Archives of Biochemistry and Biophysics 508 (2011) 64–71

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

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

Normal mode analysis with molecular geometry restraints: Bridging molecular mechanics and elastic models

Mingyang Lu [a], Jianpeng Ma [a,b,*]

[a] The Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, BCM-125, Houston, TX 77030, United States
[b] Department of Bioengineering, Rice University, Houston, TX 77005, United States

ARTICLE INFO

Article history:
Received 21 October 2010
and in revised form 15 December 2010
Available online 4 January 2011

Keywords:
Normal mode analysis
Energy minimization
All-atom normal modes
Coarse-graining
Molecular geometry restraints
Tip effect

ABSTRACT

A new method for normal mode analysis is reported for all-atom structures using molecular geometry restraints (MGR). Similar to common molecular mechanics force fields, the MGR potential contains short- and long-range terms. The short-range terms are defined by molecular geometry, i.e., bond lengths, angles and dihedrals; the long-range term is similar to that in elastic network models. Each interaction term uses a single force constant parameter, and is determined by fitting against a set of known structures. Tests on proteins/non-proteins show that MGR can produce low frequency eigenvectors closer to all-atom force-field-based methods than conventional elastic network models. Moreover, the "tip effect", found in low frequency eigenvectors in elastic network models, is reduced in MGR to the same level of the modes produced by force-field-based methods. The results suggest that molecular geometry plays an important role, in addition to molecular shape, in determining low frequency deformational motions. MGR does not require initial energy minimization, and is applicable to almost any structure, including the one with missing atoms, bad contacts, or bad geometries, frequently observed in low-resolution structure determination and refinement. The method bridges the two major representations in normal mode analyses, i.e., the molecular mechanics models and elastic network models.

Published by Elsevier Inc.

Introduction

Normal mode analysis (NMA)[1] is an important computational tool in studying vibrational motions of biomolecules [1–4]. During the past decades, many mode calculation methods have been developed, including force-field-based methods and network-based methods.

In the force-field-based methods, the molecular potential function is given by a force field, such as CHARMM [5–7] and AMBER [8–11]. Some coarse-grained schemes have also been developed to reduce computational costs, such as the Rotations–Translations of Blocks (RTB) [12] and all-atom-derived methods [13,14]. Recently, Hendrickson group developed a coarse-grained force-field for NMA and molecular dynamics [15]. These methods can keep detailed molecular interactions, but they usually require initial energy minimization step, which in many cases distorts structures [16].

In the network-based methods, such as the coarse-grained elastic network models (ENM) [17–31], the potential functions are solely harmonic terms with equilibrium positions reside on the studied structure. Therefore, they can bypass the initial energy minimization step. However, it was found that the low frequency modes produced by these methods may not be as accurate as force-field-based methods [16,26,32]. In particular, some low frequency modes contain abnormally localized motions, known as "tip effect" [33]. Methods have been developed to alleviate this problem by strengthening local stiffness [33–35], or by specific coarse-graining scheme [36]. However, in node-based (or network-based) mode calculation methods, it may become subjective for the selection of nodes, especially for non-protein components, let alone the stiffness between nodes.

Recently, we developed the minimalist network model (MNM) [16,32] to cope with issues in both conventional force-field-based methods and network-based methods. MNM utilizes a force field to calculate potential, and slightly modifies the Hessian to bypass the initial energy minimization step. Tests show that MNM outperforms both CHARMM normal mode method and all-atom elastic network model in fitting experimental anisotropic displacement parameters of crystal structures. However, in some cases, one may need to deal with low-resolution structures, which commonly contain missing atoms, bad contacts and bad geometries. In those

* Corresponding author at: Baylor College of Medicine, One Baylor Plaza, BCM-125, Houston, TX 77030, United States. Fax: +1 713 796 9438.
E-mail address: jpma@bcm.tmc.edu (J. Ma).

[1] *Abbreviations used:* NMA, normal mode analysis; MGR, molecular geometry restraints; ENM, elastic network model; MNM, minimalist network model; RTB, rotations-translations of blocks; CHARMM, chemistry at HARvard molecular mechanics; AMBER, assisted model building with energy refinement.

0003-9861/$ - see front matter Published by Elsevier Inc.
doi:10.1016/j.abb.2010.12.031

cases, the accurate potentials and Hessian matrices are usually hard to be calculated by a molecular force field. Moreover, many structures contain some components (e.g., organic ligands) that are not conveniently defined in a molecular force field.

To solve this problem, in this study, we developed a new mode calculation method for all-atom structure using molecular geometry restraints (MGR). In addition to non-bonded interactions commonly used in the conventional ENM, the MGR Hamiltonian has harmonic restraining terms on molecular geometry, i.e., bond lengths, bond angles and dihedral angles. Each potential term in MGR uses a single force constant parameter to represent the characteristic stiffness of each interaction type. Unlike force-field-based methods, MGR immediately satisfies the harmonic approximation, allowing one to bypass the lengthy initial energy minimization step. On the other hand, unlike conventional ENM, MGR includes molecular interactions defined by molecular geometry, allowing better description of low frequency modes influenced by the details of molecular structure.

Since the calculation of MGR only requires the information of atomic positions and bond connectivity of the molecules, it is widely applicable to all kinds of macromolecules, such as proteins, DNA, lipids, small molecule ligands, etc. Moreover, MGR can be applied to low quality structures, which commonly contains missing atoms, bad contacts and bad geometries. MGR can also be applied to supramolecular complexes, by combining it with coarse-graining schemes such as RTB method [12].

Our tests show that MGR can systematically produce low frequency eigenvectors closer to the force-field-based methods than ENM. This suggests that, molecular geometry plays an important role, in addition to molecular shape [37], in determining low frequency deformational motions. The MGR method provides a bridge between two major representations in normal mode analyses, i.e., the molecular mechanics models and elastic network models.

In the following, we first describe the methodology of MGR, including the potential functions defined by the molecular geometry. Then, we show the procedure of optimization of parameters and the way to test MGR.

Material and methods

MGR

The potential function used in MGR method has four terms

$$V = V_{bond} + V_{angle} + V_{dihedral} + V_{non\text{-}bond}. \quad (1)$$

Unlike force-field potentials, each potential term of MGR is a harmonic potential with energy minimum on the studied structure.

The first two terms, V_{bond} and V_{angle}, denote bond length and bond angle potential. They have forms of

$$V_{bond} = \sum k_l (l - l^0)^2, \quad (2)$$

where the summation is over all chemical bonds, l and l^0 are the instantaneous and initial bond lengths; and

$$V_{angle} = \sum k_\theta (\theta - \theta^0)^2, \quad (3)$$

where the summation is over all bond angles, θ and θ^0 are the instantaneous and initial bond angles. The constants k_l and k_θ are the force constants for the bond and bond angle interactions. Here, an assumption is made that, like the non-bonded interactions in conventional ENM, all the bond length (or bond angle) interactions share the same stiffness. It should be noted that although these potential terms in MGR have the same functional forms as those in some force-fields, Eqs. (2) and (3) are conceptually different from the traditional bonded potentials. Specifically, l^0 and θ^0 in the MGR potential term are not equilibrium values used in molecular mechanics force fields, but values directly taken from the studied structures. It automatically ensures the MGR potential has its energy minimum at any given starting configuration, i.e., MGR does not require initial energy minimization. Such feature also provides an advantage for MGR over regular force-field-based methods when applied to low quality structures, such as the ones reconstructed from low-resolution X-ray crystallographic data, the ones modeled by structure prediction, or the ones modeled in early stages of X-ray crystallographic refinement. Since these structures usually contain defects, such as unphysical bond lengths, missing atoms, or bad contacts, MGR method would have no problem in dealing with those defective structures because all interactions of each type are modeled as uniformly distributed harmonic potentials.

The third term in Eq. (1), $V_{dihedral}$, is the dihedral angle potential in a form of

$$V_{dihedral} = \sum k_\phi (\phi - \phi^0)^2, \quad (4)$$

where the summation is over all dihedral angles and improper dihedral angles, ϕ and ϕ^0 are the instantaneous and initial dihedral angles, k_ϕ is the force constant. Although the dihedral angle potential, in some cases, is modeled with a periodic function, we used a simple harmonic term to model dihedral interactions. This treatment is fine in the current study because only one potential minimum is required in the normal mode calculations. In this study, dihedral angles and improper dihedral angles were modeled with the same force constant k_ϕ, because no apparent improvements were found when modeled with separate parameters.

The last term $V_{non\text{-}bonded}$ is the non-bonded potential in a form of

$$V_{non\text{-}bonded} = \sum_{r^0 \le r_c} k_r (r - r^0)^2, \quad (5)$$

where r and r^0 are the instantaneous and initial distance of the non-bonded atom pairs, r_c is the cutoff distance and k_r is the force-constant. The non-bonded term is similar to the potential function of the conventional ENM except that MGR excludes the atom pairs that are involved in a chemical bond or a bond angle (1–2, 1–3 interactions). Like many other studies, 1–4 interactions are included in the potential function. Following the literature [17], the cutoff distance r_c for the non-bonded interaction is chosen as 6 Å.

One may wonder whether the non-bonded interactions should be modeled heterogeneously, i.e., to use several parameters to characterize various non-bonded interaction types. To answer this question, we tested MGR on several protein structures with a version in which hydrogen bonds were modeled with a stronger force constant. Hydrogen bond interactions were selected in the test mainly because they are crucial in stabilizing protein structures. However, no apparent improvement was found for the version with stiffened hydrogen bond interactions. Therefore, we believed that it is appropriate to model the non-bonded interactions homogenously for the purpose of current study.

Once the potential terms are defined in Eq. (1), it is then quite straightforward to perform NMA. As already mentioned, MGR does not require an initial energy minimization step. The Hessian matrix of MGR can be obtained by calculating second derivatives of each energy term according to the formulae derived in Ref. [38]. Although the Hessian matrices are usually highly sparse, diagonalization of the matrices is still computationally expensive for supramolecular structures. To deal with this problem, one can reduce the dimension of Hessian matrix by a coarse-graining scheme, e.g., RTB [12] and sub-structure based method [13,14,39]. In this study, we mainly implement and test a version of MGR coarse-grained by the RTB method, in which case each residue is modeled as a rigid-body. By implementing ARPACK package (http://www.caam.rice.edu/software/ARPACK) to perform matrix diagonalization, we are able to