Describing protein conformational ensembles: beyond static snapshots

George N Phillips Jr

Address: Center for Eukaryotic Structural Genomics, Great Lakes Bioenergy Research Center, Department of Biochemistry, 433 Babcock Drive, University of Wisconsin-Madison, Madison, WI 53706, USA

Email: phillips@biochem.wisc.edu

F1000 Biology Reports 2009, **1**:38 (doi:10.3410/B1-38)

The electronic version of this article is the complete one and can be found at: <http://F1000.com/Reports/Biology/content/1/38>

Abstract

Protein molecules are not static but are in varying degrees of motion. Of course, the breadth of structural arrangements in an ensemble will vary tremendously. Some proteins are like 'rocks' and are very tightly constrained in terms of deviations of their coordinates from their averages. Others are 'writhing snakes', with little recognized persistent structure, and then there is everything in between. Through the work of many, there is an increasing awareness of the role of dynamics in the biological function of proteins, and new computational and experimental methods are allowing us to make these connections. A personal perspective on the state of affairs is offered with examples primarily from the author's own work.

Introduction and context

The roles of the internal dynamics of protein molecules in supporting their biological functions are becoming increasingly apparent. Proteins are not static in living organisms; rather, their structures have complex dependencies on their environment and are also subject to thermally driven motions. The energy landscapes of proteins are rich and highly variable. The structural rigidity of a protein can range from being able to stay active in boiling water to essentially having no regular structure under any known circumstances. ('Intrinsically unstructured' or 'natively unstructured' are common labels for this class of proteins.) Furthermore, some proteins are known to be highly ordered only under certain conditions. With reference to the more ordered proteins, Ansari *et al.* [1] have talked about FIMs (functionally important motions) and, by extension, BUMs (biologically unimportant motions), both of which are present in all proteins, indeed one goal of research in this area is to tell them apart!

Since the earliest discussions about allostery [2], it was easy to accept that a 'conformational change' could occur

when a small molecule bound to a protein and that this energy of binding could change the structure, and hence the functional properties, of the protein. Biochemists freely invoke 'conformational change' just as molecular biologists invoke 'trans-acting factors' and geneticists say 'there must be a gene'. Each discipline has its reductionist arm-waving explanation waiting to be tested.

One extreme of this view of changing structures, which is quite reasonable, is that allostery can occur from changes in populations; this differs from the idea that individual proteins have to undergo some sort of concerted conformational change to achieve a functional difference [3]. This view raises the possibility that all proteins are allosteric at some level, or at least potentially so [4].

Crystallographers and other structural biologists have sought out and revealed many such conformational changes, putting forth switching mechanisms to explain protein function. Yet the crystallographers are inadvertently guilty of referring to *the* structure, as if there were only one. Even the fluctuations present within the constraints of the crystal lattice are larger than often

appreciated. The root mean square deviations for the best behaved part of the protein in a crystal structure are about one-tenth of the nominal resolution, so typically a couple of tenths of an angstrom (Å).

Ansari *et al.* and others, through the use of infrared and other spectroscopies, have also articulated the concept of tiers of substates. These can range from tiny vibrational states of bonds (or smaller if you want to go into atomic or subatomic physics) to complete unfolding/refolding transitions. Natural selection determines which of these matter to the function of the protein. Some proteins, such as hemoglobin, appear to be fairly rigid and yet have different functional states, and others have completely different folding patterns depending on the biological environment (for example, prions).

The estimated mean square deviations derived from crystallography might be thought of as being close to the lower limits of an ensemble distribution. Those from nuclear magnetic resonance (NMR) might then be an upper limit. It is common to refer to proteins without clear heteronuclear single-quantum coherence (HSQC) spectra as 'unfolded'. A better (though overly technical) description might be 'too broad of an ensemble to satisfy the proper exchange regime for an NMR determination'. Many times, these proteins can be crystallized, so it is unlikely that they are unfolded in the literal sense. Of course, the protein environment also always matters. Studies of the variation of the structure of proteins crystallized in different forms also point to the existence of conformational ensembles, as they can be correlated with other methods such as NMR. These studies support the notion that crystal lattices do not so much distort the structure as select amongst members of the ensemble that exist in the freely tumbling phase [5,6].

At first, there were perceived discrepancies between the atomic coordinates that crystallographers determined and those resulting from the application of molecular mechanics force fields [7]. Crystallographers were distrustful of the simulations, and molecular dynamicists were quick to blame the crystal environment as an artificial condition. The more recent point of view holds that both camps were right in a sense. Simulations are still difficult to run for long enough to get adequate sampling over time periods typical of experiments and the force fields suffer from necessary approximations, and it is now recognized that crystal structures vary to some degree in different crystal forms. These different camps now seem to be converging on a similar view. In molecular dynamics simulations, evidence for tiers of states can also be seen [8].

The concept of a folding funnel also embodies the idea, with a gradual narrowing of the distribution as proteins relax into lower energy states after being synthesized or unfolded [9]. Some substates in crystals can be frozen out by lowering the temperature, and some not [10]. This quenching of the structure into different substates also argues for there being many states with quite similar energies.

Major recent advances

In a more microscopic way, the bond lengths and angles that give rise to atomic arrangements and their diffraction patterns can vary, giving rise to an ensemble with an average electron density distribution that does not conform to any representation of a single structure, even if modeled as one [11]. Crystallographers usually recognize this effect only in higher resolution structures where discrete conformations of atoms can be distinguished, but there are promising ways of modeling an ensemble directly from the diffraction data even at lower resolutions [12].

NMR methods of protein structure solution have ensembles embedded in them. They are necessary, as a stochastic approach is used to generate models that satisfy the experimental observations, typically many more than are selected to represent the final set of models. The extent to which these ensembles represent the true ensemble is debatable as they may be under-constrained by the experimental data. Certainly, the depiction of the coordinates as a set instead of a single structure gives a more realistic mental image than a single structure. Modern NMR methods also continue to develop new relaxation methods for observing protein dynamics, including hydrogen deuterium exchange experiments [13,14] (both reviewed by *Faculty of 1000 Biology*).

Models for what the ensemble of protein conformations might be when *not* constrained by a crystal lattice have recently been developed and used to illustrate likely modes of motions of the proteins [15] (reviewed by *Faculty of 1000 Biology*). Furthermore, reducing the extent of the ensemble possibilities in proteins by empirically changing the amino acid sequence can result in fewer states and thus a narrower and sometimes more stable folded state [16]. This characteristic is important for developing longer shelf lives for protein-based pharmaceuticals or more robust catalysts for industrial processes.

Future directions

Biophysical techniques for observing or describing the flexible nature of proteins and their resulting ensembles continue to improve in both speed and accuracy. X-ray

diffraction and magnetic resonance methods both promise to deliver single-molecule structure determinations, the former by 'diffract and destroy' free electron laser illumination [17] and the latter by nanoscale magnetic resonance imaging [18]. If these new ideas come to fruition, scientists would be able to look at individual examples of conformations and at the ensembles directly. This would really accelerate our explorations of the molecular world of proteins.

Abbreviation

BUM, biologically unimportant motion; FIM, functionally important motion; HSQC, heteronuclear single-quantum coherence; NMR, nuclear magnetic resonance.

Competing interests

The author declares that he has no competing interests.

Acknowledgements

This report was supported by the Great Lakes Bioenergy Research Center (Madison, WI, USA) through the Department of Energy grant DE-FC02-07ER64494, and the Center for Eukaryotic Structural Genomics (Madison, WI, USA) through the National Institutes of Health (Bethesda, MD, USA) grant GM GM074901.

References

1. Ansari A, Berendzen J, Bowne SF, Frauenfelder H, Iben IE, Sauke TB, Shyamsunder E, Young RD: **Protein states and proteinquakes**. *Proc Natl Acad Sci U S A* 1985, **82**:5000-4.
 2. Monod J, Changeux JP, Jacob F: **Allosteric proteins and cellular control systems**. *J Mol Biol* 1963, **6**:306-29.
 3. Tsai CJ, del Sol A, Nussinov R: **Allostery: absence of a change in shape does not imply that allostery is not at play**. *J Mol Biol* 2008, **378**:1-11.
 4. Gunasekaran K, Ma B, Nussinov R: **Is allostery an intrinsic property of all dynamic proteins?** *Proteins* 2004, **57**:433-43.
 5. Zoete V, Michielin O, Karplus M: **Relation between sequence and structure of HIV-1 protease inhibitor complexes: a model system for the analysis of protein flexibility**. *J Mol Biol* 2002, **315**:21-52.
 6. Kondrashov DA, Zhang W, Aranda Rt, Stec B, Phillips GN Jr: **Sampling of the native conformational ensemble of myoglobin via structures in different crystalline environments**. *Proteins* 2008, **70**:353-62.
 7. Phillips GN Jr: **Comparison of the dynamics of myoglobin in different crystal forms**. *Biophys J* 1990, **57**:381-3.
 8. Andrews BK, Romo T, Clarage JB, Pettitt BM, Phillips GN Jr: **Characterizing global substates of myoglobin**. *Structure* 1998, **6**:587-94.
 9. Miyashita O, Onuchic JN, Wolynes PG: **Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins**. *Proc Natl Acad Sci U S A* 2003, **100**:12570-5.
- F1000 Factor 3.0 Recommended
Evaluated by Stuart S Licht 17 Nov 2003
10. Hartmann H, Parak F, Steigemann W, Petsko GA, Ponzi DR, Frauenfelder H: **Conformational substates in a protein: structure and dynamics of metmyoglobin at 80 K**. *Proc Natl Acad Sci U S A* 1982, **79**:4967-71.
 11. Stec B, Phillips GN Jr: **How the CO in myoglobin acquired its bend: lessons in interpretation of crystallographic data**. *Acta Crystallogr D Biol Crystallogr* 2001, **57**:751-4.
 12. Levin EJ, Kondrashov DA, Wesenberg GE, Phillips GN Jr: **Ensemble refinement of protein crystal structures: validation and application**. *Structure* 2007, **15**:1040-52.
 13. Lindorff-Larsen K, Best RB, Depristo MA, Dobson CM, Vendruscolo M: **Simultaneous determination of protein structure and dynamics**. *Nature* 2005, **433**:128-32.
- F1000 Factor 6.5 Must Read
Evaluated by Volkhard Helms 02 Feb 2005, Gabriele Varani 22 Feb 2005, Torleif Härd 05 Sep 2008
14. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KF, Becker S, Meiler J, Grubmüller H, Griesinger C, de Groot BL: **Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution**. *Science* 2008, **320**:1471-5.
- F1000 Factor 10.5 Exceptional
Evaluated by Matthias Buck 16 Jun 2008, Gottfried Otting 26 Jun 2008, Antonio Rosato 26 Jun 2008, Torleif Härd 02 Jul 2008, Moosa Mohammadi 03 Jul 2008, Thomas Szyperski 14 Jul 2008
15. Kondrashov DA, Cui Q, Phillips GN Jr: **Optimization and evaluation of a coarse-grained model of protein motion using x-ray crystal data**. *Biophys J* 2006, **91**:2760-7.
- F1000 Factor 3.0 Recommended
Evaluated by Jeffry Madura 07 Nov 2006
16. Bae E, Bannen RM, Phillips GN Jr: **Bioinformatic method for protein thermal stabilization by structural entropy optimization**. *Proc Natl Acad Sci U S A* 2008, **105**:9594-7.
 17. Chapman HN, Barty A, Marchesini S, Noy A, Hau-Riege SP, Cui C, Howells MR, Rosen R, He H, Spence JC, Weierstall U, Beetz T, Jacobsen C, Shapiro D: **High-resolution ab initio three-dimensional x-ray diffraction microscopy**. *J Opt Soc Am A Opt Image Sci Vis* 2006, **23**:1179-200.
 18. Degen CL, Poggio M, Mamin HJ, Rettner CT, Rugar D: **Nanoscale magnetic resonance imaging**. *Proc Natl Acad Sci U S A* 2009, **106**:1313-7.