PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Tensorial elastic network model for protein dynamics: Integration of the anisotropic network model with bond-bending and twist elasticities

Amit Srivastava,[1†] Roee Ben Halevi,[1†] Alexander Veksler,[1] and Rony Granek[1,2*]

[1] The Stella and Avram Goren-Goldstein Department of Biotechnology Engineering, Ben-Gurion University of The Negev, Beer Sheva 84105, Israel

[2] The Ilse Katz Institute for Meso and Nanoscale Science and Technology, Ben-Gurion University of The Negev, Beer Sheva 84105, Israel

ABSTRACT

We present a tensorial elastic network model (TNM) to describe the equilibrium fluctuations of proteins near their native fold structure. The model combines the anisotropic network model (ANM), bond bending elasticity, and backbone twist elasticity, and can predict both the isotropic fluctuations, similar to the Gaussian network model (GNM), and anisotropic fluctuations, similar to the ANM. TNM performs equally well for B-factor predictions as GNM and predicts the anisotropy of B-factors better than ANM. The model also outperforms the ANM in its predictability of the complete anisotropic displacement parameters.

Proteins 2012; 80:2692–2700.

Key words: elastic network models; normal mode analysis; Gaussian network model; anisotropic network model; protein data bank; bond bending; twist elasticity; anisotropic displacement parameters.

INTRODUCTION

Proteins are the most abundant macromolecules, occurring in all cells and all parts of cells, and their flexibility is believed to play an important role in their function. It is well known that a protein in its native fold state is not static. Instead, residues fluctuate in the neighborhood of their native state position, a simple manifestation of equilibrium fluctuations. Various methods, such as molecular dynamics simulations[1–4] and normal mode analysis[5,6] have been used to study protein dynamics, but these methods use complicated potential functions and are computationally expensive. Instead of using an all atom potential[3], Trinion[7] showed that the normal mode analysis of a network of pairwise harmonic potentials, with identical spring constants, gives almost similar results to those calculated by the more complicated potentials.

In the last decade, various coarse-grained elastic network models[8–24] have been used to study the large scale collective motion of proteins near their native fold structure. These models originated from the Flory theory of polymer networks[25] and use the seminal idea of Trinion.[7] Among these models the Gaussian network model (GNM)[8,9] and the anisotropic network model (ANM)[10,11] have been widely used. The GNM is a topology based model[9] where each residue is represented by its α-carbon (C_α), and the interactions with its neighbors are replaced by harmonic springs. The GNM has been successful in predicting x-ray crystallographic temperature factors (B-factors),[8,9,26–29] H/D exchange free energies,[30] and order parameters of NMR-relaxation measurements.[31] Kundu *et al.*[32] calculated the B-factors for a set of 113 proteins and found that the GNM yields good correlations with the experimental isotropic B-factor data. Sen *et al.*[33] showed that the correlations between the experimental and GNM predicted B-factors are very similar to those of atomic resolution and coarse-grained level models. Despite its enormous success, the GNM has some inherent weaknesses. GNM predicts that residue fluctuations are isotropic in nature. The protein is viewed as a collection of N beads, one for each residue,

Additional Supporting Information may be found in the online version of this article.

[†]Amit Srivastava and Roee Ben Halevi contributed equally to this work.

*Correspondence to: Rony Granek, The Stella and Avram Goren-Goldstein Department of Biotechnology Engineering, Ben-Gurion University of The Negev, Beer Sheva 84105, Israel. E-mail: rgranek@bgu.ac.il

Received 6 May 2012; Revised 3 July 2012; Accepted 16 July 2012
Published online 28 July 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24153

resulting with an ensemble of $3N-3$ nontrivial independent modes ($N-1$ for each axis), instead of $3N-6$ modes that should be obtained in a 3-dimensional description. In other words, GNM cannot predict a possible enhancement or suppression of fluctuations in a particular direction. However, in reality the fluctuations are anisotropic[34] which can strongly influence protein function.

To study the anisotropy of fluctuations, Atilgan *et al.*[10] proposed the ANM. The ANM potential depends exclusively on the magnitude of inter-residue distances and does not penalize any change in the orientation of the inter residue vectors. ANM has been widely used to describe the collective dynamics of biomolecular systems such as hemoglobin,[35] motor proteins,[36] and membrane proteins.[37] Using the ANM, Jernigan and coworkers[38] computed anisotropic displacement correlation functions for a large set of high resolution proteins with moderate compatibility with experimental data.

Although ANM has been proven quite successful in predicting the anisotropy of fluctuations,[11,38] it performs worse than GNM in reproducing the isotropic B-factors.[38] More importantly, while a distance cutoff of about 6–8 Å[8] is usually used for GNM based calculations – a reasonable value that accounts only for the first shell neighbor interactions – for ANM these cutoff values often lead to more than six modes with vanishing frequency, and to a few or several negative B-factors. Such results are clearly not physically acceptable and, in fact, signify an instability of the protein. A much higher cutoff, in the range 12 to 24 Å,[39] was needed to obtain positive B-factors that correlate well with experimental data.[40] The latter cutoff is much beyond the first shell neighbors and the range of strong interactions such as hydrogen bonds and screened electrostatics, which makes it unphysical.

Some insight on this ANM limitation is gained by solving ANM on lattices. For example, consider a 2D square lattice having $N = M \times M$ masses. Using the ANM – limited to nearest-neighbor springs – for this lattice yields about $2M = 2\sqrt{N}$ normal modes having a vanishing frequency (on top of the four trivial modes associated with rotation and translation). These modes are essentially harmonics of coherent shear of rows or columns (with no strain along the rows/columns) for which the nearest-neighbor ANM gives no energy penalty. When next-nearest-neighbor (diagonal) springs are added, which is equivalent to increasing the cutoff distance just above the next-nearest-neighbor distance, these modes now have finite frequency due to the energy penalty of the diagonal springs. Similarly, in the ANM that includes next-nearest-neighbor springs the huge degeneracy existing in finite frequency modes (with only nearest-neighbors springs) is removed. While proteins are not made of simple lattices, β-sheet strands are lattice-like, suggesting increasing failure of ANM – used with 6–8 Å cutoff distance – with increasing content of β-sheets in the protein. Moreover, according to the Maxwell counting criteria,[41] a random network is mechanically stable if the average coordination number (the average number of connectivity per site) exceeds $z = 2d$, where d is the dimensionality of the network (i.e. $z = 6$ in 3D). As surface residues are lightly packed, using ANM with a small cutoff distance 6–8 Å leads to an average connectivity of surface residues per α-carbon less than $z = 6$, which can lead to one or several vanishing normal mode frequencies. When a higher cutoff, in the range 12 to 24 Å is used, a highly packed stable network is formed with no vanishing normal mode frequency (besides the six trivial ones). However, as mentioned above, this is an artificial solution to a rather fundamental problem associated with the ANM.

In this article, we introduce a tensorial elastic network model (TNM), which combines ANM, bending elasticity, and backbone twist elasticity, to study the large scale collective motion of proteins. Our model removes completely the above mentioned inherent instability associated with the ANM, even for cutoff distances as small as 6 Å, due to the extra elastic contributions. ANM includes only two body interactions between connected residues, but the bending and twist elasticities produce three and four body interaction terms that stabilize the network. We test our model for isotropic B-factor dataset of 108 proteins, and for anisotropic (x-ray deduced) B-factor entries of five proteins.[42] We find that TNM performs equally well as GNM for the isotropic B-factors, and much better than ANM for the anisotropic B-factors.

METHOD

Tensorial elastic network model

The TNM describes the three dimensional structure of a protein as an elastic network of α-carbons (C_α), in which C_α pairs residing below a cutoff distance R_c are elastically connected (bonded). The TNM penalizes bond stretch-compress, bond-bending, and backbone-twist fluctuations. Similar to GNM and ANM, the energy parameters are chosen to be uniform between residues within one protein, thus leading to three independent energy parameters per protein. The TNM potential can be written as

$$\begin{aligned} H_{\mathrm{TEM}} &= H_{\mathrm{ANM}} + H_{\mathrm{BB}} + H_{\mathrm{twist}} \\ &= \frac{1}{2}\gamma \sum_{\{i,j\}} \Gamma_{ij}[(\mathbf{u_i} - \mathbf{u_j}).\hat{\mathbf{r}}_{\mathbf{ij}}]^2 + \frac{1}{2}B \sum_{\{i,j,k\}} \Gamma_{ik}\Gamma_{jk}(\delta\theta_{ikj})^2 \\ &\quad + \frac{1}{2}K \sum_{\{i,j,k,l\}} (\delta\phi_{ijkl})^2 \end{aligned} \tag{1}$$

where the first sum runs over all residue pairs (each pair counted once), the second sum runs over all residue triplets (each triplet counted once), and the third sum runs over all consecutive residue quartets along the backbone (each quartet counted once). In Eq. (1), the first term, H_{ANM}, is the ANM potential that penalizes stretch-com-

press fluctuations, whereas H_{BB} and H_{twist} penalizes bond angle and backbone-torsional fluctuations, respectively; Γ_{ij} is a connectivity matrix: $\Gamma_{ij} = 1$ if the distance between the C_α of residues i and j is less than R_c, otherwise $\Gamma_{ij} = 0$; $\{\mathbf{u_i}\}$ are C_α-residue displacement vectors; $\delta\theta_{ikj}$ is the angle-shift formed by any three residues i, j, and k; $\delta\phi_{ijkl}$ is the shift in dihedral angle formed by four consecutive residues along the backbone i, j, k, and l (hence always connected); γ, B, and K are the uniform energy constants associated with stretching, bending and twist elasticities, respectively. Out of these three energy parameters, γ, B, and K, one parameter can be used for choosing the energy scale of the resulting B-factors, henceforth, chosen to be γ. Thus there are in fact only two energy parameters, B/γ and K/γ, that can be tuned to maximize the Pearson's correlation coefficient between theoretical and experimental B-factors.

Expression of $\delta\theta_{ikj}$

In order to find the normal modes of the Hamiltonian Eq. (1), we require to rewrite it using a single set of variables, and these are chosen to be the displacement vectors $\{\mathbf{u_i}\}$. Consider first the bond bending angle variables. Assume three residues i, k, and j, and the bonds α and β that connect them, $\alpha = \langle i,k\rangle$ and $\beta = \langle k,j\rangle$. Let θ_0 be the angle between the bonds α and β, when the protein is in its native state. Under a small but otherwise arbitrary deformation of the protein, the change in this angle $\delta\theta_{ikj}$ can be expressed to linear order in the displacements $\{\mathbf{u_i}\}$ in the following way

$$\delta\theta_{ikj} = \begin{cases} \cot\theta_0 \left[\frac{\hat{\mathbf{b}}_\beta \cdot \Delta u_{jk}}{|\mathbf{b}_\beta|^2} + \frac{\hat{\mathbf{b}}_\alpha \cdot \Delta u_{ik}}{|\mathbf{b}_\alpha|^2} - \frac{\mathbf{b}_\alpha \cdot \Delta \mathbf{u_{jk}} + \mathbf{b}_\beta \cdot \Delta \mathbf{u_{ik}}}{\mathbf{b}_\alpha \cdot \mathbf{b}_\beta} \right], & \text{for } \theta_0 \neq 90^0, 180^0 \\ -\frac{\hat{\mathbf{b}}_\alpha \cdot \Delta u_{jk}}{|\mathbf{b}_\beta|^2} - \frac{\hat{\mathbf{b}}_\beta \cdot \Delta u_{ik}}{|\mathbf{b}_\alpha|^2}, & \text{for } \theta_0 = 90^0 \\ \left| \frac{\Delta \mathbf{u_{ik}}}{|\mathbf{b}_\alpha|^2} \times \hat{\mathbf{b}}_\beta - \frac{\Delta \mathbf{u_{jk}}}{|\mathbf{b}_\beta|^2} \times \hat{\mathbf{b}}_\alpha \right|, & \text{for } \theta_0 = 180^0 \end{cases} \quad (2)$$

Here $\mathbf{b}_\alpha$ and $\mathbf{b}_\beta$ are the bond vectors in the unperturbed protein, $\hat{\mathbf{b}}_\alpha$ and $\hat{\mathbf{b}}_\beta$ are the corresponding unit vectors, $\Delta\mathbf{u_{ik}} = \mathbf{u_i} - \mathbf{u_k}$ and $\Delta\mathbf{u_{jk}} = \mathbf{u_j} - \mathbf{u_k}$.

Expression of $\delta\phi_{ijkl}$

Consider four consecutive residues along the backbone $\{i,j,k,l\}$, with the corresponding (peptide) bonds $\alpha = \langle i,j\rangle$, $\beta = \langle j,k\rangle$, and $\nu = \langle k,l\rangle$. Let ϕ_0 be the dihedral angle formed by these residues, when the protein is in its native state. The change in dihedral angle can be expressed as

$$\begin{aligned} \delta\phi_{ijkl} &= \cot\phi_0 \left[\frac{\mathbf{b}_\alpha \times \mathbf{b}_\beta}{|\mathbf{b}_\alpha \times \mathbf{b}_\beta|^2} \{\mathbf{b}_\alpha \times \mathbf{u_k} - \mathbf{b}_\alpha \times \mathbf{u_j} + \mathbf{u_j} \times \mathbf{b}_\beta - \mathbf{u_i} \times \mathbf{b}_\beta\} \right. \\ &+ \frac{\mathbf{b}_\beta \times \mathbf{b}_\nu}{|\mathbf{b}_\beta \times \mathbf{b}_\nu|^2} \{\mathbf{b}_\beta \times \mathbf{u_l} - \mathbf{b}_\beta \times \mathbf{u_k} + \mathbf{u_k} \times \mathbf{b}_\nu - \mathbf{u_j} \times \mathbf{b}_\nu\} \\ &- \frac{(\mathbf{b}_\alpha \times \mathbf{b}_\beta)\{\mathbf{b}_\beta \times \mathbf{u_l} - \mathbf{b}_\beta \times \mathbf{u_k} + \mathbf{u_k} \times \mathbf{b}_\nu - \mathbf{u_j} \times \mathbf{b}_\nu\}}{(\mathbf{b}_\alpha \times \mathbf{b}_\beta) \cdot (\mathbf{b}_\beta \times \mathbf{b}_\nu)} \\ &\left. - \frac{(\mathbf{b}_\beta \times \mathbf{b}_\nu)\{\mathbf{b}_\alpha \times \mathbf{u_k} - \mathbf{b}_\alpha \times \mathbf{u_j} + \mathbf{u_j} \times \mathbf{b}_\beta - \mathbf{u_i} \times \mathbf{b}_\beta\}}{(\mathbf{b}_\alpha \times \mathbf{b}_\beta) \cdot (\mathbf{b}_\beta \times \mathbf{b}_\nu)} \right] \end{aligned} \quad (3)$$

where $\mathbf{b}_\alpha$, $\mathbf{b}_\beta$, and $\mathbf{b}_\nu$ are the corresponding bond vectors in the native state.

Mean square displacement

The mean square displacement (MSD) of residue i can be written as

$$< \mathbf{u}_i^2 >= \frac{k_B T}{\gamma} \sum_{\beta=1}^{3N-6} \frac{1}{\omega_\beta^2} \left[(V_\beta^{i,x})^2 + (V_\beta^{i,y})^2 + (V_\beta^{i,z})^2 \right] \quad (4)$$

where ω_β is the eigen-frequency of mode β, $V_\beta^{i,x}$, $V_\beta^{i,y}$, and $V_\beta^{i,z}$ are elements of the mode β eigenvector associated with displacements along the x,y, and z axes, respectively, of the i^{th} residue.

Isotropic and anisotropic B-factors

Residue fluctuations are usually not isotropic in nature mainly due to the inherent anisotropy of the native fold. One can still define an isotropic B-factor according to

$$B_i = 8\pi^2 \langle \mathbf{u}_i^2 \rangle = 8\pi^2 \langle u_{x,i}^2 + u_{y,i}^2 + u_{z,i}^2 \rangle \quad (5)$$

Isotropic B-factors data are known essentially for all proteins whose structure is available, as they can be deduced from the x-ray Debye-Waller factors or from NMR structure studies. In GNM all directions are equivalent, which leads to the expression $B_i = (8\pi^2/3) < u_i^2 >$ where $u_i \equiv u_{\nu,i}$ is the residue displacement along one of the axes $\nu = x,y,z$. As most PDB files do not contain anisotropy data, GNM is usually sufficient to describe the data.

Several PDB files do contain, however, data describing the anisotropy of fluctuations. A set of six anisotropic displacement parameters (ADP) is usually available per atom in high resolution PDB files. The ADPs are expressed as 3×3 symmetric tensor $\tilde{U}$,

$$\tilde{U} = \begin{vmatrix} U^{11} & U^{12} & U^{13} \\ U^{21} & U^{22} & U^{23} \\ U^{31} & U^{32} & U^{33} \end{vmatrix} \quad (6)$$

Here, the diagonal elements are the MSDs along the three axes (i.e., essentially the B-factors), $U^{11} = \langle u_{x,i}^2 \rangle$, $U^{22} = \langle u_{y,i}^2 \rangle$, and $U^{33} = \langle u_{z,i}^2 \rangle$. The off-diagonal elements are the cross correlation between the displacements of atom i along the different axes, $U^{12} = U^{21} = \langle u_{x,i} u_{y,i} \rangle$, $U^{13} = U^{31} = \langle u_{x,i} u_{z,i} \rangle$, and $U^{23} = U^{32} = \langle u_{y,i} u_{z,i} \rangle$.

Pearson's correlation coefficient

The linear correlation coefficient (CC) between the computed B-factors and the experimental data is given by

$$Corr(B_{\text{exp}}, B_{\text{theo}}) = \frac{\sum_{i=1}^{N} (B_i^{\text{theo}} - < B^{\text{theo}} >)(B_i^{\text{exp}} - < B^{\text{exp}} >)}{\left[\sum_{i=1}^{N} (B_i^{\text{theo}} - < B^{\text{theo}} >)^2 \sum_{i=1}^{N} (B_i^{\text{exp}} - < B^{\text{exp}} >)^2\right]^{\frac{1}{2}}} \quad (7)$$

The CC gives a value 1 for perfect correlation, and −1 for perfect anti- correlation.

X-ray data set

A set of 108 proteins, with x-ray data resolution better or equal to 2.0 Å, was taken from the larger set of Kundu *et al.*[32] This set was used to compare between the theoretically predicted isotropic B-factors of GNM, ANM, and TNM, and their experimental values. For the ADP studies, we used five high resolution PDB files (PDB id: 1K6U, 1G6X, 1G66, 1KT5, and 1A6M) that have ADP entries and we compare the experimental values with ANM and TNM based calculations.

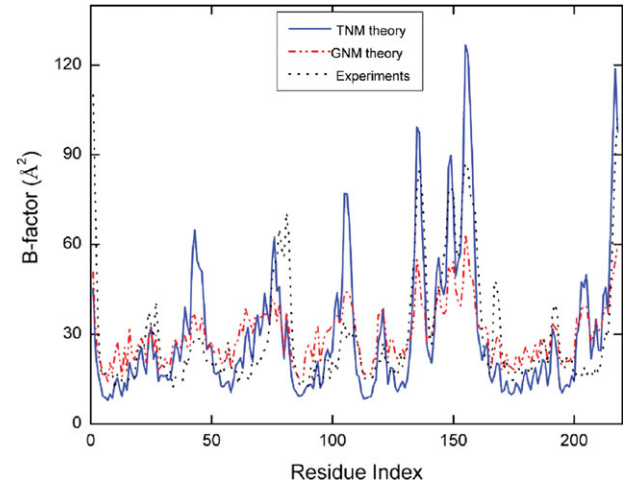

Figure 1

Isotropic B-factors. Plot of isotropic B-factors for adenylate kinase against residue index. Shown are experimental data (dotted line) and theoretical results from: (i) GNM (dash-dot-dot-dash line) and (ii) TNM (solid line). An identical cutoff distance, R_c = 7 Å, was chosen for both models. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

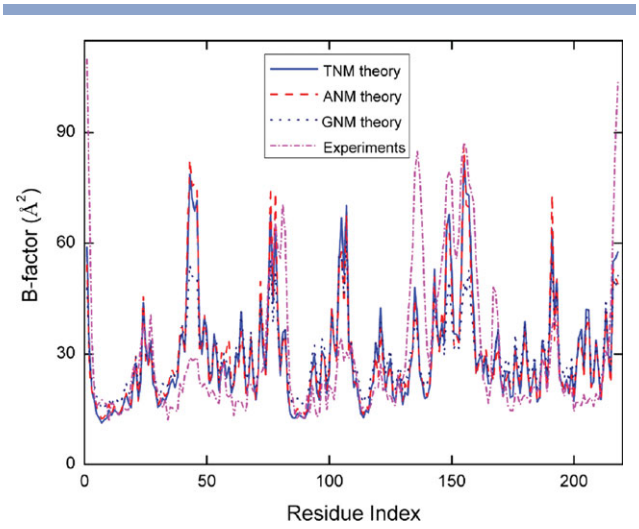

Figure 2

Isotropic B-factors. Plot of isotropic B-factors for adenylate kinase against residue index. Shown are experimental data (dash-dot line) and theoretical results from: (i) GNM (dotted line), (ii) ANM (dashed line), and (iii) TNM (solid line). An identical cutoff distance R_c = 12 Å was chosen for all three models. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

RESULTS AND DISCUSSION

Isotropic B-factors

We first show an example where we computed the isotropic fluctuations using the TNM with a 7 Å cutoff. Figure 1 shows the isotropic B-factors as a function of residue index for adenylate kinase. The B-factors were computed theoretically using the GNM (dash-dot-dot-dash line) and TNM (solid line) with 7 Å cutoff in both. The experimental data are shown by the dotted line. The CC between the experimental and the TNM predicted B-factors is 0.768, which is slightly better than that of the GNM (0.742). ANM was excluded from this analysis as it gives negative values for the B-factors calculated with 7 Å cutoff, as discussed above. We further used for all models a 12 Å cutoff, albeit unphysical, which enabled to include the ANM in the analysis. Figure 2 shows the B-factors against residue index for adenylate kinase for the three different models (TNM, ANM, and GNM) and experiment. The CC between the experiment and TNM is 0.61, better than that of ANM (0.567). The CC between the experiment and GNM (0.541) is slightly worse than those of TNM and ANM. To visualize the "flexibility" (i.e., the fluctuations), we show in Figure 3 the adenylate kinase structure in new cartoon representation, drawn using visual molecular dynamics (VMD).[43]

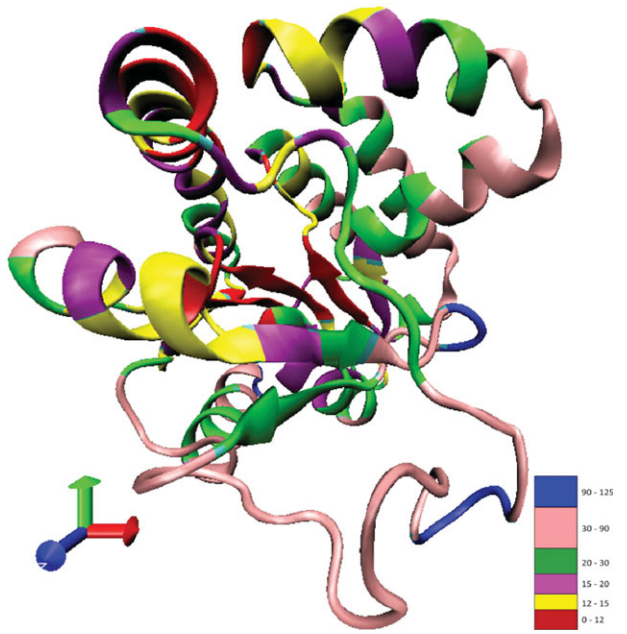

Figure 3

New cartoon representation of adenylate kinase drawn by using VMD.[43] The color coded representation shows the flexibility of different residues as indicated by their TNM calculated B-factors shown in Figure 1.

The different colors in the diagram correspond to the relative magnitude of B-factors computed by TNM with distance cutoff 7 Å. The most constrained region is shown in red color and correspond to the minima in Figure 1 (solid line), and most flexible regions in blue and correspond to the maxima in Figure 1.

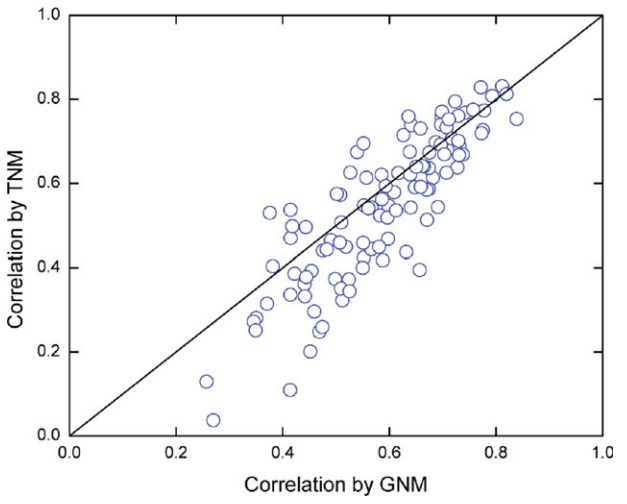

Figure 4

Scatter plot of correlation coefficients obtained by GNM and TNM for 7 Å cutoff. The solid line marks where the correlation coefficients obtained by GNM and TNM are equal. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table I

The Pearson's Correlation Coefficients between the Experimental B-Factors and Theoretically Predicted B-Factors Using Different Models, for a Set of 20 Proteins

PDBID	Length	$Corr_{GNM}$	$Corr^{a}_{TNM}$	$Corr_{TNM}$	B/γ	K/γ
1ARU	344	0.775	0.727	0.768	25.0	10
1AMM	174	0.720	0.684	0.714	19.2	10
1BKF	107	0.422	0.386	0.410	17.8	10
1BPI	50	0.490	0.465	0.494	9.98	10
1CEM	363	0.682	0.613	0.685	14.92	10
1CUS	200	0.732	0.687	0.746	18.0	10
1EDE	310	0.665	0.641	0.684	12.25	10
1EZM	298	0.609	0.579	0.615	20.06	10
2ERL	40	0.777	0.773	0.793	13.0	10
1IAG	202	0.454	0.392	0.440	25.0	10
1LIT	131	0.552	0.548	0.553	2.2	10
2MCM	112	0.819	0.813	0.845	19.7	10
1ONC	104	0.703	0.669	0.710	28.0	10
1PLC	99	0.507	0.459	0.515	22.28	10
1POA	118	0.718	0.707	0.722	28.04	10
1PPN	212	0.675	0.673	0.687	20.0	10
1RA9	159	0.475	0.440	0.484	15.12	10
2RN2	155	0.707	0.626	0.678	30.2	10
1THG	544	0.560	0.541	0.562	13.0	10
1TML	286	0.650	0.639	0.676	19.96	10
Mean		0.635 ± 0.026	0.603 ± 0.028	0.639 ± 0.027		

The 20 proteins are the subset of the complete reported data-set (see Supporting Inforamtion). Here, a uniform cutoff distance 7 Å is chosen for both models. $Corr^{a}_{TNM}$ is the CC between the experimental and TNM-predicted B-factors without using B/γ and K/γ as fit parameters, $B/\gamma = K/\gamma = 1$. $Corr_{TNM}$ is the CC between the experimental and best fitted TNM B-factors, using the best parameters B and K reported in the last two columns.

Similar calculations of the isotropic B-factors were performed for a set of 108 proteins and CCs between the different models and experiment are listed in Table S1 (see Supporting Information) for different cutoff distances. The average values of CCs between experiment and GNM and TNM results are 0.588 and 0.547, respectively, for the 7 Å cutoff (see Table S1, Supporting Information), similar to previous GNM[32] or modified GNM[19,24] studies. CCs range between 0.257(pdb: 1RIS) and 0.838 (pdb: 3PTE) for GNM, and from 0.037(pdb: 1WHI) to 0.832 (pdb: 8ABP) for TNM. The higher values of CCs are thus the same in the both models, but there is a difference in the smallest values in favor of GNM. Moreover, Figure 4 shows that GNM performs better than TNM for 65 percent of proteins in the data set.*

These results suggest that GNM yields better CC for isotropic B-factors than TNM. However, we would like to stress that we have computed the theoretical B-factors for TNM by arbitrarily choosing B/γ and K/γ equal to 1. Tuning these parameters allows us to optimize the CC between the TNM-predicted and experimental B-factor (see footnote*). Table I shows the maximum CC ($Corr_{TNM}$) obtained in this fit for a set of 20 proteins, and the corresponding best values of B/γ and K/γ for each protein. The

*Better CC values for TNM can be obtained by varying the model parameters B/γ and K/γ. See for example, Table S3 in the Supporting Information showing results for the four worse proteins shown in Figure 4.

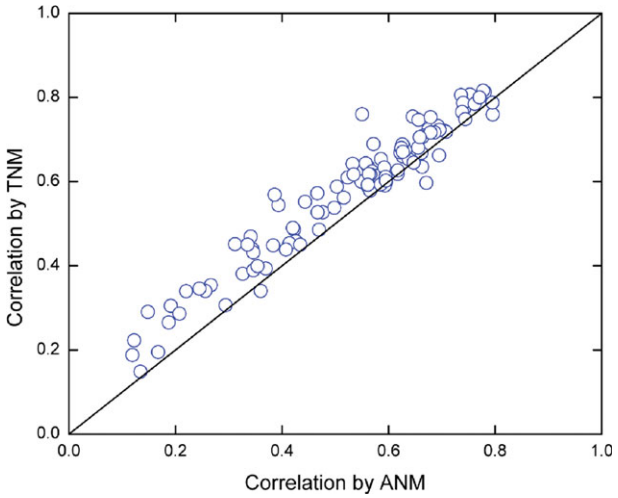

Figure 5

Scatter plot of correlation coefficients obtained by ANM and TNM for 12 Å cutoff. The solid line marks where the correlation coefficient obtained by ANM and TNM are equal. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

mean value of these 20 highest CCs is 0.639, which is almost same as that of the GNM (0.635) for the same data set, although the use of GNM did not require a fit.

Similar comparison has been performed between ANM and TNM, for the identical cutoff distance, $R_c = 12$ Å (e.g., Fig. 2 for adenylate kinase). The mean value of CC for TNM and ANM are 0.580 and 0.528, respectively (see Table S1 in Supporting Information), and values range form 0.15 to 0.82 for TNM and 0.12 to 0.795 for ANM. Figure 5 shows that 94 percent of proteins in the data set perform better for TNM relative to ANM. These results clearly indicate that TNM completely outperform the ANM.[†]

As discussed above, TNM yields anisotropic displacement fluctuations. In Figure 6 we plot the anisotropy of fluctuations for each residue along *x*, *y*, and *z* axes, for the 7 Å cutoff. The anisotropy of fluctuation is measured by subtracting the MSD of a residue along an axis, that is, $\langle u_{x,i}^2 \rangle$, $\langle u_{y,i}^2 \rangle$, or $\langle u_{z,i}^2 \rangle$, from the average MSD ($\langle \mathbf{u}^2 \rangle/3$, essentially the isotropic B-factor) of the same residue. Near zero values indicate residues with nearly isotropic fluctuations. The peaks in the figure correspond to strong anisotropy of fluctuations. Positive peaks along one axis correlate with negative peaks along another axis, as they obviously should. The abundance of these peaks clearly demonstrates that most fluctuations are in fact anisotropic, and we next check the accuracy of these TNM predictions.

To compare the computing efficiency of the GNM, ANM, and TNM, we have calculated the computational time for GNM, ANM and TNM for a dataset of eight proteins with different sizes (see Table S6 in the Supporting Information). We find that TNM takes almost the same time as ANM, whereas GNM takes much less time than both ANM and TNM. Thus, although TNM has additional terms (bending and twist) relative to ANM, it is not more computationally expensive than ANM as its matrix size is the same.

Anisotropic displacement parameters

Due to the improvement in experimental facilities, a large number of high resolution protein structures are now available in the PDB[42] with ADP entries. Jernigan and co-workers[19,38] and Eyal *et al.*[11] have used the ANM to predict the ADP values for a set of 93 proteins. To test the TNM thoroughly, we performed calculations for the minor high resolution set of five proteins. As an example, we consider acetylxylan esterase (pdb: 1G66). Figures 7(a–c) show the comparison between the experimental data and theoretically predicted MSDs along *x*, *y*, and *z* axes. Figure 7(d) shows the comparison between experimental and theoretically predicted overall MSD $\langle \mathbf{u}^2 \rangle$. The CCs are presented in each figure. We find a good correlation between the experimental and theoretical data.

Similar calculations have been performed for the other four proteins in this minor set. The CCs for each of the five proteins, calculated for ANM, GNM, and TNM, are given in Table S2 (see Supporting Information). In addition, the CCs among the experimental and theoretical MSDs along *x*, *y*, and *z* axes are given in Table S4 (see Supporting Information). The CCs in between the *x*, *y*,

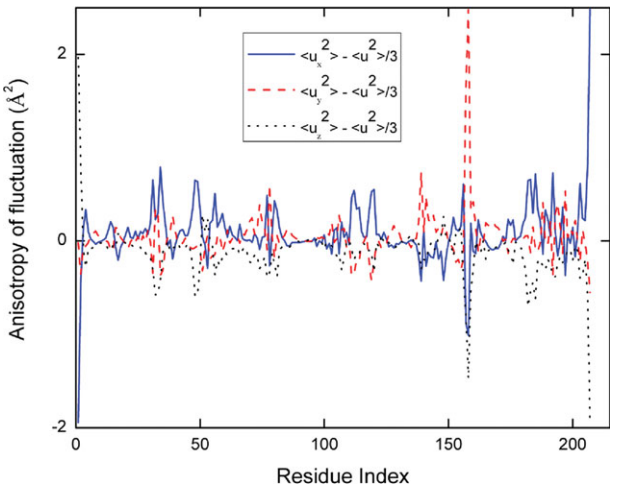

Figure 6

Anisotropy of fluctuations of each residue, along the *x*, *y*, and *z* axes, plotted against residue index for adenylate kinase. The anisotropy of fluctuations are calculated by the TNM with a cutoff distance 7 Å. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

[†]The *p*-value associated with TNM-GNM and TNM-ANM CC data set have been calculated using the *t*-tests for paired sample. The *t*-test rejects the null hypothesis in both cases and gives a *p*-value for TNM-GNM 7×10^{-6}, and for TNM-ANM 2×10^{-21}.

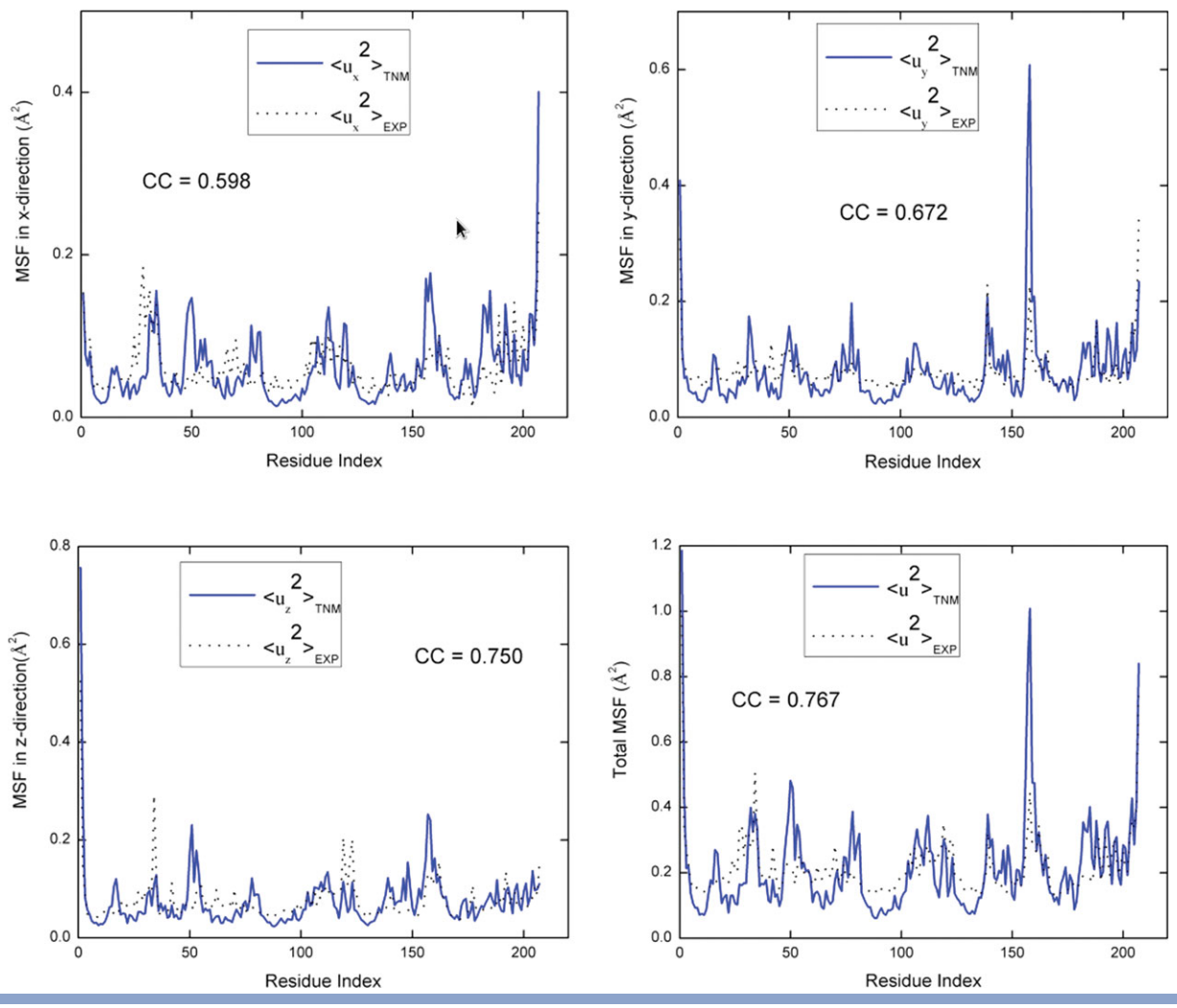

Figure 7

Comparison between the theoretical and experimental MSD of each residue along the three axes: (a) *x*, (b) *y*, (c) *z* axes. In (d) the overall MSDs, associated with the isotropic B-factors, are plotted against residue index. Theoretical results were calculated using the TNM with a cutoff of 7 Å. The CC between experimental and theoretical data sets are shown in each plot. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

and *z* anisotropic B-factors, and between them and the isotropic B-factor, are listed in Table S5 (see Supporting Information). These CC values demonstrate the absence of isotropy, which would yield all values close to 1. The average value of CCs (for all ADP values, diagonal ADPs only, and off diagonal ADPs only) calculated for the three different models are listed in Table II. The average value of the CC of all six ADP entries between experimental data and TNM predictions is 0.862 for the 12 Å cutoff and 0.768 for the 7 Å. This can be considered as a good agreement in light of the experimental uncertainties and theoretical approximations. For ANM with 12 Å distance cutoff, the mean value of CCs of all six ADP entries is 0.751, which is less than 15% of the TNM value for the same cutoff. These results indicate that for these five proteins, TNM outperforms the ANM in its predictions for anisotropy fluctuations.

Table II

Average Correlation Coefficients[a] Between the Experimentally and Theoretically Predicted ADP_S using Different Models

Models	$Corr_{iso}$	$Corr^{all}_{ADP}$	$Corr^{diag}_{ADP}$	$Corr^{offdiag}_{TNM}$
GNM[b]	0.697	–	0.591	–
TNM[b]	0.732	0.768	0.609	0.327
ANM[c]	0.676	0.751	0.557	0.341
TNM[c]	0.714	0.862	0.617	0.429

[a]The values reported here are the average of the correlation coefficients obtained for five different proteins. Correlation coefficients for each protein are listed in Table S2 (see Supporting Information).
[b]distance cutoff = 7 Å.
[c]distance cutoff = 12 Å.
$Corr_{iso}$ is the CC between experimental and theoretically predicted isotropic B-factors. $Corr^{all}_{ADP}$, $Corr^{diag}_{ADP}$, and $Corr^{offdiag}_{TNM}$ are the CCs between the experimental and theoretically computed diagonal, off-diagonal, and all elements of ADP, respectively.

CONCLUSION

In summary, we have presented a tensorial elastic network model which combines the ANM , bond bending elasticity, and backbone twist elasticity. This model can predict both isotropic fluctuations like GNM, and fluctuations anisotropy similar to ANM. We found that TNM performs equally well as GNM for isotropic B-factor predictions, and outperform ANM significantly. We found a good agreement between the theoretical TNM and experimental anisotropic displacement parameters. More importantly, ANM performs reasonably well only when it is used with an unphysical force range cutoff $R_c \gtrsim 12$ Å, while our TNM can be used with the same R_c values as used by the GNM, $R_c \simeq 6 - 8$ Å. We conclude that our TNM can be a potentially useful tool for theoretical studies of anisotropic effects, as can be, for example, in allosteric phenomena.

Moreover, we believe that the model has high potential especially when considering the response of a protein to forces, where it is clear that a tensorial elastic network model is an advantage for a realistic description of protein deformations. This excludes the use of the GNM for such a problem, while our TNM is a much better choice than the ANM. This problem will be explored in the near future.

REFERENCES

1. McCammon JA, Gelin BR, Karplus M. Dynamics of folded protein. Nature 1977;267:585–596.
2. Phillips JC, Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. J Comp Chem 2005;26:1781–1802.
3. Brooks BR, Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculation. J Comp Chem 1983;4: 187–217
4. MacKerell AD, MacKerell Jr AD, Bashford D, Bellott M, Dunbrack Jr RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, McCarthy DJ, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher III WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Kuczera JW, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998; 102:3586–3616.
5. Cui Q, Bahar I. Normal mode analysis. London/Boca Raton; Chapman and Hall/CRC: 2006.
6. Bahar I, Rader AJ. Coarse-grained normal mode analysis in structural biology. Curr Opin Struct Bio 2005;15:586–592.
7. Tirion MM. Large amplitude elastic motions in protein from a single parameter. Phys Rev Lett 1996;77:1905–1908
8. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in protein using a single-parameter harmonic potential. Fold Des, 1997;2:173–181.
9. Halilgolu T, Bahar I, Erman B. Gaussian dynamics of folded proteins. Phys Rev Lett 1997;79:3090–3093.
10. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of protein with an elastic network model. Biophys J, 2001;80:505–515.
11. Eyal E, Eyal E, Chennubhotla C, Yang LW, Bahar I. Anisotropic fluctuations of amino acids in protein structure: insight from X-ray crystallography and elastic network models. Bioinformatics 2007;23:i175–i184
12. Ma J. Usefulness and limitations of normal mode analysis in modeling dynamics of bio-molecular complexes. Structure, 2005;13:373–380.
13. Ming D, Bruschweiler R. Reorientational contact-weighted elastic network model for the prediction of protein dynamics: comparison with NMR and relaxation. Biophys J 2006;90:3382–3388.
14. Song G, Jernigan RL. An enhanced elastic network model to represent the motions of domain-swapped proteins. Proteins 2006; 63:197–209.
15. Song G, Jernigan RL. vGNM: a better model for understanding the dynamics of proteins in crystals. J Mol Bio 2007;369:880–893.
16. Lu M, Poon B, Ma J. A new method for coarse-grained elastic normal-mode analysis. J Chem Theory Comput 2006;2:464–471.
17. Thorpe MF. Comment on elastic network models and proteins. Phys Bio 2007;4:60–63.
18. Soheilifard R, Makarov DE, Rodin GJ. Critical evaluation of simple network models of protein dynamics and their comparison with crystallographic B-factors. Phys Bio 2008;5:026008–0260020.
19. Yang L, Song G, Jernigan RL. Protein elastic network models and the ranges of cooperatively. Proc Natl Acad Sci USA 2009; 106:12347–12352.
20. Zheng W. A unification of the elastic network model and the Gaussian network model for optimal description of protein conformational motions and fluctuations. Biophys J 2008;94:3853–3857.
21. Stembler JN, Wriggers W. Bend-twist-stretch model for coarse elastic network simulation of biomolecular motion. J Chem Phys 2009;131:074112–9.
22. Lin TL, Song G. Generalized spring tensor model for protein fluctuations dynamics and confirmation changes. BMC struct Bio. 2010;10 (supp1):S3(1–12).
23. Mendez R, Bastolla U. Torsional network model: normal modes in torsion angle space better correlate with confirmation changes proteins. Phys Rev Lett 2010;104:228103–4.
24. Yang LW. Models with energy penalty on inter-residue rotation address insufficiency's of conventional elastic network models. Biophys J 2011;100:1784–1793.
25. Flory PJ. Statistical thermodynamics of random networks. Proc R Soc Lond A Math Phys Sci 1976;351:351–378.
26. Bahar I, Atilgan AR, Demirel MC, Erman B. Vibrational dynamics of folded proteins: significance of slow and fast modes in relation to function and stability. Phys Rec Lett 1998;80:2733–2736.
27. Bahar I, Erman B, Jernigan RL, Atilgan AR, Covell DG. Examination of collective motions in HIV-1 reverse transcriptase. Examination of flexibility and and enzyme function. J Mol Bio 1997;266: 195–214.
28. Demirel MC, Demirel MC, Atilgan AR, Bahar I, Jernigan RL, Erman B. Identification of kinetically hot residues in proteins. Protein Sci 1998;7:2522–2532.
29. Keskin O, Jernigan RL, Bahar I. Protein with similar architecture exhibit common large scale dynamics. Biophys J 2000;78:2093–2106.
30. Bahar I, Wallquist A, Covell DG, Jernigan RL. Correlation between native state hydrogen exchange and cooperative residue fluctuations from a simple model. Biochemistry 1998;37:1067–1075
31. Halilgolu T, Bahar I. Structure-based analysis of protein dynamics. Comparison of theoretical results for hen lysozyme with x-ray diffraction and NMR relaxation data. Proteins 1999;37:654–667.
32. Kundu S, Melton JS, Sorensen DC, Phillips GN. Dynamics of proteins in crystals: comparison of experiment with simple models. Biophys J 2002;83:723–732.
33. Sen TZ, Sen TZ, Feng Y, Garcia JV, Kloczkowski A, Jernigan RL. The extent of cooperatively of protein motions observed with elastic network models is similar for atomic and coarse-grained models. J Chem Theory Comput 2006;2:696–704.
34. Ichiye T, Karplus M. Anisotropy and anharmonicity of atomic fluctuations in proteins: analysis of molecular dynamics simulations. Proteins 1987;34:369–382.

35. Xu C, Tobi D, Bahar I. Allosteric changes in protein structure computed by a simple mechanical model: hemoglobin T ⟨–⟩ R2 transition. J Mol Bio 2003;2333:153–168.
36. Zheng W, Doniach S. A comparative study study of motor-protein motions by using a simple elastic network model. Proc Natl Acad Sci 2003;100:13253–13258.
37. Bahar I, Bahar I, Lezon TR, Bakan A, Shrivastava IH. Normal mode analysis of biomolecular structures: functional mechanism of membrane proteins. Chem Rev 2010;110: 1463–1497.
38. Yang L, Song G, Jernigan RL. Comparison of experimental and computed protein anisotropic temperature factors. Proteins 2009; 76:164–175.
39. Eyal E, Yang LW, Bahar I. Anisotropic network model: systematic evaluation and a new web interface. Bioinformatics 2006;22:2619–2627
40. Cieplak M, Hoang TX. Universality class in folding times of proteins. Biophys J 2003;84:475–488.
41. Maxwell JC. On the calculation of the equilibrium and stiffness of frames. Phil Mag. 1864;27:294–299.
42. Protein Data Bank, available at: http://www.pdb.org. Accessed March 20, 2012.
43. Humphrey W, Dalke A, Schulten K. VMD - visual molecular dynamics. J Mol Graphics 1996;14:33–38.