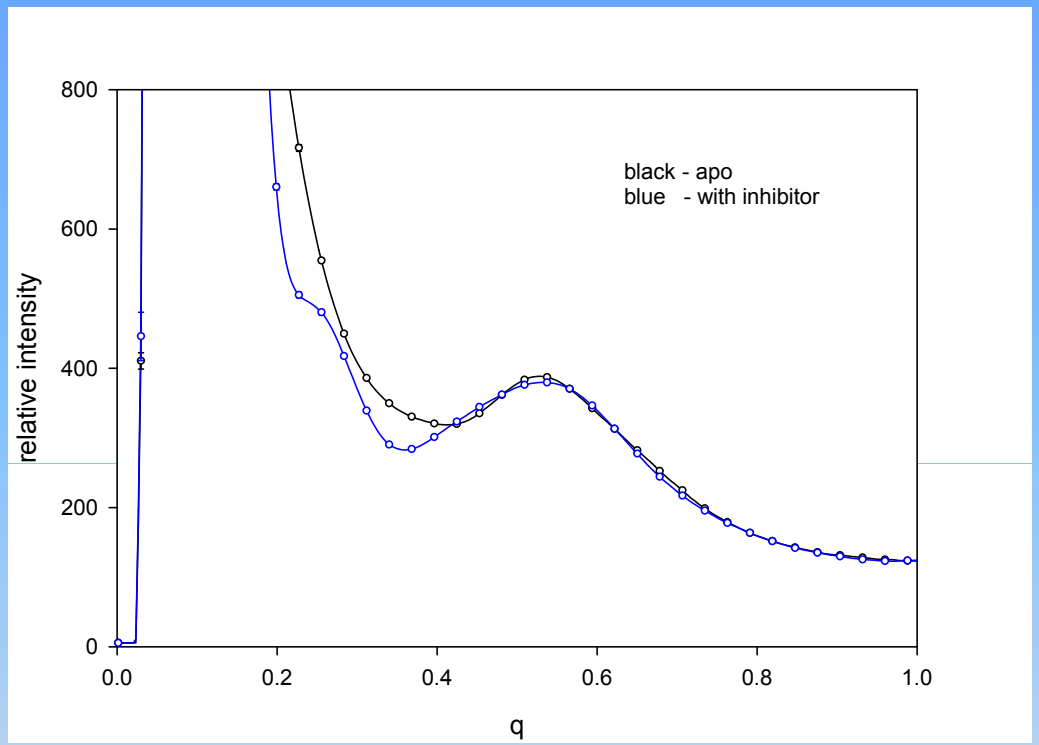
Observed WAXS differences are not consistent with crystallography

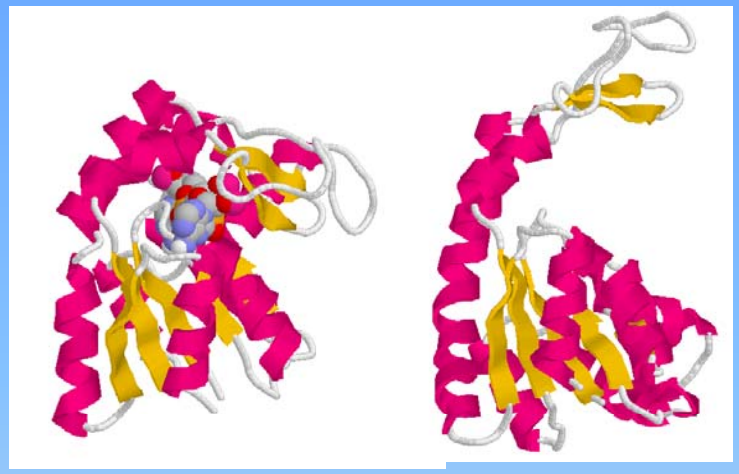


Scattering from ADH in presence of NAD at low concentration is almost identical to apo ADH – suggesting that the major difference between apo and NAD-ADH involves fluctuations – change in relative abundances of ensemble members

Ligand binding inhibits structural fluctuations

Adenylate kinase +/- inhibitor *(with George Phillips)*





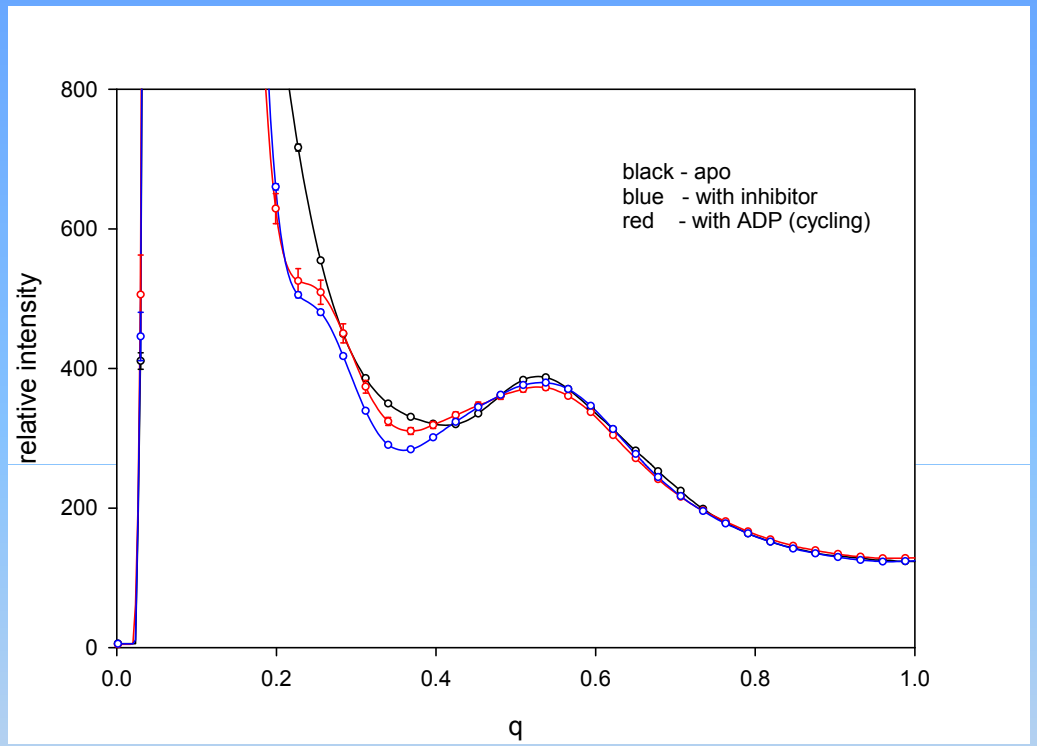
Catalyzes the interconversion of AMP; ADP and ATP

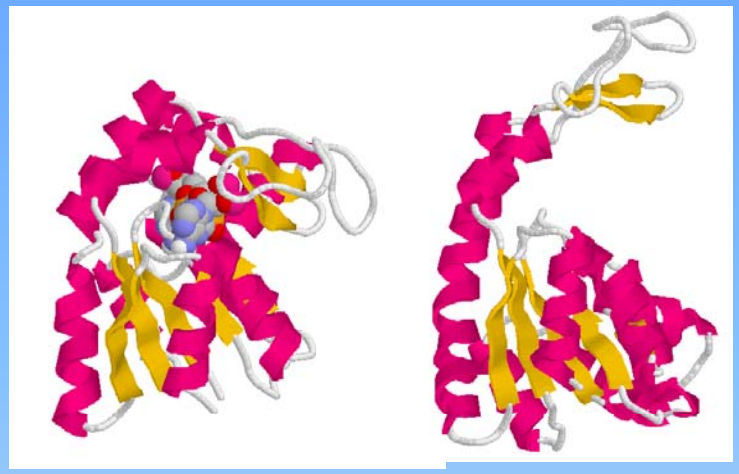
2 ADP $\longleftrightarrow$ *AMP + ATP*

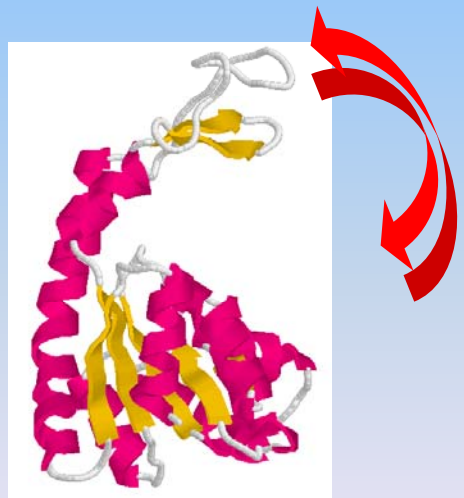
Flaps cover the reagents once in the active site

Form of scattering suggests flaps very flexible in unliganded form

Adenylate kinase +/- inhibitor … and during catalysis







During catalysis closely resembles inhibited form

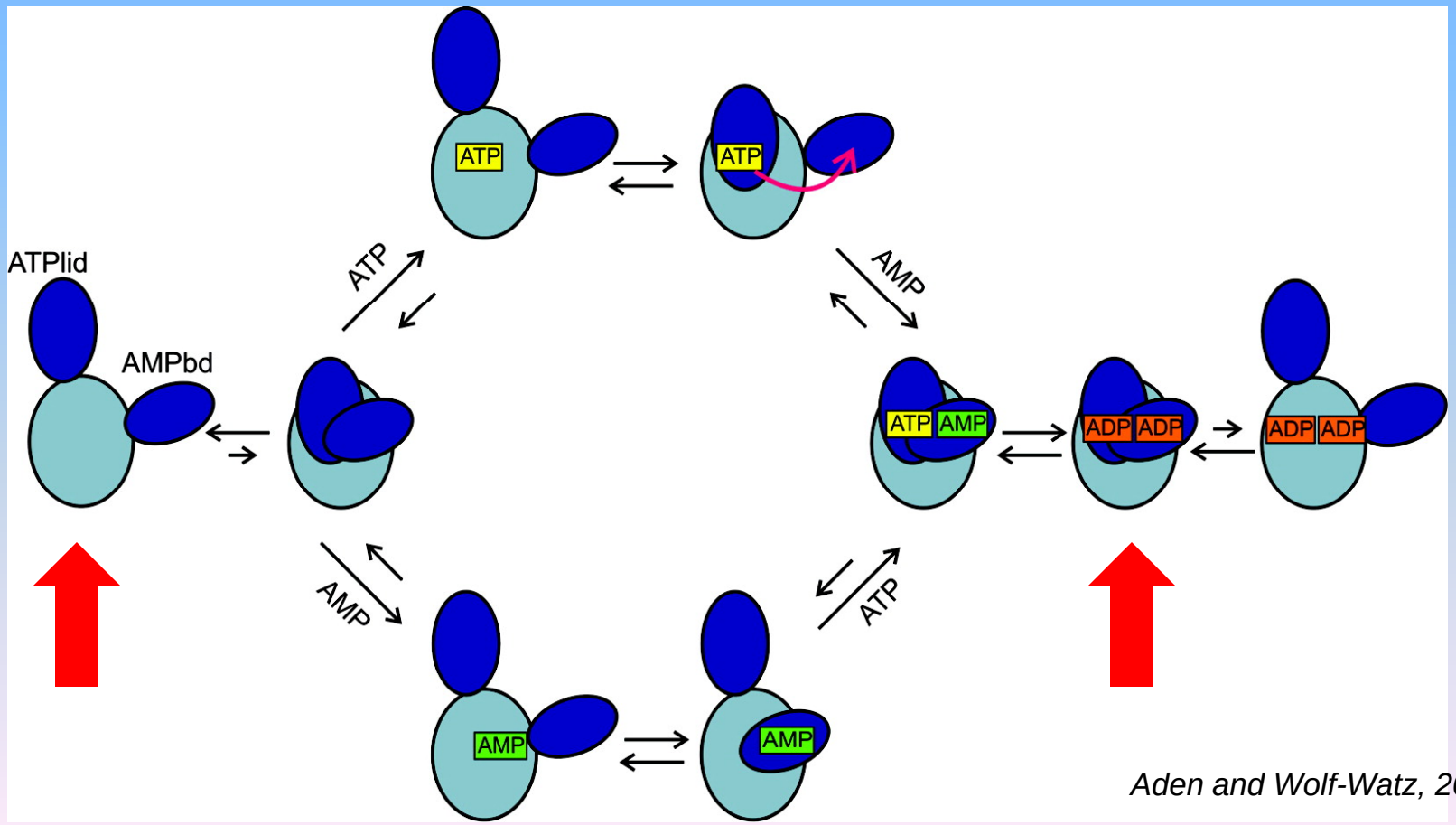
but cannot be constructed as linear combination of the two endpoints

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

Collect WAXS data from AdK under many different conditions

- AMP, binds both sites - apparent K_d of 3.2 mM
- ATP, binds LID domain - K_d of 43 μM
- AP5A, binds both sites - tight binding (nM)
- ADP, enzyme is catalytic - apparent K_d of 267 μM

From thermodynamic measurements can predict relative abundances of different species at any given ligand concentrations

Heterogeneous Solutions

- Biophysical characterization of specific structural ensembles difficult
- Solution: linear algebra

- Scattering patterns are linear combinations of all species in solution

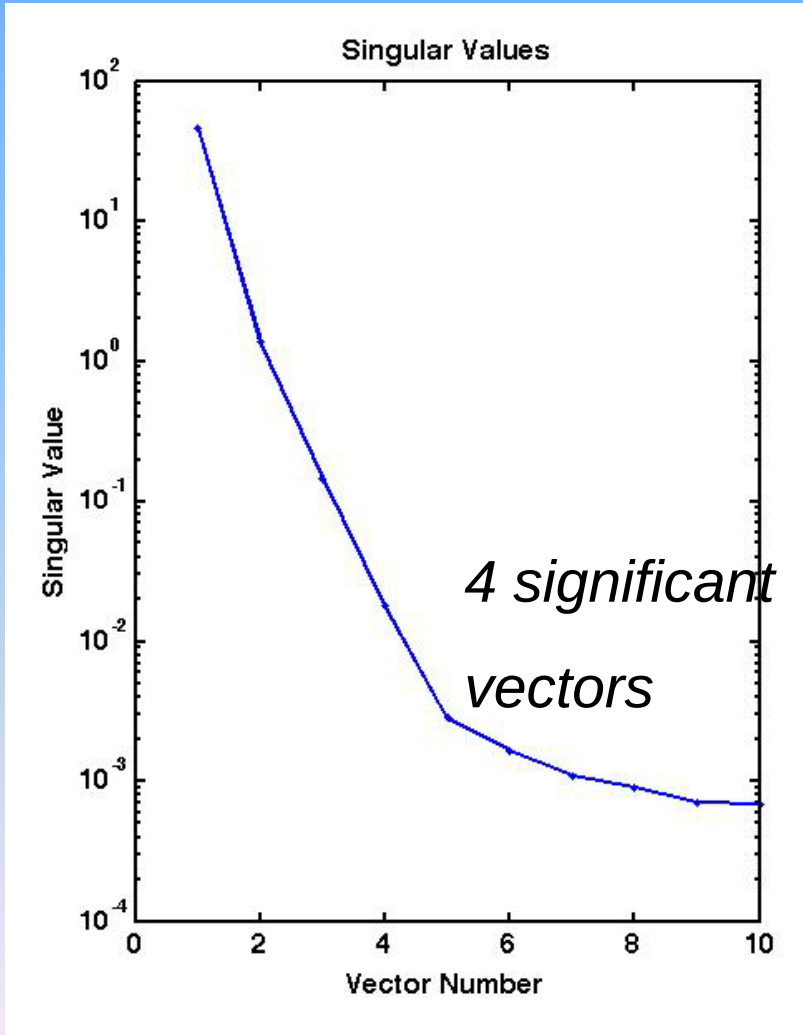
- For simple 1:1 binding, mixture of apo and holo

- With a ligand titration, they are present at different levels
- Concentration model can be inverted to obtain patterns from pure species (SVD + global analysis)

- Validation of stoichiometry and thermodynamic parameters

- Scattering patterns from individual species can thereby be extracted from the data

Singular Value Decomposition



D = USV'

A = number of angles

N = number of experiments

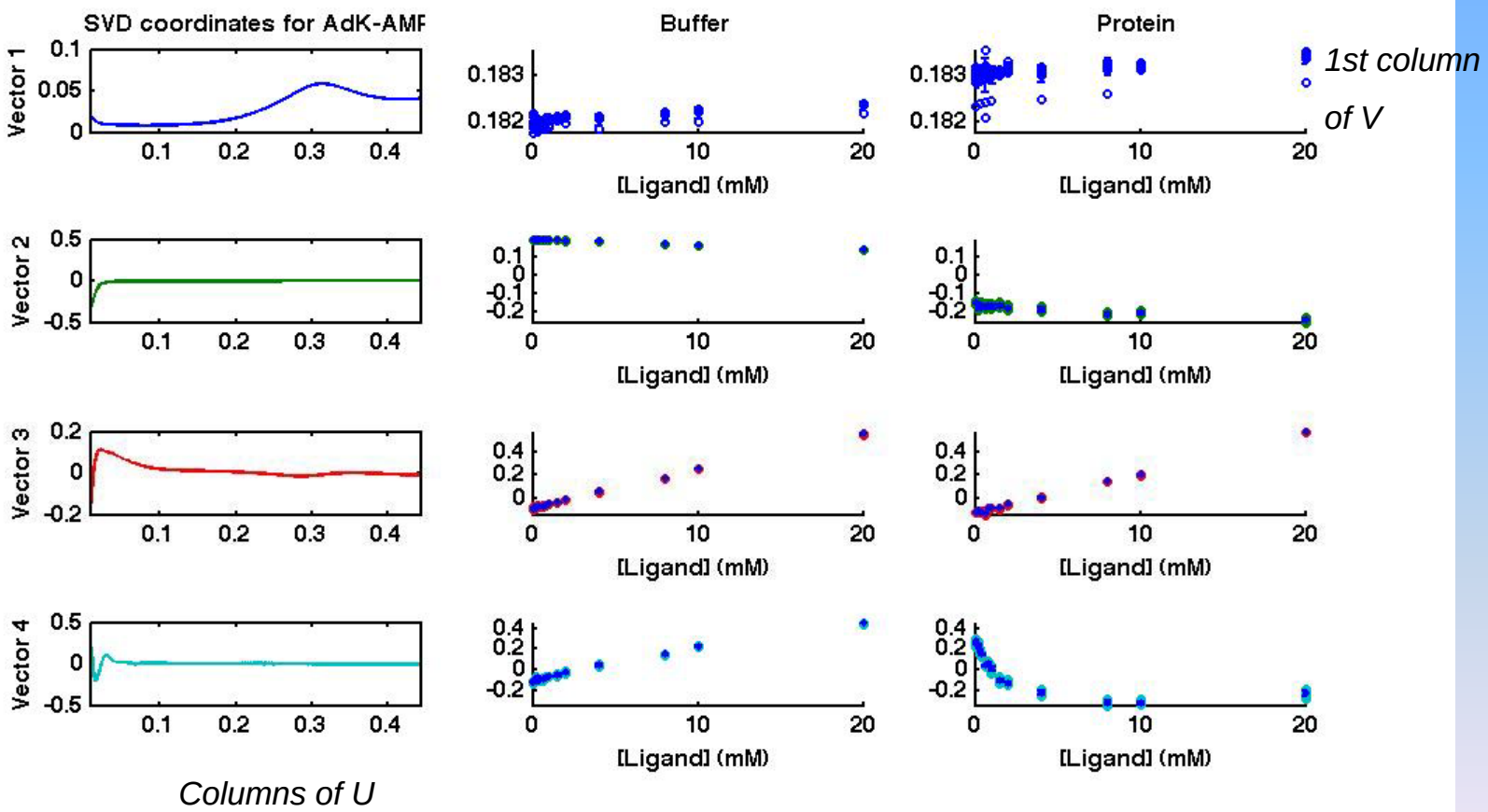
*Data Matrix, D (A*N)*

*Orthonormal Basis, U (A*N)*

*Singular Values, S (N*N)*

*Coordinates, V (N*N)*

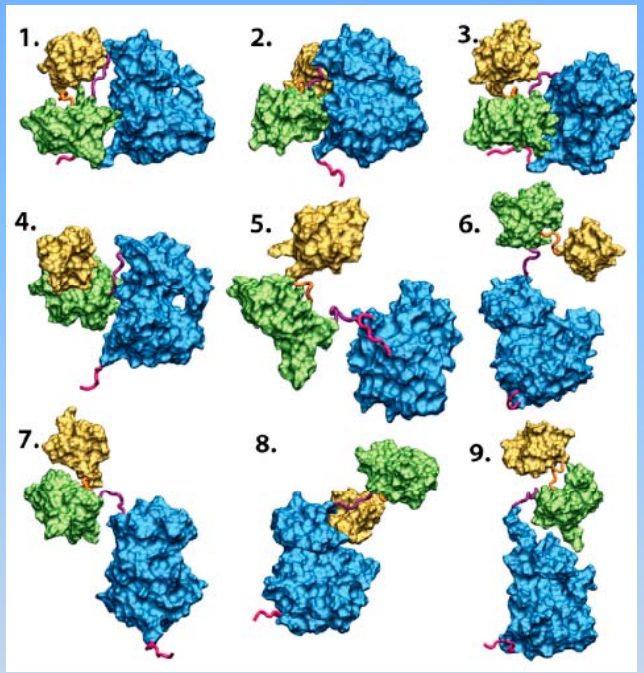
Basis set



Mathematical constructs, no physical significance

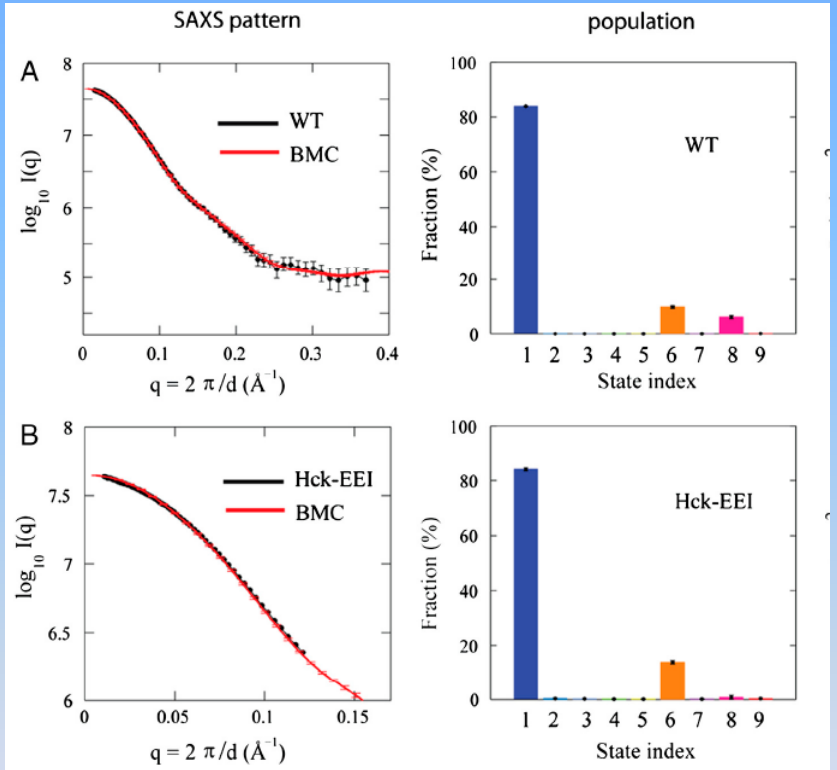
Multidomain assembled states of Hck tyrosine kinase in solution

Sichun Yang[a,2], Lydia Blachowicz[a], Lee Makowski[b], and Benoît Roux[a,b,1]



Representatives of families of conformations abundant under different conditions

Catalytic domain – blue
SH2 domain green
SH3 domain yellow



Hck-YEEI – high-affinity mutant

the future…

- Coordination with thermodynamic studies to characterize the components present under each condition studied
- Use of TR-WAXS to characterize intermediates that are not abundant in steady-state
- Decomposition of scattering from ensembles into terms associated with each individual species has the potential to provide detailed structural information on multiple intermediates
- MD modeling to predict structural ensembles under different conditions
- Comparison of WAXS and MD to evaluate the MD models and refine structures of functionally important intermediates that cannot be captured crystallographically
- Use of intermediate structures to inform drug design

Thanks

Bob Fischetti
Dave Gore

Diane Rodi
Suneeta Mandava
Amina Aziz
Lynda Dieckman

Sanghyun Park
Jaydeep Bardhan
David Minh
Sichun Yang (Case)
Benoit Roux (UCh)

Chien Ho (CMU)
George Phillips (UW)

Steve Kent (U Ch)
Vladimir Torbeev (U Ch)
Celia Schiffer (UMass)

DOE
ANL
NIH