JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

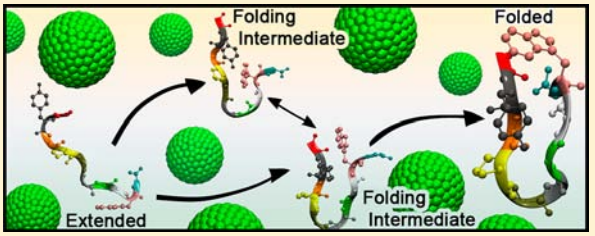
Crowding Alters the Folding Kinetics of a β-Hairpin by Modulating the Stability of Intermediates

Nicholas A. Kurniawan,[†] Søren Enemark,[†] and Raj Rajagopalan*

Singapore-MIT Alliance, National University of Singapore, Singapore 117576

Supporting Information

ABSTRACT: Crowded environments inside cells exert significant effects on protein structure, stability, and function, but their effects on (pre)folding dynamics and kinetics, especially at molecular levels, remain ill-understood. Here, we examine the latter for, as an initial candidate, a small de novo β-hairpin using extensive all-atom molecular dynamics simulations for crowder volume fractions ϕ up to 40%. We find that crowding does not introduce new folding intermediates or misfolded structures, although, as expected, it promotes compact structures and reduces the accessible conformational space. Furthermore, while hydrophobic-collapse-mediated folding is slightly enhanced, the turn-directed zipper mechanism (dominant in crowder-free situations) increases many-fold, becoming even more dominant. Interestingly, ϕ influences the stability of the folding intermediates (FI_1 and FI_2) in an apparently counterintuitive manner, which can be understood only by considering specific intrachain interactions and intermediate (and hierarchical) structural transitions. For ϕ values <20%, native-turn formation is enhanced, and FI_1, characterized by a hairpin structure but slightly mismatched hydrophobic contacts, increases in frequency, thus enhancing eventual folding. However, higher ϕ values impede native-turn formation, and FI_2, which lacks native turns, re-emerges and increasingly acts as a kinetic trap. The change in the stability of these intermediates with ϕ strongly correlates with the hierarchical folding stages and their kinetics. The results show that crowding assists intermediate structural changes more by impeding backward transitions than by promoting forward transitions and that a delicate competition between reduction in configuration space and introduction of kinetic traps along the folding route is key to understanding folding kinetics under crowded conditions.



INTRODUCTION

The folding process and the folded state of a protein are generally believed to be determined by its primary structure (amino acid sequence). However, previous studies on large proteins have shown that macromolecular crowding of the type or level typical in the interior of cells can affect folding in multiple ways. For example, crowding can modify the amount of secondary structures in the protein,[1] can influence the folding mechanism,[2] and can even induce misfolding and abrogation of the native structure.[3] These effects are crucial in protein functions and are closely related to protein stability, association, self-assembly, and aggregation as well as chaperonin activity and diffusion rates.[4−6] These findings also raise an interesting question: How much and in what way does crowding offset the deterministic influence of the primary structure during protein folding?

Several theoretical studies[6−8] have discussed the crowding-induced excluded-volume effects on folding, typically in terms of the so-called two-state folding mechanism, and the extent of protein compaction. In general, it has been predicted[8,9] and qualitatively verified in experiments[10−13] that crowding favors compact structures as they take up less space. While compaction can thermodynamically cause smaller free energy differences between unfolded and folded states, it can also accentuate the contribution of intrachain interactions in the folding mechanics. However, the underlying statistical mechanics or thermodynamic models used to describe the proteins have typically oversimplified residue-level details and the interactions among the residues, thereby obscuring direct and clearer quantitative comparisons with experiments. Consequently, the resulting effects of crowding on the native structure and on the structural transitions leading to the native structure remain debated.[14−18] To shed light on this issue, recent studies have begun to focus attention on the effects of excluded volume on the overall folding rate and modifications to the bulk viscosity and their consequences.[19−21] However, the detailed dynamics behind these macroscopic manifestations have not been investigated thoroughly. More realistic descriptions and quantitative pictures on the role of residue-specific interactions are required to further advance the understanding of how the primary sequence relates to a specific native structure in the crowded milieu within the cell.[22] The first step in addressing the above question is to begin with an examination of the effect of crowding on short polypeptides that have well-defined folded structures. Since conventional molecular dynamics (MD) simulations have proved to be successful for characterizing detailed folding mechanisms of

Received: March 27, 2012
Published: May 30, 2012

dx.doi.org/10.1021/ja302943m | J. Am. Chem. Soc. 2012, 134, 10200−10208

proteins under *noncrowded* conditions[23] and have enabled unequivocal determination of the folding process and comparison with experimental results,[24] a similar approach offers a promising route to initiate investigations of the effects of crowding on well-defined systems consisting of short polypeptide chains.

To this end, we have performed extensive atomistic, unbiased MD simulations, totaling more than 50 μs, to follow the complete course of the folding dynamics of a short de novo polypeptide within the range of in vivo crowding conditions (volume fractions of crowders, ϕ, up to 40%). As a model system, here we study a short β-hairpin, chignolin (GYDPETGTWG), which is considered the smallest protein to date, having only 10 residues.[25] β-Hairpins have received extensive attention in the literature since they exhibit certain behaviors similar to those of larger proteins, e.g., cooperative folding,[26] and have been thought to act as nucleation sites for protein folding.[27] In addition, their small size allows extensive computational analyses and implies reduced complexity in the relation among the primary structure, folding pathways, and native structure, making β-hairpins convenient vehicles for in-depth examinations of some of the fundamental issues in protein folding.[26]

We examine the effects of crowder-induced reductions in conformational space not only on the overall folding time but also on the folding mechanisms, the hierarchy of intermediate structural transitions, and the populations and the roles of folding intermediates. The results for the effect of crowding on folding intermediates and their populations as functions of ϕ reveal that the populations of the intermediates may not be affected either similarly or monotonically by increasing the ϕ values and that the net effect depends on the folding path and on the intermediate folding events such as native-turn formation and correct or mismatched hydrophobic contacts, etc. The results also show that crowding enhances folding more by impeding reverse transitions in the evolution of intermediate states than by assisting forward transitions. Furthermore, the overall folding times at large ϕ values (corresponding to those in the interior of cells) cannot be extrapolated from results obtained even at moderate crowder concentrations because of potential, nonmonotonic effects of crowding on partially folded or semistable, misfolded structures. The results also illustrate that specific intrachain interactions can significantly complicate excluded-volume effects and must be taken into account in analyzing crowding effects on protein folding.

METHODS

Crowder Model and Conditions. The excluded-volume effect was generated in the simulations by insertion of crowders, each consisting of 204 oxygen-like atoms on the surface of a sphere of radius 0.75 nm, chosen to match the length of the polypeptide hairpin measured from the end to the central turn residues (see Figure S1 in the Supporting Information). The individual atoms in the crowders were given the Lennard-Jones properties of TIP4P oxygen atoms for interactions with water and Na^+ ions but were made to repel atoms on other crowders as well as on the polypeptide to prevent crowder–crowder and crowder–polypeptide aggregation (see the Supporting Information for further details). As a result and as required for ideal crowders,[28] the crowder molecules both were highly water-soluble, despite the fact that the atoms were electrically neutral, and had no interactions with the polypeptide other than steric repulsion. The range of crowder concentrations (or, equivalently, volume occupancy) was chosen on the basis of in vivo crowding conditions, which may extend up to 40% of the total volume.[5,29]

Simulation Setup and Parameters. Five different crowding conditions were simulated corresponding to the crowder volume fractions, ϕ, of 0%, 10%, 20%, 30%, and 40% by including zero, two, four, six, and eight crowder molecules in the simulated system. (The exact volume fractions are within about 1% of the numbers indicated; e.g., 10% is actually 10.02%, etc. See the Supporting Information for further details on volume fractions and how they are estimated.) For each of these 5 conditions, 10 simulations with a total simulated time of more than 10 μs were performed starting from an extended all-trans conformation to study the folding mechanisms, pathways, and intermediate conformational transitions during the prefolding stages leading to folded structures. The peptide was initially placed in a $4 \times 4 \times 4$ nm^3 box with periodic boundaries together with a number of crowders, corresponding to the studied volume fractions, at random positions. Subsequent analyses confirmed that there are negligible finite-size effects on the polypeptide structure and interactions with the crowders (see the Supporting Information). TIP4P water molecules were added to the box in addition to two Na^+ ions that counterbalance the negative charges of Asp3 and Glu5. Before each MD simulation, a short steepest descent energy minimization was carried out to remove any high-energy interactions between atoms. Both energy minimizations and MD simulations (T = 298 K, p = 1 bar) were done in GROMACS 4.0,[30] using the OPLS-AA force field, with a twin cutoff range (0.9 and 1.4 nm) for van der Waals interactions and the particle mesh Ewald (PME) method (with a real-space transition at 0.9 nm) for electrostatic interactions. In comparison to other force fields, this force field provides native-state stability that matches the experimental results best.[31] Bond lengths in water and the polypeptide were constrained using SETTLE and LINCS-P,[30] respectively. Each of the 50 simulations performed was carried out for about 1 μs, and atom positions were sampled every 10 ps.

Parametrization and Analyses of Structures of the Polypeptide. Structural analyses of the peptide included calculation of heavy-atom root-mean-squared fluctuations (RMSFs), root-mean-squared deviations (rmsd's) of all distances between heavy atoms measured with respect to the corresponding distances in the native (reference) structure. The rmsd measure has been used[32] and detailed[33] in previous studies on chignolin. From rmsd-based structure analysis, we find that rmsd < 0.16 nm provides a suitable criterion for determination of a stable native conformation at all concentrations of crowders (Figure S5, Supporting Information). For all tested ϕ values, we find that the achieved folded states are identical to the native structure under a dilute environment, as depicted in Figure 1, providing a convenient common reference structure for further analysis. The residue-level dynamics of the folded polypeptide is also unaffected by crowding (Figure S6, Supporting Information). The lack of an effect of crowding on the folded state of chignolin, in contrast to

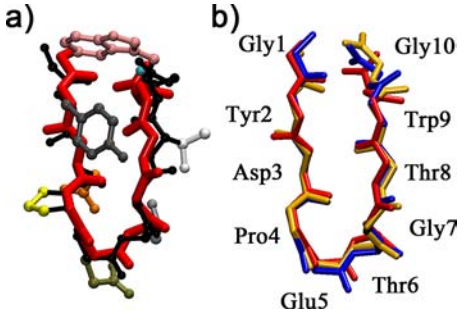


Figure 1. Comparison of folded structures obtained from NMR and from our simulations. (a) Comparison of the native-state conformation from crowder-free simulations (thick red) with an NMR structure from experiments (backbone shown in black and side chains in various colors). (b) Comparison of randomly chosen folded structures obtained at different crowder concentrations. (ϕ = 0%, red; ϕ = 20%, blue; ϕ = 40%, orange).

dx.doi.org/10.1021/ja302943m | J. Am. Chem. Soc. 2012, 134, 10200–10208

the above-reported effects on large proteins, can be attributed to its small size and the already highly ordered structure of β-hairpins.

In addition to the above criterion, we also use the following criteria to classify whether a structure is hairpin-like, or compacted, or has formed the central turn region. To quantify the hairpin-type shape of a conformation, we use the parameter $R_i = d_i^N/d_i$, where d_1, d_2, and d_3 are the cross-strand $C_\alpha-C_\alpha$ distances between Tyr2 and Trp9, Asp3 and Thr8, and Pro4 and Gly7, respectively, and d^N is the corresponding distance in the native NMR structure (see Figure S7 in the Supporting Information). We consider a conformation to be hairpin-like when all R_i values are >0.7. The native-turn region (Pro4-Glu5-Thr6-Gly7) is considered to be formed when the $C_\alpha-C_\alpha$ distance between the two corner residues, Pro4 and Gly7, $d_{PG,C\alpha}$, is smaller than 0.7 nm, as is usually assumed in the literature.[34] A structure is considered to be compact when its radius of gyration, R_g, is smaller than 0.66 nm, for reasons described in the Results.

RESULTS

Reduction in the Conformational Phase and Implications to Conformational States. Crowding leads to a structural compaction of the polypeptide (and a consequent reduction in the conformational space) as can be seen from Figure 2a, which shows the variation of the density of all

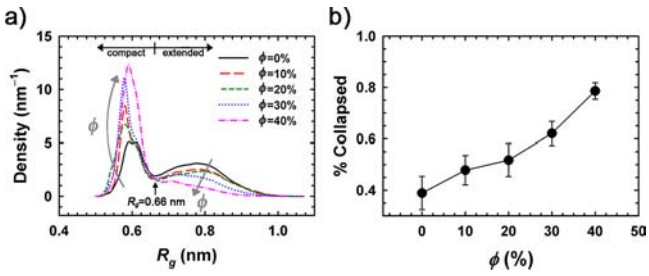


Figure 2. Effect of crowding on the compaction of the structure of chignolin. (a) Probability density as a function of R_g at different crowder concentrations ϕ. The minimum at $R_g \approx 0.66$ nm separates the compact states from the relatively extended ones. (b) Fraction of compact states (i.e., for $R_g < 0.66$ nm) as a function of ϕ. The error bars denote the standard error ($n = 10$).

unfolded conformations as a function of the radius of gyration, R_g, for various crowder volume fractions ϕ. The figure shows a clear density minimum at $R_g \approx 0.66$ nm for all ϕ values, with the minimum separating the increasing populations of compact conformations from the open ones as ϕ increases. Furthermore, compact conformations appear below the same magnitude of R_g (=0.66 nm) for all ϕ values, indicating that the region of conformational space corresponding to compact states remains essentially unaltered from the one for the crowder-free situation (although the densities of certain intermediate structures could be affected differently by the crowder concentration, as we shall see later). This observation and the fact that the density distributions of the compact states remain qualitatively the same (i.e., all are "bell-shaped") also suggest that the compact states at crowded conditions are structurally similar to those in crowder-free conditions (which are driven largely by entropic forces). These results suggest that crowding primarily contributes to the increase in the population of compact states by restraining the states from re-entering expanded conformations. We shall examine this hypothesis later using kinetic data collected from the simulations.

The fraction of compact structures among the unfolded conformations, shown in Figure 2b, exhibits a 2-fold increase from ~35% at $\phi = 0\%$ to as much as 80% at $\phi = 40\%$, implying that the unfolded polypeptide spends twice as much time in the compact states under high levels of crowding as it does under dilute conditions. It is also worth noting that the peaks in the conformational densities in the compact region remain essentially at the same R_g and that the increases in the population of compact structures arise largely from the peak region as the crowder concentration increases, suggesting that the population increases are dominated by certain types of structures, which, as we shall see later, correspond to two folding intermediates. In fact, the magnitude of R_g corresponding to the peak in the distribution shifts slightly toward smaller values first as ϕ increases (from 0.6 nm at $\phi = 0\%$ to about 0.57 nm at $\phi = 20\%$) and then shifts back to 0.6 nm at larger ϕ values. We shall see later that these observations are caused by the differing influence of crowding on the two intermediates.

Influence of Crowding on Folding Intermediates and Their Stability. To further examine the nature of the compact conformations, we construct the conformational distributions and the corresponding free energy surfaces as functions of R_g and rmsd. Figure 3a shows the existence of two prominent intermediates in the compact, but unfolded, conformations. We have previously reported[33] the occurrence of these intermediates, identified as folding intermediates FI_1 and FI_2, at crowder-free situations and have linked them to specific interactions among Tyr2, Pro4, and Trp9. In particular, FI_1 invariably has a fully formed native turn and is hairpin-like but has a mismatched side-chain arrangement, with Tyr2 departing significantly from its native position; in contrast, FI_2 lacks the native turn (and hence also the hairpin structure) and is dominated by Pro4−Trp9 contact.[33] These intermediates, illustrated in the insets in Figure 3a, which exist also under crowded conditions (Figure S8, Supporting Information), are shown in the sample trajectories in Figure 3b. A visual examination of the trajectories shows that the Pro4−Trp9 interaction initially assists native-turn formation, but slows FI_1 formation or folding when it persists longer as the polypeptide seeks the native state. The trajectories further show that the existence of the intermediates becomes more persistent as the time spent by the polypeptide in compact conformations increases, suggesting that the intermediates increasingly act as kinetic traps. A closer scrutiny of the conformational space (Figures 3c; Figures S8−S11, Supporting Information) shows that crowding significantly affects the stability of these intermediates but does not cause the formation of any new intermediates, thus offering direct evidence that crowding-induced polypeptide compaction can lead to highly structured compact states characterized by discrete intermediates, in contrast to creating random and nonspecific structures as previously hypothesized for full proteins.[15]

Interestingly, the results also show that the relative stability of the two folding intermediates varies in a nonmonotonic manner with ϕ, as illustrated in Figure 3d. The 1:1 ratio between the populations of FI_1 and FI_2 under dilute conditions is replaced by a preference for FI_1 under moderate crowding, but is restored at higher crowder concentration ($\phi \approx 40\%$). The reappearance of FI_2 at $\phi > 20\%$ is surprising because conformations in FI_2 are less compact than those in FI_1 and excluded-volume effects have been classically thought to entropically destabilize less compact structures. The initial decrease in FI_2 as ϕ increases appears to arise from the elimination of wasteful searches for the native state with the reduction in conformational space with increasing excluded volume. With further increases in ϕ, the constraining effect of

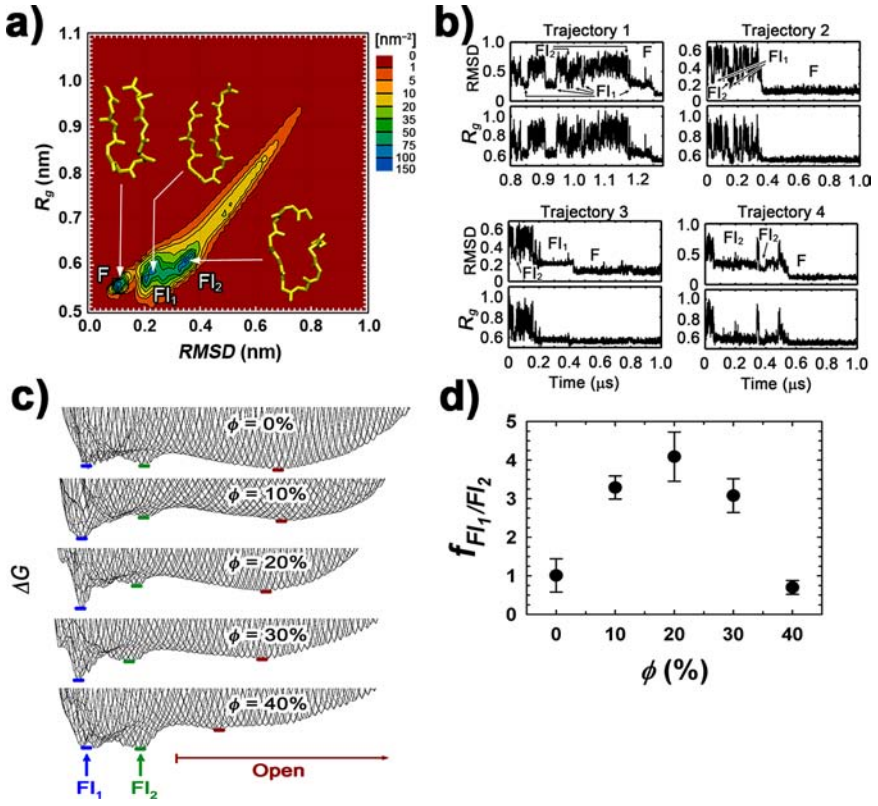


Figure 3. Folding intermediates in the dynamics of compact, unfolded structures. (a) Conformational states in the rmsd−R_g space, showing the two frequently occurring folding intermediates, FI_1 and FI_2, in addition to the folded (F) state. The insets show representative structures of the three states. (b) Sample trajectories illustrating passage through FI_1 and/or FI_2 on the way to F. Trajectories 1 (ϕ = 0%) and 2 (ϕ = 20%) show the frequent and quick transitions among unfolded, FI_1, and FI_2 states before the polypeptide folds. Extended residence in FI_1 (~200 ns) and relatively stable dynamics in FI_2 can be observed in trajectory 3 (ϕ = 20%) and trajectory 4 (ϕ = 40%), respectively. Note that not all occurrences of FI_1 and FI_2 are marked in the figures. (c) Free energy profiles based on the five simulated crowding conditions, illustrating that crowding influences the relative stability of the folding intermediates and the denatured states. See also Figures S8−S11 in the Supporting Information. (d) Ratio between the populations of FI_1 and FI_2 as a function of ϕ. The error bars denote the standard error (n = 4).

the crowders appears to restrict the conformational freedom of the polypeptide more severely, causing re-emergence of FI_2 for ϕ > 20%. The kinetics of native-turn formation extracted from the trajectories supports these observations, as discussed later. We shall also comment on the implications of these intermediates and the variation of the energy barriers between them with ϕ on the overall folding time later in the paper.

To better understand the effect of crowding on the stability of the intermediates and on the folding process, it is important to take into account the specific residue-level interactions underlying the two intermediates noted above. We separate and contrast these two interactions using a parameter $R_{Hpb} = (d_{PW}/d_{PW}^*) - (d_{YW}/d_{YW}^*)$, where d_{PW} denotes the center-of-mass-to-center-of-mass side-chain distance between Pro4 and Trp9 and d_{YW} is the corresponding distance between Tyr2 and Trp9. These distances are normalized by the respective cutoff distances d_{PW}^* and d_{YW}^* corresponding to the distances below which the respective pairs of side chains begin to interact with each other. Previous studies show that d_{YW}^* = 1 nm[33,35] and d_{PW}^* = 0.8 nm (Figure S12, Supporting Information). In short, for $R_{Hpb} > 0$ the hydrophobic interaction between Tyr2 and Trp9 is stronger than the non-native hydrophobic interaction between Pro4 and Trp9, and $R_{Hpb} < 0$ implies the reverse.

The conformational space (Figure 4) reveals that the stronger excluded-volume effect with increasing ϕ drives the polypeptide quickly to one of the relatively stable compact structures, as evident from the deepening of the local minima in the compact region. Up to $\phi \approx$ 20%, there is an increasing enrichment of FI_1, the more compact and nearly folded intermediate, at the expense of the other. However, the same effect also restricts the polypeptide's ability to freely explore other compact structures. At higher ϕ values, the population in FI_2 begins to increase, and FI_2 increasingly acts as a kinetic trap. The configuration spaces at various crowder concentrations also show that the folding trajectories often pass through both FI_2 and FI_1 at ϕ = 0% (see Figure 3b for an example), and the folding pathway becomes less well-defined at higher ϕ, thus hinting at a possible impact of crowding on the folding mechanism through a shift in the roles of the intermediates. This shift in role, as we shall show, indeed has important consequences to both the folding mechanism and kinetics.

Effect of Crowding on Intermediate Structural Transitions and Kinetics. As well-known, for β-hairpin folding under dilute conditions, two main mechanisms have been studied and debated extensively, with the mechanisms distinguished by whether the initial stages of folding are driven by formation of a central turn structure[26] or by structural collapse as a result of hydrophobic interactions.[36] To examine the contributions from these two pathways, we performed a cluster analysis of the conformations in the R_g−d_{YW} space (as we had reported earlier for the crowder-free situation[33]) to clearly separate the transitions from an extended state to a collapsed state along the above two pathways (Figures 5a;

dx.doi.org/10.1021/ja302943m | J. Am. Chem. Soc. 2012, 134, 10200−10208

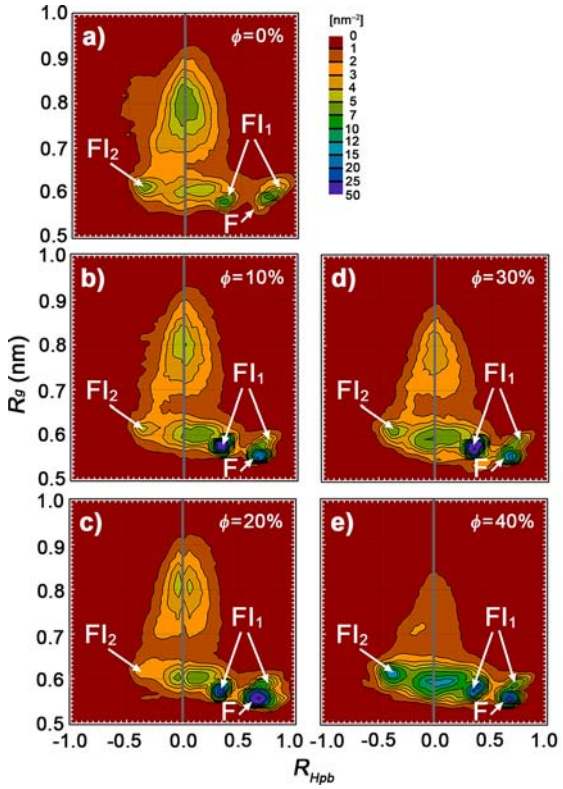


Figure 4. Conformational density distribution as a function of R_g and R_{Hpb} for various ϕ values, showing the different natures of side-chain interactions for various states.

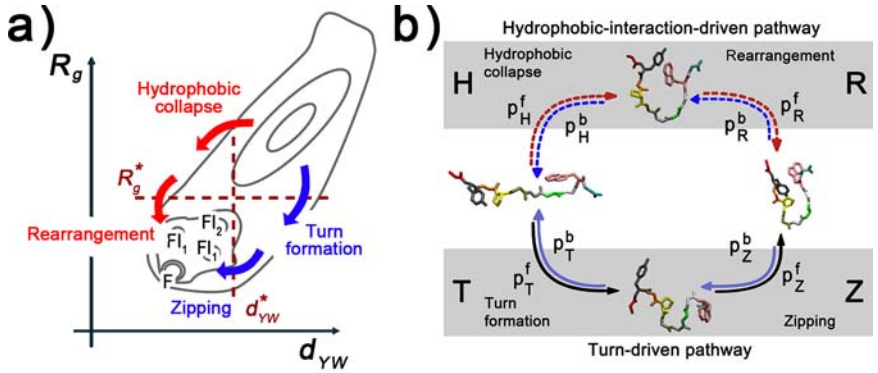
Figure S13, Supporting Information). These pathways can be examined in terms of four relevant structural transitions, namely, hydrophobic collapse (H) from an open state and the subsequent (structural) rearrangement (R), both along the hydrophobic-interaction-driven pathway toward folding, as well as turn formation (T) and subsequent "zipping" (Z) along the turn-driven pathway (Figure 5b). The rationale for this type of analysis is described in detail in the earlier work.[33] The corresponding forward (p_i^f) and backward (p_i^b) transition probabilities are shown as functions of ϕ in Figure 5c, where the subscript i denotes each structural transition.

Probabilities of forward transitions, p_H^f and p_T^f, from the extended state show that the initial steps (i.e., either hydrophobic collapse or turn formation) of the two folding mechanisms are equally likely (i.e., p_H^f and p_T^f are roughly equal (1−2%)). Nevertheless, when one looks at subsequent transitions and their relative rates, a different picture emerges. To examine these relative rates along each of the two pathways, we introduce the transition ratio, $r_{i/j} = p_i^f/p_j^b$, which describes the probability of a conformation moving forward toward folding relative to the probability that it returns to its previous state. Thus, $r_{i/j} < 1$ implies that a conformation is more likely to return to a previous state, whereas $r_{i/j} > 1$ indicates a higher likelihood of the conformation moving forward toward folding. Along the hydrophobic-collapse-driven pathway, one finds that $r_{R/H}$ increases with ϕ, from a net backward transition $r_{R/H} \approx 0.5$ at ϕ = 0% to an almost equiprobable transition $r_{R/H} \approx 0.9$ at ϕ = 40% (Figure 5d), implying that crowding not only induces more frequent hydrophobic contacts, as we noted earlier, but also promotes the effectiveness of hydrophobic contacts in facilitating folding by impeding reverse transitions. This enhancement of the hydrophobic-collapse-driven pathway could in principle also lead to a switch in the dominant folding mechanism in β-hairpins, which are driven by native-turn formation, as emerging evidence has established.[33,37,38] However, a similar analysis for the turn-driven pathway reveals that $r_{Z/T}$ also increases consistently with ϕ (Figure 5c). In fact, the transition along the turn-driven pathway is enhanced more than 3-fold from what is already a strongly net forward transition (i.e., from $r_{Z/T} \approx 1.9$ at ϕ = 0% to $r_{Z/T} \approx 6.3$ at ϕ = 40%). Hence, our results show that, while all pathways are concomitantly enhanced by crowding, the general folding mechanism remains the same and is, in fact, strengthened.

The roles played by the two intermediates, FI_1 and FI_2, on folding are most evident in the zipping stage. Both p_Z^f and p_Z^b vary nonmonotonically with ϕ, reminiscent of the similar but opposite variation of the relative occurrences of the two states with ϕ (Figure 3d). In particular, the unzipping process (i.e., p_Z^b) appears to be influenced noticeably by transitions from FI_1 to FI_2, and the transition probability shows a minimum at ϕ = 20% because of the increased energy barrier to a chain going from FI_1 to FI_2 (see Figure 3c). It is important to note that, for all other processes (i.e., R, H, and T), the enhancement in the transitions toward folded states with increasing ϕ stems primarily from reductions in p^b values rather than from increases in p^f values. This observation implies that, with increasing crowder concentration in the range expected in vivo, the polypeptide is less prone to transitions that lead to more open structures, and the volume exclusion due to crowding thus serves more as an obstruction of backward transitions rather than promotion of forward transitions in the intermediate structural transitions along the folding pathway.

Hierarchical Structural Transitions and Their Kinetics and Time Scales. The (turn-directed) folding process can be characterized in terms of four key events, namely, structural compaction, native-turn formation, hairpin arrangement, and eventual folding, which represent a highly hierarchical folding mechanism, allowing the events to act as temporal marks in the folding course.[39] The characteristic times of the four key events (see the Methods for the quantitative definition of the events), obtained through the decay of the lifetimes of the opposing events as described in Figure S15 in the Supporting Information, are listed in Table 1. The gradual increase of time scales associated with each event, from 4−5 ns for compaction, to 25−50 ns for native-turn formation, to 50−140 ns for hairpin arrangement, and to 650−1000 ns for folding, clearly demonstrates a hierarchical folding process with separate time scales. It can be observed that the time constant for structural compaction decreases with ϕ, while that for the reverse process, which we refer to as *restretching*, increases with ϕ. This variation provides a kinetic measure for the previously reported entropic stabilization of compact states and destabilization of denatured states due to crowding[12,13,15] and is in line with the observed reduction in backward transitions shown above.

For rates other than compaction and restretching rates, the observed effects of crowding become less intuitive and can only be understood by taking specific intrachain interactions into account. For example, the trend of faster compaction and slower restretching with increasing ϕ translates to faster formation of the native turn, as might be expected, but only up to moderate crowding (i.e., below $\phi \approx 20\%$). Beyond this range of ϕ, the native-turn formation rate evidently decreases with increasing ϕ, and this coincides with the reappearance of FI_2 (Figure 4). Indeed, FI_2, the predominant compact state



c)

ϕ / %	p_H^f	p_H^b	ϕ / %	p_R^f	p_R^b
0	1.5 (0.1)	24 (1)	0	14 (2)	1.9 (0.3)
10	1.5 (0.1)	23 (<1)	10	13 (1)	1.1 (0.2)
20	1.5 (0.1)	23 (1)	20	12 (1)	1.0 (0.3)
30	1.9 (0.1)	22 (1)	30	13 (1)	0.8 (0.2)
40	2.0 (0.3)	18 (1)	40	17 (1)	0.5 (0.1)
ϕ / %	p_T^f	p_T^b	ϕ / %	p_Z^f	p_Z^b
0	1.7 (0.5)	9.2 (0.9)	0	18 (1)	5.6 (0.7)
10	1.4 (0.2)	7.9 (0.7)	10	18 (<1)	3.4 (0.5)
20	1.2 (0.1)	5.8 (1.0)	20	15 (1)	3.2 (0.6)
30	1.8 (0.2)	5.6 (0.6)	30	18 (1)	3.8 (0.8)
40	2.9 (0.4)	3.2 (0.2)	40	19 (1)	5.1 (0.8)

d)

ϕ / %	$r_{R/H}$	$r_{Z/T}$
0	0.6 (0.1)	2.1 (0.3)
10	0.6 (<0.1)	2.5 (0.4)
20	0.5 (<0.1)	2.9 (0.7)
30	0.6 (<0.1)	3.7 (0.6)
40	1.0 (0.1)	6.1 (0.5)

Figure 5. Folding paths and corresponding transition probabilities. (a) Illustration of cluster analysis of structures, in terms of the distance d_{YW} between the hydrophobic side chains Tyr2 and Trp9 and the radius of gyration, R_g. The former represents the extent of native hydrophobic association. The latter was chosen because of the strong correlation between R_g and the occurrence of turns in the central four residues of the polypeptide (Figure S14, Supporting Information). The cutoffs for R_g at 0.66 nm and d_{YW} at 1 nm lead to the separation of the configuration space into four regions (see Figure S13 in the Supporting Information for details) with four prominent structural progressions, namely, hydrophobic collapse and a rearrangement process in the hydrophobic-collapse-driven pathway and turn formation and a zipping process in the turn-driven pathway. (b) Schematic illustration of structural transitions and the corresponding transition probabilities along the hydrophobic-collapse-driven pathway (red dashed arrow) and the turn-driven pathway (black solid arrow) in both the forward (unfolded → folded) and backward (folded → unfolded) directions. (c) Transition probabilities as functions of ϕ. (d) Effectiveness of forward transitions along the two pathways as a function of ϕ, as quantified using the transition ratios. The standard error is shown in parentheses (n = 10).

Table 1. Characteristic Times of Key Events in the Folding and Unfolding Pathways, Shown as the Mean (Standard Error) (n = 10)

event		characteristic time (ns)				
		ϕ = 0%	ϕ = 10%	ϕ = 20%	ϕ = 30%	ϕ = 40%
compaction	compacting	5.3 (0.4)	5.3 (0.5)	5.2 (0.3)	4.3 (0.3)	4.2 (0.5)
	restretching[a,b]	2.3 (0.5)	4.5 (0.6)	6.5 (1.2)	6.7 (2.3)	10.3 (3.8)
native turn	formation	28.4 (2.6)	22.1 (2.0)	21.6 (2.3)	32.8 (3.8)	45.2 (7.1)
	dissolution[b]	2.7 (0.5)	3.4 (1.8)	2.9 (0.5)	5.6 (0.8)	8.1 (2.2)
hairpin	arrangement	129 (17)	86 (8)	54 (7)	89 (12)	84 (18)
folding[c]		~1.0–1.3 × 10^3	1.05 × 10^3	0.67 × 10^3	0.71 × 10^3	0.65 × 10^3

[a]Restretching refers to the reverse of compaction. A structure is considered to restretch when it goes from a compact state (R_g < 0.66 nm) to an extended state ($R_g \geq$ 0.66 nm). [b]For restretching and native-turn dissolution, the distribution of residence time follows a double exponential in the form of $P(t) = A_1 \exp(-t/\tau_1) + A_2 \exp(-t/\tau_2)$. Only τ_1, corresponding to the faster relaxation, is listed, as τ_2 is associated with long-term residence in the intermediate and/or folded states. [c]Folding times were obtained from the combined data from all 10 trajectories to improve statistics. Here, the time constants corresponding to the best fits are shown ($R^2 \approx 0.9$).

without a native-turn structure (with $\langle d_{PG,C\alpha} \rangle \approx 0.85$ nm, above the upper limit of 0.7 nm for native turns; see the Methods), becomes more favored because of the increased degree of compaction and the accompanying decrease in the ease of search for the folded structure or for the near-folded FI_1 state, as we had hypothesized earlier. The reduction in the rate of native-turn formation evident from the kinetic data in Table 1 is consistent with our hypothesis.

Effect on the Overall Folding Time. The individual trajectories as the polypeptide searches for the native state, the conformational space obtained from the collection of trajectories, and the kinetic data on the hierarchical structural evolution collectively provide a clear picture of the folding dynamics and the variation in the overall folding time shown in Table 1. Increases in the crowder concentration up to ~20% enhance folding by decreasing the conformational space and wasteful searches for the native state and making the search

more efficient. In particular, the rates of compaction and turn formation, the two initial folding steps, increase thanks largely to the reduction in backward transitions. These also contribute to a reduction in FI_2 and digressions from FI_1 to FI_2. However, with further increases in crowder concentrations, although one observes slightly faster compaction and significantly slower restretching, no further enhancement of the folding rate occurs (see Table 1) for two related reasons. First, native-turn formation is restricted by the reduced conformational freedom. Second, the ease of native-turn dissolution and of unzipping increases transitions from FI_1 to FI_2, with the re-emergence of FI_2 contributing to a slowing of folding and apparently compensating for the increased efficiency of the search due to the further reductions in conformational space with increasing ϕ. In summary, while the early kinetics of β-hairpin folding is directly governed by a crowder-assisted entropic effect, the later stages of folding are only indirectly affected by crowding via the modified stability of the folding intermediates.

FURTHER DISCUSSION

Macromolecular crowding is known to affect the *equilibrium* properties of proteins, e.g., in terms of shape, stability, and compaction,[4−6] but the *dynamical* behavior and properties behind these equilibrium changes remain unclear, partly because of the time- and length-scale limitations inherent in the available experimental techniques as well as the coarse-grained descriptions of proteins often used in theoretical models. In particular, the question of how crowding-induced compaction, shown repeatedly both theoretically[9] and experimentally,[12,13,15] impacts the folding dynamics and kinetics remains open. Moreover, previous studies of the effect of crowding on the folding process have largely been interpreted in terms of a simple two-state folding mechanism. In contrast, the picture that emerges from our results, as schematically illustrated in Figure 6, is one in which crowding modifies the

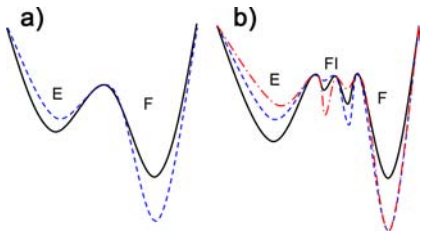


Figure 6. Effect of crowding on the free energy profiles: black lines, crowder-free conditions; blue and red lines, at two different ϕ values. (a) Schematic of the two-state model typically used in the literature to discuss the effects of crowding (E, extended states; F, folded states). (b) Schematic of our model with folding intermediates (FIs), which are affected by crowding, not necessarily similarly.

stability of a discrete number of semistable, non-native, and compact conformations to take precedence over open conformations. Furthermore, specific intrachain interactions play a crucial role in keeping the compact conformations highly structured and in determining the folding mechanism and kinetics.

The crowding-induced modification to the energy landscape affects the folding kinetics in opposing ways. One obvious, positive effect contributing to faster folding is due to the reduced size of the configuration space. This effect is well-known and represents the essence of the excluded-volume effect.[8] It has even been hypothesized that the reduction in configuration space *simplifies* the search for native contacts, resulting in a smoother folding funnel, fewer folding traps, and faster folding.[40] While our results agree that crowding leads to a simplified search for the native state and faster folding, these effects were not achieved through a smoother folding funnel. We find that a more complex and previously overlooked effect can be attributed to the alteration of the stability of the compact intermediates. This alteration, which accompanies the configuration space reduction, may affect folding rates negatively, if the stability of these intermediates becomes so strong that they act as kinetic traps, or positively, since the mere presence of folding intermediates can actually have a positive effect on the folding rate.[41] In summary, our results suggest that the effect of crowding is a tug-of-war between the positive effects of configuration space reduction and introduction of folding intermediates on the one hand and the negative effect of excessive stabilization of folding intermediates on the other. This competition not only results in the observed non-monotonic variation of the folding rate with ϕ,[11] but may also explain the discrepancy in the observed effect of macromolecular crowding on different proteins.[16,18,42] Moreover, the competition between the opposite effects provides a specific example of one mechanism by which the cell may actively use crowding to tune the dynamics and speed of the folding process, as suggested previously.[43]

It has also been suggested[7,20] that the decrease of the folding rate beyond a certain crowder concentration arises mainly from an increase in viscosity caused by addition of crowders, which slows the overall motions in the system. There are two reasons why we believe that viscosity does not influence the folding rate in the case of chignolin. First, an increase in viscosity would impact primarily the large-scale motion of the polypeptide, reflected in decreases in the rates of both compacting and restretching, whereas we find a monotonic increase in the compaction rate with increasing ϕ. Moreover, as we have argued, the folding rates are governed mostly by the dynamics subsequent to the compaction, where structural rearrangements require only small-scale motions, which would be little affected by the viscosity of the environment.

A striking conclusion of our study is that crowding can affect folding at remarkably many levels. First, the native state and the overall folding mechanism of the β-hairpin remain unaffected by crowding, implying that the primary sequence plays the dominant role in structure formation on this secondary structure level even under crowded conditions. Furthermore, the primary sequence also determines the existence of intermediate states through the specific ordering of interacting residues or microdomains in the polypeptide; in turn, the stability of these intermediate states can be regulated by ϕ, and as we have shown, the modification of the stability of these intermediates can strongly perturb β-hairpin folding dynamics and kinetics. These effects can have implications in the downstream folding processes, since the formation of β-hairpins, the minimal β-structure element,[26] has been suggested to play an important role in the early kinetics and nucleation of (large-scale) protein folding.[44] In fact, crowding-induced compaction of large proteins may bring different domains closer together, creating stronger local excluded-volume effects in the secondary structure's microenvironment. In addition, our results suggest that high levels of crowding can lead to a slowing of the overall folding mechanism on the tertiary structure level by the introduction of kinetic traps along the pathways of secondary structure formation, thus providing an

alternative explanation for the nonmonotonic dependence of the folding rate with ϕ in full proteins.

CONCLUDING REMARKS

In summary, we have presented, using a small β-hairpin peptide as a first step, a detailed perspective on how crowding can alter the conformational space as well as the underlying dynamics and thus could interfere with the primary sequence's determining role in the dynamics and kinetics of folding. First, the results show how high crowder concentrations can restrict a polypeptide to essentially compact conformations and how this restriction to compact states affects semistable, structurally well-defined, nonfolded intermediates. For the polypeptide we have studied here, crowding does not lead to formation of any additional folding intermediates, due to the shortness of the polypeptide and the limited number of combinations of interactions contributing to intermediates; here, these interactions are restricted to hydrophobic contacts between Tyr2 and Trp9 and between Pro4 and Trp9. As full proteins consist of much longer and more complex primary sequences, a plethora of intermediates are possible, some of which are likely to be stabilized to the extent that they may influence the folding dynamics under crowded conditions. Second, we find that the kinetics of specific steps in the folding/unfolding processes can vary nonmonotonically with ϕ in relation to the relative stability of the intermediates and that crowding generally impedes backward transitions rather than facilitating forward transitions. In chignolin, the changed stability of the folding intermediate FI_2 leads to an increased diversion of near-folded conformations in the on-path folding intermediate FI_1, causing FI_2 to act as a kinetic trap for $\phi >$ 20%. Our results thus show that folding rates based on moderate crowding conditions cannot be simply extrapolated to larger crowder concentrations.

While further studies of longer and more complex polypeptides are required for a complete perspective on the effect of crowding, it would also be instructive to investigate the role of other crowding-associated factors in the cellular environment, such as crowder−protein interactions and heterogeneous crowding, which have not been taken into account explicitly in this study,[20,45] but which are equally important for expanding one's understanding of crowding on the basis of increasingly more realistic models. In particular, it remains to be seen to what extent specific crowder sizes, shapes, composition, and interactions with the surroundings, all of which have been shown to play a role in the folding mechanics,[8,15,16,46] would bring about quantitative differences in the general behavior.

ASSOCIATED CONTENT

Supporting Information

Detailed simulation setup and methods, description of the crowder model, convergence of the simulations, conformational distribution and free energy landscape under various levels of crowding, detailed quantitative criteria for conformational classifications and structural transitions, residence time analysis for the kinetics of four key folding events, and Figures S1−S15. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

raj@nus.edu.sg, raj.private.email@gmail.com

Author Contributions

[†]N.A.K. and S.E. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Singapore-MIT Alliance for financial support.

REFERENCES

(1) Perham, M.; Stagg, L.; Wittung-Stafshede, P. *FEBS Lett.* **2007**, *581*, 5065−5069.
(2) Homouz, D.; Stagg, L.; Wittung-Stafshede, P.; Cheung, M. S. *Biophys. J.* **2009**, *96*, 671−680.
(3) van den Berg, B.; Ellis, R. J.; Dobson, C. M. *EMBO J.* **1999**, *18*, 6927−6933.
(4) Ellis, R. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 114−119.
(5) Chebotareva, N. A.; Kurganov, B. I.; Livanova, N. B. *Biochemistry (Moscow)* **2004**, *69*, 1239−1251.
(6) Zhou, H. X.; Rivas, G. N.; Minton, A. P. *Annu. Rev. Biophys.* **2008**, *37*, 375−397.
(7) Zhou, H.-X. *J. Mol. Recognit.* **2004**, *17*, 368−375.
(8) Minton, A. P. *Biophys. J.* **2005**, *88*, 971−985.
(9) Tsao, D.; Dokholyan, N. V. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3491−3500.
(10) van den Berg, B.; Wain, R.; Dobson, C. M.; Ellis, R. J. *EMBO J.* **2000**, *19*, 3870−3875.
(11) Cheung, M. S.; Klimov, D.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 4753−4758.
(12) Sasahara, K.; McPhie, P.; Minton, A. P. *J. Mol. Biol.* **2003**, *326*, 1227−1237.
(13) Tokuriki, N.; Kinjo, M.; Negi, S.; Hoshino, M.; Goto, Y.; Urabe, I.; Yomo, T. *Protein Sci.* **2004**, *13*, 125−133.
(14) Stagg, L.; Zhang, S. Q.; Cheung, M. S.; Wittung-Stafshede, P. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18976−18981.
(15) Hong, J. A.; Gierasch, L. M. *J. Am. Chem. Soc.* **2010**, *132*, 10445−10452.
(16) Christiansen, A.; Wang, Q.; Samiotakis, A.; Cheung, M. S.; Wittung-Stafshede, P. *Biochemistry* **2010**, *49*, 6519−6530.
(17) Qin, S.; Zhou, H.-X. *Biophys. J.* **2009**, *97*, 12−19.
(18) Schlesinger, A. P.; Wang, Y. Q.; Tadeo, X.; Millet, O.; Pielak, G. *J. Am. Chem. Soc.* **2011**, *133*, 8082−8085.
(19) Samiotakis, A.; Wittung-Stafshede, P.; Cheung, M. S. *Int. J. Mol. Sci.* **2009**, *10*, 572−588.
(20) Dhar, A.; Samiotakis, A.; Ebbinghaus, S.; Nienhaus, L.; Homouz, D.; Gruebele, M.; Cheung, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 17586−17591.
(21) Samiotakis, A.; Cheung, M. S. *J. Chem. Phys.* **2011**, *135*, 175101.
(22) Elcock, A. H. *Curr. Opin. Struct. Biol.* **2010**, *20*, 196−206.
(23) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. *Science* **2011**, *334*, 517−520.
(24) Snow, C. D.; Nguyen, N.; Pande, V. S.; Gruebele, M. *Nature* **2002**, *420*, 102−106.
(25) Honda, S.; Yamasaki, K.; Sawada, Y.; Morii, H. *Structure* **2004**, *12*, 1507−1518.
(26) Muñoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196−199.
(27) Du, D. G.; Zhu, Y. J.; Huang, C. Y.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15915−15920.
(28) Ellis, R. J. *Trends Biochem. Sci.* **2001**, *26*, 597−604.
(29) Ellis, R. J.; Minton, A. P. *Nature* **2003**, *425*, 27−28.
(30) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(31) Matthes, D.; de Groot, B. L. *Biophys. J.* **2009**, *97*, 599−608.

(32) van der Spoel, D.; Seibert, M. M. *Phys. Rev. Lett.* **2006**, *96*, 238102.
(33) Enemark, S.; Rajagopalan, R. *Phys. Chem. Chem. Phys.* **2012**, DOI: 10.1039/c2cp40285h.
(34) Venkatachalam, C. M. *Biopolymers* **1968**, *6*, 1425−1436.
(35) Hatfield, M. P.; Murphy, R. F.; Lovas, S. *J. Phys. Chem. B* **2010**, *114*, 3028−3037.
(36) Dinner, A. R.; Lazaridis, T.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9068−9073.
(37) Thukral, L.; Smith, J. C.; Daidone, I. *J. Am. Chem. Soc.* **2009**, *131*, 18147−18152.
(38) Best, R. B.; Mittal, J. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11087−11092.
(39) "Native-turn formation" discussed here is not to be confused with the "turn formation" discussed earlier in relation to Figure 5. The former refers to the formation of the full native-turn structure, while the latter refers to residue-specific formation of turn structures along the chain.
(40) Stagg, L.; Christiansen, A.; Wittung-Stafshede, P. *J. Am. Chem. Soc.* **2011**, *133*, 646−648.
(41) Wagner, C.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6716−6721.
(42) Mukherjee, S.; Waegele, M. M.; Chowshury, P.; Guo, L.; Gai, F. *J. Mol. Biol.* **2009**, *393*, 227−236.
(43) Jefferys, B. R.; Kelley, L. A.; Sternberg, M. J. E. *J. Mol. Biol.* **2010**, *397*, 1329−1338.
(44) McCallister, E. L.; Alm, E.; Baker, D. *Nat. Struct. Biol.* **2000**, *7*, 669−673.
(45) Dhar, A.; Girdhar, K.; Singh, D.; Gelman, H.; Ebbinghaus, S.; Gruebele, M. *Biophys. J.* **2011**, *101*, 421−430.
(46) Miklos, A. C.; Li, C. G.; Sharaf, N. G.; Pielak, G. J. *Biochemistry* **2010**, *49*, 6984−6991.