ELSEVIER

Polymer 43 (2002) 403–408

polymer

www.elsevier.com/locate/polymer

Distribution of cooperative interactions in barnase at different time windows by coarse-grained simulations ☆

Nese Kurt, Türkan Haliloglu*

Polymer Research Center and Chemical Engineering Department, Bogaziçi University, Bebek 80815, Istanbul, Turkey

Abstract

The backbone dynamics of barnase has been studied by a recently developed off-lattice Monte Carlo (MC)/Metropolis simulation technique, where a low-resolution model (virtual-bond model) is used together with knowledge-based potentials, with the main emphasis on its cooperative motions at different time windows. The conformations generated around the native state are analysed by time-dependent auto- and cross-conformational correlation functions of the virtual bonds. There exists a correlation between the long time auto-correlated behaviour of the bond rotations and the potential stability of the respective regions. The analysis at different time windows reveals that there are cooperative motions between the bond rotations, which are only near neighbours and basically local motions at all time windows. However, as the time window widens, a progressive increase in the number of correlated pairs, which are separated far along the sequence and are not necessarily close in space, is observed. The structural distribution of these motions shows that the cooperative interactions are not bi-directional and that different residues have a different role within the network of interactions. Thus, the conditions yielding global motion coherence can be accounted for by the existence of anisotropic cooperative long-range interactions among the units in cooperation with the short-range interactions.

Keywords: Low resolution model; Monte Carlo simulations; Knowledge-based potentials

1. Introduction

Proteins in solution undergo various motions at different time scales; a broad hierarchy of functionally significant timescales, ranging from local atomic fluctuations and bond oscillations to hinge bending motions, helix coil transitions and local and global unfolding processes, which depend, in detail, on the conformational energy landscape. To characterise these dynamic motions and to understand their role in protein function, stability and folding are of great interest. On the other hand, the stability and the folding are strongly associated with the cooperative interactions, which also provide a key to understanding function and regulations of function. The collective dynamics in relation to function was recently reviewed [1] with references to many efforts in this area. The consensus among many investigators is that functionally relevant protein motions involve correlated displacement of many residues. For example, binding of small [2] and other molecules [3] to proteins occur via induced fit of concerted displacements of several structural elements. Thus, dynamic correlations between distant sites were shown to have important implications in the design of protease inhibitors [4]. The normal mode analysis employed [5] to the classical model system for allosteric regulation led to the understanding of how the information is transmitted from the allosteric binding site to the active site which is 60 Å away. In fact, the cooperative interactions do not extend uniformly throughout the entire protein molecule, and so some residues have a more important role than the others in defining the cooperativity, which is marked by the conformational heterogeneity. The latter was evidenced from several hydrogen exchange experiments [6–10] that yielded heterogeneity in the magnitude of the hydrogen exchange protection factors measured under native conditions, which indicated local unfolding rather than global unfolding. On the other hand, the secret in the stability is the effective cooperation of intramolecular interactions, which are individually too weak to withstand the dissipative action of the thermal motion. The cooperativity of intermolecular interactions in proteins seems to be achieved only in a molecule with a tight and unique packing of groups, i.e. the cooperativity is the peculiarity of an

☆ This paper was originally submitted to *Computational and Theoretical Polymer Science* and received on 11 November 2000; received in revised form on 23 March 2001; accepted on 20 April 2001. Following the incorporation of *Computational and Theoretical Polymer Science* into *Polymer*, this paper was consequently accepted for publication in *Polymer*.

* Corresponding author. Tel.: +90-212-358-1540 ext. 2001; fax: +90-212-257-5032.

E-mail address: turkan@prc.bme.boun.edu.tr (T. Haliloglu).

0032-3861/02/$ - see front matter
PII: S0032-3861(01)00431-1

aperiodic structure [11,12] which can provide complex interlacing of short- and long-range interactions involving different time-scales.

Studying the detailed mechanism of the cooperation of the weak intermolecular forces and their ability to cooperate is an important physical problem in current studies of the protein research, which has a great importance in the framework connecting structure, energy landscape, dynamics and function.

Experiments do not offer direct access to an atomic description of the correlated motions and furthermore the experimental data need to be interpreted in the framework of an appropriate model. Nevertheless, computational means is more straightforward. They can be predicted from one representative structure using a description of intramolecular interactions at different levels of detail. Molecular dynamics (MD) simulation is one means of generating ensembles; but however it is not at present feasible to extend beyond the nanosecond time scale for most cases, and furthermore it poses sampling problems on the time scale of large concerted motions. In those motions, multiple independent trajectories are required rather than having only one even long simulation trajectory, for assuring the enough sampling of the conformational space, as pointed out in a recent study [13] on the problem of undersampling in nanosecond MD simulation studies. Principal component analysis (PCA) which facilitates the study of long time protein dynamics from MD trajectories may fail in many cases because of the sampling errors due to the finite simulation time [14]. On the other hand, low-resolution models and methods can be effectively utilised to explore motions on time scales ranging from the order of nanoseconds to milliseconds with several independent trajectories.

In the present study, the generated ensemble of conformations around the native basin of barnase, by the recently developed coarse-grained simulation method (off-lattice Monte Carlo (MC)) [15–18], have been analysed by means of correlation functions. Barnase was chosen as the subject because it was extensively studied by experiments and simulations [19–28]. The technique employed here has been shown to adequately simulate protein dynamics for different time scales with results which are in agreement with those from NMR measurements, such as order parameters [17] and hydrogen exchange data [17,18]. It is based on the virtual bond-model [29] with structure-derived potential functions and geometry parameters [30,31]. The specificity of the residues is addressed with the latter potentials. The usefulness of the statistical potentials was reviewed in a recent study [32]. In the present analysis, the fluctuations of Cα atoms, the conformational auto- and cross-correlations between rotations are calculated. The correlations calculated at different time windows map the interaction scheme of the structure at different levels, which lead to the elucidation of the intricate network of interactions laced together by short and long-range interactions.

2. Overview of model and simulation characteristics

A given conformation of the coarse-grained model protein, which has two interaction sites for each residue, Cα atom and side chain centroid, is represented by a set of $3N-6$ generalised coordinates (N is the number of residues and thus of Cα atoms): $N-1$ backbone virtual bonds l_i connecting α-carbons $i-1$ and i, $N-2$ bond angles θ_i, the angle between l_i and l_{i+1}, and N-3 bond torsional angles Φ_i. The virtual bond lengths are kept within a value 3.81 ± 0.03 Å by a harmonic potential with a force constant of 10 (mass/time^2). Side chain conformation, on the other hand, is conveniently expressed by the set $\{l_i^s, \theta_i^s, \phi_i^s\}$, l_i^s being the bond length connecting backbone and side chain interaction sites, θ_i^s is the bond angle between l_i and l_i^s, and ϕ_i^s the torsion angle defined by l_{i-1}, l_i and l_i^s. The atoms used for defining the side-chain interaction centres for each type of amino acid are presented in Ref. [31].

Barnase is a small, monomeric, 110-residue ribonuclease that is secreted by *Bacillus amyloliquefaciens*. The solution structure of barnase (PDB entry 1bni) [33] is used as the starting conformation. This α + β protein contains four helices followed by a five-stranded β-sheet. In accordance with the PDB file, the secondary structural units are named as follows: α-helices: α1(7–17), α2(27–33), α3(37–39), α4(42–45) and β-strands: β1(52–56), β2(71–75), β3(87–91), β4(96–99), β5(107–110). There are three hydrophobic cores in barnase and these are depicted on the later ribbon diagrams.

A series of low-resolution conformations or substates in the neighbourhood of the native basin, are generated by MC/Metropolis simulation scheme. In this technique, at each 'step', hereafter referred to as the 'MC step', each residue is given a chance to move. A randomly chosen site, either a Cα or a sidechain site, is subjected to a differential perturbation, using a uniformly distributed random number generator. The strength of the perturbation, Δx is controlled by the formula $\Delta x = k(2r-1)$, where r is the random number variable in the range $0 \le r \le 1$, and k is a constant 'proportionality factor' that may be adjusted to mimic the response at a given temperature. The new conformation is accepted or rejected according to the Metropolis criterion [34]. The energy of the conformation at each step is found using database-extracted residue-specific potentials composed of two major contributions: (i) bond rotation and bond angle distortion energies [35] for pairwise-coupled virtual bonds, and (ii) non-local interactions including all-side-chain (S–S), side chain–backbone (S–B), and backbone–backbone (B–B) pairs separated by at least five virtual bonds [36]. The detailed description of the simulation method can be found in Refs. [15,16,37].

The analysis is carried out over 10 independent runs, the length of each being 3000 MC steps, to ensure an efficient conformational sampling. The root-mean-square (rms) deviation of the simulated structure is calculated as

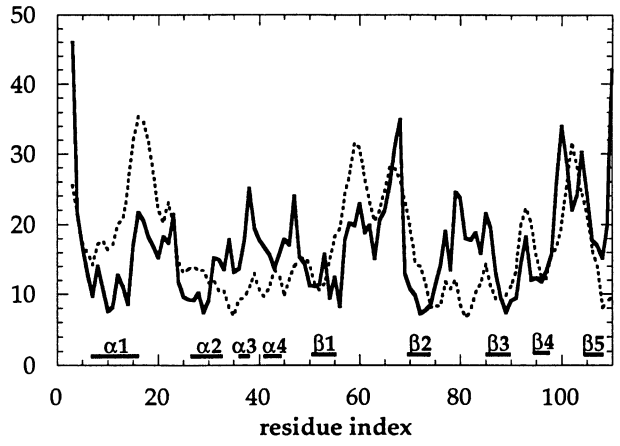

Fig. 1. Mean-square fluctuations $\langle \Delta \mathbf{R}_i^2 \rangle$ in α-carbon positions as a function of residue numbers for barnase. The solid curve is obtained from simulations and normalised, whereas the dashed curve represents the results from crystallographic temperature factors [33].

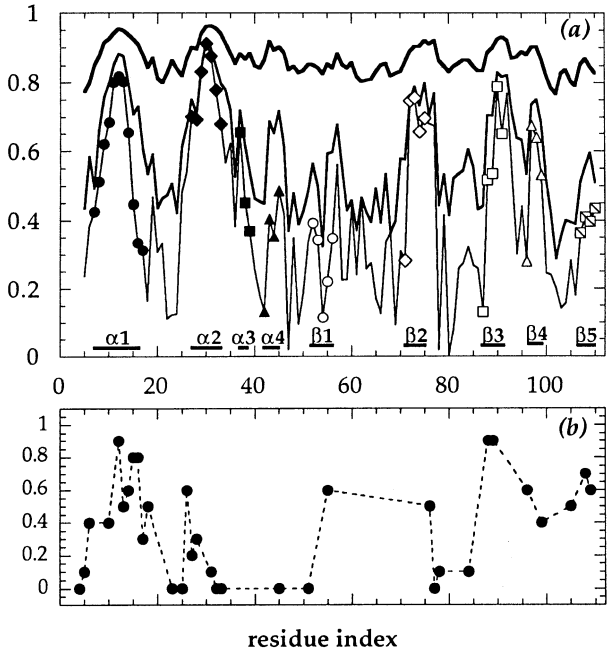

Fig. 2. (a) Time-delayed auto-correlations, $G_i(\tau)$, evaluated for all C^{α}–C^{α} virtual bonds at $\tau = 50$ (thick), 500 (medium) and 1500 (thin) MC steps. The secondary structural units are given along the *x*-axis and the residues belonging to these units are depicted by different symbols for clarity. (b) Experimental Φ values for the major transition state of barnase obtained by protein engineering method [20,23].

4 ± 0.5 Å at the end of each run. Thus, the generated conformations can be viewed as substates in the vicinity of the native basin, within the resolution of the present coarse-grained simulations.

3. Results and discussions

The mean-square fluctuations $\langle \Delta \mathbf{R}_i^2 \rangle$ in the position vectors $\mathbf{R}_i$ of the backbone sites of barnase are compared with the results from the crystallographic temperature factors [33] $B_i = 8\pi^2 \langle \Delta \mathbf{R}_i^2 \rangle / 3$ in Fig. 1. The results from the simulations (solid line) are normalised to enable a direct comparison with the experimental results (dashed line). A satisfactory agreement for the distribution of the mean-square fluctuations of Cα atoms between the experiments and simulations is observed, excluding the regions between α2 and β1, and the loop between β2 and β3, where higher fluctuations compared to those of experiments are predicted. In fact, MD simulations by Wong et al. [19] as well yielded similar results.

3.1. Correlation functions

The generated conformations around the native state during a simulation time of 3000 MC steps are analysed by means of conformational correlation functions of the virtual bonds. The conformation of each virtual bond is the result of collective motions of three real backbone bonds determined by the rotations of ψ and ϕ angles in each residue.

The time-delayed torsional auto-correlation function for each virtual bond is defined as

$$G_i(\tau) = \langle \cos(\phi_i(t) - \phi_i(t+\tau)) \rangle \quad (1)$$

where $\phi_i(t)$ and $\phi_i(t+\tau)$ are the rotational angles of the virtual bond i at times t and $t+\tau$, respectively. The angle brackets refer to the ensemble average over all initial times t, and over all independent runs. Within a time interval of τ from 0 to ∞, the correlation of a bond torsional angle may vary from 1, where there is complete correlation, to the equilibrium correlation value, $G_i(\infty)$, which may carry any value between 0 and 1 and, which indicates the equilibrium local characteristic differences in the conformational behaviour of substates.

Fig. 2a presents $G_i(\tau)$ values of all virtual bonds for different values of τ (50, 500 and 1500 MC steps), which reflect the conformational dynamic heterogeneity within the structure. Some regions, in particular, of α1, α2 and β2, β3, β4 regions, experience long-time surviving auto-correlated rotational motions. On the other hand, the small helices α3, α4 and the edge strands of the β-sheet, β1 and β5 exhibit relatively less correlated rotational motions. These edge strands are detected to be less stable than the central strands by both experiment [20] and MD simulations of high-temperature [21] and low-pH [22] denaturation. On the other hand, in Fig. 2b, the experimental Φ values for the major transition state of barnase obtained by protein engineering method are given [20,23]. Φ value is a measure of the stability and change in folding kinetics due to mutation performed on a particular amino acid. The similar behaviour in the distributions of Φ values and auto-correlations at long times supports the idea that stability at native state is related to the unfolding characteristics. It is interesting to note that the virtual bonds in the middle of the α2

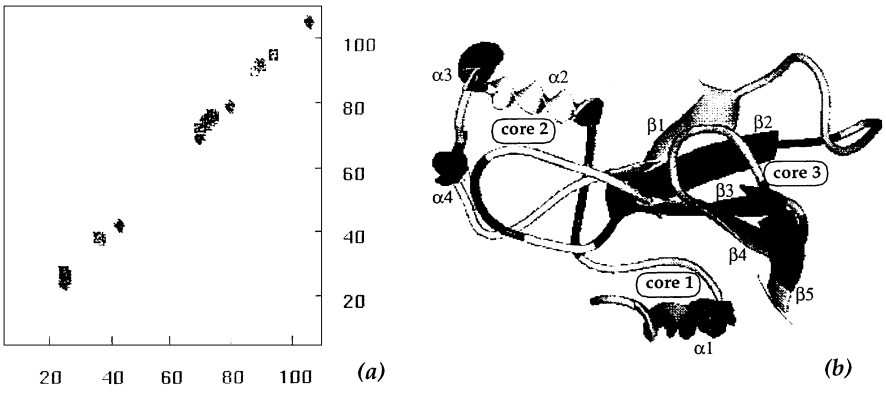

Fig. 3. (a) The contour map for time-delayed cross-correlations $C_{ij}(\tau)$, between the rotations of the virtual bonds i and j at $\tau = 50$ MC steps. (b) Ribbon diagram of barnase on which the regions displaying short-range correlations corresponding to the diagonal contours in part (a) are coloured black.

helical region carry highly correlated rotational motions even at long times. However, the N- and C-caps of α1 and α2 display a higher decay in their correlation values compared to their middle residues. This agrees with the instability of the caps seen in high-temperature denaturation simulations [21]. The N- and C-caps of α1 are unstable even in the room temperature MD simulation [22]. Moreover, the loop and the turn regions between the secondary structural elements undergo large internal motions by the least correlated rotational motions as revealed by lower $G_i(\tau =$ 1500 MC steps) values. This, on the other hand, indicates that the time interval of 1500 MC steps is adequate for the relaxation of the most flexible regions of the structure. Regarding the correspondence between MC step and real time, the relaxation time of the orientational behaviour of a hypothetical bond that is appended perpendicular to the vertex of two successive virtual bonds was calculated. The latter hypothetical bond was considered for different parts of the chain and the average relaxation time calculated is ~100 MC steps. On the other hand, average correlation time for N–H bond relaxation in general ranges form 50 to 250 ps.

Apparently, the relative structural stability and/or flexibility of different parts of the protein is linked to conformational dynamic heterogeneity and thus, an assessment of which regions of the protein are most susceptible to coming apart can be supported from the long-time auto-correlated behaviour of the backbone bonds. Along this line, the results in recent applications [17,18] of this technique to other proteins further support the latter argument.

The time dependent cross-correlations between the rotation of the virtual bonds i and j have been evaluated from

$$C_{ij}(\tau) = \langle \Delta\phi_i(t)\Delta\phi_j(t+\tau)\rangle / \langle \Delta\phi_i(t)^2\rangle^{1/2}\langle \Delta\phi_j(\tau)^2\rangle^{1/2} \qquad (2)$$

C_{ij} values are normalised, hence they lie in the range $-1 \le C_{ij} \le 1$; where the upper and lower limits refer to fully correlated and fully anti-correlated bonds, respectively. Here, the time interval τ within which the cross-correlation is searched is taken as 50, 500 and 1500 MC steps, the same as the times considered in the evaluation of the auto-correlation functions in Fig. 2a.

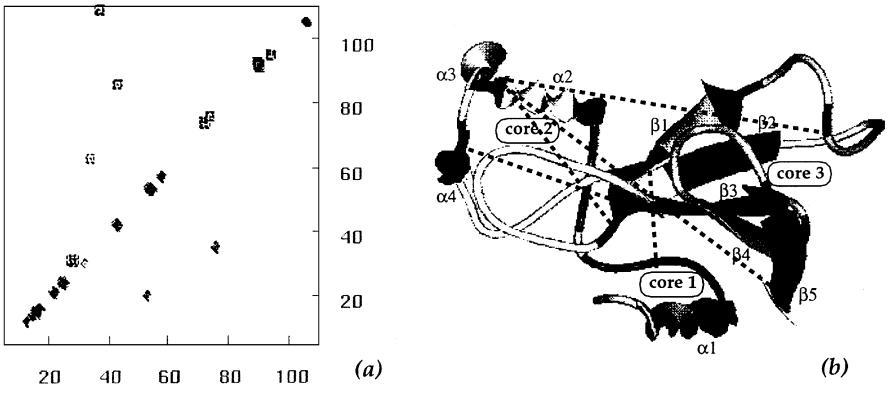

Fig. 4. (a) The contour map for cross-correlations $C_{ij}(\tau)$, between the rotations of the virtual bonds i and j at $\tau = 500$ MC steps. The contours on the upper and lower triangular portions of the correlation map are for negative and positive correlations, respectively. (b) Ribbon diagram of barnase on which the regions displaying short-range correlations corresponding to the diagonal contours in part (a) are coloured black; and the long-range rotational cooperativities corresponding to the off-diagonal contours are depicted by dashed lines between the two respective regions.

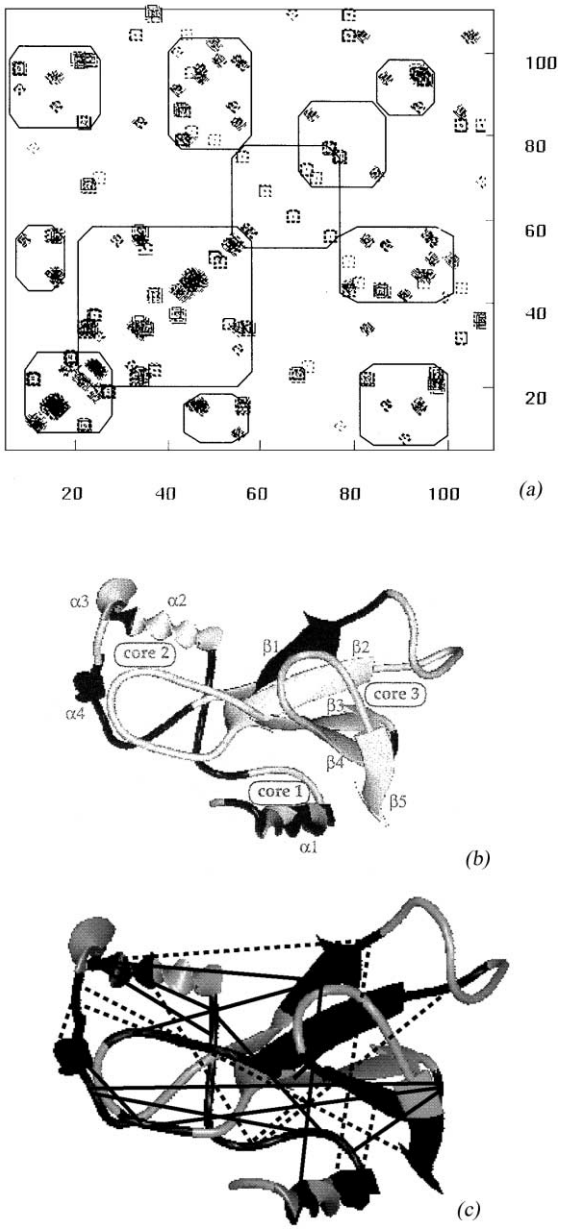

Fig. 5. (a) The contour map for $C_{ij}(\tau)$ at $\tau = 1500$ MC steps. The negative and positive correlations are given together in a symmetric map and the correlations are grouped by encircling contours. (b) Ribbon diagram of barnase on which the regions displaying short-range correlations corresponding to the diagonal contours in part (a) are coloured black. (c) Ribbon diagram of barnase on which the regions exhibiting long-range correlations corresponding to the off-diagonal contours in part (a) are coloured black. Positive and negative cooperativities are depicted by solid and dashed lines, respectively, between the two regions.

In Fig. 3a, the cross-correlation map for $\tau = 50$ MC steps is displayed as contours of highest correlation. For the diagonal when $i = j$, it is clear that $C_{ij} = 1$ but they are not displayed on the figure for clarity. At this short time, the regions of high correlation lie only near the diagonal, meaning that they are correlations in the rotations of the bonds, which are near in sequence and thus near in space. These regions are depicted in black on the ribbon diagram in Fig. 3b. It is noticed that these regions comprise the residues surrounding cores 2 and 3. The short-range correlated rotations at this time scale serve to preserve the local structure and may contribute to the stability of hydrophobic cores. On the other hand, a similar pattern is not observed for core 1. This is the major hydrophobic core formed by the packing of α1 against the β-sheet. This helix is capable of forming native-like structure even in the absence of tertiary interactions [24] and it is thought to be the major stabilising element of barnase [25]. However, the separation of α1 from the β-sheet by their relative movement is observed as an early event in unfolding simulations [22,26].

Fig. 4 is for $\tau = 500$ MC steps and analogous to Fig. 3. At this longer time, off-diagonal contours appear on the correlation map, indicating that the virtual bonds, which are more distant in sequence, may display cooperativity in their rotations at this time scale (Fig. 4a). The ribbon diagram in Fig. 4b depicts regions with short-range cooperative rotations coloured in black [corresponds to diagonal contours in Fig. 4a]; and long-range cooperatively moving units with dashed lines drawn between them [off-diagonal contours in Fig. 4b]. Comparison with the similar diagram for $\tau = 50$ MC steps (Fig. 3b) reveals that cooperatively moving regions are readjusted, yet some correlations are sustained. The contours on the upper and lower triangular portions of the correlation map are for negative and positive correlations, respectively.

As τ is further increased to 1500 MC steps, more sites, farther apart along the sequence but not necessarily close in space, display long-range cooperative motions. The positive and negative correlations are given together in a symmetric map in Fig. 5a in order to have a global view of these interactions and enable a comparison with the mutual perturbation/response analysis map of Hilser et al. [27]. The correlations are grouped by encircling contours that can be gathered for the ease of presentation. It is interesting to observe that there is a great similarity with these grouped regions and the cooperativity given by the mutual perturbation/response analysis of barnase under native conditions. The regions corresponding to the correlations appearing on the diagonal are shown in Fig. 5b and c as in respective parts of Figs. 3 and 4. When all three analogous ribbon diagrams are investigated, it is seen that there are three regions displaying short-range cooperativity at all times: C-terminal end of the loop connecting α1 and α2, α4 and the turn between β3 and β4. These regions, being the most interacting units with the other parts of the structure at $\tau =$ 1500 MC steps, may have an important contribution to the stability of the structure, by maintaining their short range cooperative motions at all times. In Fig. 5c, the positve and negative long-range correlations between different regions are depicted by solid and dashed lines, respectively, drawn between the two regions. The regions participating in these interactions are coloured black on the ribbon diagram. If we

elaborate Fig. 5a, the residues 23–55 behave in a mutually cooperative manner with themselves and with residues 78–100, and with some residues from the N-terminus of the structure. On the other hand, the similar behaviour is observed with the residues 55–74, which also exhibit mutually cooperative manner mostly with themselves. These correlations are in accord with the definitions of the two lobes of barnase, where lobe A: residues 23–51 and 74–82, and lobe B: residues 50–73. However, there are some residues in both lobes interacting with most of the residues in the structure, in addition to C-terminal end of the loop, $\alpha 4$ and the turn which are mentioned above. Thus, they link the behaviour of the two lobes, which behave in relatively independent fashion with each other and with the rest of the molecule. These residues are within the group 52–58 (lobe B) and 78–83 (lobe A). In a recent mutation study [28], the residues Gly53 and Asp54, belonging to one of the above groups, were identified as residues of vital importance to the structural integrity of barnase and its function. Apparently, some residues are more important than the others because of their involvement in the anisotropic collective motions of the structure.

4. Conclusions

In the present work, a recently developed off-lattice MC/Metropolis simulation technique has been employed to elucidate the conformational dynamics of barnase. The conformations generated around the native basin have been analysed by means of time-delayed conformational auto- and cross-correlation functions for different time windows to study the conformational peculiarities of different regions. There is a correlation between the extent of the torsional auto-correlated motions of virtual bonds at long times and the potential stability to unfolding of the respective regions. The cross-correlations between rotations of virtual bonds investigated at different times enables us to monitor the short- and long-range cooperativity distributed over time as well as in space. At short times, there are only short-range correlations between residues, which are close in space. At longer times, long-range correlations between residues, which are far along the sequence and which are not necessarily close in space, appear. There are regions with strong short-range cooperativity at all times, and those regions are found to be participating in the long-range cooperative motions of the structure, as well. This indicates that the dynamics of the local structure may have an important role in determining the overall structural dynamics and functional properties.

The distribution of the cooperative interactions indicates that some residues have a more important role with respect to the others and further that the interactions are all not bi-directional. The global motion coherence is supported by anisotropic collective motions, which apparently relates the structure, energy surface, dynamics and function.

Acknowledgements

It is a great pleasure for TH to dedicate this paper to Wayne Mattice, who has had a major contribution to the development of her career during her postdoctoral studies with him. Their collaborations over several years have been very enjoyable and inspiring for her.

Partial support from Bogaziçi University Research Funds, Project No: 00HA502 is gratefully acknowledged. NK thanks TÜBİTAK-BAYG for her BDP fellowship.

References

[1] Berendsen HJC, Hayward S. Curr Opin Struct Biol 2000;10:165.
[2] Remington S, Wiegand G, Huber R. J Mol Biol 1982;158:111.
[3] Boriack-Sjodin PA, Margarit SM, Bar-Sagi D, Kuriyan J. Nature 1998;394:337.
[4] Rose RB, Craik CS, Stroud RM. Biochemistry 1998;37:2607.
[5] Thomas A, Hinsen K, Field MJ, Perahia D. Proteins 1999;34:96.
[6] Clarke J, Fersht AR. Fold Des 1996;1:243.
[7] Bai Y, Sosnick TR, Mayne L, Eglander SW. Science 1995;269:192.
[8] Swint-Kruse L, Robertson AD. Biochemistry 1996;35:171.
[9] Hilser VJ, Freire E. J Mol Biol 1996;262:756.
[10] Hilser VJ, Freire E. Proteins 1997;27:171.
[11] Privalov PL. Adv Protein Chem 1979;33:167.
[12] Privalov PL. Adv Protein Chem 1982;35:1.
[13] Clarage B, Romo T, Andrews BK, Pettitt BM, Philips GA. Proc Natl Acad Sci USA 1995;92:3288.
[14] Balsera MA, Wriggers W, Oono Y, Schulten K. J Phys Chem 1996;100:2567.
[15] Haliloglu T, Bahar I. Proteins 1998;31:271.
[16] Bahar I, Erman B, Haliloglu T, Jernigan RL. Biochemistry 1997;36:13512.
[17] Haliloglu T. Proteins 1999;34:533.
[18] Kurt N, Haliloglu T. Proteins 1999;37:454.
[19] Wong KB, Clarke J, Bond CJ, Neira JL, Freund SMV, Fersht AR, Daggett V. J Mol Biol 2000;296:1257.
[20] Matouschek A, Serrano L, Fersht AR. J Mol Biol 1992;224:819.
[21] Caflisch A, Karplus M. In: Merz KM, LeGrand SM, editors. The protein folding problem and tertiary structure prediction. Boston: Birkhauser, 1994. p. 193–230.
[22] Caflisch A, Karplus M. J Mol Biol 1995;252:672.
[23] Matthews JM, Fersht AR. Biochemistry 1995;34:6805.
[24] Kippen AD, Arcus VL, Fersht AR. Biochemistry 1994;33:10013.
[25] Mauguen Y, Hartley RW, Dodson EJ, Dodson GG, Bricogne G, Chothia C, Jack A. Nature 1982;297:162.
[26] Li A, Daggett V. J Mol Biol 1998;275:677.
[27] Hilser VJ, Dowdy D, Oas TG, Freire E. Proc Natl Acad Sci USA 1998;95:9903.
[28] Axe DD, Foster NW, Fersht AR. J Mol Biol 1999;286:1471.
[29] Flory PJ. Statistical mechanics of chain molecules. New York: Hanser Publishers, 1988.
[30] Jernigan RL, Bahar I. Curr Opin Struct Biol 1996;6:195.
[31] Bahar I, Jernigan RL. Fold Des 1996;1:357.
[32] Lazaridis T, Karplus M. Curr Opin Struct Biol 2000;10:139.
[33] Buckle AM, Henrick K, Fersht AR. J Mol Biol 1993;234:847.
[34] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller EJ. J Chem Phys 1953;21:1087.
[35] Bahar I, Kaplan M, Jernigan RL. Proteins 1997;29:292.
[36] Bahar I, Jernigan RL. J Mol Biol 1997;266:195.
[37] Haliloglu T. Comp Theor Polym Sci 1999;9:255.