Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

High anisotropy and frustration: the keys to regulating protein function efficiently in crowded environments

Akio Kitao and Kazuhiro Takemura

Highly anisotropic protein dynamics in equilibrium can be observed experimentally or through structural bioinformatics and molecular simulations. This anisotropic nature causes a response, to an external perturbation, along a small number of intrinsic large-amplitude directions as expected from the fluctuation–dissipation theorem. It is also key for controlling specific reactions as stochastic processes in macromolecular crowded environments. Protein anisotropy can be exploited for the calculation of physical properties, such as entropy, which can be employed for binding affinity studies. Energy frustration along soft modes including both global large-amplitude and localized small-amplitude movements is another key feature, as conformational transitions along soft modes, triggered by external perturbations such as the binding of other molecules, can act as a switch to control function.

Address
Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo, Tokyo 113-0032, Japan

Corresponding author: Kitao, Akio (kitao@iam.u-tokyo.ac.jp)

Current Opinion in Structural Biology 2017, **42**:50–58

This review comes from a themed issue on **Proteins: bridging theory and experiment**

Edited by **Igor N Berezovsky** and **Ugo Bastolla**

For a complete overview see the Issue and the Editorial

Available online 2nd November 2016

http://dx.doi.org/10.1016/j.sbi.2016.10.014

0959-440X/ The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

What properties have native proteins acquired during evolution? How can native proteins conduct a specific regulated function in macromolecular crowded environments? Classically, this ability is attributed to the structure of the protein in the native state. Proteins in the native state typically adopt compact structures compared to that in the denatured state [1]. The packing density of the interior of a native protein is high and uniform if surface water molecules are included [2]. The radii of gyration of native proteins are almost proportional to (molecular weight)$^{1/3}$, the power law of a sphere [3]. Consistently, protein radius of gyration normalized by the radius of gyration of a sphere with the same volume is independent of the size of the protein [4]. On the other hand, artificial proteins with random amino acid sequences tend to be larger in size and do not form stable secondary structure [3,5], suggesting that compactness and higher-order structures are properties of native proteins acquired through molecular evolution.

The highly anisotropic nature of protein dynamics

A compact protein structure in the native state is closely related to the highly anisotropic nature of protein dynamics which utilizes compact and relatively rigid structural elements (such as domains), or flexible elements exposed to solvent (such as loops and linkers) as moving units. Systematic analysis of protein structure variations in crystal structures has revealed typical protein conformational changes. For example, pioneering work by Gerstein *et al.* described protein domain movements as hinge and shear motions [6,7]. Recently, Taylor *et al.* classified domain movements into five types: free, open-closed, anchored, sliding-twist and see-saw [8•] and applied to the assignment of hinge and shear movements, showing that a relative translation of domains is rare and that there is no difference between hinge and shear mechanisms [9••]. Significant domain movements are observed in many proteins. The analysis of a comprehensive and non-redundant dataset of structures differing by more than 0.5 Å indicated that more than half of the proteins in the dataset exhibit domain motions [10]. Proteins also conduct conformational transitions with smaller amplitude. Analysis of an equivalent database showed that main-chain dihedral angle transitions occur in 82% of the proteins [11]. Many of these dihedral angle transitions are responsible for global and local hinge motions and the flap motion of loops, but 24.3% of the transitions are involved in so-called 'path-preserving' motions, in which the localized collective dihedral transitions occur to preserve the main-chain path and which correlate with functional events such as ion bindings. It is difficult to detect this type of motion by analyzing atomic fluctuations because the amplitude of the fluctuations is very small. Therefore, the analysis of dihedral angles is also important. The high anisotropy of protein dynamics is also observed in structure ensembles determined by solution NMR. The conformational differences observed between the solution NMR structures and their crystal structure counterparts are consistent with the collective motion identified by principal component analysis (PCA) and the anisotropic network model (ANM) [12].

The high anisotropy of protein dynamics in equilibrium has been well characterized by collective coordinate sets determined by normal mode analysis (NMA), PCA and multidimensional scaling (MDS) [13–19]. In PCA, the collective coordinates are introduced from a variance-covariance matrix of a given coordinate systems (typically Cartesian coordinates of atoms) as:

$$\mathbf{A} = \langle \Delta\mathbf{q}\Delta\mathbf{q}^t \rangle \quad (1)$$

where $\Delta\mathbf{q}$ represents the column vector of the displacement of coordinates from the average, and $\langle \cdots \rangle$ shows the ensemble average. The superscript 't' indicates the matrix transpose. The axes of the collective coordinates in PCA (principal axes) are determined as the eigenvectors of $\mathbf{A}$:

$$\mathbf{AV} = \mathbf{V}\lambda \quad (2)$$

with the orthonormalized condition,

$$\mathbf{VV}^t = \mathbf{V}^t\mathbf{V} = \mathbf{I} \quad (3)$$

where $\mathbf{V}$ and $\boldsymbol{\lambda}$ are the matrices of the eigenvectors and eigenvalues, and $\mathbf{I}$ is a unit matrix. The ith column vector of $\mathbf{V}$, $\mathbf{v}_i$, indicates the ith principal axis. Since the ith diagonal element of $\boldsymbol{\lambda}$, λ_i, is the variance of the ith principal component, its contribution to the total variance,

$$s_i = \frac{\lambda_i}{tr\lambda}, \quad (4)$$

shows the anisotropy of the system. If s_i is much larger than the others, the component is considered a 'soft mode' because a larger fluctuation occurs compared to other components. Proteins intrinsically have a small number of large-amplitude modes. For example [20], s_1 is equal to 0.35 (35%) among 1002 internal degrees of freedom and the accumulated value for the first ten components is 0.89 (89%) in the case of C^{α}-atom PCA of a 20 ns molecular dynamics (MD) simulation of FlhAc protein (Figure 1a). Consistent with this, the s_1 and accumulated values for the first ten components were 32 and 81% in a recent unpublished 1 μs MD simulation. These values are typical for proteins. Another good measure to understand the anisotropy is 'anharmonicity factor', which is defined as the root-mean-square-fluctuation along a PC axis divided by that expected from normal mode along the same axis [21]. It should be also noted that the anhamonicity factor also reflects the effect of multiple minima, which will be discussed later. This factor is typically more than two for large-amplitude modes [15] and can be more than 10 for the largest-amplitude PC mode [22]. The dominancy of a small number of collective degrees of freedom clearly indicates the high anisotropic nature of protein dynamics. The important concept here is that a subset spanned by a small number of collective coordinates is robust, and thus useful for investigating both simulation data and experimental data [13–19].

A recent trend is the consideration of time dependence in the analysis of MD simulations [23,24,25•,26]. Time-structure based independent component analysis (tICA) determines statistically independent components from a time-lagged covariance matrix [23,25•,27], and these independent components were also applied to build Markov state model (MSM) [28,29]. In ICA, all the modes are conceptually uncoupled. However, to understand the mechanisms of protein function, a more important goal is the investigation of the relationship between trigger and response. Independent subspace analysis (ISA) determines a set of subspaces as follows: The collective variables in each subspace are significantly correlated and correlation between the variables from distinct subspaces is insignificant [30]. Interestingly, only five subspaces were identified and all other collective variables are independent in T4 lysozyme. Cross correlation function analysis of the modes in the same subspace quantified the time delay and advance among the correlated modes, and showed that only small number of movements can have the relationship of trigger and response. The largest block consists of six modes and clearly showed the propagation of movements from a global motion mode to a local mode, and then on to other global modes. ISA is useful for identifying a series of correlated events including domain and local motions.

Fluctuation–dissipation theorem and protein function

The high anisotropy of protein dynamics in equilibrium is closely correlated with specific protein response to a weak external perturbation as predicted by the fluctuation–dissipation theorem [31]. This statement is clearer if the response $\Delta\mathbf{q}_R$ to the perturbation force $\mathbf{f}$ is described by the time-independent linear response theory (ti-LRT):

$$\Delta\mathbf{q}_R = \beta\mathbf{Af} \quad (5)$$

where β is the thermodynamic beta. The concept of LRT is applied to investigate protein dynamics [32,33,34•,35]. Ikeguchi *et al.* clearly demonstrated that ti-LRT explains and predicts structural changes in some proteins upon ligand binding [32]. In that work, they determined $\mathbf{A}$ from MD simulations of unliganded protein and reproduced the response of the liganded protein induced by $\mathbf{f}$ mimicking the protein–ligand interaction. Recently Yang and coworkers used ti-LRT and time-dependent (td-) LRT to investigate the response of myoglobin upon CO binding and showed agreement of the time response between LRT, ultraviolet resonance Raman spectroscopy, and time-resolved X-ray crystallography, suggesting that the primary response can be described by LRT [34•].

If the response is observed in the principal component space, we obtain

Download English Version:

https://daneshyari.com/en/article/5510890

Download Persian Version:

https://daneshyari.com/article/5510890

Daneshyari.com