Journal of Molecular Graphics and Modelling 47 (2014) 25–36

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Efficient prediction of protein conformational pathways based on the hybrid elastic network model

Sangjae Seo[a,1], Yunho Jang[b,1], Pengfei Qian[c], Wing Kam Liu[d], Jae-Boong Choi[a,c], Byeong Soo Lim[c], Moon Ki Kim[a,c,*]

[a] SKKU Advanced Institute of Nanotechnology, Sungkyunkwan University, Suwon 440-746, Republic of Korea
[b] Department of Mechanical and Industrial Engineering, University of Massachusetts, Amherst, MA 01003, USA
[c] School of Mechanical Engineering, Sungkyunkwan University, Suwon 440-746, Republic of Korea
[d] Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA

ARTICLE INFO

Article history:
Accepted 22 October 2013
Available online 1 November 2013

Keywords:
Protein dynamics
Elastic network model
Normal mode analysis
Elastic network interpolation
Pathway generation

ABSTRACT

Various computational models have gained immense attention by analyzing the dynamic characteristics of proteins. Several models have achieved recognition by fulfilling either theoretical or experimental predictions. Nonetheless, each method possesses limitations, mostly in computational outlay and physical reality. These limitations remind us that a new model or paradigm should advance theoretical principles to elucidate more precisely the biological functions of a protein and should increase computational efficiency. With these critical caveats, we have developed a new computational tool that satisfies both physical reality and computational efficiency. In the proposed hybrid elastic network model (HENM), a protein structure is represented as a mixture of rigid clusters and point masses that are connected with linear springs. Harmonic analyses based on the HENM have been performed to generate normal modes and conformational pathways. The results of the hybrid normal mode analyses give new physical insight to the 70S ribosome. The feasibility of the conformational pathways of hybrid elastic network interpolation (HENI) was quantitatively evaluated by comparing three different overlap values proposed in this paper. A remarkable observation is that the obtained mode shapes and conformational pathways are consistent with each other. Our timing results show that HENM has some advantage in computational efficiency over a coarse-grained model, especially for large proteins, even though it takes longer to construct the HENM. Consequently, the proposed HENM will be one of the best alternatives to the conventional coarse-grained ENMs and all-atom based methods (such as molecular dynamics) without loss of physical reality.

1. Introduction

The great developments of biological experiments have led to a rapid increase in the solved structures of macromolecules [1]. Also the development of computational models has enabled us to relate static structures of macromolecules to its functions by predicting conformational change [2,3]. Molecular dynamic (MD) simulation [4,5] is one of the most common computational tools utilized to elucidate macromolecular motions at the atomic level. However, although MD simulation provides time-dependent behaviors of a molecular system, including the effect of the surrounding solvent, current computer power limits the investigation of the relevant dynamics solely to an early stage [6]. In attempt to reduce computational cost, various coarse-grained (CG) approaches have been developed [7]. Among these CG approaches, the elastic network model (ENM) having a single-parameter or simplified potential function has proved to be a useful tool for investigating global dynamics of proteins [8,9]. The ENM represents the system as multiple degrees-of-freedom (DOF) linear mass-spring systems. For instance, a protein structure is modeled as a spring network among representative Cα atoms.

Early studies have shown that the normal mode analysis (NMA) based on ENM can effectively predict low frequency motions, which are relevant to the functional motion [10–13]. NMA basically describes thermal fluctuations of a macromolecule around its equilibrium energy. By combining several lowest normal modes, one can reproduce the early stage of the functional collective motions of the given structure. The results of NMA based on ENM correlate well with MD simulation results and experimentally observed

* Corresponding author at: SKKU Advanced Institute of Nanotechnology, Sungkyunkwan University, Suwon 440-746, Republic of Korea.
Tel.: +82 31 299 4840; fax: +82 31 290 5889.
E-mail address: mkkim@me.skku.ac.kr (M.K. Kim).
[1] These authors contributed equally to this work.

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.10.009

conformational changes in both the magnitude and direction of the motion [14–19]. Another advance in NMA technique is the consideration of rigidity of protein. The rigidity and flexibility are important aspects to account for the protein motions. The large-scale movements in macromolecules, such as domain and hinge-like motion, are highly engaged in relative motion among those rigid domains [20,21]. A rigid domain motion can be analyzed by rigid body motion; translation and rotation. The first attempt was the rotations–translations of blocks approach (RTB) proposed by Sanejouand group [22]. RTB considered consecutive amino-acid residues as a pure rigid body block and six DOF are projected onto the Hessian matrix. The block normal mode (BNM) method [23] was proposed to reduce the storage size in diagonalization step of RTB method by the projection matrix. Recently, a new domain decomposition method for RTB, called density-cluster RTB was developed [24]. The cluster NMA (cNMA) [25,26] was developed to overcome several limitations in RTB and BNM method. cNMA can directly convert Cartesian coordinates into rigid clusters without any projection, vice versa, whereas RTB and BNM methods transform the full Hessian matrix into a reduced subspace. One incomparable advantage of rigid-clustering is the dramatic reduction of computational cost. The computational complexity of the conventional NMA is the order of n^3 and the storage of a Hessian matrix is proportional to n^2. However, in case of rigid cluster based NMAs, the computational cost depends on the number of rigid domains into which the system is structurally decomposed [22].

In addition to the study of the intrinsic motion of macromolecules, investigating the transition pathway is also important to elucidate the relationship between structure and function. Firstly, elastic network interpolation (ENI) iteratively obtains the intermediate conformations by interpolating two corresponding sets of interatomic distances in each elastic network [27,28]. By comparing the transition pathway with molecular dynamics simulation result, ENI is also able to compromise between a physically oversimplified linear interpolation method and computationally expensive MD simulation [29]. Also Feng et al. simulated the pathway of adenylate kinase (AK) in atomic details by incorporating CHARMM force field to ENI [30]. Furthermore, various methods based on the free energy surface have been introduced. Plastic network model (PNM) generated pathway of AK based on the free energy surface of two ENM potentials [30] and mixed-ENM extended the free energy based technique by mixing the Boltzmann-weighted Go potentials [31]. To improve the accuracy of mixed-ENM, Tekpinar and Zheng generalized mixed-ENM by searching the saddle point of double-well potential functions [32]. Even though earlier studies have successfully predicted transition pathways, there has been no attempt to predict transition pathway holding rigidity information.

In order to preserve the rigidity and flexibility of protein and consider both for rigid clusters and point masses, there is clearly a need for methods that generally treat the macromolecules for NMA calculation and pathway generation. In this study, the hybrid ENM (HENM) is introduced for both the hybrid NMA (HNMA) and the hybrid ENI (HENI). The HENM consists of point-mass parts and rigid-cluster parts that correspond to flexible regions and rigid domains, respectively. Thus, the HENM was proposed as a tradeoff between Cα coarse graining and rigid-cluster modeling in which rigid clusters and point masses are linked to one another with linear springs [12,13,33]. HENM's computational advantage over the other methods is also discussed. The comparison of overlap of the computed pathways between the two end conformations with normal modes from each end conformation shows excellent consistency. Moreover, a comparison of the source code run time proves that the HENI method provides computational efficiency for large proteins (residue number > 300). Ribosome was chosen as the representative to test reliability of the proposed HENM as a simulator of protein conformational transitions. HNMA was applied to the 70S ribosome and the first six normal modes showed its major dynamic motions, which facilitate the translocation of the tRNA inside the 70S ribosome. The generated translocation pathway of ribosome correlates well with the current understanding in the literature. Consequently, the proposed HENM can not only reduce computational cost but also generate reliable dynamic motions of large macromolecules without loss of generality.

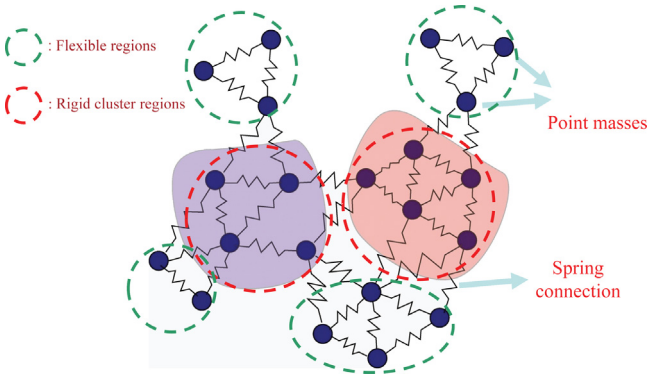

Fig. 1. Overview of HENM. Interactions between atoms are represented as a set of elastic springs. Rigid regions are described with rigid body motion, while the other flexible regions are represented as point masses.

2. Methods

In this section, we derive a full mathematical description of HENM. The HENM concept is based on the fact that most conformational changes in protein dynamics can be resolved into hinge and shear motions that are associated with the collective behavior of atoms in the rigid domains. An overview of HENM (Fig. 1) is as follows: (i) all representative atoms are networked with linear springs in a conventional way; (ii) rigid regions (e.g., secondary structures) are re-modeled as a set of rigid clusters; and (iii) the other regions, which are flexible (e.g., hinges and loops) are represented as point masses. The mathematical description of HENM is derived based on the potential energy between point to point mass, cluster to cluster, and point to cluster.

2.1. Determination of rigid clusters

Although various rigidity algorithms and theories [34–37] have been introduced, there is still no unique way to define rigid clusters and point masses with given structures. The secondary structures, such as alpha helices or beta sheets are usually assumed to be rigid domains. For the efficient and systematic detection of rigid regions, computational approaches are widely used. The graph-theoretical approach named FIRST identifies flexible and rigid regions of the proteins based on the mathematical rigidity theory [38]. The windowed root mean square deviation (WRMSD) method was developed from the conventional static comparison between two end conformations [39]. While the conventional root mean square deviation (RMSD) method calculates the RMSD value over the whole structure as a single representative metric to quantify how much two given structures are close to each other, the WRMSD method continuously computes the RMSD value of the local region captured by a finite window size along the backbone. The lower WRMSD value indicates that the corresponding region moves collectively like a rigid body during the conformational change. It also suggests that the window size should be small enough to avoid missing functionally important (even locally flexible) structures. For example, in the case of adenylate kinase (AK) presented in Fig. 2, there are four rigid clusters marked as green boxes when the window size is set to be 20 residues. After passing the local