Review: Elastic Network Model for Protein Structural Dynamics

Min Hyeok Kim¹ and Moon Ki Kim¹,²*

¹SKKU Advanced Institute of Nanotechnology (SAINT), Sungkyunkwan University, South Korea
²School of Mechanical Engineering, Sungkyunkwan University, South Korea

Abstract

Proteins and their complexes undergo conformational changes, which are closely related to their unique biological functions. However, it is of great challenge for both theoretical and experimental studies to resolve the protein conformational changes due to the limitations regarding the time scale, data size and computational cost. In recent years, normal mode analysis based on coarse-grained elastic network model has been proven to be suitable for the study of the collective vibration motions in macromolecules. Based on the topology of native contacts, this coarse-grained analysis can provide the global motions effectively, thus getting insights into the mechanical aspects of proteins dynamics. In this short review, the basic theory and fundamental features of elastic network models are introduced and a wide variety of examples and applications are then discussed.

ABBREVIATIONS


INTRODUCTION

In the last several decades, more than 90,000 macromolecules structures have been revealed according to the great development in biological experiments [1]. Based on these three dimensional structures, macromolecules perform various functions such as catalysis [2,3], regulation [4-6], transport [7], ligand binding [8,9], and so on. Since the relationship between molecular structure and its function is tightly involved with conformational change [10,11], several experimental and computational methods have been evolved in this field [12-16]. However, the experimental methods undergo the difficulty in direct observation of protein motions [17]. The nuclear magnetic resonance (NMR) spectroscopy is usually used to determine both the statics and dynamics of protein, but it is limited on the size of protein and also has difficulty in discrimination of fast and slow diffusion [12]. In addition, single-molecule FRET experiment is usually insufficient to completely define the conformational change with low resolution [16]. Overall, such direct experimental methods including mass spectrometry with hydrogen/deuterium exchange and single-molecule experiments using optical trapping still are not enough to reveal protein dynamics in atomic detail. As an alternative, simulations could potentially fill in some of the details [14]. Molecular dynamics (MD) simulation, based on solving the Newton’s equations of motion and getting the time-dependent behaviors of proteins, is one of the most representative computational tools utilized to understand protein motions at atomic level [18,19]. Even though MD simulation has become more accurate enough to explain and predict experiment results, standard all-atom MD simulation with transferable force fields is still limited to the prediction of only an early event due to its computational burden [20]. In order to reduce the computational cost, various coarse-grained (CG) methods have been proposed by using much simplified description of potential and structure [21]. Among these CG methods, the elastic network model (ENM) with a single parameter harmonic potential has been widely used for studying global dynamics of proteins [22,23]. In this short review, we give a brief summary of ENM. An overview of the theoretical foundation and the basic features of ENMs are presented, focusing on the anisotropic network model (ANM) [24-26], and then the extensive applications of ENM are followed.

Elastic network model theory

In 1996, Tirion[22] suggested that a quadratic potential with a uniform constant for all atomic interactions would be sufficient to describe low-frequency collective motions of macromolecules. The corresponding potential is as follows:

\[ U = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=i+1}^{n} k_{i,j} (R_{i,j} - R_{0,i,j})^2, \]
where \( R_{i,j}, R_{i,j}^0 \) are the instantaneous and equilibrium difference between \( i \)th and \( j \)th atom, respectively, and \( k_{i,j} \) is a spring constant between them, which equals to 1 if \( R_{i,j} \) is within a cutoff distance and zero otherwise. At first, Bahar and Jernigan suggested 7Å as the reliable distance cutoff value in their Gaussian network model (GNM), from which one could predict the accurate B-factors compared to the experimental ones [23,27-28]. Zheng empirically proposed that the adequate cutoff distance for optimal description in ANM is 8Å [29]. However, a distance cutoff of less than 10Å could generate the mathematically unstable results in normal mode analysis (NMA) having more than six zero eigen values, which physically senses that the ENM is too sparsely modelled so that a single collective system unrealistically moves like several independent pieces. In order to overcome this problem, Jeong and his coworker [30] proposed the bond-cutoff method which determines the spring constants corresponding to the chemical interactions. This method cannot only reduce computational burden, but also generate plausible conformational changes, especially for opening motion [31]. As an example, (Figure 1) shows the procedure of ENM of a protein.

Once ENM is constructed, the vibrational characteristics of a target protein can be investigated by NMA [32-34]. The equation of motion is derived from the Lagrangian mechanics such that

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\delta}_i} \right) - \frac{\partial L}{\partial \delta_i} = 0, \quad \text{(1.2)}
\]

where \( L = T - V \) and \( V \) are the general kinetic energy and potential energy, respectively. \( \dot{\delta}_i \) is the \( i \)th component of displacement vector. Substitution of these two energy terms into Eq. (1.2) yields the following equation of motion (EOM).

\[
M \ddot{\delta} + K \delta = 0, \quad \text{(1.3)}
\]

Where \( M \) is the global inertia matrix consisting of sub-diagonal 3x3 matrices, \( M_{i,j} \), which represent the specific mass values of each representative atom and \( K \) is the global stiffness matrix having sub-stiffness matrices, \( K_{i,j} \) such that

\[
\begin{align*}
K_{i,i} &= -G_{i,i} \\
K_{i,j} &= \sum_{k=1}^{n} G_{k,j} + \sum_{k=i+1}^{n} G_{i,k}, \quad \text{if } i \neq j \\
G_{i,j} &= k_{i,j} (R_{i,j}^0)^T (R_{i,j}^0)^T \\
&= -\sum_{i,j,k} \left( \begin{array}{c} R_{i,j}^0 \end{array} \right)^T \left( \begin{array}{c} R_{i,j}^0 \end{array} \right) K_{G_{i,j}} \\
&= -\sum_{i,j,k} \left( \begin{array}{c} R_{i,j} \end{array} \right)^T \left( \begin{array}{c} R_{i,j} \end{array} \right) G_{i,j}
\end{align*}
\]

\[
(1.4)
\]

The full mathematical derivation of EOM is available in Ref. [35].

In order to get more precise analysis, Kim et al. recently proposed a mass-weighted chemical ENM (WMCENM) that includes not only the chemical interaction but also total masses of each residue according to types of residue [31]. Substitution of \( \delta \) by \( M^{-1/2} v \) in Eq. (1.3) yields the mass-weighted stiffness matrix as follows:

\[
\ddot{v} + M^{-1/2} K M^{1/2} v = 0, \quad \text{(1.5)}
\]

Once NMA is performed with respect to the transformed vector \( v \) in Eq. (1.5), the eigenvector set would be inversely transformed into \( \delta \) by multiplication of \( M^{-1/2} \). In this eigen problem, eigen values and eigenvectors of the target protein represent vibration frequencies and corresponding vibration mode shapes, respectively [36]. By combining several lowest modes, one could represent functionally collective motions of the given protein.

**Applications of elastic network models**

The most attractive feature of ENM is its simplicity and robustness. Despite reduced structural information of coarse-grained masses simply connected by harmonic springs, the combination of conventional normal modes forms an orthogonal basis set. Interestingly, global dynamics involving the collective motion could be represented through these several lowest frequency modes within the normal mode spectrum, thus

---

**Figure 1** Elastic network representations of for the periplasmic lysine-, arginine-, ornithine-binding protein (PDB ID: 2LAO). (a) Ribbons diagram of target protein colored depending on secondary structures. (b) Elastic network connections between the C-alpha atoms (orange sphere). The interactions within the cutoff distance of 11Å are illustrated as blue solid lines. (c) Example of the WMCENM. Various chemical interactions are represented by different colored lines. Black, green, cyan, yellow, and blue lines indicate backbone, hydrogen bonds, ionic bonds, disulfide bonds, and van der Waals interactions, respectively.
dynamics behaviors of large systems such as ribosome [37] and virus capsids [38,39] that are hardly accessible by MD simulations could be elucidated by coarse-grained ENM with the advantage of computational efficiency. ENM based dynamics studies can provide the comprehensive description of functional motions in proteins, which is also consistent with experimental results. For example, Wang studied the functional motion of ribosome complex using the ENM [37]. The high correlation of motion between A-tRNA and P-tRNA indicates that their translocations would occur simultaneously. On the other hands, E-tRNA shows the weak correlation with other two tRNAs, which represents the independent exiting motion from E-site. In short, the comparison of several the lowest modes at each subunit provided the insight of the translocation mechanism in the ribosome. From these studies including other large protein cases, the beauty of ENM-based NMA is that lowest frequency modes involving the collective motions are sensitive only to structural (i.e. geometric) information, not chemical properties of proteins [40]. Therefore, ENM based NMA can be widely used for analyzing many biological problems.

As shown in (Figure 2), one of representative studies based on ENM is a molecular docking (ligand to protein and protein to protein) simulation [41-49]. For instance, Tobi and Bahar have showed that structural changes caused by relevant ligand binding are strongly correlated with intrinsic motions of proteins in their unbound state [48]. This work confirms, in selecting/ rearranging complex formation, the roles of a preexisting equilibrium called Monod-Wyman-Changeux model (MWC). Keskin recently enlarged this hypothesis into the enzymes and antibodies case [44]. Regardless of classes of proteins, the set of the lowest normal modes, although different combination with or without ligand, could cover the limited range of conformational states which are adequate for the structural change inbinding occurrence. The simulation results showed that an ensemble of similar conformations driven by intrinsic motions of native state could bind to different antigens or ligand. Despite MWC model leads to the more complicated complex combination than other two models (i.e. rigid adaptation and induced fit model), ENM could be utilized to restrict the number of candidates for molecular docking.

Another interesting application of ENM is the refinement of low-resolution structural data using the lowest-frequency normal modes [38,50-53]. Tama et al. have proposed the normal mode flexible fitting (NMFF) where the flexible fitting of high-resolution structures into the low-resolution cryo-electron microscopy (cryo-EM) data is performed by deforming the structure along a few low-frequency normal modes [50,51]. This method enables us to build the feasible atomic structure and determines the most important mode for its functional motion. The similar methodology using a set of the lowest normal modes for refinement in low resolution of X-ray crystallography has been successively performed [54]. In order to improve the quality of low-frequency modes and remove the tip effect, a quite small set of collective variables were used as refinement parameter [55]. By focusing one assumption of harmonicity of protein motions in X-ray crystallography, the proposed refinement protocol results in improvements of the resolution, especially in case of large and mobile complexes at moderate resolutions [54]. Indeed, structure refinement based on ENM and NMA not only replaces the traditional homology modeling method based on sequence comparison to template, but also steps forward to the abundant practical needs such as drug designs [56].
ENM is also used for investigating the transition pathway [35, 55-59]. In case that molecular transition is accompanied by large structural change, a harmonic model presumably fails to describe it. Instead, a different class of an harmonic structure-based model called Go model has been widely used for this purpose [60]. Alternatively, an ENM-based pathway generation method called elastic network interpolation (ENI) has been introduced [35, 59]. The intermediate conformations are generated by interpolating two corresponding sets of interatomic distances. This ENI overcomes the limitation of ENM-based NMA which only utilizes a harmonic potential, thus cannot describe the conformational change crossing over the energy barrier. It is also expected that ENI can be utilized as a refinement method by filling the incomplete information obtained from NMR experiment [63]. Moreover, the proposed transition pathway of ENI can act like an ensemble of MD data by capturing most of them with a span of a few lowest normal modes of each intermediate conformation [64].

The further researches have been introduced based on the free energy surface. Maragakis and karplus [65] proposed the plastic network model (PNM) in which the pathway of adenylate kinase (AK) is generated based on free energy surface. Also, the similar but more improved approach, mixed-ENM, has been proposed by solving the double-well potential function [66].

For convenience, several web servers are now available to calculate and visualize NMA results of various proteins. In ANM [25] and oGNM [67], one can similarly select either C-alpha only or all-atoms for ENM. In order to reduce the computational cost, NOMAD-Ref [68] and ElNemo [69] use the building block approximation which groups several residues into a single block. This grouping method is less effective on observation of lowest-frequency modes, but could save substantial amount of computing time. However, these coarse-graining methods would fail to address the atomic details of protein motions unless any other consideration of the rigidity of protein is provided. To overcome this limitation, cluster-NMA [70] and hybrid ENM [71] have been introduced, in which user can adjust the degrees of coarse-graining level from single atom to rigid cluster, regardless of the number of atoms that belong to a rigid cluster. More details on the complexity of various ENMs can be found elsewhere [72].

Most recently, KOSMOS has been launched by integrating various ENM-based dynamics analysis methods including both NMA and ENI [73]. This fully automated web server cannot only provide various coarse-grained ENMs from all-atom model to rigid-cluster model, but also offer chemical information based cutoff method for better simulation accuracy.

CONCLUSION

ENM-based simulation methods have shown great success in understanding of biological functions of macromolecules based on...
on their structural information. The fundamental philosophy of ENM is that the topological features play a dominant role in defining the global and collective motions of proteins. Hence, coarse-grained ENMs have been widely used to solve a variety of biological problems including functional motions of protein complexes, ligand binding mechanism, refinement of low resolution structural data, and transition pathway generation. Despite ENM sometimes shows the limitation owing to its modest resolution structural data, and transition pathway generation.

REFERENCES


40. Petrovne P, Vandes P. Can conformational change be described by only a few normal modes? Biophys J. 2006; 90: 1583-1593.


