Chemical Physics 396 (2012) 84–91

Contents lists available at SciVerse ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

Large-scale motions in the adenylate kinase solution ensemble: Coarse-grained simulations and comparison with solution X-ray scattering

Michael D. Daily [a,b], Lee Makowski [f,g], George N. Phillips Jr. [c,d], Qiang Cui [a,e,*]

[a] Department of Chemistry, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, United States
[b] Computation and Informatics in Biology and Medicine Training Program, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, United States
[c] Department of Biochemistry, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, United States
[d] Department of Computer Sciences, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, United States
[e] Theoretical Chemical Institute, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, United States
[f] Department of Chemistry and Chemical Biology, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, United States
[g] Department of Electrical and Computer Engineering, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, United States

ARTICLE INFO

Article history:
Available online 24 August 2011

Keywords:
Protein flexibility
Solution ensembles
Conformational selection
Enzyme dynamics
Computational mutation

ABSTRACT

While coarse-grained (CG) simulations provide an efficient approach to identify small- and large-scale motions important to protein conformational transitions, coupling with appropriate experimental validation is essential. Here, by comparing small-angle X-ray scattering (SAXS) predictions from CG simulation ensembles of adenylate kinase (AK) with a range of energetic parameters, we demonstrate that AK is flexible in solution in the absence of ligand and that a small population of the closed form exists without ligand. In addition, by analyzing variation of scattering patterns within CG simulation ensembles, we reveal that rigid-body motion of the LID domain corresponds to a dominant scattering feature. Thus, we have developed a novel approach for three-dimensional structural interpretation of SAXS data. Finally, we demonstrate that the agreement between predicted and experimental SAXS can be improved by increasing the simulation temperature or by computationally mutating selected residues to glycine, both of which perturb LID rigid-body flexibility.

1. Introduction

Many recent works have provided evidence that not just the average structure, but also motions or "dynamics" around this structure, are important to protein functions including catalytic rate control [1–4], macromolecular recognition [5], and/or allosteric regulation [6–9]. For therapeutic and engineering applications, it is important to understand physical allosteric mechanisms in specific proteins [10]. Recent studies have been building evidence to support the hypothesis that evolution has selected well-defined motions in allosteric proteins. For example, motions in apo-proteins tend to parallel closure pathways associated with ligand binding [11–13]. In many cases the bound conformation has a substantial minority population in the absence of ligand [5,14], though this "population shift" behavior does not necessarily indicate that the conformational transition precedes ligand binding in the presence of ligand [15–17]. For simulation of conformational transitions, coarse-grained (CG) models provide a useful first approximation for capturing gross motional features, possibly because such features are primarily determined by structural topology [18–20]. They are also a useful approach to simulate large-scale conformational transitions without constraints so that many different reaction coordinates can be simultaneously assessed [16,21]. However, due to the approximations inherent in CG simulations, it is important to validate these models with experiments. While detailed testing requires high-resolution experiments like NMR HSQC [22] and relaxation [23], hydrogen exchange [24], these methods can be less sensitive to large-scale global motions. Thus, techniques that provide even low-resolution validation of these motions are important adjuncts.

For example, small-angle X-ray scattering (SAXS) provides a useful approach to assess the flexibility (motional amplitude) of rigid bodies in different allosteric states. Specifically, since SAXS is the Fourier transform of the interatomic distance distribution, scattering can be predicted from structural coordinates using the Debye formula [25,26] and compared to the experiment. While calculating average scattering from ensembles of full-atom structures [25,26] is prohibitively costly, scattering can be estimated accurately at low angles from large ensembles of coarse-grained structures through the use of effective residue structure factors [27]. Specifically, the Debye formula is applied to the atoms within res-

* Corresponding author at: Department of Chemistry, University of Wisconsin – Madison, 1101 University Avenue, Madison, WI 53706, United States.
E-mail address: cui@chem.wisc.edu (Q. Cui).

0301-0104/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemphys.2011.08.015

idues and averaged over different conformations from known crystal structures. With this approximation, hundreds of C_α-scale conformations can be analyzed per hour on an 8-core AMD computer.

Although SAXS does not provide three-dimensional structural information due to orientational averaging, simulations can be used to refine the structural interpretation of SAXS data. Here, we show that while experimental SAXS curves can guide optimization of simulation conditions to approximate a given solution ensemble, these curves can be structurally interpreted with more detail by calculating correlations between predicted scattering and various structural properties over a large ensemble of structures.

Here, we apply this approach to the allosteric-like protein adenylate kinase (AK), which recycles AMP to ADP by phosphorylating it with ATP. Crystal structures of apo and bound AK show that upon substrate binding, two small domains (LID and NMP) close over the larger CORE domain. The substrates ATP and AMP bind at the CORE-LID and CORE-NMP interfaces, respectively, and the two predominant states are open (O) and closed (C) [28]. NMR experiments have shown that the thermophilic adenylate kinase from *Aquifex aeolicus* (AKthermo) exhibits an opening rate six-times-slower than that of *E. coli* AK (AKmeso), thereby limiting the catalytic turnover rate at room temperature [3]. This conformational gating of a chemical reaction is analogous to allostery.

Recently, we have performed CG conformational transition (double-well Gō) simulations of AKmeso, AKthermo, and high-temperature and mutational variations of each protein [16,29]. These simulations showed that LID rigid-body flexibility is higher in AKmeso, and that in AKthermo, such flexibility can be increased to that of AKmeso by mutating some key hinge prolines to glycine [29]. We also found that the glycine mutants increased flexibility of some key hinges and destabilized parts of the contact network (e.g. various interactions involving the CORE-LID connector helices).

Here, we investigate several key questions. First, while X-ray crystallography provides detailed information about selected conformations of AK in the absence and presence of ligand, SAXS data can be used to estimate the extent of Cartesian fluctuations within the solution structural ensemble. To estimate this global flexibility for AKmeso, we compare experimental curves to those calculated from CG simulation ensembles of AKmeso. We show the relation between predicted scattering and varying strengths of interresidue interactions, which modulates the global flexibility. Second, to identify possible population shifts, we fit SAXS curves measured in both the absence and the presence of ligand to linear combinations of predicted scattering from the O and C state simulations. Such a population-based approach has previously been demonstrated in prediction of SAXS data through CG simulations of Hck tyrosine kinase [30].

Third, to structurally interpret the scattering data, we calculate the correlation between predicted scattering and various structural metrics over large simulation ensembles as a function of ($q = 2\pi/d$), where d is the Bragg spacing. These metrics include CORE-LID and CORE-NMP center of mass distances, which reflect rigid-body motions, and root mean square deviations (rmsds) of the flexible LID and NMP domains to the closed crystal structure, which reflect the compactness of these domains. These correlations can also suggest explanations for how differences in predicted scattering between simulations arise from differences in the structural properties of the simulation ensembles. To expand upon the connection between scattering and global flexibility, we predict scattering from AK simulations at a range of nominal temperatures. Finally, to isolate the effects of individual structural features upon the scattering curve, we perform computational glycine mutations [29] designed to selectively perturb individual features.

2. Results and discussion

2.1. Simulation and scattering calculation approach

To generate ensembles from which to predict scattering in O and C states, we perform 150-ns single-well Gō simulations based on Karanicolas–Brooks potentials [31,32]. These potentials include sequence-specific dihedral and contact energies, which compared to "vanilla" Gō potentials enable characterization of important small-scale motions and performing computational mutants. Since we are interested in ground-state O and C properties, rather than the conformational transition mechanism described in our prior AK Gō models [16,29], we simulate O and C separately rather than unify them into a double-well potential. In addition, as in our previous simulations [16,29], we simulate ligand binding to the closed state by adding selected ligand-mediated interactions to the C-state contact potentials.

We calibrate the simulated flexibility of AK by varying the contact energy scale (S_{con}), by which we scale the Karanicolas/Brooks [31,32] energies to compensate for extra backbone conformational entropy induced by the generic bond angle potential [33]. For our O/C conformational transition simulations, we calibrated S_{con} to 2.5 so that the C simulation averaged about 2.0 Å C_α rmsd with respect to the closed crystal structure (rms_C) [16] to reproduce prior atomistic simulations of AK [34]. This significantly exceeded the S_{con} of 1.7 that used in double-well Gō simulations of smaller conformational transitions [33,35]). This high S_{con} of 2.5 may best approximate the behavior of AK *in vivo*, where crowding from other biomolecules reduces the population of expanded structures [36]. However, since SAXS is measured in more dilute solution, we simulated AK with S_{con} between 1.5 and 2.5 to optimize the computational/experimental fit.

For each Fast-SAXS prediction, we randomly select an ensemble of 1000 structures from the 150-ns simulation. In addition, as described in the methods, we augment Fast-SAXS [27] to include CG ligand atoms for "ligand-bound" structures for which the rmsd of AMP and/or ATP binding sites from the C crystal structure is small. For experimental data, we compare to *Bacillus globisporus* AK since scattering data is not available for *E. coli* AK in all relevant states. SAXS patterns from *E. coli* and *B. globisporus* AKs correspond closely under conditions where data from both species are available (data not shown). This is not surprising since their two sequences are 51% identical and the closed crystal structures differ by only 1.08 Å rmsd according to the MultiProt structural alignment server [37] (203/214 residues aligned).

2.2. Estimating flexibility in solution

Fig. 1A shows that at $S_{con} = 2.5$, the C simulation predicts a more curved profile than the O simulation, especially near $q \sim 0.22$, which is consistent with C being more ordered than O as expected. Fig. 1B shows that with $S_{con} = 1.9$, the predicted O scattering profile is substantially less inflected near $q \sim 0.22$ than at $S_{con} = 2.5$. Conversely, the predicted C curves are similar for ensembles generated using the two S_{con}, both exhibiting a small dip near $q \sim 0.22$.

Panels C–E compare fits of predicted scattering from simulation ensembles at different S_{con} to experimental scattering measured under apo (ligand-free) conditions. Since both apo and liganded conditions may comprise a mixed population of O and C [34,38], according to the conformational selection hypothesis, we fit to the data varying linear combinations of the predicted open and closed scattering as follows: