PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Deriving protein dynamical properties from weighted protein contact number

Chih-Peng Lin, Shao-Wei Huang, Yan-Long Lai, Shih-Chung Yen, Chien-Hua Shih, Chih-Hao Lu, Cuen-Chao Huang, and Jenn-Kang Hwang*

Institute of Bioinformatics, National Chiao Tung University, HsinChu 30050, Taiwan, Republic of China

ABSTRACT

It has recently been shown that in proteins the atomic mean-square displacement (or B-factor) can be related to the number of the neighboring atoms (or protein contact number), and that this relationship allows one to compute the B-factor profiles directly from protein contact number. This method, referred to as the protein contact model, is appealing, since it requires neither trajectory integration nor matrix diagonalization. As a result, the protein contact model can be applied to very large proteins and can be implemented as a high-throughput computational tool to compute atomic fluctuations in proteins. Here, we show that this relationship can be further refined to that between the atomic mean-square displacement and the weighted protein contact-number, the weight being the square of the reciprocal distance between the contacting pair. In addition, we show that this relationship can be utilized to compute the cross-correlation of atomic motion (the B-factor is essentially the auto-correlation of atomic motion). For a nonhomologous dataset comprising 972 high-resolution X-ray protein structures (resolution <2.0 Å and sequence identity <25%), the mean correlation coefficient between the X-ray and computed B-factors based on the weighted protein contact-number model is 0.61, which is better than those of the original contact-number model (0.51) and other methods. We also show that the computed correlation maps based on the weighted contact-number model are globally similar to those computed through normal model analysis for some selected cases. Our results underscore the relationship between protein dynamics and protein packing. We believe that our method will be useful in the study of the protein structure-dynamics relationship.

Proteins 2008; 72:929–935.
© 2008 Wiley-Liss, Inc.

Key words: protein dynamics; protein contact number; B-factors.

INTRODUCTION

Protein dynamics is dictated by protein structure. The dynamic properties of proteins result from a network of complex interactions like covalent bonding and nonbonded electrostatic or van der Waals interactions. To compute the dynamical properties of proteins, one usually resorts to molecular dynamics simulation,[1–5] which consists of integrating long time trajectories of protein structure using empirical force field. Though molecular dynamics is a powerful method, it is computationally expensive. Because of the recent progress of structural biology research, the number of protein structures deposited in Protein Data Bank has nearly quadrupled since 2000. Hence, there is increasing interest in developing efficient methods to compute protein dynamic properties from protein structures in a high-throughput fashion. The elastic network model (ENM) or Gaussian network model (GNM)[6–8] provides an alternative for molecular dynamics in computing average dynamical properties. In the GNM, each Cα atom is connected through a single-parameter harmonic potential to its neighboring atoms that are within a certain cut-off distance, usually in the range of 7–10 Å. The GNM then builds a connectivity matrix (also called the Kirchhoff matrix), from which cross-correlations and auto-correlations of fluctuations of residues can be obtained through matrix diagonalization. Micheletti *et al.*[9] have developed a model based on a mean field theory to study the dynamics of a protein. The shape of the protein is specified by the locations of the Cα atoms with two types of interactions: simple harmonic potential functions describing bonded interactions and Go-like functions describing nonbonded interactions. This model was applied to protein–protein interactions.[10] Zhou and coworkers[11] later extended this Cα-based model protein to a model based on all heavy atoms and they were able to make a more accurate prediction of protein flexibility, i.e., the B-factors, using this extended model.

A recent study[12] showed that the atomic mean-square displacement (or B-factor) is closely related to the number of noncovalent

The Supplementary Material referred to in this article can be found online at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsors: National Science Council, MoE ATU plan, Taiwan, Republic of China.
*Correspondence to: J.-K. Hwang, Institute of Bioinformatics, National Chiao Tung University, HsinChu 30050, Taiwan, Republic of China. E-mail: jkhwang@faculty.nctu.edu.tw
Received 6 September 2007; Revised 24 December 2007; Accepted 31 December 2007
Published online 25 February 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.21983

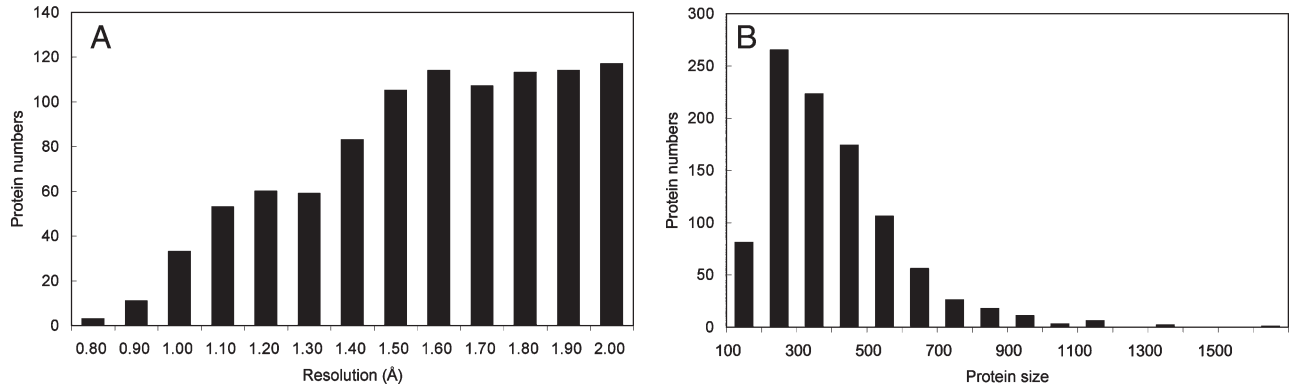

Figure 1

*The distribution of (**A**) protein size and (**B**) structural resolution of the data set comprising 972 protein structures with resolution ≤2.0 Å and R-factors ≤0.2 selected from PDB-REPRDB.*

neighboring atoms. For convenience, we will refer to this method as the protein contact number (CN) model. The CN model is appealing, since it predicts the B-factor profiles directly form protein structures without either trajectory integration or matrix diagonalization. However, despite its simplicity, the CN model has been shown to be superior to the GNM for a small set of 38 structure.[12] Here, we show that the CN model, which relates the B-factors to protein CN, can be further improved if the protein CN is scaled down by the square of the distance between the contacting pair. In addition, we show that cross-correlation between residues can also be computed in the framework of the weighted CN model.

METHODS

Protein contact number

The CN v_i of the ith residue is defined as the number of the neighboring residues whose Cα atoms are within a cut-off radius r_0 of that of the ith residue.

$$v_i = \sum_{j \neq i}^{N} H(r_0 - r_{ij}) \tag{1}$$

where r_{ij} is the distance between Cα atoms of residue i and j, and $H(x) = 1$ if $x \geq 0$ and $H(x) = 0$ if $x < 0$. Equation (1) defines an integral CN and gives an equal unitary weight to every contacting atom regardless of its distance to the central atom. The distance-dependent CN v_i' of the ith residue is defined as $v_i' = \sum_{j \neq i}^{N} H(r_0 - r_{ij})/r_{ij}^2$, which defines a real-valued CN, i.e., the integral CN weighted by the square of the reciprocal distance between the contact pair. Because of the fast decay of the factor $1/r_{ij}^2$ at large separation r_{ij}, the real-valued CN can be simplified to

$$v_i' = \sum_{j \neq i}^{N} \frac{1}{r_{ij}^2} \tag{2}$$

We will refer to v as the contact number (or CN), while v' [Eq. (2)] as the weighted contact number (or WCN). The CN (or WCN) profile of a protein of N residues is defined as:

$$\mathbf{w} = (\omega_1, \omega_2, \ldots \omega_N) \tag{3}$$

where ω_i is defined as the reciprocal CN, i.e., $\omega_i = 1/v_i$ or $\omega_i = 1/v_i'$. The X-ray B-factor profile is denoted as $\mathbf{b} = (b_1, b_2, \ldots b_N)$, where b_i is the B-factor of the Cα atom of the ith residue taken from the PDB file. For easy comparison, we will normalize both the CN (or WCN) and the B-factor profiles to the corresponding z-scores: $z_{x_i} = (x_i - \bar{x})/\sigma_x$, where $\bar{x}$ and σ_x are the mean and standard deviation of x. Here x designates b or ω. The normalized CN (or WCN) and the B-factor profiles are denoted by the vectors $\mathbf{z}_\omega$ and $\mathbf{z}_b$, respectively. In the CN model, the cut-off distance is set to 7.35 Å, which corresponds to the second minimum of the average contact-pair distribution of protein structures.[12] For the prediction assessment, we use the Pearson's linear correlation coefficient between the profiles, $c = \mathbf{z}_b \cdot \mathbf{z}_\omega$. If $c = 1$, two profiles are perfectly correlated; if $c = 0$, they are completely independent of each other; if $c = -1$, they are perfectly anticorrelated. Two indices of prediction assessment are used to compare the global performances of different methods for a data set: $\bar{c}$, the average correla-

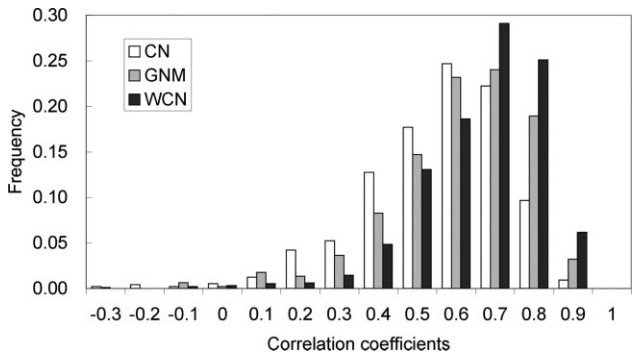

Figure 2

Comparison of the correlation coefficients between experimental and the computed B-factor profiles bases on the CN model (white), the GNM (grey), and the WCN model (black) for the nonhomologous data set.

tion coefficient, and $p_{0.5}$, the fraction of number of structures with $c \geq 0.5$.

Cross-correlation between residues

The normalized correlation between fluctuations of atom i and j is defined as

$$C_{ij} = \frac{\langle \delta \mathbf{r}_i \cdot \delta \mathbf{r}_j \rangle}{\sqrt{\langle \delta \mathbf{r}_i \cdot \delta \mathbf{r}_i \rangle \langle \delta \mathbf{r}_j \cdot \delta \mathbf{r}_j \rangle}} \quad (4)$$

where $\delta \mathbf{r}_i$ and $\delta \mathbf{r}_j$ are the fluctuations of the atom i and j, respectively, around their equilibrium positions. In the framework of the WCN model, we formulate the correlation term W_{ij} between residue i and residue j as

$$W_{ij} = \left(\sum_{k \neq j,k}^{N} \frac{1}{r_{ik} r_{jk}} \right)^{-1} \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j \quad (5)$$

where $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}_j$ are the unit vectors in the direction of $\sum_k^N \mathbf{r}_i - \mathbf{r}_k$ and $\sum_k^N \mathbf{r}_j - \mathbf{r}_k$, respectively. Note that $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}_j$ can be cast into another forms: $(\mathbf{r}_i - \bar{\mathbf{r}})/|\mathbf{r}_i - \bar{\mathbf{r}}|$ and $(\mathbf{r}_j - \bar{\mathbf{r}})/|\mathbf{r}_j - \bar{\mathbf{r}}|$, respectively, where $\bar{\mathbf{r}}$ is the protein centroid given by $\sum_k^N \mathbf{r}_k/N$. It is not hard to check that when $i = j$, W_{ii} reduces to ω_i.

Dataset

We selected from PDB-REPRDB[13] 972 protein chains of length $\geq$60. Their structures are solved by X-ray crystallography with resolution $\leq$2.0 Å and R-factors $\leq$0.2. All chains are of pair-wise sequence identity $\leq$25%. The chains of the data set are listed in Table S1 in the supplementary material. In the data set, the protein size ranges from 60 to 1520 with an average protein size around 300 residues. The resolution of the X-ray structures ranges from 0.73 to 2.0 Å with an average structural resolution 1.78 Å. The distribution of protein size (i.e., the number of residues) and structural resolution of the data set are shown in Figure 1.

RESULTS

The computed B-factor profiles using different methods

Figure 2 shows the histogram of the correlation coefficients between the X-ray B-factors and those computed by

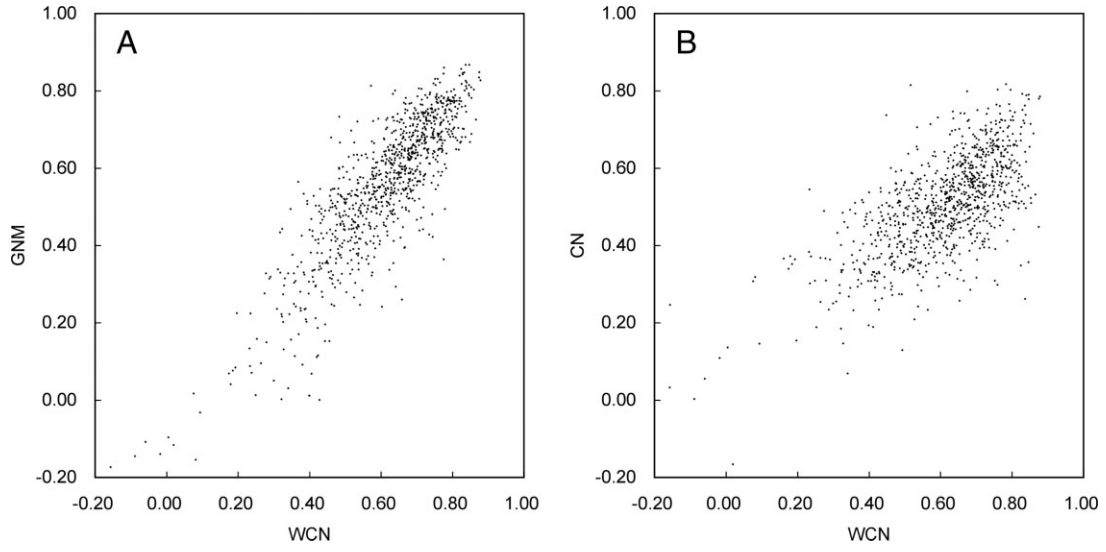

Figure 3

*The correlation of model predictions (**A**) between the WCN model and the GNM and (**B**) between the WCN model and the CN model for the 972-structure data set.*

the WCN model, the CN model and the GNM model. In the WCN model, the average correlation coefficient is $\bar{c}$ = 0.61 and the fraction of structures with a correlation coefficient $\geq$0.5 is $p_{0.5}$ = 79%. The CN model yields poorer results: $\bar{c}$ = 0.51 and $p_{0.5}$ = 54%. The effect of the term $1/r_{ij}^2$, which is missing in the CN model, on the results is significant. The GNM yields $\bar{c}$ = 0.56 and $p_{0.5}$ = 69%. This is in contrast with the previous study[12] that the CN model performs better than the GNM. It should be noted that, however, the previous study was conducted on a much smaller data set of 38 structures. Though the correlation–coefficient distributions of these models seems to look quite different, we perform additional Student *t*-test to check these distributions using the statistical package R.[14] The *P*-values of the WCN-GNM, GNM-CN and WCN-CN are all smaller than 2.20×10^{-16}, indicating that the distributions are significantly different from each other. On the other hand, we notice that a better correlation between the WCN model and the GNM (0.86) than that between the WCN and the CN model (0.67). These results are shown in Figure 3. Thought both the CN model and the GNM consider the contributions from any atoms to be identical as long as they are within the cut-off distance, the CN model completely ignores those atoms that are out of the cut-off range, while the GNM takes them into account implicitly through the network. The WCN model considers the contributions from any atoms with a weighting factor $1/r_{ij}^2$. In Figure 4, we compare the B-factor profiles of flavocytochrome c3 (1Y0P:A) computed by three methods with each compared with the X-ray B-factor profile. The WCN and the GNM B-factor profiles agree relatively well with the X-ray B-factors, but the CN B-factor profile appears to be much more rugged, probably due to the artificial cut-off effect.

The GNM program[15] uses only Cα atoms in the calculation of the B-factors, and therefore, for the sake of comparison, the previously calculated results are based on the Cα atoms in both the WCN and CN models. If the average B-factors for the entire residue are used, the WCN model yields $\bar{c}$ = 0.60 and $p_{0.5}$ = 79%, and the CN model yields $\bar{c}$ = 0.50 and $p_{0.5}$ = 54%. These results are not much different from those based on the Cα atoms.

Comparison of models based on Cα and all atoms

To study the effects of the atoms other than the Cα atom on the computed B-factors, we calculate the WCN and CN B-factor profiles using all nonhydrogen atoms (i.e., C, N, O, and S atoms) of proteins. We do not calculate the GNM all-atom B-factor profiles, since, as mentioned before, the currently available GNM program[15] uses only Cα atoms for proteins. In the WCN model, if all heavy atoms are included in calculation, the results are $\bar{c}$ = 0.62 and $p_{0.5}$ = 85%, while the all-atom CN model yields $\bar{c}$ = 0.56 and $p_{0.5}$ = 77%. Both results are better than those based on only Cα atoms.

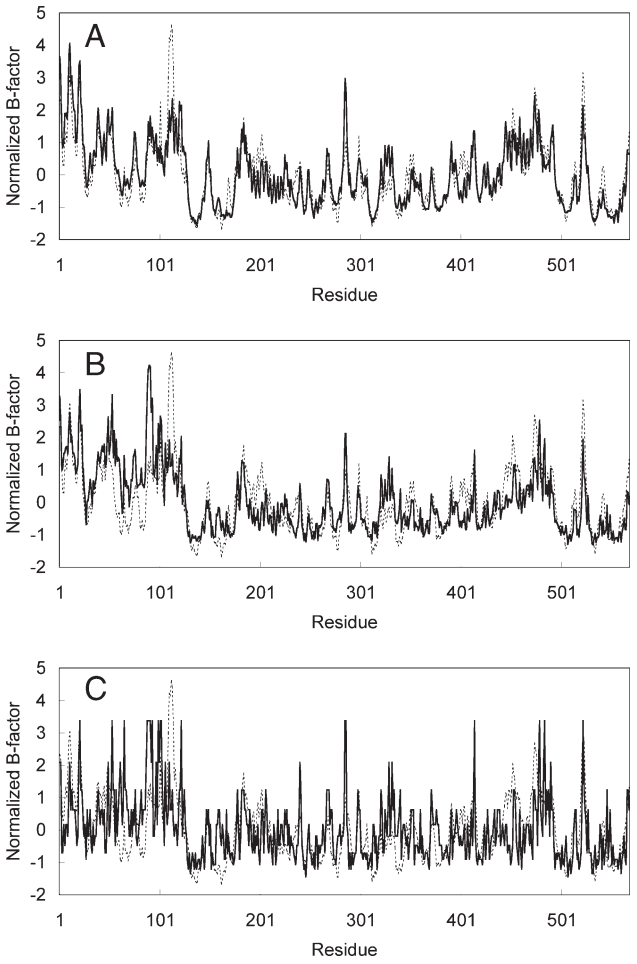

Figure 4

*The X-ray B-factor profile (dotted line) of flavocytochrome c3 (1Y0P:A) compared with the computed B-factor profile (solid line) by (**A**) the WCN model (c = 0.85), (**B**) the GNM (c = 0.69), and (**C**) the CN model (c = 0.51).*

The breakdown analysis for the accuracy of different models

Classified in terms of the SCOP classes, the structures in the dataset has 111 all-α proteins, 181 all-β proteins, 245 α/β proteins, 193 α + β proteins, 15 multi-domain proteins, 11 membrane and cell surface proteins and peptides, 22 small proteins, five coiled coil proteins, one designed protein and 188 undefined in SCOP. In Table I, we compare the statistics of the performances of different models for the four major SCOP classes: all-α proteins, all-β proteins, α/β proteins and α + β proteins, since the other classes have much smaller sample size (1–22). All three models perform best for the all-β proteins. In general, the trends of the performance of these methods appear to be similar. We compare in Figure 5 the performance of all three methods as a function of protein

Table I
The Performance Breakdown of the WCN Model, the CN Model and the GNM for the Structures Classified According to the SCOP Classes

	WCN		CN		GNM	
SCOP classes	$\bar{c}$	$p_{0.5}$	$\bar{c}$	$p_{0.5}$	$\bar{c}$	$p_{0.5}$
All-α proteins	0.59	0.73	0.47	0.43	0.54	0.68
All-β proteins	0.64	0.82	0.51	0.58	0.58	0.73
α/β proteins	0.62	0.82	0.49	0.51	0.57	0.75
α + β proteins	0.60	0.77	0.49	0.51	0.54	0.65

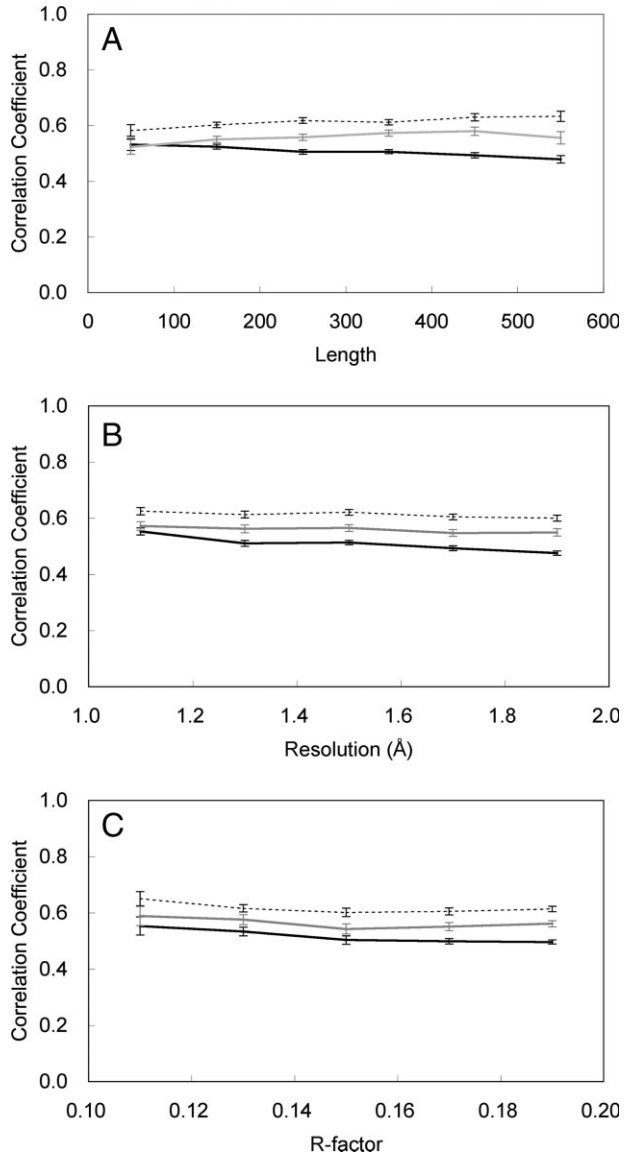

Figure 5
*The correlation coefficient between the computed and the X-ray B-factors for the WCN model (dotted line), the CN model (solid line), and the GNM (grey line) as a function of the protein size in terms of (**A**) the residue number, (**B**) the X-ray structure resolution in Å, and (**C**) the R-factor.*

size, X-ray resolution and R-factor. We remove the proteins in the marginal regions (for example, protein size 600–1550) due to their smaller number in those regions. One notices that, while the average performance of the CN model shows slightly downward trend for proteins of larger size or lower resolution, the average performances of the WCN model and the GNM appear to be relatively unchanged in the range of protein properties studied.

Applications to large proteins and protein complexes

The WCN model can be readily applied to large proteins, since its memory requirement is of the order $O(N)$, where N is the size of protein. An example is the 50S ribosomal subunit (1YJW) comprising 3774 residues. Figure 6 shows its computed B-factor profile using the WCN model. On the other hand, the oGNM,[15] the web version of the GNM, is unable to return the B-factor profile of 1YJW. Some proteins in our data set are in fact part of larger biological units, which are the assumed functional form of the macromolecule. In the PDB, the biological units are built from the crystallographic space group using symmetry operation. Currently, the coordinates of biological units can be obtained from either PDB or PQS.[16] The PDB and PQS biological units agree on 82% of entries.[17] In this work, we use the PDB biological units for computation. We computed the WCN B-factor profiles of the same proteins of the data set with other parts of the whole biological units (if any) taken into consideration. The Cα WCN model yields $\bar{c} = 0.65$ and $p_{0.5} = 86\%$, which are better than the previous results.

The cross-correlation of fluctuations between residues

The knowledge of correlated motion between residues is useful in understanding long-range communication[18,19] and large domain movements relevant to protein function.[20,21] The correlation matrix can be com-

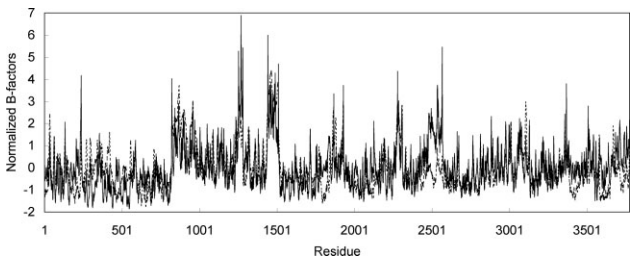

Figure 6
The X-ray B-factor profile (dotted line) of the 50S ribosomal subunit (1YJW) compared with the computed B-factor profile (solid line). The correlation coefficient between the X-ray and the computed B-factors is 0.60.

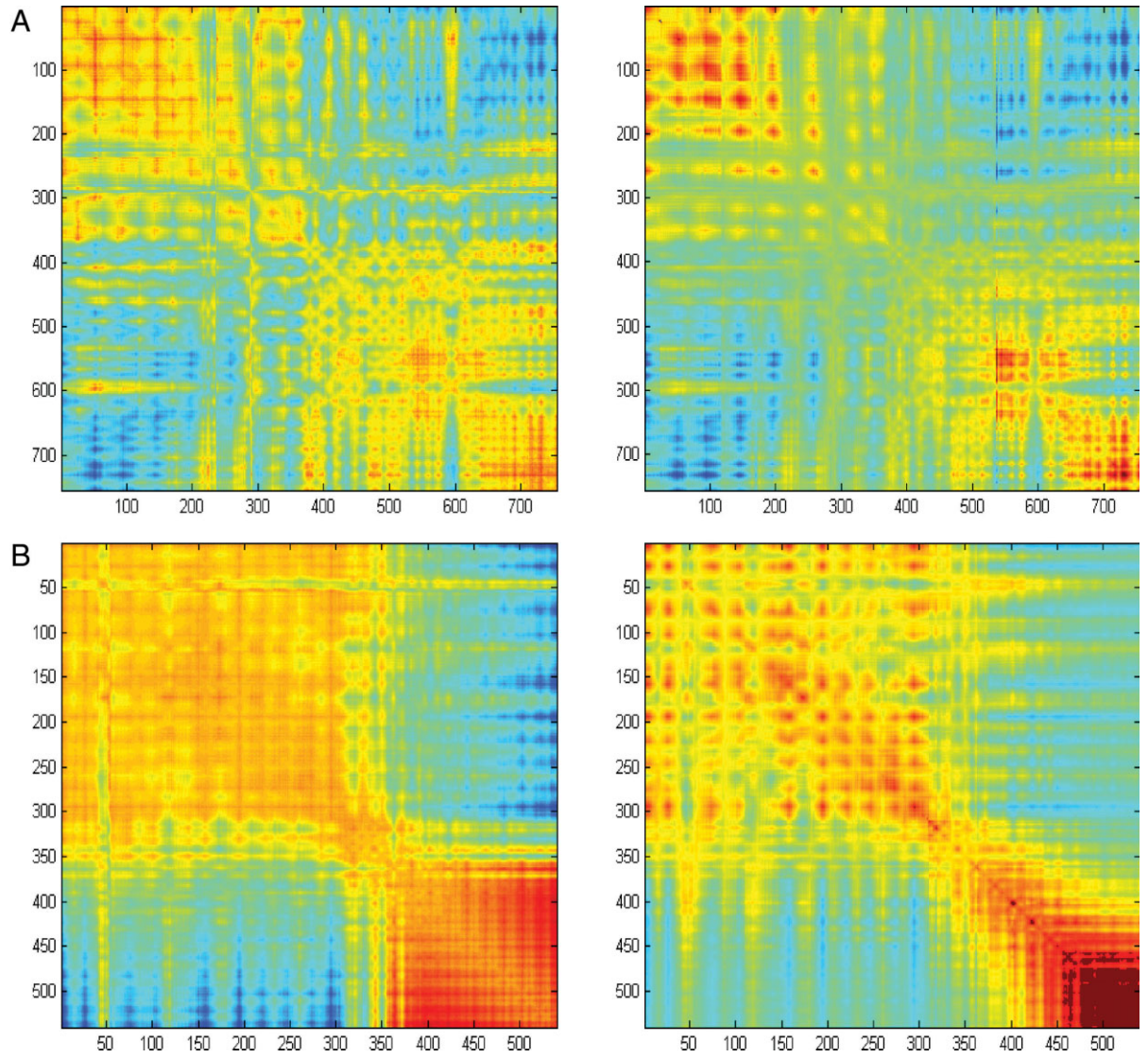

Figure 7

*The cross-correlation maps for (**A**) chondroitin AC lyase (1RWH:A) and (**B**) complete modular teichioic acid phosphorylcholine esterase PCE (2BIB:A). For each protein, the map on the left is computed by Eq. (5) and the map on the right by the NMA. The colors of the map ramp from red (positive correlation) to blue (negative correlation). NMA was performed using the simplified force field of ENZYMIX.*[25,26]

puted through the normal mode analysis (NMA).[22–24] In this method, the protein structure is first optimized through energy minimized. The second derivative matrix of the total potential function (also called the Hessian matrix) is computed from the optimized structure. The correlation between fluctuations is computed from the eigenvalues and the eigenvectors of the Hessian matrix using $\langle \delta \mathbf{r}_i \cdot \delta \mathbf{r}_j \rangle \sim \sum_k U_{ik} U_{jk} / \lambda_k$, where λ_k is the eigenvalues of the kth mode and U_{ik} is the ith component of the eigenvector of the kth mode. Instead of going through procedures of energy minimization and matrix diagonalization as in the case of NMA, we can compute the correlation map directly from protein structure using Eq. (5). Figure 7 shows the computed correlation maps of 1RWH:A and 2BIB:A, and compares them with those of NMA. Currently, there is no "experimental" correlation map (except for the diagonal terms, which correspond to the X-ray B-factors) as a reference standard. However, the similarity of both types of computed correlation maps indicates that the WCN model provides a quick alternative to NMA to compute the correlation of motions in proteins. The correlation of 1RWHA is 0.924 and that of 2BIBA is 0.817. It should be noted that Eq. (5) is essentially an empirical one verified by results instead of by derivation; further study is needed to understand the physics behind it.

DISCUSSION

The close relationship between the thermal fluctuations of proteins and the distance-dependent protein CN allows one to compute dynamic properties of proteins. This method, i.e., the WCN model, does not presuppose a mechanical model[6–8] as well as the potential functions associated with that model as other methods: molecular dynamics is based on sophisticated molecular force field,[1–3] while the GNM assumes a harmonic oscillator model for proteins with their structures described in terms of a collection of masses connected to each other through a spring of a uniform force constant. We showed that the WCN model can produce more accurate B-factor profiles than other methods. In addition, we show for the first time that cross-correlation between residues can be computed directly from protein contact number. We have recently showed that[27] the atoms in proteins lying on the same spherical shell centered at the fixed point tend to have similar thermal fluctuations. We will refer to this model as the protein fixed-point (PFP) model. The PFP model assumes that the protein centroid in the simple single-domain protein is the position of the smallest fluctuations, i.e., the fixed point. The PFP model, like the WCN model, provides a simple way to compute both auto-correlation and cross-correlation between residues in reasonable accuracy. It is not hard to show that the WCN model can reduce to the PFP model; the diagonal term of Eq. (5) can be simplified as

$$W_{ii} = \left(\sum_{k}^{N} \frac{1}{r_{ik} r_{ik}} \right)^{-1} \hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_i = \left(\sum_{k}^{N} \frac{1}{r_{ik}^2} \right)^{-1} \qquad (6)$$

Using the approximation,

$$\left(\sum_{k \neq i} \frac{1}{r_{ik}^2} \right)^{-1} \sim R_i^2$$

where R_i is the distance of the ith Cα atom from the protein centroid. We obtain from Eq. (6) $W_{ii} = R_i^2$, which is the PFP B-factor.[27]

ACKNOWLEDGMENTS

We are grateful to both hardware and software supports of the NRPGM Structural Bioinformatics Core at National Chiao Tung University.

REFERENCES

1. Levitt M, Warshel A. Computer simulation of protein folding. Nature 1975;253:694–698.
2. Warshel A. Bicycle-pedal model for the first step in the vision process. Nature 1976;260:679–683.
3. McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. Nature 1977;267:585–590.
4. Warshel A. Molecular dynamics simulations of biological reactions. Acc Chem Res 2002;35:385–395.
5. Rueda M, Ferrer-Costa C, Meyer T, Perez A, Camps J, Hospital A, Gelpi JL, Orozco M. A consensus view of protein dynamics. Proc Natl Acad Sci USA 2007;104:796–801.
6. Tirion MM. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Phys Rev Lett 1996;77:1905–1908.
7. Bahar I, Atilgan AR, Erman B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. Fold Des 1997;2:173–181.
8. Ming D, Kong Y, Lambert MA, Huang Z, Ma J. How to describe protein motion without amino acid sequence and atomic coordinates. Proc Natl Acad Sci USA 2002;99:8620–8625.
9. Micheletti C, Banavar JR, Maritan A Conformations of proteins in equilibrium. Phys Rev Lett 2001;87:088102.
10. Canino LS, Shen T, McCammon JA. Changes in flexibility upon binding: application of the self-consistent pair contact probability method to protein-protein interactions. J Chem Phys 2002;117:9927–9933.
11. Pandey BP, Zhang C, Yuan X, Zi J, Zhou Y. Protein flexibility prediction by an all-atom mean-field statistical theory. Protein Sci 2005;14:1772–1777.
12. Halle B. Flexibility and packing in proteins. Proc Natl Acad Sci USA 2002;99:1274–1279.
13. Noguchi T, Akiyama Y. PDB-REPRDB: a database of representative protein chains from the protein data bank (PDB) in 2003. Nucleic Acids Res 2003;31:492–493.
14. Team RDC. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2007.
15. Yang LW, Rader AJ, Liu X, Jursa CJ, Chen SC, Karimi HA, Bahar I. oGNM: online computation of structural dynamics using the Gaussian network model. Nucleic Acids Res 2006;34:W24–W31.
16. Henrick K, Thornton JM. PQS: a protein quaternary structure file server. Trends Biochem Sci 1998;23:358–361.
17. Xu Q, Canutescu A, Obradovic Z, Dunbrack RL, Jr. ProtBuD: a database of biological unit structures of protein families and superfamilies. Bioinformatics 2006;22:2876–2882.
18. Budiman ME, Knaggs MH, Fetrow JS, Alexander RW. Using molecular dynamics to map interaction networks in an aminoacyl-tRNA synthetase. Proteins 2007;68:670–689.
19. Zheng W, Liao JC, Brooks BR, Doniach S. Toward the mechanism of dynamical couplings and translocation in hepatitis C virus NS3 helicase using elastic network model. Proteins 2007;67:886–896.
20. Ming D, Kong Y, Wakil SJ, Brink J, Ma J. Domain movements in human fatty acid synthase by quantized elastic deformational model. Proc Natl Acad Sci USA 2002;99:7895–7899.
21. Yang LW, Bahar I. Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes. Structure 2005;13:893–904.
22. Brooks B, Karplus M. Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor. Proc Natl Acad Sci USA 1983;80:6571–6575.
23. Levitt M, Sander C, Stern PS. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. J Mol Biol 1985;181:423–447.
24. Kidera A, Go N. Normal mode refinement: crystallographic refinement of protein dynamic structure. I. Theory and test by simulated diffraction data. J Mol Biol 1992;225:457–475.
25. Lee FS, Chu ZT, Warshel A. Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. J Comp Chem 1993;14:161–185.
26. Fan Z-Z, Hwang J-K, Warshel A. Using simplified protein representaion as a reference potential for all-atom calculations of folding free energy. Theor Chem Acc 1999;103:77–80.
27. Shih CH, Huang SW, Yen SC, Lai YL, Yu SH, Hwang JK. A simple way to compute protein dynamics without a mechanical model. Proteins 2007;68:34–38.