How does a β-hairpin fold/unfold? Competition between topology and heterogeneity in a solvable model

Chinlin Guo*, Herbert Levine*†, and David A. Kessler‡

*Department of Physics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0319; and ‡Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel

We study the competition between topological effects and sequence inhomogeneities in determining the thermodynamics and the un/folding kinetics of a β-hairpin. Our work utilizes a new exactly solvable model that allows for arbitrary configurations of native contacts. In general, the competition between heterogeneity and topology results in a crossover of the dominant transition state. Interestingly, near this crossover, the single reaction coordinate picture can be seriously misleading. Our results also suggest that inferring the folding pathway from unfolding simulations is not always justified.

In their important work (1), Muñoz *et al.* discovered a sharp un/folding transition between two dominant thermodynamic ensembles during hairpin formation in the small peptide, the 16-residue C-terminal fragment of streptococcal protein G (sequence GEWTYDDATKTFTVTE). In such a system, it is often assumed that the formation of a folding droplet (2, 3) starting from the β-turn (4, 5) is the determining factor in transition kinetics. However, given that peptides are inherently inhomogeneous, nucleation sites might develop preferentially at strong heterogeneities and thereby change the transition kinetics. This was predicted to be the case for this specific system by a recent unfolding simulation (6). Thus, we are motivated to study the factors that govern the competition between topological effects such as turns and heterogeneity.

To accomplish this, we need a model that allows for all of these partially folded configurations to occur and compete. The popular "single sequence stretch approximation" (4) does not suffice in this respect. As introduced by Schellman (7) and used, for example, in refs. 8–10, this approach assumes that one need keep only configurations with a contiguous native peptide conformation and furthermore use only a discrete degree of freedom for native vs. nonnative backbone states. It is clear, however, that this approach rules out the possibility of heterogeneity dominance. For the aforementioned fragment, requiring contiguous native conformations disallows early formation of the hydrophobic cluster with contacts W-V and Y-F.

One could update this model as suggested in refs. 3 and 8–10 and include multiple sequences. But we prefer instead to use an exactly solvable hairpin model, which we introduce below. Aside from its capability of allowing nucleation at the heterogeneous "hot spots," our model has the added advantage of correctly treating the fact that the backbone lives in three-dimensional continuous space. Within this model, we study the influence of heterogeneity on the overall cooperativity of the transition as well as on the aforementioned transition kinetics. We also comment on the validity of the assumption that folding and unfolding use related pathways and on the validity of determining kinetic barriers from a landscape based on a single reaction coordinate.

The Model

We consider a polymer composed of two interacting Gaussian strands connected by a β-turn. Our basic assumption (11) is that the effective Hamiltonian can be written in terms of the spatial coordinates of the interacting residues that form native interstrand H-bonds, $\vec{x}_i^{(s)}$; here i is a residue index counted from the β turn ($i = 0$), and the superscript $s = N, C$ stands for the N-, C-terminal directed polymer. An exact model requires, of course, including nonnative interactions as well as loop self-avoidance effects; we expect these to be relatively unimportant for properly designed sequences (that favor native terms), which fold into a linear hairpin with no need for packing together multiple pieces of the peptide (limiting the role of self avoidance). We can decompose the coordinates into the (irrelevant) mean coordinate $\tilde{X}_i \equiv 1/2\,(\vec{x}_i^{(N)} + \vec{x}_i^{(C)})$ and the interstrand separation $\vec{x}_i \equiv (\vec{x}_i^{(N)} - \vec{x}_i^{(C)})$. Also, we use the notation $\Delta_i = 1$ or 0 to indicate whether the ith residue pair is in contact. Specifically, we define $\Delta_i = 1$ if the local interstrand distance $|\vec{x}_i|$ falls into an effective attraction window $|\vec{x}_i| \leq r_{0,i}$ and 0 otherwise. This "box" approximation has also been used elsewhere (11, 12).

The first component of our Hamiltonian accounts for the entropy of unfolded segments. Specifically, for a completely unfolded hairpin from residue $i = n$ (distal end) to $i = 0$, the loop entropy is governed by a Gaussian Hamiltonian $1/2 \sum_{i=1}^{n} \kappa_i\,|\nabla \vec{x}_i|^2 + \kappa_0/2\,|\vec{x}_0^{(N)} - \vec{x}_0^{(C)}|^2$, where $\nabla \vec{x}_i = \vec{x}_i - \vec{x}_{i-1}$ is the vector connected nearest-neighbor residue pairs, and κ_i is the unfolded local backbone stiffness. On the other hand, for any residue i that is in contact, we approximate $r_{0,i} \sim 0$ (as compared to $(\beta \kappa_i)^{-1/2}$) and replace the Gaussian term of contiguous unfolded residues by the term $\lambda_i\,|\vec{x}_{i \pm 1}|^2$. Finally, we assume that possible H-bonds start from $i = 1$ and integrate out $\vec{x}_0^{(N,C)}$. Thus, we obtain the loop Hamiltonian

$$H_{\text{loop}} = \sum_{i=2}^{n} \frac{\kappa_i}{2}(1 - \Delta_i)(1 - \Delta_{i-1})|\nabla\vec{x}_i|^2$$

$$+ \sum_{i,j=1}^{n} \frac{\gamma_{ij}}{2}\Delta_i(1 - \Delta_j)\kappa_j J_{ij}|\vec{x}_j|^2 + \frac{\kappa_1\sigma}{2}(1 - \Delta_1)|\vec{x}_1|^2, \quad [1]$$

where $J_{ij} = 1$ if i, j are nearest neighbors and 0 otherwise, $\sigma = \kappa_0/(\kappa_0 + \kappa_1)$, and $\gamma_{ij} = \lambda_i/\kappa_j$.

The second ingredient of the Hamiltonian is the interstrand interactions H_{int}. First, there is a pairwise interaction to mimic the native H-bond formation and side-chain packing between the two strands. This has the form $-V_{1,i}\Delta_i$ with $V_{1,i} \geq 0$. Second, there is a term representing the possible formation of hydrophobic side-chain clusters. Because the side-chain orientations
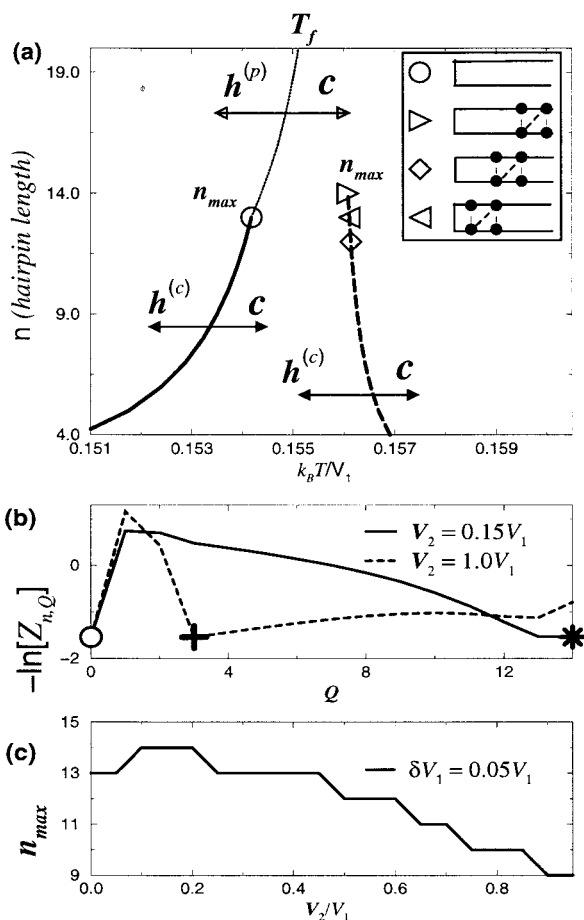
Fig. 1. (a) The typical folding temperature T_f for different polymer lengths with a two-state-like transition. The solid curve indicates T_f for peptides without heterogeneities. Above n_{max} (indicated by ○), the curve is replaced by a thin one where the transition is between a partially folded ensemble and the coil state. Here c represents the "coil" state, and $h^{(c)}$, $h^{(p)}$ stand for the "complete," "partial" hairpin structure, respectively. In the presence of a hydrophobic cluster (with locations indicated in the box), T_f is increased (the dashed curve) with modified n_{max} (◇, central; ▷, distal-end; ◁, β-turn); the transition curves above n_{max} for these cases are not shown. Here $\sigma = 1$, $V_2 = 0.15 V_1$, $\delta V_1 = 0.05 V_1$, $\gamma = 3$, and $v_0 = 10^{-4}$. (b) The free energy $-\ln[Z_{n,Q}]$ of different ensembles at T_f for $n = 14$ polymers with a distal-end hydrophobic cluster. Here the parameters aside from V_2 are the same as in a. The symbols "○," "*," and "+" stand for the coil, completely folded, and partially folded ensembles. Note that in the strong heterogeneity case, there is a highly folded ensemble ($Q = 13$) as an intermediate. Curves of this kind allow for the determination of n_{max}. (c) The dependence of n_{max} on the variation of V_2, in the case of having a distal-end hydrophobic cluster.

Fig. 2. Two typical heterogeneity-dominant unfolding simulations averaged over 50 hairpins for the single (*Top*) and double (*Middle*) heterogeneity case. The system starts in the native state, and the parameters are chosen to lie above the crossover curve in Fig. 4a ($k_BT/V_1 = 0.19$, $\delta V_1 = 0.4 V_1$) and Fig. 5a ($k_BT/V_1 = 0.192$, $\delta V_1 = V_2 = 0.3 V_1$). (*Bottom*) The typical heterogeneity-dominant folding in the double-heterogeneity case ($k_BT/V_1 = 0.13$, $\delta V_1 = V_2 = 0.4V_1$), performed by temperature quenching. The local interstrand distance at residue i is defined as the average of $\langle|\bar{x}_i|\rangle/r_0$ and indicated by the color map (deep blue, folded; red, totally unfolded; dark red, $\langle|\bar{x}_i|\rangle/r_0 \geq 3.0$). Time is in arbitrary units.

The Partition Function

The partition function of our model is given by

$$Z_n = \prod_{k=1}^{n} \int \left[\frac{\beta\kappa_i}{2\pi}\right]^{\frac{d}{2}} d^d x_k e^{-\beta H}, \qquad [3]$$

where $d = 3$ is the dimensionality of interest. It will be of interest to decompose this sum into terms with a fixed number of contacts Q (6, 13). Specifically, we define $Z_{n,Q}$ as the thermodynamic probabilities of the subensembles specified by the number of contacts Q. We will now discuss briefly how to develop a recursion relation that allows for the computation of these objects. A more detailed discussion, albeit for a homogeneous model, appears elsewhere (11).

$Z_{n,Q}$ is determined by the energy gain of Q local contacts and the entropy of the intervening unfolded segments. To compute this entropy, we define a functional M_{ij} as the path integral of a completely unfolded segment connecting residue j and i, weighted by additional terms on the two ends,

are aligned alternatively in the hairpin structure (see figure 1 in ref. 1), the typical situation leading to clusters is the presence of two hydrophobic residues at locations h and $h + 2$, with one of them carrying a large aromatic ring (tryptophan, e.g.). From the unfolding simulation (6), the compactness and hence the solvation of the hydrophobic cluster are improved if both residues are in the contact position. This leads to a many-body term $\phi \equiv -V_2\Delta_h\Delta_{h+2}$. In general, we may have more complicated high-order terms because of the presence of multiple hydrophobic residues. Our entire Hamiltonian is then

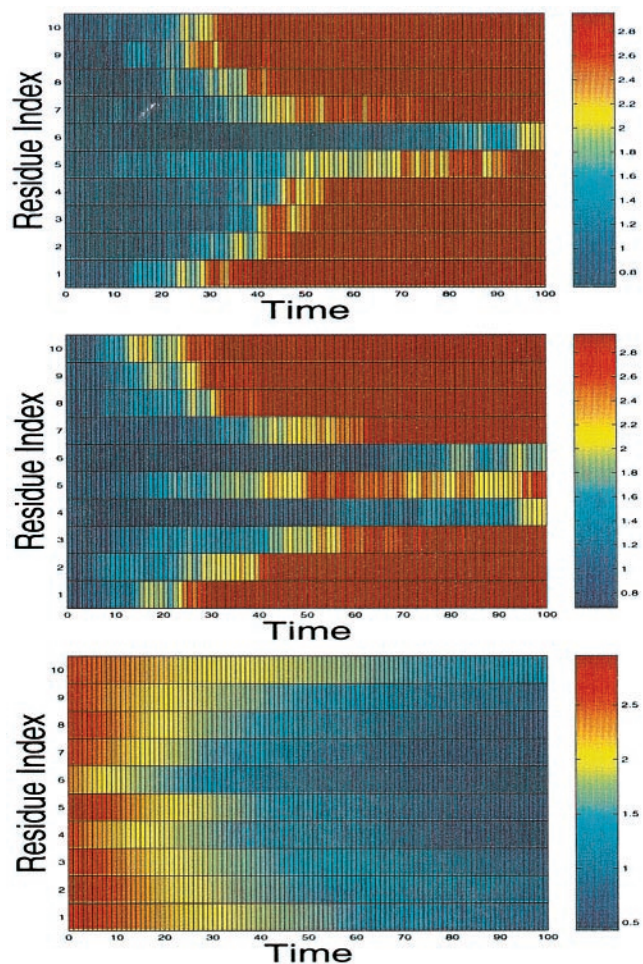$$H = H_{loop} - \sum_{k=1}^{n} V_{1,k}\Delta_k + \phi \qquad [2]$$

$$M_{i,j}(\mu_1,\mu_2) = \prod_{k=j}^{i} \int \left[\sqrt{\frac{\beta\kappa_k}{2\pi}} \right]^d d^d x_k [1 - \Delta_k]$$

$$\times\, e^{-\beta H_{\text{loop}} - \frac{\beta\kappa_i}{2\mu_1}(1-\Delta_i)|\vec{x}_i|^2 - \frac{\beta\kappa_j}{2\mu_2}(1-\Delta_j)|\vec{x}_j|^2}, \quad [4]$$

with the sums running from j to i in Eq. **1** for H_{loop}. It is easy to derive the recursive relationship

$$M_{i,j}(\mu_1,\mu_2) = \left[\frac{\mu_1}{1+\mu_1}\right]^{\frac{3}{2}} M_{i-1,j}\left(\frac{\kappa_i[\mu_1+1]}{\kappa_{i-1}},\mu_2\right) - q_i M_{i-1,j}(\gamma_{ij}^{-1},\mu_2)$$

$$[5]$$

supplemented by the boundary condition $M_{j,j}(\mu_1,\,\mu_2) = [\mu_1 \mu_2/(\mu_1 + \mu_2)]^{3/2} - q_j$, with $q_i = \frac{4\pi}{3} r_{0,i}^3 [\sqrt{\beta\kappa_i/2\pi}]^3$. By definition, $M_{i,j}(\mu_1,\mu_2) = 0$ if $j > i$. For consistency with the approximations utilized above, we should set $q = 0$ in these formulas.

Next, we define $W_{n,Q}$ as the ensemble specified by contact number Q with the further constraint that the polymer distal end $i = n$ is in the contact position; we have $W_{s,a} = 0$ if $a \leq 0$. Then, as in refs. 1, 3, 8, 10, and 11, we can use these as building blocks to construct $Z_{n,Q}$. Specifically, in the absence of any clustering effect, we have for $1 \leq s \leq n$ and $1 \leq Q \leq s$,

$$Z_{s,Q} = W_{s,Q} + \sum_{k=Q}^{s-1} M_{s,k+1}(\infty, \xi_k') W_{k,Q} \quad [6]$$

$$\frac{W_{s,Q}}{[q_s e^{\beta V_{1,s}}]} = W_{s-1,Q-1} + \sum_{k=Q-1}^{s-2} M_{s-1,k+1}(\xi_s, \xi_k') W_{k,Q-1}$$

$$+ \delta_{Q,1} M_{s-1,1}(\xi_s, \xi_0) + \delta_{s,1}\delta_{Q,1}, \quad [7]$$

with $\xi_k = \gamma_{k+1,k}^{-1}$, $\xi_k' = \gamma_{k,k+1}^{-1}$ for $1 \leq k \leq n$, $\xi_0 = 1/\sigma$, and $Z_{s,0} = M_{s,1}(\infty, \xi_0)$ (the coil state with $Q = 0$). Here, the strength of the native contacts is determined by $\alpha_s \equiv q_s e^{\beta V_{1,s}}$, which is assumed to remain finite even as we take the limit of small q.

Finally, we generalize this formula to include the clustering effect. We assume one hydrophobic cluster involving the two side-chain groups at residues h and $h + 2$. The recursion relations Eq. **7** will be unchanged for $s \neq (h + 2)$, but for $s = h + 2$, we have

$$\frac{W_{h+2,Q}}{[q_{h+2} e^{\beta V_{1,h+2}}]} = e^{\beta V_2} M_{h+1,h+1}(\xi_{h+1}, \xi_h') W_{h,Q-1}$$

$$+ e^{\beta[V_2 + V_{1,h+1}]} W_{h,Q-2} + \delta_{Q,1} M_{h+1,1}(\xi_{h+1}, \xi_0)$$

$$+ e^{\beta V_{1,h+1}} \sum_{k=Q-2}^{h-1} M_{h,k+1}(\xi_h, \xi_k') W_{k,Q-2}$$

$$+ \sum_{k=Q-1}^{h-1} M_{h+1,k+1}(\xi_{h+1}, \xi_k') W_{k,Q-1}. \quad [8]$$

We next turn to an evaluation of the equilibrium phase diagram of our model.

Phase Diagrams

It was shown in ref. 11 that one can obtain two-state behavior in homogeneous models of this class for large enough values of the folding-induced stiffness γ. Here, our first concern is the effect that the heterogeneities can have on system cooperativity. To address this issue, we consider the simple situation of constant κ, γ, V_1, r_0, and one hydrophobic cluster. We assume this cluster gives rise to a many-body term as detailed above and also, via side-chain packing, to an increase δV_1 of the (otherwise homogeneous) contact energy for the two hydrophobic residues.

To present our results, we rescale all the energies by V_1 and introduce the dimensionless parameter ν_0 via $q = \nu_0 [V_1/k_B T]^{3/2}$. In Fig. 1a, we plot for the case of $\nu_0 = 10^{-4}$ the variation of the folding temperature T_f as we change the polymer length n. In general, there is a maximal polymer length n_{max} that allows for a two-state transition between the completely folded hairpin $Z_{n,n}$ and the coil state $Z_{n,0}$; above n_{max}, the folding transition is defined as between the coil state and a dominant partially folded ensemble. We should mention that such a dominant ensemble is not fixed; as the temperature is reduced, it moves to the completely folded state. Also, a larger γ corresponds to a bigger n_{max}, because more entropy will be lost in configurations that contain interleaved folded and unfolded segments (11).

We now focus on the effect of heterogeneity. For moderate V_2 and δV_1, T_f is slightly shifted, but n_{max} is more or less unchanged. As V_2 is increased further, however, the partially folded state with an intact cluster becomes more significant, and the system loses two-state behavior between $Z_{n,n}$ and $Z_{n,0}$ at smaller n (Fig. 1 b and c). The extent to which one would be able to detect the presence of important partially folded states in the thermodynamics depends of course on the sensitivity of the probes to whether the hairpin is fully folded vs. their sensitivity to more local features, such as whether the hydrophobic cluster has been formed.

Simulations

To get some idea how folding and unfolding occur in our model, we performed Langevin simulations on 50 independent hairpins, using the Verlet algorithm proposed in ref. 14. Specifically, we have the time step $\Delta t = 10^{-4} \sqrt{mr_0^2/V_1}$ with the mass m and V_1 set to unity. The box potential $\Delta_i (|\vec{x}_i|)$ is approximated by $\exp[-(|\vec{x}_i|/r_0)^s]$ with $s = 100$. The damping coefficient is taken from ref. 14. Also, we used a stiffness $\kappa = 0.1$ to allow for good visualization of the individual residue kinetics; too large a κ would force all residues to un/fold almost simultaneously.

For small inhomogeneity, we found, as expected, that the topology determined the folding pathway, i.e., that folding took place from the β-turn and unfolding from the distal end (data not shown). In the case of strong heterogeneity, however, we observed dramatic changes to the un/folding patterns (Fig. 2). Now, the hairpin unfolds from both the β-turn and distal end (similar to the observations in ref. 6), leaving the decomposition of the hydrophobic cluster for last. Folding occurs in the reverse manner, with the cluster forming first. Thus, there is a crossover in the kinetics from being topology dominated to being heterogeneity controlled. Next, we analyze this crossover by studying the free energy landscape of our model.

The Transition Kinetics

We begin by assuming that the transition rate has an Arrhenius form $k_{tx} \sim e^{-\beta\Delta F}$, where ΔF is the free energy difference between the transition and metastable states. In our system, this is just the thermodynamic probability ratio between the transition and metastable ensembles. In the absence of heterogeneity, the metastable states are simply the coil state for folding (i.e., if $T < T_f$) or the folded state (if $T > T_f$). In the presence of heterogeneity, however, there might be additional metastable states (6). As in the simulations, we focus our discussion on the case of having a "hot-spot" near the hairpin center. Specifically, two cases are considered: (i) two hydrophobic residues at locations h and $h + 2$ with $\delta V_{1,i} = \delta V_1 (\delta_{i,h} + \delta_{i,h+2})$ plus an additional clustering energy V_2, and (ii) a single hydrophobic residue at the location $h + 2$ with $\delta V_{1,i} = \delta V_1 \delta_{i,h+2}$.
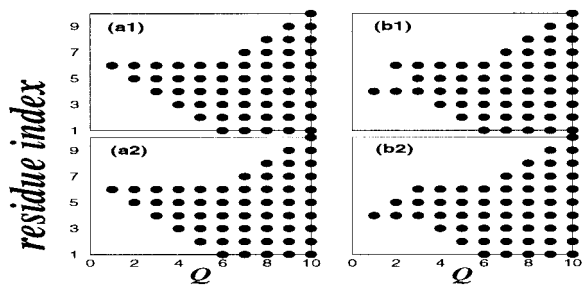
Fig. 3. Typical hetero-droplet profiles. Here (a1,2) stand for the single (residue $i = 6$), and (b1,2) represent the double (residue $i = 4, 6$ with a clustering effect) heterogeneity cases. $\delta V_1/V_1 =$ (a1) 0.4 (a2) 0.2, and $\delta V_1 = V_2 =$ (b1) 0.4 (b2) 0.2 V_1 with $k_B T/V_1 = 0.19$, stand for the strong/mild heterogeneity cases, respectively. Other parameters include $n = 10$, $\gamma = 3$, $\sigma = 1$, and $\nu_0 = 10^{-4}$. The folded residues are indicated by ●. Note that the droplet profiles are kinetically connected by add/dropping a folded residue, and in (b1), the $Q = 2$ droplet corresponds to four noncontiguous sequence stretches.

To clarify the transition rate and hence the rate-limiting structure, we use the droplet concept (2). A droplet is a configuration containing one or more sets of contiguous folded residues, and our basic method involves computing the free energy of differing droplets. For the homogeneous case, it is easy to see that we need consider only two possible droplets: one
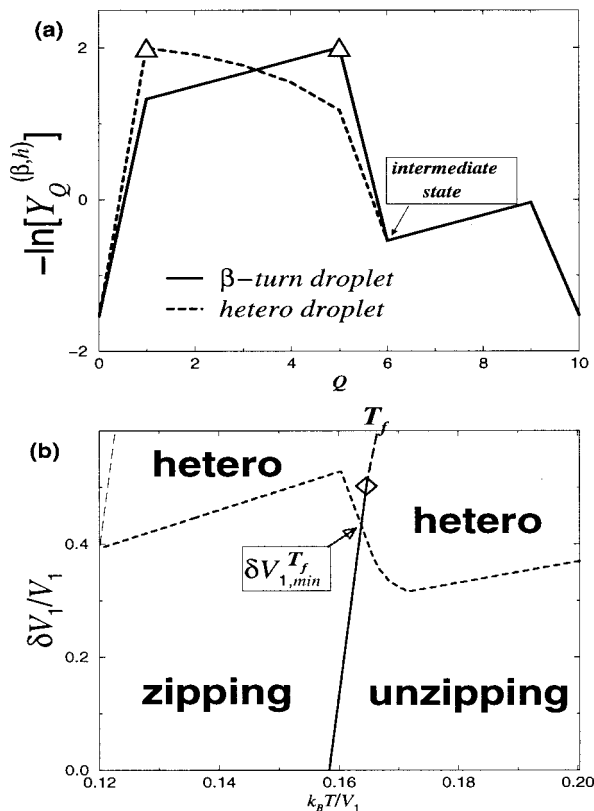

Fig. 4. (a) The typical free energy profiles for the β-turn and hetero-droplets at a T_f point indicated by the hollow arrow in b, in the single heterogeneity case. The barriers encountered in transition states are indicated by Δ. Here $n = 10$, $\gamma = 3$, $\sigma = 1$, $V_2 = 0$, and $\nu_0 = 10^{-4}$. (b) The typical crossover diagram. The solid curve is T_f for transitions between the coil and complete hairpin states; above $\diamond$, the transition is instead between the coil and a partially folded ensemble (the long-dashed curve). The crossover is indicated by the dashed curve. The thin long-dashed curve indicates where the barriers disappear.
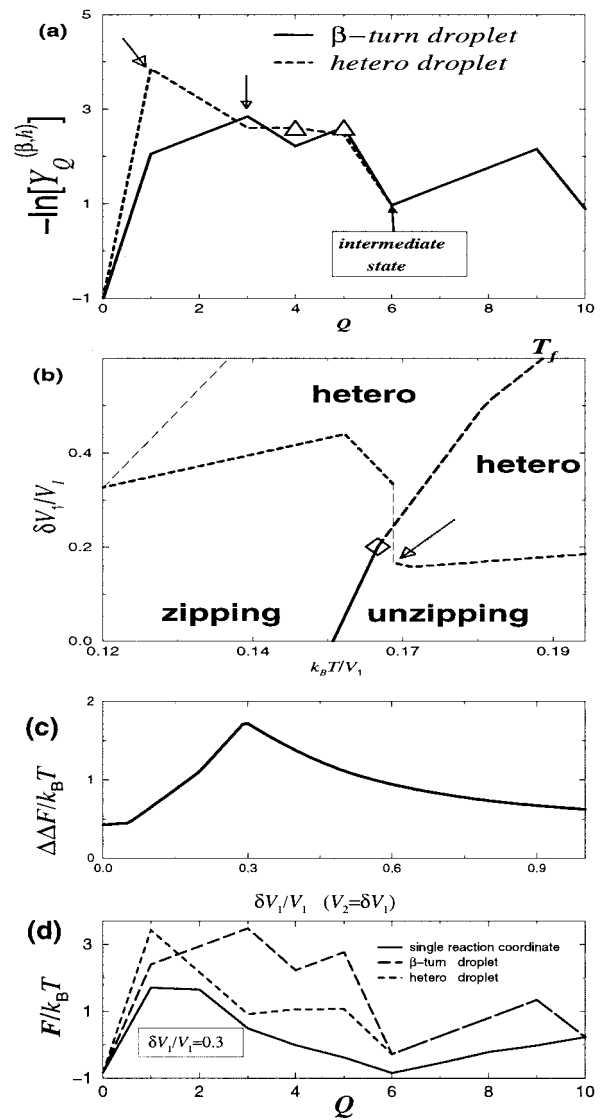

Fig. 5. (a) The typical free energy profiles for the hetero- and β-turn droplets at a T_f point indicated by the hollow arrow in b, in the double heterogeneity case. The barriers for the unfolding transition are indicated by Δ. Note that the folding barrier (the hollow arrow) is different from the unfolding one in the hetero-droplet. Here $n = 10$, $\gamma = 3$, $\nu = 10^{-4}$, and $\delta V_1 = V_2 = 0.4V_1$. (b) The typical crossover diagram. The solid curve, the $\diamond$, and the long-dashed curve have the same meanings as in Fig. 4. Because the rate-limiting barriers are different for un/folding, there is a "discontinuity" (jump) in the crossover curve (indicated by the hollow arrow). (c) The difference $\Delta\Delta F = \Delta F_d - \Delta F_Q$ between the droplet analysis (ΔF_d) and the single reaction coordinate landscape (ΔF_Q) as a function of heterogeneity strength; this curve is based on data of the type shown in d for the specific value $\delta V_1 = V_2 = 0.3V_1$. Note that the droplet analysis, although dependent on a single coordinate Q, does not lump together all states with a given number of contacts.

folded from the β-turn (the β-turn droplet) and one from the distal end (the distal-end droplet). The thermodynamic probabilities of these two droplets, defined as $Y_{n,Q}^{(\beta)}$, $Y_{n,Q}^{(d)}$ with the number of contacts Q can be obtained simply by using the M functionals Eq. **4** (11).

The situation is more complicated in the presence of heterogeneity. Now, we should also consider droplets surrounding the heterogeneity (a "hetero-droplet") with a thermodynamic probability $Y_{n,Q}^{(h)}$. These configurations contain contiguous unfolded segments between the heterogeneity and topological constraints.

In the single heterogeneity case, a hetero-droplet with Q contacts might contain one (or two) unfolded "bubbles" between either the β-turn or the distal end and the heterogeneity (or both); for the cluster case, the situation can be even more complicated, as the intervening residue $i = h + 1$ might also be in or out of contact. Obviously, there can be many configurations with the same number of contacts. Hence, we define the hetero-droplet as the configuration with the maximal thermodynamic probability. Fig. 3 a1,2 and b1,2 show the hetero-droplets in the single and double heterogeneity cases.

We determine the favored transition pathway by determining the barrier heights separating differing Q values for a set of droplets. Typical results from calculating these free energies are shown in Figs. 4a and 5a for the two different heterogeneity models. For unfolding, the model exhibits a metastable intermediate state at $Q = h + 2 = 6$. The rate-limiting step then involves starting from this state and breaking additional H-bonds between the heterogeneity and the β-turn. For the case of the more complex heterogeneity model in Fig. 5, there is an additional intermediate state at smaller Q. For this state to unfold fully, one needs to traverse an additional barrier, which for the case shown is smaller than the previous one. Hence for unfolding, both models yield similar pathways.

For folding, however, the presence of the new intermediate state has a large effect; we note in passing that the maximal number of intermediates is correlated with the number of heterogeneities. In the case of a single heterogeneity, folding proceeds exactly as the reverse of unfolding. For the clustered inhomogeneity, on the other hand, the largest barrier is the first one encountered from the coil state, namely the formation of the cluster by traversing the $Q = 1$ transition state.

Overall, there is only a very narrow range of parameters for which two-state behavior coexists with heterogeneity-dominated kinetics; we therefore expect that the experimental system lies in the zipping kinetics regime. Note that at higher temperatures, the range of heterogeneous domination broadens; this might account for the disagreement between the high T results of ref. 6 and other treatments to date (1). Two other points are worth mentioning. We found that the crossover at a smaller β-turn rigidity (σ) requires less heterogeneity, consistent with a recent 33-residue GCN4 peptide experiment (15). Also, we examined the validity of directly using the free energy profile with Q as the single reaction coordinate ($Z_{n,Q}$ in Eq. 1) to predict the kinetic barrier. Clearly, this approach lumps together all states at a given Q as opposed to the droplet configurations that have a fixed pattern of the Q native contacts. Near the crossover, we observed a systematic underestimate of the free energy barrier via the single reaction coordinate landscape method (Fig. 5c). A possible reason for this is that near the crossover, the system has two dominant transition ensembles that are kinetically unconnected, as also noticed in ref. 8.

Discussion

In this paper, we utilized a new exactly solvable β-hairpin model to investigate a competing droplet picture of folding and unfolding. In this model, folding energies occur only if the residues are very close to each other, and this allows us to give a discretized labeling of configurations without losing the ability to get reliable estimates of the loop entropies. Our approach does not suffer from the difficulties of the single sequence stretch approximation, in that it includes all possible arrangements of native structure, not merely contiguous ones (1, 3, 8, 10). Our results were obtained by detailed analysis of the free energy landscape as well as by direct simulation. The most important finding shows that there is a direct competition between pathways dominated by topological considerations vs. those dominated by heterogeneous free energy terms. This leads to a crossover in the structure of the transition state, which could be seen experimentally as one varies the strength and location of any hydrophobic clusters and/or the rigidity of the β-turn. Our results suggest that if we demand that the system have real two-state behavior, it is likely that the fold/unfolding will be topology dominated.

Interestingly, the single reaction coordinate approach becomes rather inaccurate near this crossover as it lumps together parts of configuration space that share the same number of contacts but are in fact not connected by the model kinetics. Also, we show that it can be incorrect to assume that the most significant free energy barriers are the same for folding and unfolding. Because it is common practice to simulate unfolding and to infer the computationally inaccessible folding process, this caveat will be important to keep in mind.

Of course, no model (including all-atom potentials) can be perfect. We have used a simplified Gaussian approximation for the backbone and have also neglected all nonnative contacts. Our study was for a single heterogeneous region (either one residue or one cluster), but one can extend our methodology to the multiple heterogeneity case. We are hopeful that this modeling approach will serve as a tractable way of putting in real biological complexity without sacrificing the ability to do a full analysis of the folding kinetics.

1. Muñoz, V., Thompson, P. A., Hofrichter, J. & Eaton, W. A. (1997) *Nature (London)* **390,** 196–199.
2. Wolynes, P. G. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 6170–6175.
3. Galzitskaya, O. V. & Finkelstein, A. V. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11299–11304.
4. Muñoz, V., Henry, E. R., Hofrichter, J. & Eaton, W. A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 5872–5879.
5. Kilonski, B., Ilkowski, W. & Skolnick, J. (1999) *Biophys. J.* **77,** 2942–2953.
6. Pande, V. S. & Rokhsar, D. S. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9062–9067.
7. Schellman, J. A. (1958) *J. Phys. Chem.* **62,** 1485–1494.
8. Alm, E. & Baker, D. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11305–11310.
9. Riddle, D. S., Grantcharova, V. P., Santiago, J. V., Alm, E., Ruczinski, I. & Baker, D. (1999) *Nat. Struct. Biol.* **6,** 1016–1024.
10. Muñoz, V. & Eaton, W. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 11311–11316.
11. Guo, C., Levine, H. & Kessler, D. A. (2000) *Phys. Rev. Lett.* **84,** 3490–3493.
12. Takada, S. & Wolynes, P. G. (1997) *J. Chem. Phys.* **107,** 9585–9598.
13. Hao, M.-H. & Scheraga H. A. (1998) *Acc. Chem. Res.* **31,** 433–440.
14. Veitshans, T., Klimov, D. & Thirumalai, D. (1997) *Folding Des.* **2,** 1–22.
15. Morna, L. B., Schneider, J. P., Kentsis, A., Reddy, G. A. & Sosnick, T. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 10699–10704.

BIOPHYSICS