A Unification of the Elastic Network Model and the Gaussian Network Model for Optimal Description of Protein Conformational Motions and Fluctuations

Wenjun Zheng
Department of Physics, University at Buffalo, State University of New York, Buffalo, New York 14260-1500

ABSTRACT Coarse-grained elastic models with a C_α-only representation and harmonic interactions have been increasingly used to describe the conformational motions and flexibility of various proteins. In this work, we will unify two complementary elastic models—the elastic network model (ENM) and the Gaussian network model (GNM), in the framework of a generalized anisotropic network model (G-ANM) with a new anisotropy parameter, f_{anm}. The G-ANM is reduced to GNM at $f_{anm} = 1$, and ENM at $f_{anm} = 0$. By analyzing a list of protein crystal structure pairs using G-ANM, we have attained optimal descriptions of both the isotropic thermal fluctuations and the crystallographically observed conformational changes with a small f_{anm} ($f_{anm} \leq 0.1$) and a physically realistic cutoff distance, $R_c \sim 8$ Å. Thus, the G-ANM improves the performance of GNM and ENM while preserving their simplicity. The properly parameterized G-ANM will enable more accurate and realistic modeling of protein conformational motions and flexibility.

INTRODUCTION

Understanding protein conformational dynamics holds the key to decrypting protein functions at a microscopic level. Simplified coarse-grained models (1–4) have been established as valid and efficient means to probe protein conformational motions and flexibility beyond the reach of atomistic molecular simulations (5). Here we focus on two coarse-grained elastic models: the elastic network model (ENM) (6–8) and the Gaussian network model (GNM) (9,10). Both models simplify the atomic interactions in proteins by using elastic interactions between C_α atoms within a cutoff distance R_c. GNM has been shown to perform better than ENM in describing the thermal fluctuations of protein structures measured by the isotropic crystallographic B factors (11–13). Additionally, the GNM-based calculation of B-factors is insensitive to R_c in the range 7.3 Å $\leq R_c \leq$ 15 Å (14), but for ENM, a higher R_c value (15 Å $\leq R_c \leq$ 24 Å) is needed for optimal fitting of B-factors (13), which is beyond the physical range (4.4 Å $\sim$ 12.8 Å; see Cieplak and Hoang (15)) of residue-residue contact interactions. However, the isotropic GNM cannot predict the directions of protein motions. Instead, the normal mode analysis (16) of ENM has been shown to yield a handful of lowest normal modes that quantitatively capture the conformational changes observed between different protein crystal structures (8,17–20). Therefore, GNM and ENM are complementary in describing the thermal fluctuations and conformational motions in proteins, but neither is satisfactory by itself.

In this work, we intend to unify GNM and ENM in the framework of a generalized anisotropic network model (G-ANM) with a new parameter f_{anm} that defines the extent of anisotropy between the longitudinal and transverse motions between pairs of neighboring residues (or C_α atoms). At the isotropic limit ($f_{anm} = 1$), the G-ANM is reduced to a GNM; at the fully anisotropic limit ($f_{anm} = 0$) the G-ANM is reduced to an ENM. Then we explore the intermediate values of $f_{anm} \in (0, 1)$ to quantitatively assess the performance of a G-ANM in describing both the isotropic thermal fluctuations and the observed conformational changes for a selected list of 18 test cases, each corresponding to a pair of protein structures from the Protein Data Bank (PDB). The systematic evaluation of this list allows us to understand the f_{anm}-dependence of the quality of G-ANM. We also consider a range of R_c values (7 Å $\leq R_c \leq$ 20 Å) to explore the R_c-dependence of the quality of G-ANM.

Our main findings are as follows: by parameterizing G-ANM at a small f_{anm} ($f_{anm} \leq 0.1$) and a relatively short cutoff distance $R_c = 8$Å, we are able to achieve optimal descriptions of both the isotropic thermal fluctuations and the crystallographically observed conformational changes, which are comparable with the best descriptions of the thermal fluctuations attained by GNM (for 8 Å $\leq R_c \leq$ 12 Å) and the best descriptions of the observed conformational changes attained by ENM (for 8 Å $\leq R_c \leq$ 12 Å). Therefore, this study demonstrates an effective way to improve both GNM and ENM without hurting the simplicity of these coarse-grained models.

METHODS

Generalized anisotropic network model

Given the C_α atomic coordinates of a protein crystal structure, we define the G-ANM potential energy as a weighted sum of two harmonic potentials to describe the pairwise interactions between neighboring C_α atoms:

Submitted November 15, 2007, and accepted for publication January 10, 2008.

Address reprint requests to Wenjun Zheng, E-mail: wjzheng@buffalo.edu.

Editor: Ron Elber.

0006-3495/08/05/3853/05 $2.00

doi: 10.1529/biophysj.107.125831

$$E_{\text{G-ANM}} = \frac{1}{2}\sum_{\text{i,j}:d_{\text{ij}}^0 < R_{\text{c}}} C_{\text{ij}}\{f_{\text{anm}}|\vec{x}_{\text{i}} - \vec{x}_{\text{j}}|^2 + (1-f_{\text{anm}})[(\vec{x}_{\text{i}} - \vec{x}_{\text{j}})\cdot\vec{n}_{\text{ij}}^0]^2\}, \quad (1)$$

where $0 \le f_{\text{anm}} \le 1$ is the anisotropy weight parameter (see below) and i or j is the index for a C_α atom. d_{ij}^0 is the distance between the equilibrium positions of i and j. $\vec{x}_{\text{i}}$ ($\vec{x}_{\text{j}}$) is the three-dimensional (3D) displacement of i (j). $\vec{n}_{\text{ij}}^0$ is the unit vector pointing from the equilibrium position of i to that of j. C_{ij} is the force constant of the spring between i and j: $C_{\text{ij}} = 10\,C$ if i and j are bonded, and $C_{\text{ij}} = C$ otherwise. C can be determined by fitting the crystallographic B-factors (see below). The use of two force constants for the bonded and nonbonded residue-residue interactions was shown to improve the performance of ENM (21) and GNM (12).

The physical basis for the weighted combination adopted in Eq. 1 is as follows. For a pair of contacting C_α atoms (i, j), $\vec{x}_{\text{j}} - \vec{x}_{\text{i}}$ can be partitioned into the longitudinal (parallel to $\vec{n}_{\text{ij}}^0$) and the transverse (perpendicular to $\vec{n}_{\text{ij}}^0$) components (see Fig. 1). In ENM, the stiffness for the latter component (the curvature of $E_{\text{t}}(x)$ in Fig. 1) is zero, which leads to a fully isotropic (orientation-independent) interaction between i and j. In GNM, both components have the same positive stiffness (same curvature for $E_{\text{t}}(x)$ and $E_{\text{l}}(y)$ in Fig. 1), so the interaction between i and j is anisotropic (orientation-dependent). In G-ANM, f_{anm} gives the ratio of stiffness between the transverse displacement and the longitudinal displacement. Since $f_{\text{anm}} = 0$ corresponds to the isotropic limit, f_{anm} describes the extent of anisotropy in the contact interaction between i and j (thus named an anisotropy weight parameter).

The G-ANM potential energy is reformulated as follows:

$$E_{\text{G-ANM}} = \frac{1}{2}X^{\text{T}}H_{\text{G-ANM}}X = \frac{1}{2}X^{\text{T}}\{f_{\text{anm}}K_{\text{GNM}}\otimes I_3 + (1-f_{\text{anm}})H_{\text{ENM}}\}X, \quad (2)$$

where $X = [\,\vec{x}_0 \quad \dots \quad \vec{x}_{\text{N}-1}\,]^{\text{T}}$ is the 3 N-dimensional displacement vector (N, number of residues or C_α atoms). $H_{\text{ENM}} = \nabla^2 E_{\text{G-ANM}}|_{f_{\text{anm}}=0}$ is the ENM Hessian matrix. K_{GNM} is the N by N Kirchhoff's matrix as defined in GNM, which is constructed as follows (9):

$$K_{\text{ij}} = \begin{cases} -C_{\text{ij}}\theta(R_{\text{c}} - d_{\text{ij}}^0) & i \ne j \\ -\sum_{\text{k}\ne\text{i}} K_{\text{ik}} & i = j \end{cases}, \quad (3)$$

where I_3 is a 3×3 identity matrix, and $\theta(x)$ is the Heaviside function.

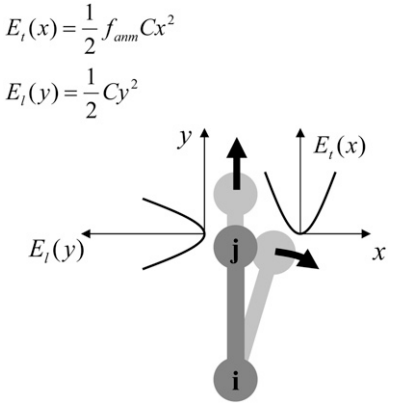

FIGURE 1 Physical basis of G-ANM potential function (see Eq. 1). The relative displacement between C_α atom j and i ($\vec{x}_{\text{j}} - \vec{x}_{\text{i}}$) is partitioned into longitudinal and transverse components that have different stiffness (see Methods for details).

At $f_{\text{anm}} = 0$, $E_{\text{G-ANM}} = (1/2)X^{\text{T}}H_{\text{ENM}}X$, and the G-ANM is reduced to an ENM. Note that the ENM potential is normally expanded in a quadratic form:

$$E_{\text{ENM}} = \frac{1}{2}\sum_{\text{i,j}:d_{\text{ij}}^0 < R_{\text{c}}} C_{\text{ij}}|d_{\text{ij}} - d_{\text{ij}}^0|^2 \approx \frac{1}{2}\sum_{\text{i,j}:d_{\text{ij}}^0 < R_{\text{c}}} C_{\text{ij}}[(\vec{x}_{\text{i}} - \vec{x}_{\text{j}})\cdot\vec{n}_{\text{ij}}^0]^2$$

(d_{ij}^0 is the distance between C_α atom i and j at equilibrium).

At $f_{\text{anm}} = 1$, $E_{\text{G-ANM}} = (1/2)X^{\text{T}}\{K_{\text{GNM}}\otimes I_3\}X = (1/2)Y^{\text{T}}K_{\text{GNM}}Y$, where $Y = [\,|\vec{x}_0| \quad \dots \quad |\vec{x}_{\text{N}-1}|\,]^{\text{T}}$, and the G-ANM is reduced to a GNM (22).

Therefore G-ANM unifies GNM and ENM as its two limits.

For the Hessian matrix $H_{\text{G-ANM}}$ in Eq. 2, we perform the normal mode analysis, which yields 3 N-3 nonzero modes and 3 zero modes (corresponding to 3 translations) for $f_{\text{anm}} > 0$, and 3 N-6 nonzero modes and 6 zero modes (corresponding to 3 translations and 3 rotations) for $f_{\text{anm}} = 0$ (the ENM limit).

Evaluation of G-ANM in describing the crystallographic B-factors

By summing the nonzero modes of G-ANM, we compute the isotropic thermal fluctuations F_{i} to simulate the isotropic crystallographic B-factor B_{i} in a crystal structure as follows:

$$\frac{F_{\text{i}}}{8\pi^2} = \frac{k_{\text{B}}T_{\text{crystal}}}{3}\sum_{\text{m}}\frac{\vec{v}_{\text{m,i}}^2}{\lambda_{\text{m}}}, \quad (4)$$

where k_{B} is the Boltzmann constant, $\vec{v}_{\text{m,i}}$ is the 3D component of the eigenvector of mode m at C_α atom i, λ_{m} is the eigenvalue of mode m, and T_{crystal} is the crystallographic temperature. The quality of G-ANM in fitting the B-factors is assessed by the cross-correlation coefficient $CC = (\sum_{\text{i}}(F_{\text{i}} - \langle F_{\text{i}}\rangle)\cdot(B_{\text{i}} - \langle B_{\text{i}}\rangle)/\sqrt{\sum_{\text{i}}(F_{\text{i}} - \langle F_{\text{i}}\rangle)^2 \cdot \sum_{\text{i}}(B_{\text{i}} - \langle B_{\text{i}}\rangle)^2})$, where $\langle F_{\text{i}}\rangle(\langle B_{\text{i}}\rangle)$ is the arithmetic average of $F_{\text{i}}(B_{\text{i}})$ over all C_α atoms.

For each test case, we compute the cross-correlation coefficient (CC) as a function of f_{anm} and R_{c} (we only fit the B-factors of the first structure of the pair of structures in each test case). To remove sample heterogeneity, CC is normalized to $CC_{\text{norm}} = CC/CC_{\text{max}}$, where $CC_{\text{max}} = Max_{\substack{0\le f_{\text{anm}}\le 1 \\ 7\le R_{\text{c}}\le 20}}\{CC(f_{\text{anm}}, R_{\text{c}})\}$. Then the average ($CC_{\text{norm}}^{\text{AVG}}$) and standard deviation ($CC_{\text{norm}}^{\text{SD}}$) are computed for CC_{norm} among a selected list of 18 test cases (Table 1). A high quality of G-ANM in fitting B-factors is reflected by a high (low) value of $CC_{\text{norm}}^{\text{AVG}}$ ($CC_{\text{norm}}^{\text{SD}}$).

Evaluation of G-ANM in describing the observed conformational changes

The quality of G-ANM in describing the observed conformational changes is assessed by the cumulative overlap (CO) for the lowest 15 modes: $CO = \sqrt{\sum_{\text{m}=1}^{15} I_{\text{m}}^2}$, where $I_{\text{m}} = (\sum_{\text{i}}\vec{v}_{\text{m,i}}\cdot\vec{x}_{\text{obs,i}}/\sqrt{\sum_{\text{i}}|\vec{x}_{\text{obs,i}}|^2})$, $\vec{v}_{\text{m,i}}$ is the 3D component of the eigenvector of mode m at C_α atom i, and $\vec{x}_{\text{obs,i}}$ is the observed structural displacement at C_α atom i. We perform a similar normalization for CO and then compute the average ($CO_{\text{norm}}^{\text{AVG}}$) and standard deviation ($CO_{\text{norm}}^{\text{SD}}$) of CO_{norm} for the 18 selected test cases (Table 1). A high quality of G-ANM in describing the observed conformational changes is embodied by a high (low) value of $CO_{\text{norm}}^{\text{AVG}}$ ($CO_{\text{norm}}^{\text{SD}}$).

Comparison between the lowest modes of G-ANM and ENM

We compute the cumulative similarity score between the lowest 15 modes of the G-ANM and that of the ENM: $SIM = (1/15)\sum_{\text{m}_2=1}^{15}\sum_{\text{m}_1=1}^{15} I(m_1, m_2)^2$,

TABLE 1 List of 22 pairs of protein structures from PDB

Protein	No. of residues	PDB codes and chains
Adenylate kinase	218	1aky, 2ak3A
Alcohol dehydrogenase	373	8adh, 6adhA
Annexin V	317	1avr, 1avhA
Calmodulin	144	**1cll, 1ctr**
Che Y protein	128	3chy, 1chn
Enolase	436	**3enl, 7enl**
HIV-1 protease	99	**1hhp, 1ajxA**
Lactoferrin	691	1lfh, 1lfg
LAO binding protein	238	2lao, 1lst
Maltodextrin binding protein	370	1omp, 1anf
Thymidylate synthase	264	3tms, 2tscA
Triglyceride lipase	265	3tgl, 4tgl
Tyrosine phosphatase	278	1yptA, 1lyts
Guanylate kinase	186	**1ex7A, 1ex6A**
Serum transferrin	328	1bp5A, 1a8e
Ras p21 protein catalytic domain	169	4q21, 5p21
Transducin-α	314	1tag, 1tndA
5-Enol-pyruvyl-3-phosphate synthase	427	1eps, 1g6sA
Oligo-peptide binding protein	517	1rkm, 2rkmA
RNA helicase	435	8ohm, 1cu1A
Myosin	730	1vom, 1mma
Rb69 DNA polymerase	897	1ih7A, 1ig9

Four pairs eliminated from the analysis are in bold (the selection removes those low-quality test cases if $CC_{\max} < 0.5$ or $CO_{\max} < 0.5$). The remaining 18 pairs are used for the evaluation of G-ANM.

where $I(m_1, m_2) = \sum_i \vec{v}^{\text{G-ANM}}_{m_1,i} \cdot \vec{v}^{\text{ENM}}_{m_2,i}$, and $\vec{v}^{\text{G-ANM}}_{m_1,i}(\vec{v}^{\text{ENM}}_{m_2,i})$ is the 3D component of the eigenvector of mode m_1 (m_2) of the G-ANM (ENM) at C_α atom i. If the two sets of modes span the same subspace, $SIM = 1$; otherwise $0 \le SIM < 1$. We note that as $f_{\text{anm}} \to 0$, $SIM \to 0.8$ instead of 1 because the lowest 3 nonzero modes of the G-ANM converge to the 3 rotational zero modes of the ENM. We compute the average (SIM^{AVG}) and standard deviation (SIM^{SD}) of SIM for the 18 selected test cases as a function of f_{anm} and R_c.

RESULTS

We quantitatively assess the performance of G-ANM in describing both the isotropic thermal fluctuations and the observed conformational changes for a selected list of 18 test cases, each consisting of a pair of protein structures from the PDB. The list (Table 1) is compiled from an early work on ENM (17) and our recent work (23,24). We only include the crystal structures that do not have extensive interface between individual structural units.

Evaluation of G-ANM in describing the crystallographic B-factors

The quality of G-ANM in fitting the crystallographic B-factors is assessed by the CC between theoretical and experimental B-factors (see Methods). We analyze the average ($CC^{\text{AVG}}_{\text{norm}}$) and standard deviation ($CC^{\text{SD}}_{\text{norm}}$) of ''normalized'' CC over a list of selected test cases (see Methods and Table 1).

$CC^{\text{AVG}}_{\text{norm}}$ and $CC^{\text{SD}}_{\text{norm}}$ as a function of f_{anm} and R_c are shown in Fig. 2, *a* and *b*. The f_{anm}-dependence of CC_{norm} at fixed R_c is as follows: for $R_c = 7$ Å (8 Å), $CC^{\text{AVG}}_{\text{norm}}$ peaks at $f_{\text{anm}} = 0.01(0.1)$. For $R_c \ge 10$ Å, the peak shifts to $f_{\text{anm}} = 1$ and its height decreases as R_c increases. For 8 Å $\le R_c \le$ 20 Å, $CC^{\text{SD}}_{\text{norm}}$ rapidly decreases as f_{anm} increases in $0 < f_{\text{anm}} < 0.1$, and it becomes flat or slightly increases as f_{anm} increases in $0.1 \le f_{\text{anm}} \le 1$. The observation that $CC^{\text{AVG}}_{\text{norm}}$ and $CC^{\text{SD}}_{\text{norm}}$ vary substantially in $0 < f_{\text{anm}} \le 0.1$ but change little in $0.1 \le f_{\text{anm}} < 1$ suggests that the introduction of small isotropic interactions (the first term of Eq .1) significantly improves the quality of G-ANM in fitting the B-factors to a level comparable with GNM. This improvement is much more pronounced for $R_c = 7$ Å or 8 Å than for $R_c \ge 10$ Å: for $R_c = 8$ Å, $CC^{\text{AVG}}_{\text{norm}}$ increases significantly from 0.64 at $f_{\text{anm}} = 0$ to 0.94 at $f_{\text{anm}} = 0.1$, and $CC^{\text{SD}}_{\text{norm}}$ decreases sharply from 0.22 at $f_{\text{anm}} =$ 0 to <0.05 at $f_{\text{anm}} = 0.1$.

Then we examine the R_c-dependence of CC_{norm} at fixed f_{anm}: for $f_{\text{anm}} = 0$ (the ENM limit), $CC^{\text{AVG}}_{\text{norm}}$ ($CC^{\text{SD}}_{\text{norm}}$) is maximal (minimal) at $R_c = 20$ Å and minimal (maximal) at $R_c =$ 7 Å, suggesting that the optimized fitting of the B-factors by ENM requires high R_c (beyond the physical interaction range: 4.4 Å ~ 12.8 Å, see Cieplak and Hoang (15)). However, the above R_c-dependence is changed for $10^{-4} < f_{\text{anm}} < 0.1$: the maximum (minimum) of $CC^{\text{AVG}}_{\text{norm}}$ ($CC^{\text{SD}}_{\text{norm}}$) is moved to $R_c =$ 7 Å or 8 Å, which is now within the physical interaction range.

When f_{anm} and R_c are both variable, the optimal fitting of the B-factors by GNM is attained at a physically realistic R_c ~8 Å and a small $f_{\text{anm}} \sim 0.1$, instead of the ENM limit or the GNM limit.

Evaluation of G-ANM in describing the observed conformational changes

The quality of G-ANM in describing the crystallographically observed conformational changes is assessed by the cumulative overlap (CO) between the 15 lowest modes and the observed changes (see Methods). We analyze the average ($CO^{\text{AVG}}_{\text{norm}}$) and standard deviation ($CO^{\text{SD}}_{\text{norm}}$) of ''normalized'' CO over a list of selected test cases (see Methods and Table 1).

$CO^{\text{AVG}}_{\text{norm}}$ and $CO^{\text{SD}}_{\text{norm}}$ as a function of f_{anm} and R_c are shown in Fig. 2, *c* and *d*. The f_{anm}-dependence of CO_{norm} at fixed R_c is as follows: for $R_c =$7 Å (8 Å), $CO^{\text{AVG}}_{\text{norm}}$ is maximal at $f_{\text{anm}} =$ 0.001 (0.003); for $R_c \ge 12$ Å, the maximum shifts toward higher f_{anm} and its height decreases gradually as R_c increases. Similarly, for $R_c = 7$ Å (8 Å), $CO^{\text{SD}}_{\text{norm}}$ is minimal at $f_{\text{anm}} =$ 0.001 (0.003); for $R_c \ge 12$ Å, the minimum shifts toward higher f_{anm} and its value increases gradually as R_c increases. Notably, with the exception of $R_c = 7$ Å, $CO^{\text{AVG}}_{\text{norm}}$ and $CO^{\text{SD}}_{\text{norm}}$ change little in $0 < f_{\text{anm}} < 0.01$, but vary substantially in $0.01 < f_{\text{anm}} \le 1$. Therefore, small isotropic interactions (the first term of Eq .1) do not significantly degrade the quality of G-ANM in describing the observed protein conformational changes when compared with the ENM. Instead, for $R_c = 7$ Å and 8 Å, an improvement in such quality is found.

Next we study the R_c-dependence of CO_{norm} at fixed f_{anm}: for $f_{\text{anm}} = 0$ (the ENM limit), $CO^{\text{AVG}}_{\text{norm}}$ ($CO^{\text{SD}}_{\text{norm}}$) is maximal (minimal) at $R_c = 8$ Å and minimal (maximal) at $R_c = 20$ Å,

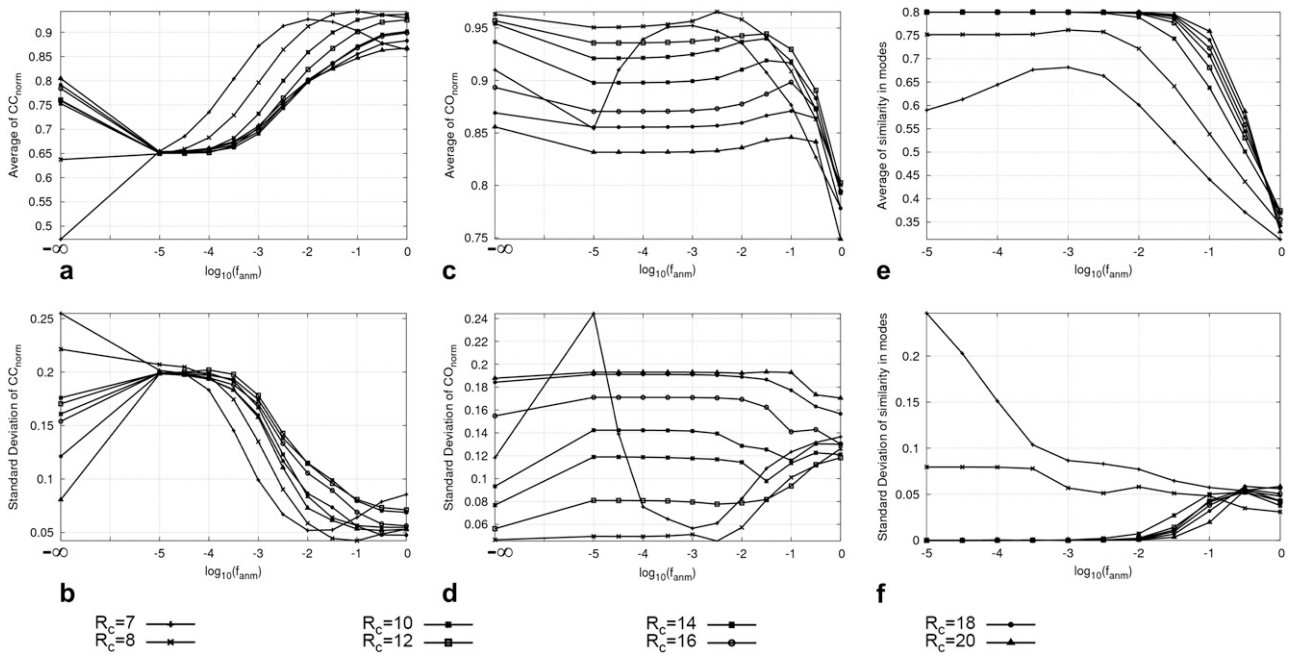

FIGURE 2 Average (*a*) and standard deviation (*b*) of the normalized cross-correlation coefficient between the experimental and theoretical B-factors. Average (*c*) and standard deviation (*d*) of the normalized cumulative overlap for the lowest 15 modes. Average (*e*) and standard deviation (*f*) of the cumulative similarity in the lowest 15 modes between the ENM and the G-ANM.

suggesting that the optimized description of the observed conformational changes by ENM requires a relatively small R_c (contrary to the fitting of B-factors). With the exception of $R_c = 7$ Å, the above R_c-dependence is essentially maintained in $0 < f_{anm} \leq 0.01$, and the maximum (minimum) of CO_{norm}^{AVG} (CO_{norm}^{SD}) remains at $R_c = 8$ Å.

When both f_{anm} and R_c are variable, the optimal description of the observed protein conformational changes is achieved at a physically realistic $R_c = 8$ Å and a small $f_{anm} \sim 0.003$, which is slightly better than at the ENM limit.

Comparison between the lowest modes of G-ANM and ENM

To further understand the f_{anm}-dependence of the quality of G-ANM in describing the observed conformational changes, we will evaluate how much the lowest modes of the G-ANM differ from that of the ENM as f_{anm} varies using a cumulative similarity score *SIM* (see Methods).

The f_{anm}-dependence of SIM^{AVG} at fixed R_c resembles that of CO_{norm}^{AVG} (Fig. 2 *e*): for $R_c = 7$ Å and 8 Å, SIM^{AVG} is peaked at $f_{anm} = 0.001$; for $R_c \geq 10$ Å, the peak disappears and the curve's right-side edge shifts toward higher f_{anm} as R_c increases. For fixed f_{anm}, SIM^{AVG} (SIM^{SD}) is lower (higher) at $R_c = 7$ Å or 8 Å than at $R_c \geq 10$ Å. Thus, for $f_{anm} < 0.01$, the lowest modes of the G-ANM differ significantly from that of the ENM only if R_c is relatively small ($R_c \leq 8$ Å). Such difference is mainly due to the occurrence of extra zero modes in ENM for $R_c \leq 8$ Å (in addition to the 6 translational and rotational zero modes), which overestimate the mobility of the sparsely connected regions in ENM (such as a surface loop). The addition of the isotropic interaction energy (the first term of Eq. 1) in the G-ANM eliminates these additional zero modes in all the 22 test cases, thus removing a major source of errors in ENM.

DISCUSSIONS AND CONCLUSIONS

This work is, to our knowledge, the first attempt to unify GNM and ENM for the simultaneous modeling of both the thermal fluctuations and conformational motions in protein structures, despite recent efforts for model improvement within the framework of either GNM (12) or ENM (25). Our optimal solution is a generalized anisotropic network model parameterized with a physically realistic cutoff distance $R_c = 8$ Å and a small anisotropy parameter $f_{anm} \leq 0.1$. The optimal values of f_{anm} for describing thermal fluctuations and conformational motions are both small, although they are numerically different: the former (~0.1) is higher than the latter (~0.003). The contradicting parameter optimizations in ENM (the B-factors' fitting demands high R_c whereas the description of observed conformational changes requires low R_c) are resolved in G-ANM: the optimal descriptions of both quantities are achieved at $R_c \sim 8$ Å.

The use of a relatively small R_c is more advantageous because:

1. It agrees with the physical range of residue-residue contact interactions (including van der Waals and screened electrostatic interactions).

2. It enables more realistic modeling of medium-range (8–20 Å) interactions and couplings, which are crucial in allostery but obscured by large R_c.
3. Smaller R_c also leads to lower computational cost in the normal mode analysis of ENM (or GNM), because the Hessian (or Kirchhoff) matrix is more sparse (note that the Hessian matrix of G-ANM is as sparse as that of ENM; thus the computational cost of the G-ANM is similar to that of ENM).

Due to the anisotropic geometry of amino acid side chains, the physical interactions between two contacting residues are intrinsically anisotropic: they depend on both the distance and the orientation between the two residues. The orientation-dependence, which is absent in the ENM potential, is incorporated in the G-ANM by introducing a new parameter $f_{anm} > 0$ that defines the extent of anisotropy between the longitudinal and transverse motions between pairs of contacting residues. Our result, in favor of a small f_{anm}, suggests that the transverse motions are far less restrained energetically than the longitudinal motions, which may be explained by the high flexibility of side chains that facilitates easy accommodation to transverse motions between residues. This finding also validates the ENM as the zero order approximation to the G-ANM.

In our future studies, through proper parameterization of G-ANM (fitting the thermal fluctuations and/or the observed conformational changes with f_{anm} and R_c), we will strive to probe several key aspects of conformational dynamics in proteins such as the allosteric couplings (21,26) and the ligand-binding induced conformational motions (27). It will be interesting to assess the performance G-ANM in describing the anisotropic displacement parameters from crystallography (28) or structural fluctuations from NMR data (29). Comparison with other efforts to improve GNM (for example, see Song and Jernigan (30) and Erman (31)) also will be useful.

REFERENCES

1. Tozzini, V. 2005. Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* 15:144–150.
2. Ma, J. 2005. Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure.* 13:373–380.
3. Bahar, I., and A. Rader. 2005. Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* 15:586–592.
4. Tama, F., and C. L. Brooks. 2006. Symmetry, form, and shape: guiding principles for robustness in macromolecular machines. *Annu. Rev. Biophys. Biomol. Struct.* 35:115–133.
5. Karplus, M., and J. A. McCammon. 2002. Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* 9:646–652.
6. Tirion, M. M. 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905–1908.
7. Hinsen, K. 1998. Analysis of domain motions by approximate normal mode calculations. *Proteins.* 33:417–429.
8. Atilgan, A. R., S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80:505–515.
9. Haliloglu, T., I. Bahar, and B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.
10. Bahar, I., A. R. Atilgan, and B. Erman. 1997. Direct evaluation of thermal fluctuations in proteins using a single parameter harmonic potential. *Fold. Des.* 2:173–181.
11. Kundu, S., J. S. Melton, D. C. Sorensen, and G. N. Phillips Jr. 2002. Dynamics of proteins in crystals: comparison of experiment with simple models. *Biophys. J.* 83:723–732.
12. Kondrashov, D. A., Q. Cui, and G. N. Phillips Jr. 2006. Optimization and evaluation of a coarse-grained model of protein motion using X-ray crystal data. *Biophys. J.* 91:2760–2767.
13. Eyal, E., L. W. Yang, and I. Bahar. 2006. Anisotropic network model: systematic evaluation and a new web interface. *Bioinformatics.* 22: 2619–2627.
14. Yang, L. W., A. J. Rader, X. Liu, C. J. Jursa, S. C. Chen, H. A. Karimi, and I. Bahar. 2006. oGNM: online computation of structural dynamics using the Gaussian Network Model. *Nucleic Acids Res.* 34:W24–31.
15. Cieplak, M., and T. X. Hoang. 2003. Universality classes in folding times of proteins. *Biophys. J.* 84:475–488.
16. Cui, Q., and I. Bahar, editors. 2006. Normal mode analysis. *In* Theory and Applications to Biological and Chemical Systems. CRC press, Taylor & Francis Group, Boca Raton, FL.
17. Tama, F., and Y. H. Sanejouand. 2001. Conformational change of proteins arising from normal mode calculations. *Protein Eng.* 14:1–6.
18. Delarue, M., and Y. H. Sanejouand. 2002. Simplified normal mode analysis of conformational transitions in DNA-dependent polymerases: the elastic network model. *J. Mol. Biol.* 320:1011–1024.
19. Zheng, W., and S. Doniach. 2003. A comparative study of motor-protein motions by using a simple elastic-network model. *Proc. Natl. Acad. Sci. USA.* 100:13253–13258.
20. Nicolay, S., and Y. H. Sanejouand. 2006. Functional modes of proteins are among the most robust. *Phys. Rev. Lett.* 96:78104–78107.
21. Ming, D., and M. E. Wall. 2005. Allostery in a coarse-grained model of protein dynamics. *Phys. Rev. Lett.* 95:198103.
22. Chennubhotla, C., A. J. Rader, L. W. Yang, and I. Bahar. 2005. Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies. *Phys. Biol.* 2:S173–S180.
23. Zheng, W., and B. R. Brooks. 2005. Normal modes based prediction of protein conformational changes guided by distance constraints. *Biophys. J.* 88:3109–3117.
24. Zheng, W., and B. R. Brooks. 2006. Modeling protein conformational changes by iterative fitting of distance constraints using re-oriented normal modes. *Biophys. J.* 90:4327–4336.
25. Song, G., and R. L. Jernigan. 2006. An enhanced elastic network model to represent the motions of domain-swapped proteins. *Proteins.* 63:197–209.
26. Zheng, W., and B. R. Brooks. 2005. Identification of dynamical correlations within the myosin motor domain by the normal mode analysis of an elastic network model. *J. Mol. Biol.* 346:745–759.
27. Zheng, W., and B. R. Brooks. 2005. Probing the local dynamics of nucleotide-binding pocket coupled to the global dynamics: myosin versus kinesin. *Biophys. J.* 89:167–178.
28. Kondrashov, D. A., A. W. Van Wynsberghe, R. M. Bannen, Q. Cui, and G. N. Phillips, Jr. 2007. Protein structural variation in computational models and crystallographic data. *Structure.* 15:169–177.
29. Yang, L. W., E. Eyal, C. Chennubhotla, J. Jee, A. M. Gronenborn, and I. Bahar. 2007. Insights into equilibrium dynamics of proteins from comparison of NMR and X-ray data with computational predictions. *Structure.* 15:741–749.
30. Song, G., and R. L. Jernigan. 2007. vGNM: a better model for understanding the dynamics of proteins in crystals. *J. Mol. Biol.* 369: 880–893.
31. Erman, B. 2006. The Gaussian network model: precise prediction of residue fluctuations and application to binding problems. *Biophys. J.* 91: 3589–3599.