7. Scheunert, A. and Theile, E., *Pharmazie*, 1952, **7**, 776.
8. Salma, R. B., *Planta Med.*, 1973, **24**, 375–377.
9. Menounds, P., Staphylakis, K. and Gegiou, D., *Phytochemistry*, 1986, **25**, 761–763.
10. Al-Vaya, M. A., *Fitoterapia*, 1986, **57**, 179–182.
11. Merfort, I., *Phytochemistry*, 1997, **46**, 359–363.
12. Shay, H., *Gastroenterology*, 1945, **5**, 43–61.
13. Hegde, D. A., Khosa, R. L. and Goel, R. K., *Ancient Sci. Life*, 1994, **14**, 77–81.
14. Kunchandy, J., Khanna, S. and Kulkarni, S. K., *Arch. Int. Pharmacodyn.*, 1985, **275**, 123–138.
15. Szabo, S. and Szienji, S., *Trends Pharma Sci.*, 1987, **8**, 149–154.
16. Aly, A. and Scand, J., *Gastroenterology*, 1987, **137**, 43–49.
17. Rainsford, K. D., *Adv. Inflamm. Res.*, 1984, **6**, 51–64.
18. Alarcon, D. L., Martin, M. J. and Motilva, V., *J. Ethnopharmacol.*, 1994, **42**, 161–170.
19. Parmar, N. S. and Shikha Parmar, *Indian J. Physiol. Pharmacol.*, 1998, **48**, 343–351.

ACKNOWLEDGEMENT. We are grateful to the Head of our institution for providing necessary facilities to carry out the present investigation.

Received 12 April 2001; revised accepted 19 November 2001

Folding and unfolding of chicken villin headpiece: Energy landscape of a single-domain model protein

Goundla Srinivas and Biman Bagchi*

Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore 560 012, India

Folding and unfolding of a thermostable chicken villin headpiece subdomain, a 36-residue protein (HP-36), is studied by using Brownian dynamics simulations. The hydropathy scale of amino acids is used to obtain the varying interactions among the amino acids. A qualitative picture of the energy landscape funnel is obtained from simulations. Although there are several states near the minimum of the folding funnel, we could identify a stable native configuration. The energy of the folded protein scales with the hydrophobic contact parameter, as found in recent analyses. The model also allows for a description of cold denaturation by the salt-induced modification of the 'effective' interactions among the various amino acids. In this model, the kinetics of denaturation is found to be considerably different from that of folding–folding, seems to face more barriers and involves a more complex pathway.

THE folding of an extended, unfolded protein to its unique three-dimensional folded native state is a highly complex problem which has attracted a great deal of interest in recent years[1–3]. Recent theoretical studies have suggested that the size, stability and the topology of a protein influence the folding rate and mechanisms[2,4]. Recently, the concept of the free-energy landscape-guided folding[5] has been introduced, which has drawn a lot of attention. However, it is fair to say that a quantitative understanding of many facts of the protein-folding problem is still not within our reach.

Early statistical mechanical theories of protein folding have considered single-domain proteins[4,6] because of their simplicity. These initial theories were followed by a series of studies, which vastly improved our understanding of protein folding[7,8]. For example, Zwanzig *et al.*[7] showed that a small energy bias (of the order of a few k_BT against the locally unfavourable configurations can reduce the Levinthal's time to a biologically significant size. Later, Zwanzig[8] extended this model by introducing the reversible accessibility of the final folded state to study the kinetics of protein folding. Onuchic *et al.*[5] have introduced and elaborated on the concept of energy landscape. According to this latter development, the folding kinetics is determined by an energy landscape and for foldable proteins this resembles a funnel with a free energy gradient towards the native structure. The introduction of the concept of folding funnel provided a much-needed breakthrough in understanding the pathways of protein folding.

Duan and Kollman[9] carried out the first ever one-microsecond simulation of a protein in aqueous solution. They have studied the thermostable chicken villin headpiece subdomain, a 36-residue protein (commonly known as HP-36 protein) in the aqueous solution by explicit representation of water molecules. They found a native-like structure with two pathways.

Earlier, Shakhnovich and coworkers[10,11] have performed a lattice model simulation study and obtained the density of states. Their 3-dimensional figures (with energy, number of topological contacts and number of configurations as the coordinates) had a nice funnel-like structure. However, for a continuum (that is, off-lattice) model like the one treated here, the problem is more difficult. In particular, it is a difficult task to simulate folding of a model protein by the conventional simulation techniques such as molecular dynamics/Monte Carlo or all-atom simulation[9] or even by the Langevin dynamic simulations[12]. To overcome these difficulties we use the off-lattice Brownian dynamics (BD) simulation technique. BD simulation has the advantage that it does not include the detailed description as all-atom simulations (that is, does not include water molecules), but includes the qualitative features by an implicit representation through the intermolecular interactions which are now modified to reflect the hydrophobicity of each amino acid. In this study we report the construction of an energy landscape by carrying out the BD simulation study of a minimalist model HP-36 protein.

*For correspondence. (e-mail: bbagchi@sscu.iisc.ernet.in)

We note in this connection that earlier, Rojnuckarin *et al.*[13] obtained the folding times of a four-helix bundle protein model by carrying out the BD simulations. Their study showed the reliability of the minimalist models in obtaining the qualitative features of the protein folding. We also study the HP-36 protein unfolding by modelling the interactions, to account for the salt effects or the effects of urea – that is, the commonly known cold denaturation.

In this work, we study the HP-36 protein using the hydropathy scale of the constituting amino acids. A representative stable structure of this protein (obtained from the Protein Data Bank) is shown in Figure 1. The reason for studying this particular protein is that it is the smallest monomeric polypeptide characterized, consisting of only naturally-occurring amino acids, that autonomously folds into a unique and thermostable structure without disulphide bonds or ligand binding[14].

We model HP-36 protein as a necklace of different kinds of beads. Each bead in the sequence represents the corresponding amino acid in the protein sequence. There are 36 beads in the chain, since the number of residues in the original protein sequence (MLSDEDFKAV FGMTRSAFAN LPLWKQQNLK KEKGLF) is 36. All the beads are assumed to be of the same mass and size.

As pointed out by Kauzmann[15] and also Tanford[16] many years ago, one of the major driving forces of protein folding in aqueous media is the hydrophobic/hydrophilic nature of amino acids. This can be best represented by the hydropathy scale[16,17]. This scale arranges the standard free-energies of transfer from aqueous solutions to pure liquid hydrocarbons and provides a measure of hydrophobicity. In a certain sense, the hydropathy scale provides a quantitative measure of the liking of a particular amino acid for water. Depending on the hydropathy values we have categorized all the amino acids present in the HP-36 sequence into three classes; (i) hydrophobic, (ii) weakly hydrophilic, and (iii) strongly hydrophilic. In Table 1, the classification of amino acids is presented. The classification is done according to the following criterion. If the hydropathy value is positive, the amino acid is hydrophobic. On the other hand, among the hydrophilic amino acids (hydropathy value is negative), if the hydropathy value is smaller than -2.5 it is strongly hydrophilic, otherwise it is weakly hydrophilic. In Figure 2*a*, a schematic representation of the hydrophobic scale is presented. Figure 2*b* shows a pictorial representation of the colour code of the hydropathy values of both the original sequence and the simplified sequence due to the present categorization.

It should be made clear at this point that the transfer of hydropathy scale to intermolecular potential is to be understood as a 'solvent-averaged potential'. This can also be considered as a potential of mean force, well-known in colloids. Also, such a transfer of hydrophobicity to inter-atomic potential was perhaps first done by Dill[18] and coworkers in their lattice simulations. As can be seen from Table 2, interaction between two strongly hydrophilic groups is least favoured because water will shield them, while that between two hydrophobic groups is strongly attractive. Thus, these potentials are all water-averaged. In Table 2 we list the

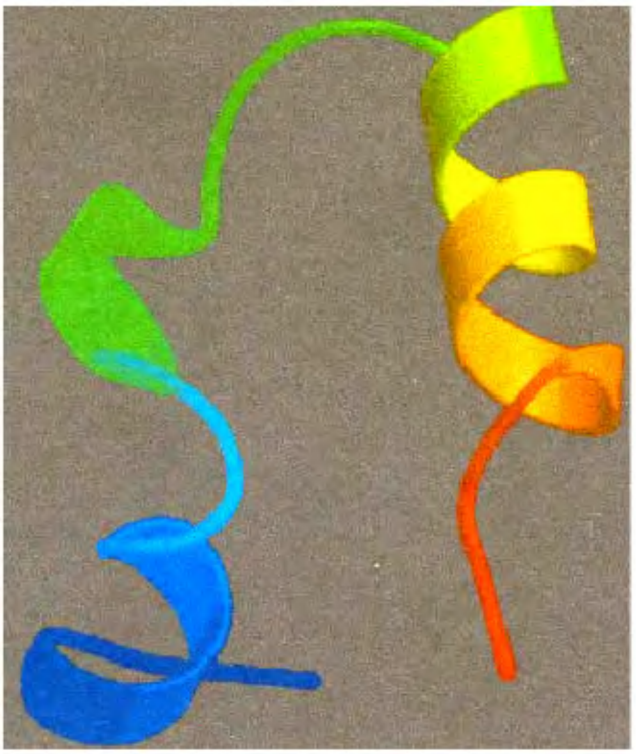

Figure 1. One of the stable structures of thermostable chicken villin headpiece subdomain, a 36-residue (HP-36) protein (PDB code: 1VII).

Table 1. Classification of the amino acids constituting the HP-36 protein, according to the hydropathy values

Amino acid	Category
AFLMPV	Hydrophobic
GSTW	Weakly hydrophilic
DEKNQR	Strongly hydrophilic

Table 2. Interaction parameter ($\varepsilon_{i,j}$) value for all the six different interactions in the folding of model HP-36 protein

Nature of the interaction	$\varepsilon_{i,j}$
Hydrophobic–hydrophobic	2.0 ε
Weakly hydrophilic–weakly hydrophilic	0.3 ε
Strongly hydrophilic–strongly hydrophilic	0.3 ε
Hydrophobic–weakly hydrophilic	1.0 ε
Hydrophobic–strongly hydrophilic	0.8 ε
Strongly hydrophilic–weakly hydrophilic	0.3 ε

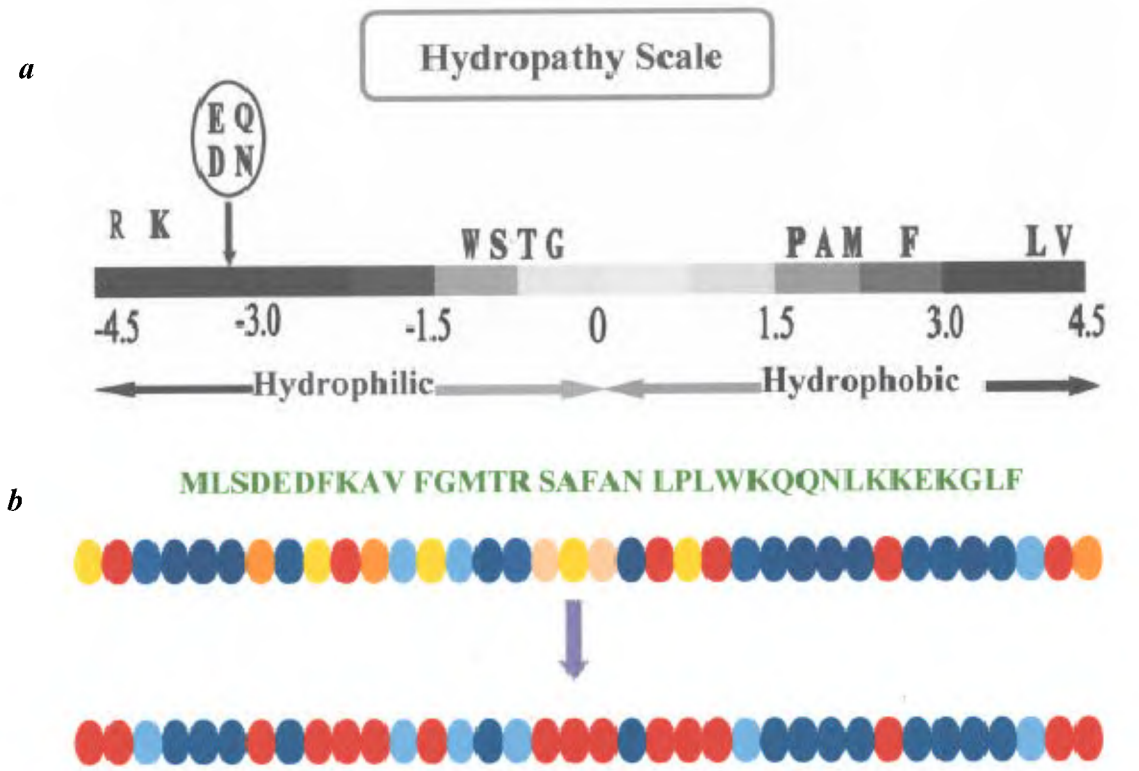

Figure 2. Schematic representation of modelling of the HP-36 protein (shown in Figure 1) by using the hydropathy values. ***a***, Schematic representation of the hydropathy scale. The hydrophilic nature decreases from blue to red; ***b***, Pictorial representation of the colour code of the hydropathy values of both the original sequence and the simplified sequence used in the present study.

Table 3. Interaction parameter ($\varepsilon_{i,j}$) value for all the six different interactions in the unfolding of model HP-36 protein

Nature of the interaction	$\varepsilon_{i,j}$
Hydrophobic–hydrophobic	0.3 ε
Weakly hydrophilic–weakly hydrophilic	1.0 ε
Strongly hydrophilic–strongly hydrophilic	1.0 ε
Hydrophobic–weakly hydrophilic	0.8 ε
Hydrophobic–strongly hydrophilic	0.8 ε
Strongly hydrophilic–weakly hydrophilic	1.0 ε

interaction strength parameter values for all the six different interactions.

It is well-known that a protein can be denatured or unfolded from its native state by adding salts (like guanadenium chloride) or chemical agents like urea. This is sometimes called cold denaturation. It is believed that these agents modify the interactions of water at the protein–water interface. Recent computer simulation studies[19] seem to suggest that the role of urea is to provide an energetically favourable environment of the hydrophobic groups in water. This provides the required driving force for unfolding. In the context of our hydropathy scale, an aqueous solution containing urea makes the hydrophobic groups less attractive to each other. Therefore, to motivate the unfolding of the folded protein, we have changed the interaction among the different residues (polymer beads) to reflect the altered scenario in the presence of urea in solution. This gives rise to a nice unfolding of the folded state, whose dynamics has been reported here. The modified interaction parameters for this case are listed in Table 3.

The beads in HP-36 interact via a site–site Lennard–Jones (LJ) potential. Neighbouring beads are connected via harmonic springs. The total potential energy of the chain can be written as,

$$U = U_b + U_{LJ} + U_s, \tag{1}$$

where U_b represents the bonding potential,

$$U_b = \sum_{i=2}^{N} \kappa(|\mathbf{r}_i - \mathbf{r}_{i-1}|)^2, \tag{2}$$

with $\kappa = 9$ in this study. N is the number of beads in chain and $\mathbf{r}_i$ is the position of bead i. The interaction between non-bonded beads is represented by the L–J-like potential,

$$U_{LJ}(r) = \varepsilon_{i,j}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \tag{3}$$

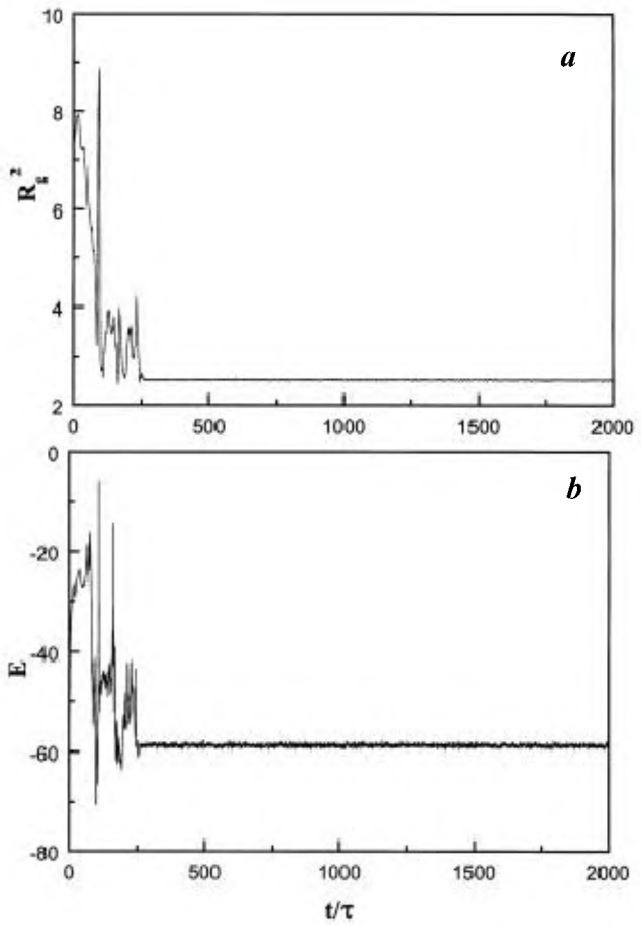

Figure 3. Variation in (*a*) mean square radius of gyration and (*b*) energy during the folding of a model HP-36 protein obtained from BD simulations is shown as a function of time.

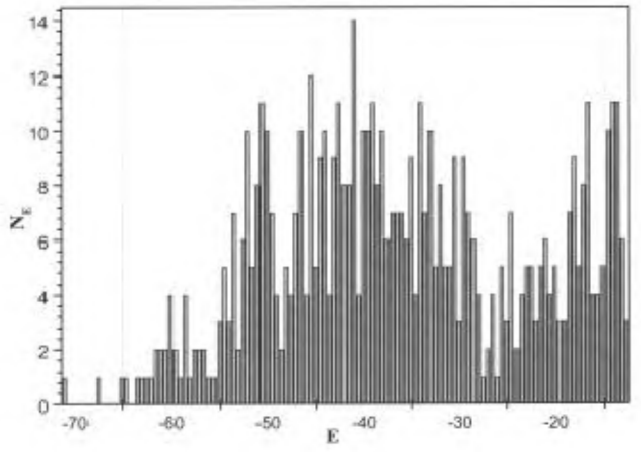

Figure 4. Histogram representation of the number of configurations found in BD simulations with the energy between E and $E + \Delta E$.

where σ is the LJ collision-diameter and $\varepsilon_{i,j}$ represents the interaction strength. The stiffness is introduced through the bending potential U_s,

$$U_s = S(\cos\theta - 1)^2, \tag{4}$$

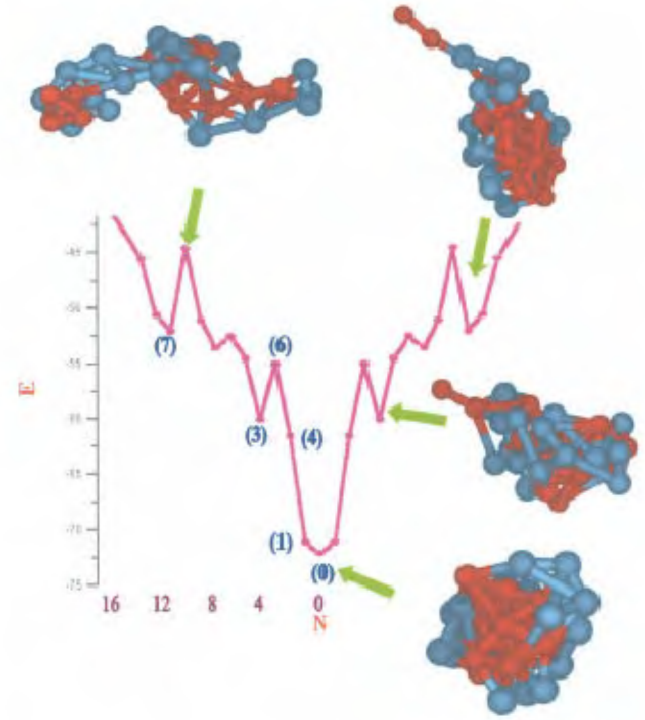

Figure 5. Energy landscape (the funnel) for the model HP-36 protein obtained from BD simulations. Configurations correspond to various energy states (unfolded, transition and native state). *X*-axis denotes the number of configurations with energy *E*.

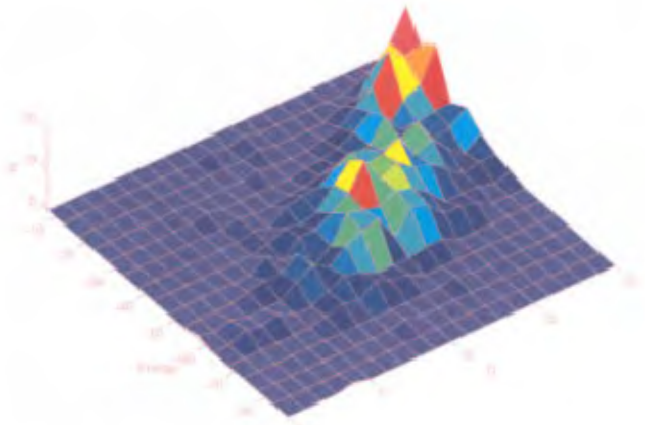

Figure 6. Full 3-dimensional plot of energy landscape obtained from BD simulations for a model HP-36. *Q* represents the distance from the native state in terms of topological contacts, while *N* represents the number of configurations.

where we set the chain stiffness $S = 1$. For convenience, we define $\varepsilon^* = \varepsilon/k_BT$, where k_BT is the thermal energy. The unit of time, τ is b^2/D_0, where the single-bead diffusion coefficient is denoted by D_0. The length is scaled by b, the bead diameter, as usual.

The time evaluation of the model protein is done according to the following equation of motion[20],

$$\mathbf{r}_j(t + \Delta t) = \mathbf{r}_j(t) + \frac{D_0}{k_BT} F_j(t)\Delta t + \Delta x^G(t), \tag{5}$$

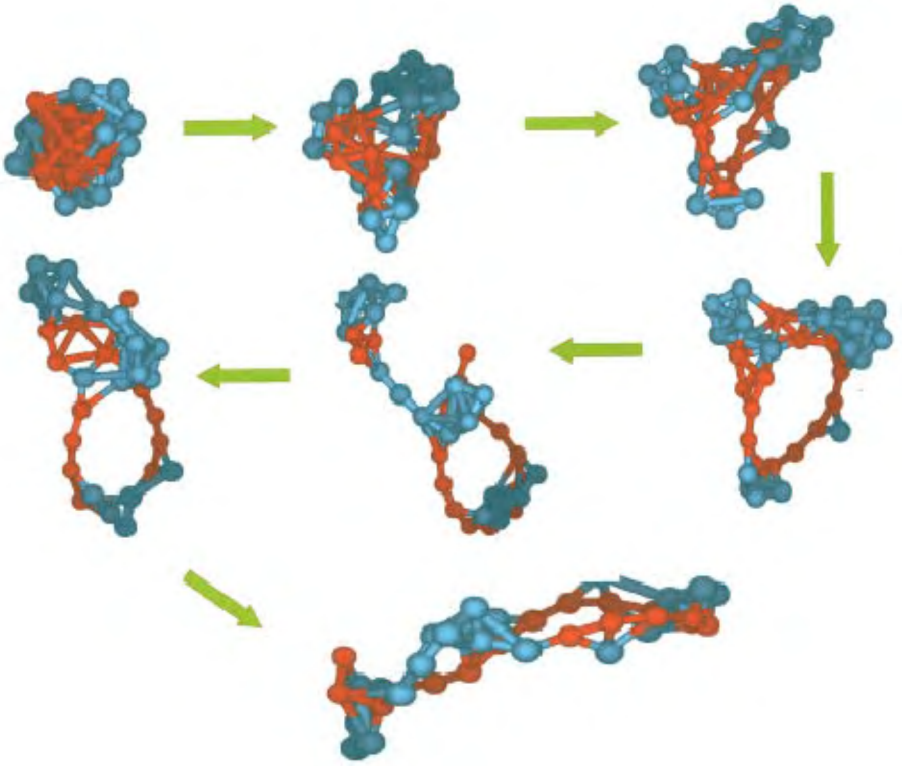

Figure 7. Snapshots of the unfolding of a model HP-36 protein as observed in BD simulations. Interaction parameters for the unfolding case are listed in Table 3.

where $\mathbf{r}_j(t)$ is the position of the jth bead at time t and the systematic force on j is denoted by $F_j(t)$. The random Brownian displacement, $\Delta X^G(t)$, is taken from a Gaussian distribution with zero mean and $2\Delta t$ variance.

For each trajectory, an initial configuration is selected from the Monte Carlo-generated equilibrium configurations at $\varepsilon^* = 0.1$. The temperature of the initial configuration is then instantaneously reduced by 0.1ε, after 2.5×10^5 BD steps. Five such quenches, each with a gap of 2.5×10^5 steps have been incorporated, to facilitate the folding. Further simulations for 2.5 million BD steps are carried out (subsequent to the quenching) to obtain the final configuration. Such a procedure is repeated for the model proteins with 1000 different configurations. More details on the simulation scheme can be found in a similar study on homopolymers[20,21].

In each simulation run, after choosing an initial configuration, the folding is followed till a stable final state is reached. It is important to note that in each simulation only one *protein* is simulated to obtain the corresponding final energy. In Figure 3, the variation in mean square radius of gyration (Figure 3 *a*) and energy (Figure 3 *b*) during the folding of a single model protein is shown. Oscillations in both energy and R_g^2 reveal that the protein experiences barriers during folding. We will discuss this point in detail later. In order to obtain a list of final energies, we had to repeat such simulations for N number of independent *single proteins* sampled from an equilibrium distribution. In other words, we have carried out N different simulation runs with independent protein configurations (with the same sequence) to obtain N number of final energies.

After the execution of the simulation, the energy landscape is obtained in the following fashion. The energies corresponding to the final states obtained from BD simulation are distributed into the bins of width 0.5 E. This gives an energy versus N_E histogram, where N_E is the number of configurations with energy E. This is shown in Figure 4. These histograms are then reorganized in the ascending order of number of configurations (N_E). The result, the evolution of funnel-like energy landscape, is shown in Figure 5. In the same figure, configurations corresponding to the native state, misfolded, unfolded and metastable states are also shown. This figure shows that the fast formation of native state requires funnel landscape. It is important to note that this funnel is multi-dimensional $E(N, Q)$, as shown earlier by Bryngelson and Wolynes[4] and also by Chan and Dill[6], and N and Q represent the number of configurations and the distance from the native state in terms of topological contacts, respectively. In Figure 5, we have also shown Q values for a few states. Figure 6 shows the full 3-dimensional picture of the energy landscape. Recently, Gutin *et al.*[10] also obtained a qualitatively similar multi-dimensional energy landscape by carrying out the lattice Monte-Carlo simulations. These

two figures clearly demonstrate that the present minimalist model is reliable in obtaining the qualitative features of protein folding. This is one of the main results of the present study.

The native configuration (corresponding to the minimum in the energy funnel) is chosen for the unfolding study. The protein unfolding simulation is carried out in an exactly opposite manner to that of the folding, but with different interactions (given in Table 3). In the study of unfolding, the interaction energies among the various amino acids are changed instantaneously at time $t = 0$. Snapshots of the unfolding of the HP-36 protein found in BD simulations are shown in Figure 7. As shown in this figure, the completely folded initial configuration gradually unfolds by breaking the native contacts to reach the fully extended state. To emphasize this point, in Figure 8, we have plotted the number of topological contacts as a function of time. For comparison, the variation in N_{topo} during the folding is also shown (inset). As shown in the main figure, N_{topo} is maximum at $t = 0$ which corresponds to the native state. In the unfolding case, variation in number of topological contacts (main figure) shows exactly the opposite

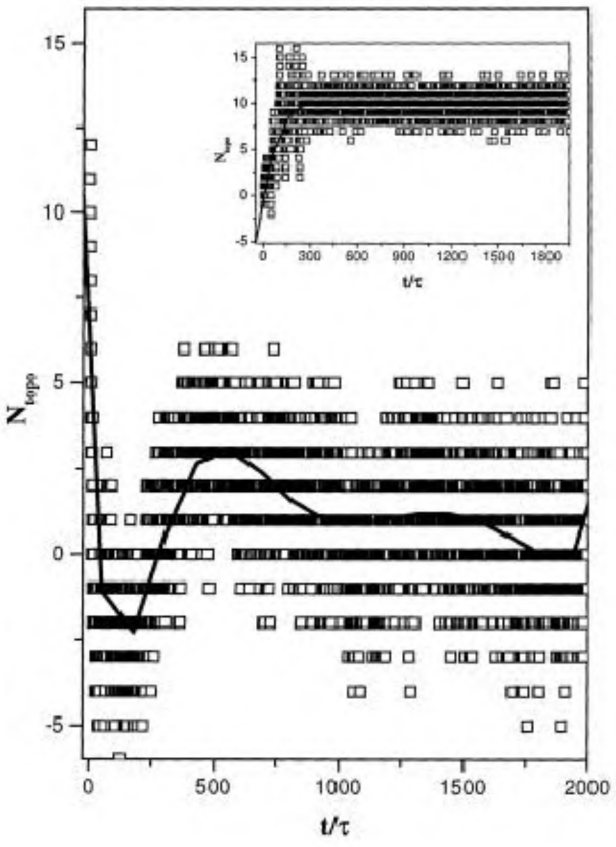

Figure 8. Variation in the number of hydrophobic topological contacts shown as a function of time. The main figure shows the result for unfolding, while the inset represents that for the folding. As can be seen, during the unfolding, N_{topo} shows an opposite trend to that observed in case of folding.

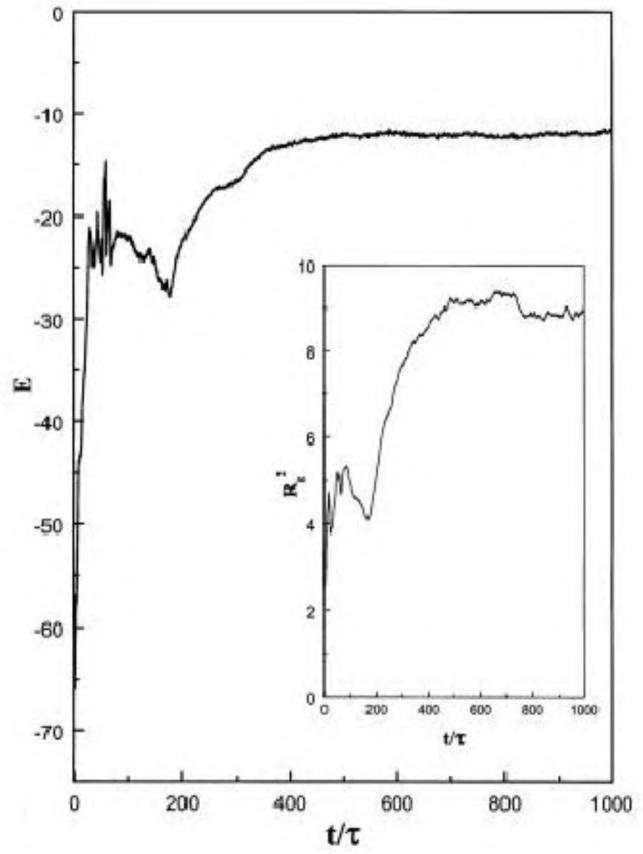

Figure 9. Time dependence of energy and mean square radius of gyration during the unfolding of a model HP-36 protein obtained from BD simulations.

trend to that of the folding (inset). On the whole, N_{topo} decreases during the unfolding, while it increases during folding. This figure clearly demonstrates the role of topological contacts intern the importance of the hydrophobic forces in protein folding/unfolding.

In Figure 9, the increase in energy during unfolding is shown. In the inset we show the same for the mean square radius of gyration. It is interesting to note the oscillatory dynamics, also recorded for topological contact formation in Figure 8. Such oscillations seem to indicate that the polymer, when unfolding, faces a barrier. Note that oscillations in the energy and in the mean square radius are larger for folding. Of course these oscillations may be very much dependent on the effective potentials used in this study. They indicate the presence of barriers along folding and unfolding, except that there seem to be more barriers during folding and also the pathway seems to be more complicated. In particular, the dynamics during folding shows considerably more oscillations in the final stage of folding, which is relatively smooth for unfolding.

Given the complexity of the real protein-folding problem, the ability of such a minimalist model to capture many of the essential features of protein folding is

quite interesting. The present study suggests that it may be possible to obtain the qualitative information on folding mechanism, folding rates and also about the stability, by modelling the more complex proteins in a similar way.

Future studies will explore the sensitivity to the potential employed and also focus on generalizing the set of potentials to accommodate more realistic potentials. Also, we have not made any study of the formation of the *specific* contacts that characterize the native state. Work in this direction is under progress.

1. Levinthal, C., in Proceedings of a Meeting held at Allerton House, Monticello, IL (eds Debrunner, P., Tsibris, J. C. M. and Munck E.), University of Illinois Press, Urbana, 1969, pp. 22–24.
2. Alm, E. and Baker, D., *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 11305–11310.
3. Zhou, Y., Hall, C. K. and Karplus, M., *Phys. Rev. Lett.*, 1996, **77**, 2822–2825.
4. Bryngelson, J. D. and Wolynes, P. G., *J. Phys. Chem.*, 1989, **93**, 6902–6915.
5. Onuchic, J. N., Wolynes, P. G., Luthey-Schulten, Z. and Socci, N. D., *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 3626–3630.
6. Chan, H. S. and Dill, K. A., *J. Chem. Phys.*, 1994, **100**, 9238–9257.
7. Zwanzig, R., Szabo, A. and Bagchi, B., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 20–22.
8. Zwanzig, R., *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 9801–9804; 1997, **94**, 148–150.
9. Duan, Y. and Kollman, P. A., *Science*, 1998, **282**, 740–744.
10. Shakhnovich, E. I., Gutin, A. M. and Abkevich, V. I., *Proc. Natl. Acad. Sci. USA*, 1995, **92**, 1282–1286.
11. Shakhnovich, E. I., *Curr. Opin. Struct. Biol.*, 1997, **7**, 29–40.
12. Honeycutt, J. D. and Thirumalai, D., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 3526–3529.
13. Rojnuckarin, A., Kim, S. and Subramaniam, S., *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 4288–4292.
14. McKnight, J. C., Doering, D. S., Matsudaria, P. T. and Kim, P. S., *J. Mol. Biol.*, 1996, **260**, 126–134.
15. Kauzmann, W., *Adv. Protein Chem.*, 1959, **14**, 1–63.
16. Tanford, C., in *The Hydrophobic Effect*, Wiley, New York, 1980, 2nd edn.
17. Stryer, L., *Biochemistry*, W. H. Freeman and Company, New York, 1995, 4th edn.
18. Dill, K. A., *Protein Sci.*, 1999, **8**, 1166–1180.
19. Cheng, Y.-K. and Rossky, P. J., *Nature*, 1998, **392**, 696–699; Cheng, Y.-K., Sheu, W. S. and Rossky, P. J., *Biophys. J.*, 1999, **76**, 1734–1743.
20. Srinivas, G., Yethiraj, A. and Bagchi, B., *J. Phys. Chem. B*, 2001, **105**, 2475–2478; *J. Chem. Phys.*, 2001, **114**, 9170–9178.
21. Srinivas, G. and Bagchi, B., cond-mat 0105138.

ACKNOWLEDGEMENTS. Financial support from DST, India is gratefully acknowledged. G.S. thanks CSIR for a research fellowship.

Received 10 October 2001; revised accepted 6 December 2001

Reduced susceptibility to deltamethrin in *Anopheles culicifacies sensu lato*, in Ramnathapuram district, Tamil Nadu – Selection of a pyrethroid-resistant strain

P. K. Mittal[†,], T. Adak[†], O. P. Singh[*], K. Raghavendra[*] and S. K. Subbarao[*]**

[†]Malaria Research Centre, Indian Council of Medical Research, 2, Nanak Enclave, Radio Colony, Delhi 110 009, India
[*]Malaria Research Centre, Indian Council of Medical Research, 22-Sham Nath Marg, Delhi 110 054, India

***Anopheles culicifacies sensu lato*, the major vector of malaria in the plain areas of rural India, has become resistant to commonly used insecticides, viz. DDT, HCH and malathion in most parts of the country. To control resistant *An. culicifacies*, synthetic pyrethroids have been sprayed during the past one decade in some areas, but so far resistance to pyrethroids has not been reported in this species. This paper reports the reduced susceptibility of *An. culicifacies* from Rameshwaram Island, Ramnathapuram district, Tamil Nadu, to deltamethrin, a synthetic pyrethroid. Knock-down bioassays revealed more than two-fold higher values of KT_{50} and KT_{90} (time for knock-down of 50% and 90% exposed mosquitoes) in *An. culicifacies* from Rameshwaram Island, than populations from other areas. Laboratory selection of the adult mosquitoes of Rameshwaram (Rmr) strain against deltamethrin for 12 generations resulted in the development of 31- and 15-fold higher resistance at LT_{50} and LT_{90} levels, respectively. The selected deltamethrin-resistant strain also showed cross-resistance to other synthetic pyrethroids, viz. lambdacyhalothrin, cyfluthrin, permethrin and bifenthrin.**

ANOPHELES culicifacies sensu lato, the major vector of malaria in most parts of the rural plain areas in India, has developed widespread resistance to DDT and HCH and also to malathion in some areas, where it has been used extensively for indoor residual spraying[1–4]. As a result, synthetic pyrethroids have been introduced as an alternative to control double (DDT and HCH) or triple (DDT, HCH and malathion) resistant *An. culicifacies* in India, either in the form of indoor residual spray or as impregnant on mosquito nets. Synthetic pyrethroids have been reported to be highly effective for the control of *An. culicifacies* populations[5,6]. Though past experiences have shown the development of resistance in *An. culicifacies* to different insecticides which have been used for indoor residual spray, resistance to pyrethroids in *An. culicifacies* from India has not been reported so

**For correspondence. (e-mail: adak@vsnl.com)