Engineering Diverse Changes in β-Turn Propensities in the N-Terminal β-Hairpin of Ubiquitin Reveals Significant Effects on Stability and Kinetics but a Robust Folding Transition State†

Emma R. Simpson, Jill K. Meldrum, and Mark S. Searle*

Centre for Biomolecular Sciences, University Park, University of Nottingham, Nottingham NG7 2RD, U.K.

Received December 7, 2005; Revised Manuscript Received February 7, 2006

ABSTRACT: Using the N-terminal 17-residue β-hairpin of ubiquitin as a "host" for mutational studies, we have investigated the influence of the β-turn sequence on protein stability and folding kinetics by replacing the native G-bulged turn (TLTGK) with more flexible analogues (TG_3K and TG_5K) and a series of four-residue type I′ β-turn sequences, commonly found in β-hairpins. Although a statistical analysis of type I′ turns demonstrates residue preferences at specific sites, the frequency of occurrence appears to only broadly correlate with experimentally determined protein stabilities. The subsequent engineering of context-dependent non-native tertiary contacts involving turn residues is shown to produce large changes in stability. Relatively few point mutations have been described that probe secondary structure formation in ubiquitin in a manner that is independent of tertiary contacts. To this end, we have used the more rigorous rate−equilibrium free energy relationship (Leffler analysis), rather than the two-point ϕ value analysis, to show for a family of engineered β-turn mutants that stability (range of ∼20 kJ/mol) and folding kinetics (190-fold variation in refolding rate) are linearly correlated ($\alpha_f = 0.74 \pm 0.08$). The data are consistent with a transition state that is robust with regard to a wide range of statistically favored and disfavored β-turn mutations and implicate a loosely assembled β-hairpin as a key template in transition state stabilization with the β-turn playing a central role.

Altering the folding kinetics of a protein, or even the sequence or timing of events by which the polypeptide chain assembles, has been shown to be a realistic possibility by changing the intrinsic stability of the various substructural elements. Both theoretical and experimental approaches concur that favorable nativelike local interactions dominate the initial conformational search while minimizing the loss of configurational entropy (*1*−*4*). As a consequence, folding rates of small two-state proteins correlate with the topology of the native state when described by a simple contact parameter which reflects the average separation between interacting residues in the sequence (*5*−*8*). Folding pathways which contain a large proportion of local interactions formed at small entropic cost should lead to the most favored trajectories down a smooth folding funnel while avoiding any possible barriers that could arise from non-native interactions.

Local interactions between residues in turns or loops constitute an important element of protein secondary structure. Although proteins permit considerable variability in the size and composition of turns, they can have a large effect on protein stability and folding kinetics (*9*−*12*), with conformational rigidity suggested as an important factor in restricting the structural heterogeneity of the transition state ensemble. The occurrence of non-Gly residues located in the energetically unfavorable left-handed α-helical region (L-α) of protein structures has attracted attention because of the potential for conformational strain and the scope for further optimization of stability and folding kinetics through mutation (*13*−*15*). In model peptide systems, it is clear that intrinsic residue conformational preferences for β-turn-compatible ϕ and ψ backbone dihedral angles can dictate stability, β-strand residue pairing, and hydrogen bonding register (*16*−*18*). Although tertiary packing interactions, rather than intrinsic secondary structure propensities, ultimately determine α-helix and β-sheet stability in native proteins (*19*−*21*), in the context of small model systems devoid of tertiary contacts, local interactions strongly influence the observable population of marginally stable elements of secondary structure that may be important folding nucleation sites (*16*−*18*).

The β-hairpin represents the simplest model for studying local interactions that nucleate antiparallel β-sheet formation. β-Hairpin peptides have been used extensively as vehicles for quantitative studies of weak interactions and intrinsic residue preferences because the population of the folded state of these marginally stable systems is sensitive to small changes in stability (*16*−*18*). Two-residue type I′ and type II′ β-turns are particularly abundant in β-hairpin peptides (*22*), since the turn geometry is highly complementary to the right-handed twist of the antiparallel β-hairpin. We have investigated the influence of the statistically abundant type I′ β-turn sequence in nucleating protein folding using the

† E.R.S. was funded by the EPSRC of the UK and J.K.M. by a University of Nottingham Research Scholarship and by the School of Chemistry.

* To whom correspondence should be addressed. E-mail: mark.searle@nottingham.ac.uk. Telephone: (+44) 115 951 3564. Fax: (+44) 115 8468002.

10.1021/bi052495g CCC: $33.50
Published on Web 03/10/2006

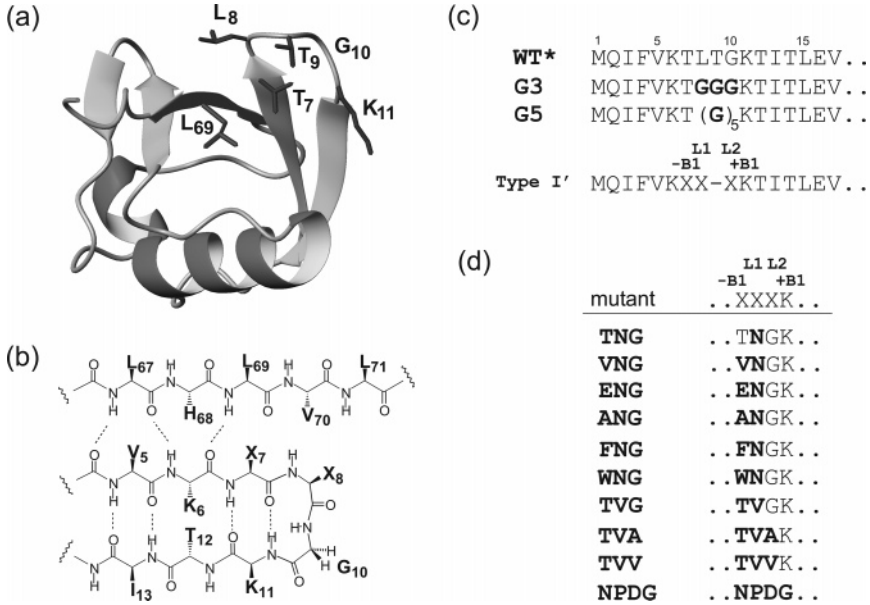

FIGURE 1: (a) X-ray structure of ubiquitin showing the solvent-exposed TLTGK type I G-bulged turn sequence within the N-terminal β-hairpin sequence, together with the position and orientation of the Leu69 side chain on the C-terminal β-strand. (b) Schematic representation of the hydrogen bonding pattern within the replacement type I′ turn ($X_7X_8G_{10}K_{11}$). The residue numbering follows that in the WT* sequence (residue 9 has been deleted in the mutants). (c) N-terminal β-hairpin sequences (residues 1−17) showing substitutions and the positions of residues −B1, L1, L2, and +B1 [according to the nomenclature of Sibanda et al. (*22*)] which define the type I′ turn. (d) List of mutants showing nomenclature and turn sequence.

N-terminal 17-residue β-hairpin of ubiquitin as a "host" for mutational studies (*23*−*25*). In the native ubiquitin sequence, the natural turn is a 3:5 type I G-bulged turn which is also commonly found in protein β-hairpins (*26*). The isolated peptide fragment corresponding to the native N-terminal hairpin sequence shows only a weak propensity to fold in solution (*24*). We have engineered a series of type I′ turns into the native sequence to quantify the effects of intrinsic residue propensities on stability (Figure 1). The effects of both stabilizing and destabilizing mutations on the refolding and unfolding kinetics have now been examined in detail.

The nature of the transition state for folding of ubiquitin has been the subject of recent theoretical and experimental investigation in identifying interactions in the TS ensemble (*27*−*31*). Φ value analysis has identified a large number of fractional values (ϕ in the range of 0.1−0.5) (*27*−*29*), which are open to different possible interpretations (*27*, *32*, *33*). On one hand, these intermediate ϕ values may arise from partial or weakly stabilized nativelike structural interactions in the TS ensemble. On the other hand, they may originate from the presence of multiple alternative TSs with the observed ϕ value representing a population-weighted average in which a particular residue may participate to very different extents in different pathways (*34*). In the case of ubiquitin, relatively few point mutations have been described that probe secondary structure formation in a manner that is independent of tertiary contacts (*29*, *35*). To investigate further the origin of these intermediate ϕ values, and the robustness of the transition state to mutations, we have used the more rigorous rate−equilibrium free energy relationship (REFER) (Leffler analysis) (*32*, *33*, *36*) to examine the role of the N-terminal β-hairpin in the folding of ubiquitin by engineering a family of type I′ β-turns into the hairpin sequence, guided by statistical analyses of residue propensities.

MATERIALS AND METHODS

Cloning and Expression. Details of the cloning and overexpression protocols have been described elsewhere (*20*, *30*, *35*, *37*). All mutations using the WT* background were introduced using the QuikChange site-directed mutagenesis protocol (Stratagene, La Jolla, CA) and confirmed by DNA sequence analysis and by NMR.

Statistical Analysis of Protein Structures. To analyze the frequency of occurrence of residues in β-turns in protein β-hairpins, we used the recently reported Nh3D data set (*38*) of 655 high-resolution nonhomologous protein structures derived from the Protein Data Bank (*39*). The secondary structure assignment program STRIDE was used to identify positions −B1, L1, L2, and +B1 in all possible two-residue turns (*40*), and those sequences with at least one β-sheet residue immediately preceding and following these four residues were selected for further analysis. We identified a total of 609 two-residue turns within this high-resolution data set with the following distribution among turn types: I = 183, I′ = 153, II = 67, II′ = 85, VI = 1, VIII = 29, and undefined = 91 (angles not defined by the other turn types). The distribution of residues within type I′ turns was examined in more detail in Figure 2. The β-turns are classified as 2:2 or 3:5 turn sequences according to the nomenclature described by Sibanda et al. (*22*).

NMR Structural Analysis. All NMR experiments were performed on a Bruker Avance600 spectrometer. Total correlation spectroscopy (TOCSY)[1] and nuclear Overhauser effect spectroscopy (NOESY) experiments were used as previously described on ∼1 mM protein samples at pH 5.0 (*20*, *30*, *35*, *37*). Spectra were internally referenced to

[1] Abbreviations: NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; GdmCl, guanidinium chloride; TS, transition state.

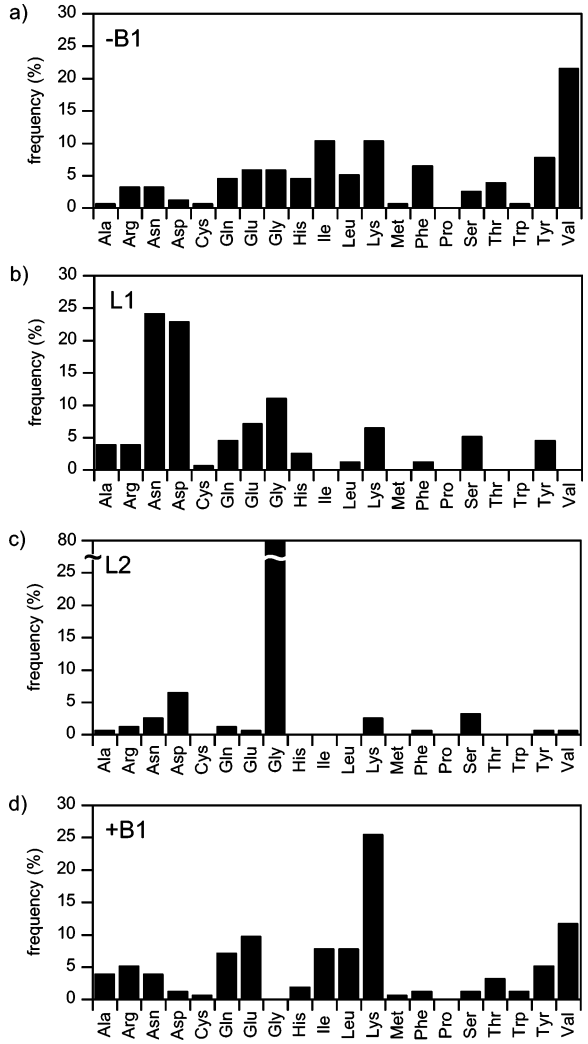

FIGURE 2: Statistical analysis of residues in type I′ β-turns in protein β-hairpins. The percentage frequency of occurrence (of 153 turns identified) is shown at each of the −B1, L1, L2, and +B1 positions. Analysis was derived from the Nh3D database of nonhomologous structures (*38*).

trimethylsilyl propionate (TSP). Data were processed and assigned using Bruker XWINNMR and ANSIG software (*41*). Structural models were visualized using MOLMOL (*42*). The structure of the FNG mutant has been modeled as previously described (*37*). Structural models of other mutants, including G3, G5, and TVV, were generated from the FNG structure or from the X-ray coordinates of ubiquitin (*43*) and subjected to several rounds of steepest descent and conjugate gradient minimization. Unrestrained molecular dynamics simulations were performed using the AMBER6 suite of programs (//http:amber.scripps.edu/doc6/install.html) and equilibrated with counterions and TIP3P water molecules to a distance of 8 Å and dynamics run for 1000 ps. The ensemble of structures was sampled over the final 500 ps to examine the conformation and dynamics of the N-terminal hairpin and loop for comparisons with experimental NOE data.

Stability and Kinetic Measurements. Protein stabilities were determined using equilibrium fluorescence measurements as previously described (*35*, *37*, *44*). Kinetic unfolding and refolding measurements were performed using an Applied Photophysics Pi-star 180 spectrophotometer. The temperature was regulated using a Neslab RTE-300 circulating programmable water bath. All kinetics experiments were performed in 25 mM acetate buffer at pH 5.0 and 298 K. Refolding experiments were performed by 1:10 dilution of unfolded protein (15 μM in 7 M GdmCl) into buffered solutions with different GdmCl concentrations, yielding a final protein concentration of 1.36 μM. For unfolding experiments, a buffered solution of native protein was unfolded by a 1:10 dilution to yield final concentrations of GdmCl near or above the midpoint of the equilibrium unfolding transition (concentrations of GdmCl in the range of 2.0−7.3 M). Kinetic measurements for both unfolding and refolding reactions were averaged four to six times at each GdmCl concentration (*30*, *35*). In all cases, the GdmCl concentration was determined using a refractometer (*44*). In no cases do we see deviations of the kinetic data from a linear chevron plot. However, to check for the possible complications of transient aggregation on refolding kinetics at low denaturant concentrations, experiments were repeated at protein concentrations in the range of 0.1−1.5 μM. All of the kinetic data were superimposable and independent of the protein concentration.

Analysis of Kinetic Data. Fluorescence-detected kinetic traces were deconvoluted using a multiexponential fitting procedure involving two or three components for refolding and one to two components for unfolding. In the main, the unfolding data fit to a single-exponential process; however, a minor slow second unfolding phase is evident at high denaturant concentrations that accounts for only ~1% of the signal amplitude. This is difficult to characterize but could be associated with a slow isomerization event in the unfolded state. The major refolding phase, which accounts for ~85% of the fluorescence amplitude, was analyzed using a linear chevron plot (ln k_{obs} vs [D]) and equations appropriate for an apparent two-state model (*45*−*47*). The observed rate constant, k_{obs}, is the sum of the folding and unfolding rates; $k_{obs} = k_{UN} + k_{NU}$, where k_{obs} is dependent on [D] according to the expression

$$\ln k_{obs} = \ln[k_{UN} \exp(-m_{UN}[D]/RT) + k_{NU} \exp(m_{NU}[D]/RT)] \quad (1)$$

The linear dependence of ln k_{obs} on [D] gives extrapolated values for k_{NU} and k_{UN} in water alone, together with the slopes of the unfolding and refolding components, m_{NU} and m_{UN}, respectively. From the kinetic parameters derived from eq 1, the stability ΔG_{kin} can be determined from the rate constants k_{UN} and k_{NU} from eq 2:

$$\Delta G_{kin} = -RT \ln(k_{UN}/k_{NU}) \quad (2)$$

and from the refolding and unfolding m values, m_{UN} and m_{NU}, respectively, and midpoint denaturant concentration $[D]_{50\%}$ derived from the chevron plot (eq 3):

$$\Delta G_{kin} = -(m_{UN} + m_{NU})[D]_{50\%} \quad (3)$$

Although the former approach (eq 2) can lead to uncertainties in ΔG_{kin} from potentially large extrapolation errors when calculating unfolding rates in water, we obtain consistent results using both approaches.

RESULTS

Statistical Analysis of Type I′ β-Turns in Proteins. Two-residue type I′ β-turns are common in protein β-hairpins (*22*),

because of the complementarity between the turn geometry and the right-handed twist of the β-sheet. The turn is described by the (−B1)−(L1)−(L2)−(+B1) sequence (*22*), where L1 and L2 are the turn residues and −B1 and +B1 are the β-strand residues preceding and following the turn, respectively. The backbone ϕ and ψ angles of L1 and L2 both lie in the left-handed α-helical (L-α) region of the Ramachandran plot. We have extended earlier analyses of the frequency of residue occurrence at the key turn positions (*48*, *49*) using the recently reported Nh3D database of high-resolution nonhomologous protein structures (38). We have specifically identified 153 type I′ turns in the context of β-hairpin motifs (Figure 2). The analysis confirms that Gly is by far the most abundant residue at position L2 since the required backbone angles ($\phi \sim 90°$ and $\psi \sim 0°$) are disallowed for all residues other than Gly (*49*). At position L1, the dihedral angles ($\phi \sim 60°$ and $\psi \sim 30°$) are accessible to other non-β-branched residues, with Asn, Asp, and Gly appearing in ∼60% of the cases, Asn being marginally preferred. In contrast, β-branched residues such as Val and Ile do not occur at this position in this data set which correlates with their strong preference for an extended β-sheet conformation. For these reasons, the Asn-Gly (L1-L2) type I′ turn has been widely used in model hairpin peptide systems for stability and folding studies and has proven to be highly effective at inducing folded structure (*16*−*18*). In other cases, the use of ^{D}Pro at position L1 to induce the correct backbone geometry has also proven to be highly successful (*16*−*18*).

Design and Stability of β-Turn Mutants of Ubiquitin. All of the mutations described have been constructed against a background F45W substitution (denoted WT*) which gives a sensitive fluorescence probe for stability and folding studies (*29*, *30*, *50*−*53*). Estimates of the change in stability were derived from kinetic measurements based on refolding and unfolding rate constants and from kinetic *m* values and transition midpoints (see Materials and Methods). Turn mutations were engineered into the N-terminal β-hairpin of ubiquitin which spans residues 1−17. The X-ray structure of ubiquitin reveals a 3:5 type I G-bulged turn sequence (TLTGK), as shown in Figure 1a (*43*). The first and last of the five residues results in closure of the turn through cross-strand hydrogen bonding. A number of contacts within the TLTGK turn sequence suggest stabilizing interactions, for example, between the β-OH group of Thr7 and the backbone NH group of Thr9, as well as a tertiary interaction between the side chain of Leu8 and Val70, the latter in the C-terminal parallel strand, and a salt bridge between Lys11 and Glu34, the latter in the main α-helix (*43*).

Initially, the native TLTGK G-bulged turn was mutated to the more flexible TGGGK turn (G3 mutant) to eliminate interactions within the turn while maintaining the same loop size. Surprisingly, these mutations had essentially no effect on the stability of the protein ($\Delta\Delta G \sim 0.1$ kJ/mol). To further examine the effects of loop flexibility, we extended the turn by a further two Gly residues (TGGGGGK, G5 mutant), which destabilized WT* by 5.9 kJ/mol.

The native G-bulged loop was subsequently replaced by with a series of 2:2 type I′ turn sequences (Figure 1d), and we proceeded to examine the relationship between the frequency of occurrence of residues at key positions within the turn sequence and their effects on stability. The residue at position +B1 following the Gly residue is frequently found to be Lys (Figure 2d), and since Lys is found at this position in the native turn and is involved in a Lys11−Glu34 salt bridge (*43*), we preserved this residue in all of the mutants and examined the effects at positions −B1, L1, and L2. To test the effects of intrinsic conformational preferences on hairpin stability, the LT residues of the native G-bulged type I turn sequence TLTGK were mutated to residue X to yield the TXGK turn sequence, with T, X, G, and K representing −B1, L1, L2, and +B1, respectively (see Figure 1). Initially, we concentrated on residues −B1 and L1, since Gly is found to be by far the most common residue at position L2 (Figure 2c). The most favored Asn-Gly and least favored Val-Gly sequences were incorporated within the TNGK and TVGK turns (TNG and TVG mutants, respectively). In the first instance, TNG produces a slight decrease in stability of 1.2 kJ/mol, suggesting that TLTGK and TNGK are equally good turns in this context, whereas TVG resulted in a further reduction in stability of 2.3 kJ/mol versus WT*, reflecting a relatively small free energy difference between the most favored and least favored residue at position L1. A detailed NMR analysis of the TVGK and TNGK mutants at the level of a full backbone assignment has enabled us to confirm that both of these sequences are accommodated with similar local backbone geometries typical of type I′ turns, with minor chemical shift perturbations confined to residues adjacent to the mutated residue (*37*).

Subsequently, we investigated the effects of substitutions at the −B1 position (XNGK) on protein stability. As a reference point, we calculated stability changes with respect to the TNGK sequence (X = T). The statistically preferred X = V (VNG) substitution, which reflects the preference for β-sheet ϕ and ψ backbone angles at the −B1 position, produces a substantial increase in stability of −6.3 kJ/mol. However, the X = A substitution (ANGK, ANG), a residue with a high α-helical propensity, destabilizes the hairpin by 1.3 kJ/mol, while the X = E substitution (ENGK, ENG) produces little effect on stability (0.3 kJ/mol). Thus, at the two extremes with regard to β-sheet forming propensities (Val vs Ala), the nature of the residue at position −B1 modulates stability by ∼8 kJ/mol.

Modeling studies suggested that the introduction of Phe and Trp (FNG and WNG mutants) within the XNGK turn sequence could lead not only to favorable hydrophobic interactions with residues within the hairpin sequence but might also allow us to engineer new stabilizing tertiary contacts with L69 and L71 on the adjacent C-terminal β-strand. These mutations produce significant further increases in stability with respect to a TNGK sequence of −11.5 kJ/mol (FNGK) and −14.5 kJ/mol (WNGK) (see Table 1). NMR analysis of the FNG and WNG mutants reveals substantial upfield shifts for the methyl resonances of L69 and L71 of between 0.4 and 0.9 ppm that confirm tertiary interactions between the aromatic side chain of Phe or Trp and residues on the adjacent C-terminal strand (*37*). Further, a significant enhancement of the change in fluorescence intensity between folded and unfolded states for WNG is consistent with the indole ring becoming at least partially excluded from solvent by the formation of tertiary contacts. We subsequently introduced the L69A mutation into WT* and FNG (WT*^{A} and FNG^{A}) and observed a significant destabilization (12−14 kJ/mol). A thermodynamic cycle constructed from the stability data for WT* and FNG with

Table 1: Kinetic Data for β-Turn Mutants of Yeast Ubiquitin

mutant[a]	k_{NU} (s^{-1})	m_{NU} (J mol^{-1} M^{-1})	k_{UN} (s^{-1})	m_{UN} (J mol^{-1} M^{-1})	ΔG_{kin} (kJ mol)[b]	β_T[c]
WT*	0.0084 ± 0.0007	2905 ± 44	263 ± 9	5684 ± 55	−25.6	0.66
G3	0.0067 ± 0.0005	2860 ± 38	200 ± 7	5670 ± 61	−25.5	0.66
G5	0.014 ± 0.0013	2628 ± 50	41 ± 3	6154 ± 141	−19.7	0.70
TNG	0.0060 ± 0.00082	2836 ± 68	115 ± 6	5039 ± 82	−24.4	0.64
VNG	0.0044 ± 0.003	3082 ± 279	1049 ± 172	4769 ± 226	−30.7	0.61
ENG	0.010 ± 0.002	2811 ± 116	168 ± 18	5939 ± 206	−24.1	0.68
ANG	0.014 ± 0.002	2787 ± 61	165 ± 11	5413 ± 113	−23.1	0.66
FNG	0.00059 ± 0.00024	3370 ± 186	1148 ± 51	4891 ± 57	−35.9	0.59
WNG	0.00012 ± 0.00006	3348 ± 207	800 ± 58	4947 ± 89	−38.9	0.60
TVG	0.0043 ± 0.0004	2915 ± 50	54 ± 3	5662 ± 95	−23.3	0.66
TVA	0.0037 ± 0.0003	2929 ± 33	35 ± 2	5824 ± 99	−22.6	0.67
TVV	0.0037 ± 0.0003	2913 ± 43	23 ± 1	5774 ± 104	−21.8	0.66
NPDG	0.030 ± 0.002	2482 ± 27	9 ± 0.5	5135 ± 134	−14.1	0.67
WT*[A]	0.98 ± 0.08	1901 ± 45	175 ± 31	6546 ± 455	−12.9	0.77
FNG[A]	0.078 ± 0.02	2470 ± 132	505 ± 77	5898 ± 289	−21.8	0.70

[a] Substitution inserted against the background F45W mutation; a superscript A represents the L69A mutation. All kinetic data were collected at a protein concentration of ~1.4 μM at 298 K in 25 mM acetate buffer (pH 5.0). [b] Calculated assuming two-state folding and $\Delta G_{kin} = -RT \ln(k_{UN}/k_{NU})$; errors in stabilities lie in the range of ±0.3−0.6 kJ/mol. [c] From $\beta_T = m_{UN}/(m_{UN} + m_{NU})$.

and without the L69−F7 interaction contributes only ~2 kJ/mol to the enhanced stability of FNG, suggesting that the major contribution involves the side chain of L71.

Finally, we also considered mutations at the L2 position to investigate the tolerance to residues other than the statistically highly preferred Gly residue (Figure 2). From the database, both Ala and Val have a very low frequency of occurrence at this position, representing residues with a strong preference for ϕ and ψ angles characteristic of α-helix and β-sheet, respectively, rather than the required α-L backbone geometry. Substitution at position L2 (X) in the already destabilizing TV<u>X</u>K turn sequence (X = A, TVA mutant, and X = V, TVV mutant, respectively) surprisingly resulted in only small further decreases in stability in both cases (<2 kJ/mol; see Table 1), despite the poor representation of these residues at the L2 position in the statistical data. We carried out a detailed NMR analysis of the TVV mutant to establish whether inserting sterically bulky Val residues in adjacent positions led to significant structural changes (such as an expansion of the loop size) that may help to partially offset the energetic cost of turn formation. We compared Hα chemical shift perturbations from random coil values for residues within the β-hairpin sequence (residues 1−16) of the TNGK, TVAK, and TVVK turn mutants (Figure 3), representing the statistically most abundant and one of the least abundant substitutions at positions L1 and L2. Modest differences in Hα shift perturbations ($\Delta\delta^{H\alpha} \leq$ 0.2 ppm) for TNGK versus TVAK and TVVK are confined to the residues immediately adjacent to the turn sequence in positions −B1 and +B1, suggesting that conformational changes are not propagated significantly beyond the turn residues. The pattern of NOEs within the turn and between residues in cross-strand positions ($T7^{NH}$−$T12^{H\alpha}$ and $T7^{NH}$−$K11^{NH}$), together with the magnitude of $^3J_{NH-H\alpha}$ values estimated from TOCSY cross-peak intensities, confirms the strand alignment and loop size are those of a 2:2 β-turn in all three cases, consistent with earlier studies using model peptides (*54*, *55*). Thus, the NMR data suggest that the TVAK and TVVK sequences are accommodated with a backbone geometry similar to that of TNGK.

Kinetic Analysis of Ubiquitin Folding. The role of the β-hairpin motif in nucleating the folding of ubiquitin has been examined in kinetic detail from fluorescence-detected stopped-flow experiments at 298 K on the family of β-turn mutants described herein. The refolding data at low denaturant concentrations (0−3 M GdmCl) show clear evidence for multiphase kinetics and are resolved into three folding processes (*30*, *35*). The fastest phase accounts for ~85% of the change in fluorescence amplitude when extrapolated to 0 M denaturant, with the minor phases previously attributed to slower isomerization-limited events (*27*, *29*, *50*−*53*), although these remain to be characterized in detail. The observation that the variation in the rate of folding of the major phase as a function of denaturant concentration fits to a linear chevron plot (ln k_{obs} vs [D]) without deviations from linearity at low denaturant concentrations suggests that the majority of molecules fold along this apparent two-state pathway. For the purpose of this analysis, we have focused on this dominant fast phase whose transition state we now analyze (Figure 4) (*30*, *35*, *52*, *53*).

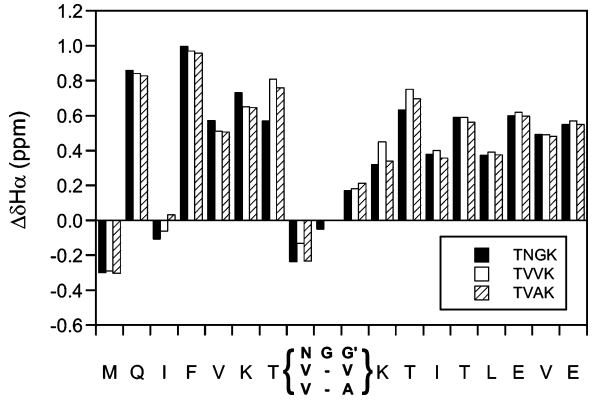

FIGURE 3: Perturbations to Hα chemical shifts (measured values vs random coil values, $\Delta\delta^{H\alpha}$, parts per million) of residues 1−17 in the N-terminal β-hairpin. NMR data were collected at 600 MHz, 298 K, and pH 5.0. $\Delta\delta^{H\alpha}$ values are compared for the TNG, TVA, and TVV mutants, with TNG and TVV representing the most favored (Asn-Gly) and one of the least favored (Val-Val) type I′ turns.

Initially, we examined the consequences of the Gly substitutions on folding kinetics. Chevron plot analysis of the kinetic data of WT* and G3 (Figure 4a) showed that the small difference in apparent stability is manifested in similarly small effects (<2-fold) on both the refolding and

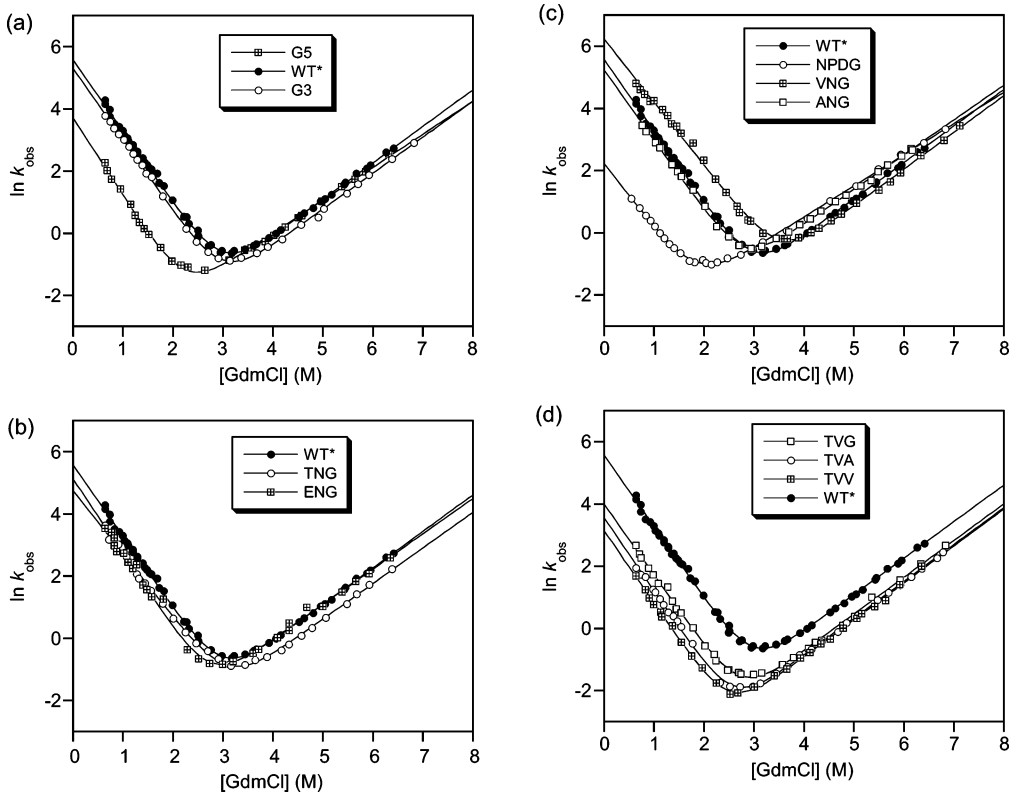

FIGURE 4: Chevron plots (a−d) showing refolding and unfolding kinetic data as a function of denaturant concentration for WT* ubiquitin and the full set of turn mutants highlighted in Figure 1 (all data collected at pH 5.0 in 25 mM sodium acetate buffer at 298 K). Data were fitted to a two-state model according to eq 2, and all kinetic paramaters are listed in Table 1.

unfolding rates. However, the data for G5, with the destabilizing flexible loop sequence, show a significant (~6-fold) deceleration of the refolding rate compared to that of WT* but a relatively small effect (<2-fold) on the unfolding kinetics. The observation of only small changes in the denaturant dependence of the refolding and infolding rates (m_{UN} and m_{NU}, respectively) indicate that expanding the loop size is not significantly affecting the structure or solvent accessibility of the transition state as manifested in Tanford β_T values (see Table 1).

Turning to the data on the type I′ turn mutants, where the effects on stability are small (TNG, ENG, and ANG mutants vs WT*), we observe equally small perturbations to both the refolding and unfolding kinetics (Figure 4b,c). However, where more pronounced stability changes are apparent we see specific effects on the kinetics that lead to a much clearer interpretation in terms of the effects on the folding pathway of ubiquitin. In Figure 4d, the data for WT* are compared with those for the type I′ TVGK turn substitution. This turn is less stable than the native G-bulged type I turn. The difference in stability is clearly manifested in substantial reductions in the refolding rate (~5-fold) of ubiquitin, with relatively small perturbations to the unfolding kinetics (~2-fold). The stabilizing effect of the VNGK turn (Figure 4c), which represents the statistically preferred type I′ turn sequence, is manifested largely in accelerated refolding kinetics (Table 1). Importantly, we see only relatively small variations in m_{UN} and m_{NU} values, which suggest that in these cases we are not substantially affecting the hydrophobic surface burial in the transition state for folding.

We also examined the L2 type I′ turn mutants in which the preferred Gly residue was replaced with Ala and Val, which occur with a low frequency at this position. The effects on stability of the X = Gly to X = Ala/Val substitutions (in the context of TVXG) are surprisingly small, and subsequently, we see only small effects on refolding and unfolding kinetics (Figure 4d). However, most of the effect is clearly manifested in the refolding rate. In Figure 4c, we have also included previously published data on the NPDG mutant in which the native TLTGK bulged turn sequence was replaced by the highly destabilizing NPDG type I turn (*20*). Although the latter is a common type I turn in proteins, it is significantly less common in β-hairpins because of the apparent incompatibility of the geometry of the type I turn sequence with the right-handed twist of the antiparallel β-strands (*25*). The substantial destabilization of the hairpin that this turn substitution introduces (11.5 kJ/mol) is manifested in a large (~30-fold) deceleration of the refolding kinetics (Figure 4c).

Engineering Stabilizing Tertiary Interactions Decelerates Unfolding. Substitution with a bulky aromatic residue at position −B1 (FNG and WNG mutants) uniquely produces a large change in the unfolding kinetics (Figure 5a). While we observe an ~3−4-fold acceleration of the extrapolated refolding rate compared to that of WT*, the largest contribution from the stabilizing effects of these substitutions is apparent in the substantial deceleration of the unfolding rate (~14-fold for FNG and ~70-fold for WNG). Thus, while hydrophobic contacts involving the aromatic side chains of F7 and W7 appear to play some role in stabilizing the hairpin conformation in the transition state, with a subsequent acceleration of the refolding rate, the larger part of the observed increase in stability is manifested not in transition state stabilization but in the stabilization of the native state,

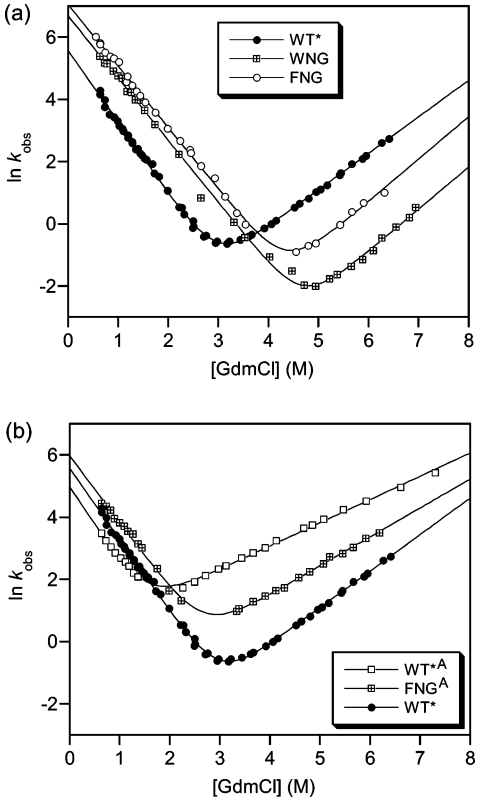

FIGURE 5: Chevron plots showing the logarithm of refolding and unfolding rate constants as a function of denaturant concentration for (a) WT* ubiquitin and the FNG and WNG mutants. In panel b, the effects on the folding and unfolding kinetics of the L69A mutation are shown in the context of WT* and FNG (assigned superscript A). All data were collected at pH 5.0 in 25 mM sodium acetate buffer at 298 K. Data were fitted to a two-state model according to eq 2, and all kinetic paramaters are listed in Table 1.

implying that tertiary contacts, in this case within the C-terminal β-strand, are consolidated at a late stage after the main barrier crossing event has taken place. While the NMR analysis of the FNG mutant confirms the interactions between F7 and L69/L71 in the C-terminal β-strand, the kinetic data provide the details of the timings of these events.

To dissect the kinetic data for FNG into the relative contributions from local interactions within the hairpin, and those from tertiary contacts with the C-terminal β-strand, we introduced the L69A mutation into WT* and FNG (WT*^{A} and FNG^A, respectively). In both cases, the L69A mutation decreases the stability (12−14 kJ/mol) reflecting extensive hydrophobic interactions involving L69. We examined the FNG^A mutant by NMR and found that perturbations to the Hα chemical shifts were confined to residues close in space to the mutation site. In contrast to the large perturbations observed for FNG and to the chemical shifts of the L69 methyl resonances from F7, the methyl of A69 is relatively unperturbed ($\Delta\delta^{H\alpha} = 0.20$ ppm), suggesting a much weaker interaction of this truncated side chain with F7. The perturbations to the methyl signals of L71 were unaffected by the L69A substitution. In the case of FNG^A, the stability changes associated with the L69A mutation are manifested in substantial changes to the unfolding rate (~130-fold) rather than refolding rate (~2-fold) (Figure 5b), in good agreement with earlier data from Went and Jackson on the L69A mutation in WT* (*29*).

Tolerance of the Folding Transition State to β-Turn Substitutions. Rate−equilibrium free energy relationships, describing effects of structural modification on equilibria and reaction rates, are the classical method of analysis of chemical reaction transition states (*56*, *57*). This multipoint analytical approach has been successfully extended in examining noncovalent interactions present in protein folding transition states using protein engineering methods and has advantages over the traditional two-point ϕ value analysis because the former monitors smooth changes in kinetics over a wide range of stabilities (*32*, *33*, *36*, *58*, *59*). In the current context, we have analyzed local interactions within the solvent-exposed β-turn sequence using the same approach based upon empirical Leffler (or Brønsted) relationships:

$$\ln k_{UN} = \ln k_{UN}^{\circ} - \alpha_f \Delta\Delta G/RT \quad (4)$$

$$\ln k_{NU} = \ln k_{NU}^{\circ} + (1 - \alpha_f)\Delta\Delta G/RT \quad (5)$$

where $\Delta\Delta G$ represents the effect of the mutation on protein stability, k_{UN}° and k_{NU}° are the refolding and unfolding rates of WT*, respectively, k_{UN} and k_{NU} are the rate constants of the mutant, and α_f is a proportionality constant describing the sensitivity of the TS relative to the ground state to structural perturbations. Thus, α_f measures the position of the transition state along the reaction coordinate so that an α_f of ~1 indicates a highly structured TS and an α_f of ~0 indicates a TS that is largely unstructured. If the data for FNG and WNG are excluded, a linear correlation ($R = 0.94$) is apparent for the plot of ln k_{UN} versus $\Delta\Delta G/RT$ over a ~20 kJ/mol stability range and a ~190-fold range of the refolding rate constant (Figure 6a). From the linear Leffler plots, we calculated an α_f value of 0.74 ± 0.08. We have also included in this analysis data from our earlier investigation of a family of mutants in which the N-terminal hairpin has been elongated by expanding the loop sequence with an autonomously folding β-hairpin peptide sequence (*35*). The stability and kinetic data for this mutant (Uβ4), and several analogues substituted with Ala in the solvent-exposed hairpin extension, further substantiate the analysis. The observation of a linear correlation suggests that all of the mutations are probing the same homogeneous region of structure in the transition state, the position of which is not changing significantly with respect to the folded and unfolded states. This is also evident from the observation that the denaturant dependence of the refolding and unfolding rate constants (*m* values) are also relatively invariant with stability. When the data for FNG and WNG are included in this plot, they deviate from the linear fit (Figure 6a) which reflects the fact that in these two cases the analysis is more complex because new interactions have been generated that are affecting other substructural elements, as already demonstrated. When we apply the Leffler analysis to residues at various positions within the C-terminal strand of ubiquitin using the L69A mutations, additional data generated in previous studies involving mutations to solvent accessible residues T66 and H68 that are also found on this β-strand (*35*), and the data for FNG and WNG, linear correlations are also observed for refolding and unfolding. A slope α_f of 0.21 ± 0.06 (Figure 6b) confirms that interactions between the C-terminal β-strand and the N-terminal β-strand of the hairpin are at best only weakly formed in the folding transition state and are consolidated at a late stage in the folding process.

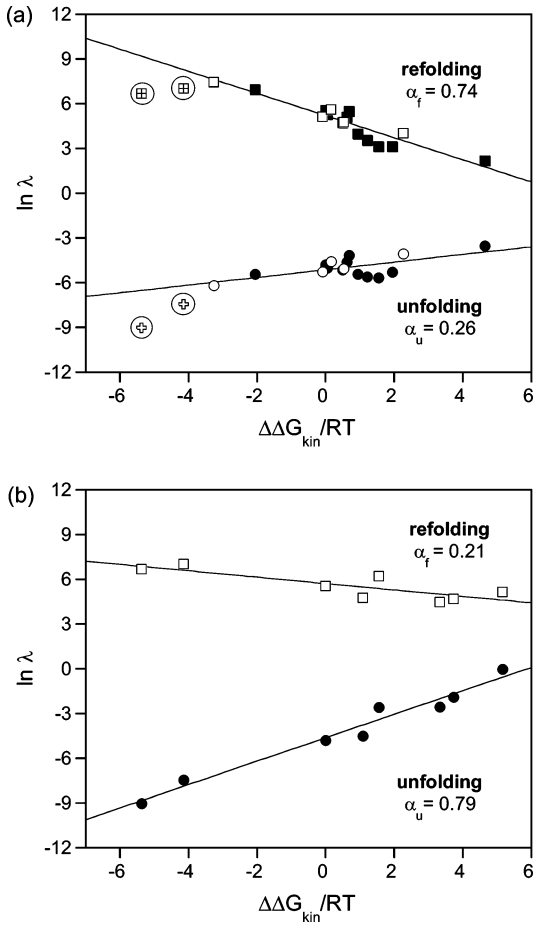

FIGURE 6: Leffler plots showing the logarithm of the observed rate (refolding or unfolding) vs the change in stability ($\Delta\Delta G/RT$) for the family of ubiquitin mutants. In panel a, the effects of mutations in the β-turn fit a linear correlation ($R = 0.94$) giving a slope (α_f) of 0.74 ± 0.08, suggesting that the β-turn is significantly structured in the transition state for folding. Included in this plot are data for the ubiquitin Uβ4 protein and several Ala mutants that contain an extension of the N-terminal hairpin with an autonomously folding 14-residue β-hairpin motif (*35*) (□ for refolding data and ○ for unfolding data). In addition, the two outlying points marked with circles correspond to data for the FNG and WNG mutants (not included in the fit), where effects on stability and folding kinetics also contain a component corresponding to new tertiary interactions between the turn residues and the C-terminal β-strand (Leu69). In panel b, the data for FNG and WNG are shown along with data for the L69A mutants (WT^{*A} and FNG^{A}) and the T66A/H68A and T66E/H68E mutants that represent the substitution of residues within the C-terminal β-strand or involving tertiary contacts with this strand. A linear correlation is observed ($R = 0.83$) with a Leffler α_f value of 0.21 ± 0.06, showing that the contacts with the C-terminal β-strand are at best only very weakly formed in the folding transition state. Thus, adjacent β-strands within the N-terminal and C-terminal parts of the sequence show distinctly different levels of participation in stabilizing interactions in the folding transition state which is highly polarized toward structural elements within the N-terminal half of the sequence (*28*, *29*).

The observation of high correlation coefficients in the Leffler plots for both data sets confirms that the mutated side chains within each family of mutants are structured to a similar extent in the transition state. Significant scatter, or deviation of the data from linearity (biphasic behavior), has been suggested as an indication of competing parallel pathways with multiple transition states (*32*). In principle, a large number of mutations in a small region of the protein analyzed by two-point ϕ value analysis and by Leffler analysis should enable us to draw similar conclusions. There are important differences between ϕ and α values, where ϕ can depend on very specific interactions that are present in a folding transition state, so mutations must be selected carefully for nondisruptive side chain deletions (*33*). Thus, ϕ values are capable of delivering atomic level information about the structure of folding transition states, whereas the multipoint Leffler plot can be used to analyze a region of structure such as a helix and, in this context, lends itself well to a solvent-exposed β-turn. However, as discussed in several recent papers, the two-point analysis can lead to large uncertainties in ϕ values when dealing with mutations that result in only small differences in stability (*58*, *59*). In the data presented here, we see the same variability in ϕ values, where those that correlate well with α_f values from Leffler plots have correspondingly large stability changes, while stability changes of ≤ 3 kJ/mol produce, in some cases, ϕ values of >1 or <0 (*59*). Rather than focusing on a few data points that result in unusual ϕ values that are not well-determined by the small stability changes, we find the same data sets result in smooth linear free energy relationships using the multipoint Leffler analysis of the same homogeneous region of structure.

DISCUSSION

Correlating Turn Sequence with Stability and Folding Kinetics. Using the N-terminal 17-residue β-hairpin of ubiquitin as a "host" for mutational studies, we have investigated the influence of the β-turn sequence on protein stability and folding kinetics. Interestingly, replacing the native TLTGK G-bulged turn with the equivalent Gly-rich TGGGK sequence appears to have no effect on stability or folding kinetics. We can speculate that the removal of backbone steric strain through Gly substitution may compensate for the loss of any side chain specific stabilizing contacts in the native sequence. However, expanding the turn to a TG_5K loop results in a 6 kJ/mol reduction in stability and slower refolding kinetics that may reflect a substantial increase in the entropic cost of folding the hairpin template. Subsequently, we replaced the native G-bulged turn (TLTGK) with a four-residue sequence corresponding to a type I′ β-turn that is found to be abundant in protein β-hairpins (*22*). Statistical analysis shows that certain residues are preferred at specific sites and that the frequency of occurrence appears to broadly correlate with turn stability (*49*, *60*). It is evident that turn sequence selection, involving the substitution of a relatively small number of residues, can lead to quite significant effects on protein stability (~25 kJ/mol in the current context) between turn-destabilized mutants and those with engineered tertiary contacts. This stability range is in general agreement with earlier studies of the B1 domain of protein G in which stabilizing two-, three-, and four-residue loop sequences were identified through in vitro selection (*60*). However, the data suggest that additional evolutionary factors other than intrinsic backbone conformational preferences may contribute to the frequency of occurrence, including the importance of side chain polarity in β-turns in encoding surface accessibility and avoidance of non-native contacts. We analyzed the frequency of occurrence of polar residues in 609 two-residue β-turns of all types and found that a subset of polar residues (Arg, Asn, Asp, Glu, Lys and Ser) or the flexible Gly (data not shown)

could account for >70% of residues at positions L1 and L2. Thus, the low frequency of occurrence of hydrophobic residues in these positions could be ascribed partly to unfavorable intrinsic ϕ and ψ preferences and also to eliminate the possibility of non-native hydrophobic interactions that could decelerate folding. However, in some contexts, it has been suggested these may be beneficial to productive folding by sequestering at an early stage hydrophobic residues that might otherwise make the protein prone to misfolding and aggregation (*61*, *62*). In small fast folding proteins, compact non-native states are likely to be counterproductive in the initial conformational search for nucleating elements of structure (*35*, *63*−*67*). In other studies, the role of β-turns in nucleating folding has been shown to be strongly context-dependent and that the choice between topologically similar folding mechanisms must be defined by local interactions. Mutations within the two symmetrically disposed β-turns of the B1 domain of protein L show that they have quite different effects on folding kinetics with the transition state polarized toward a structured first turn but disrupted second turn, as evident from contrasting effects of residue substitutions on refolding and unfolding kinetics (*68*). However, re-engineering stability at the level of local interactions within the two symmetrically disposed β-hairpins has shown that it is possible to alter the folding mechanism or nucleation point, reflecting differences in free energy between the isolated β-hairpins in the unfolded state (*11*).

It is interesting to compare these studies with recent investigations of the folding and dynamics of isolated β-hairpin peptides that are amenable to T-jump fast-kinetic methods. Folding rates appear to be substantially accelerated by reducing the distance of the turn sequence from the stabilizing hydrophobic cluster of interstrand interactions, with the entropic penalty associated with turn formation being an important determinant of the free energy barrier for β-hairpin folding (*69*−*72*). A more recent ϕ value analysis of β-hairpin folding using the trpzip4 template has shown that a strong turn-promoting sequence increases the stability of a β-hairpin primarily by increasing the folding rate; however, a strong stabilizing hydrophobic cluster between strand residues increases the stability of the hairpin primarily by decreasing the unfolding rate (*73*). The implication here is that cross-strand interactions are instrumental in preventing unfolding, forming mainly after the rate-limiting barrier crossing event has occurred, and that the turn is a key nucleating element in directing folding. Although there appear to be differences in the nature of the unfolded state of peptide β-hairpins induced thermally or with chemical denaturants, in both cases the "denatured" state appears to retain a high degree of compactness as evident from FRET studies (*73*), and as described previously in NMR studies (*74*).

Linear Rate−*Equilibrium Free Energy Relationships and Folding Transition States.* The observation of fractional ϕ values in protein engineering studies of folding pathways has led to speculation regarding their interpretation in terms of a heterogeneous population-weighted folding landscape with canonical ϕ values at the two extremes, or a single TSE where there is weak partial formation of structure (*3*, *27*, *28*, *32*, *33*). In recent studies, few ϕ values were determined that directly probed secondary structure formation in the N-terminal hairpin sequence that were independent of tertiary contacts (*29*). Using the more rigorous linear rate−equilibrium free energy relationship approach to study a family of mutants spanning a significant range of stabilities (~20 kJ/mol), we have focused on structure formation in the solvent-exposed turn sequence. There are many examples of well-characterized systems in which the nature of the folding transition state varies considerably from the diffuse structures found for nucleation−condensation mechanisms to more compact structures approaching the framework folding model (*33*, *34*). In the current context, the β-hairpin of ubiquitin is folded in the TSE, but as demonstrated from peptide fragment stability studies (*23*, *24*), this structure is not compact but stabilized by a loose network of a few hydrophobic side chains that is probably only consolidated at a late stage in the folding reaction (*75*). Indeed, little NH protection against solvent exchange occurs early in the folding of ubiquitin; rather, secondary structure appears to be significantly stabilized late in a major cooperative folding event (*76*, *77*). The protein engineering approach described here, coupled with statistics on type I′ β-turn propensities, NMR structural analysis, and the more rigorous approach of using rate−equilibrium free energy relationships to probe structure in the folding transition state, shows the latter to be robust with regard to mutations and that the β-turn sequence in the N-terminal β-hairpin peptide is substantially structured ($\alpha_f = 0.74 \pm 0.08$), playing a central role in mediating protein stability and in templating the folding of ubiquitin.

ACKNOWLEDGMENT

We thank Dr. H. Williams for assistance with NMR experiments and database analysis and E. Holmes for some preliminary mutagenesis work.

REFERENCES

1. Jackson, S. E., elMasry, N., and Fersht, A. R. (1993) Structure of the hydrophobic core in the transition state for folding of chymotrypsin inhibitor 2: A critical test of the protein engineering method of analysis, *Biochemistry 32*, 11270−11278.
2. Abkevich, V. I., Gutin, A. M., and Shakhnovich, E. I. (1994) Specific nucleus as the transition state for protein folding: Evidence from the lattice model, *Biochemistry 33*, 10026−10030.
3. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: Evidence for a nucleation-collapse mechanism for protein folding, *J. Mol. Biol. 254*, 260−288.
4. Shaknovich, E. I., Abkevich, V., and Ptitsyn, O. (1996) Conserved residues and the mechanism of protein folding, *Nature 379*, 96−98.
5. Plaxco, K. W., Simons, K. T., and Baker, D. (1998) Contact order, transition state placement and refolding rates of single domain proteins, *J. Mol. Biol. 277*, 985−994.
6. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M., and Dobson, C. M. (1999) Mutational analysis of acylphosphatase suggests the importance of topology and contact order in protein folding, *Nat. Struct. Biol. 6*, 1005−1009.
7. Alm, E., and Baker, D. (1999) Matching theory and experiment in protein folding, *Curr. Opin. Struct. Biol. 9*, 189−196.
8. Paci, E., Lindorff-Larsen, K., Dobson, C. M., Karplus, M., and Vendruscolo, M. (2005) Transition state contact orders correlate with protein folding rates, *J. Mol. Biol. 352*, 495−500.
9. Dyson, H. J., Rance, M., Houghton, R. A., and Lerner, R. (1998) Folding of immunogenic peptide fragments of proteins in water solution. I. Sequence requirements for the formation of reverse turns, *J. Mol. Biol. 201*, 161−200.
10. Klimov, D. K., and Thirumalai, D. (2002) Stiffness of the distal loop restricts the structural heterogeneity of the transition state ensemble in SH3 domains, *J. Mol. Biol. 317*, 721−737.

11. Nauli, S., Kuhlman, B., and Baker, D. (2001) Computer-based redesign of a protein folding pathway, *Nat. Struct. Biol. 8*, 602−605.
12. McCallister, E. L., Alm, E., and Baker, D. (2000) Critical role of β-hairpin formation in protein G folding, *Nat. Struct. Biol. 7*, 669−673.
13. Kim, J., Brych, S. R., Lee, J., Logan, T. M., and Blaber, M. (2003) Identification of a key structural element for protein folding within β-hairpin turns, *J. Mol. Biol. 328*, 951−961.
14. Nicholson, H., Soderlind, E., Tronrud, D. E., and Matthews, B. W. (1989) Contributions of left-handed helical residues to the structure and stability of bacteriophage T4 lysozyme, *J. Mol. Biol. 210*, 181−193.
15. Stites, W. E., Meeker, A. K., and Shortle, D. (1994) Evidence for strained interactions between side chains and the polypeptide backbone, *J. Mol. Biol. 235*, 27−32.
16. Gellman, S. H. (1998) Minimal model systems for β-sheet secondary structure in proteins, *Curr. Opin. Chem. Biol. 2*, 717−725.
17. Searle, M. S. (2001) Peptide models of protein β-sheets: Design, folding and insights into stabilising weak interactions, *J. Chem. Soc., Perkin Trans. 2*, 1011−1020.
18. Searle, M. S. (2004) Insights into stabilising weak interactions in designed peptide β-hairpins, *Biopolymers 76*, 185−195.
19. Cregut, D., Civera, C., Macias, M. J., Wallon, G., and Serrano, L. (1999) A tale of two secondary structure elements: When a β-hairpin becomes an α-helix, *J. Mol. Biol. 292*, 389−401.
20. Platt, G. W., Simpson, S. A., Layfield, R., and Searle, M. S. (2003) Stability and folding kinetics of a ubiquitin mutant with a strong propensity for non-native β-hairpin conformation in the unfolded state, *Biochemistry 42*, 13762−13771.
21. Prieto, J., Wilmans, M., Jimenez, M. A., Rico, M., and Serrano, L. (1997) Non-native local interactions in protein folding and stability: Introducing an helical tendency in the all β-sheet α-spectrin SH3 domain, *J. Mol. Biol. 268*, 760−778.
22. Sibanda, B. L., Blundell, T. L., and Thornton, J. M. (1989) Conformation of β-hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering, *J. Mol. Biol. 316*, 759−777.
23. Cox, J. P. L., Evans, P. A., Packman, L. C., Williams, D. H., and Woolfson, D. N. (1993) Dissecting the structure of a partially folded protein. CD and NMR studies of peptides from ubiquitin, *J. Mol. Biol. 234*, 483−492.
24. Zerrela, R., Evans, P. A., Ioniodes, J. M. C., Packman, L. C., Trotter, B. W., Mackay, J. P., and Williams, D. H. (1999) Autonomous folding of a peptide corresponding to the N-terminal β-hairpin of ubiquitin, *Protein Sci. 8*, 1320−1331.
25. Searle, M. S., Williams, D. H., and Packman, L. C. (1995) A short linear peptide derived from the N-terminal sequence of ubiquitin folds into a water-stable non-native β-hairpin, *Nat. Struct. Biol. 2*, 999−1006.
26. Jourdan, M., Griffiths-Jones, S. R., and Searle, M. S. (2000) Folding of a β-hairpin peptide derived from the N-terminus of ubiquitin, *Eur. J. Biochem. 267*, 3539−3548.
27. Sosnick, T. R., Dothager, R. S., and Krantz, B. A. (2004) Differences in the folding transition state of ubiquitin indicated by ϕ and ψ analysis, *Proc. Natl. Acad. Sci. U.S.A. 101*, 17377−17382.
28. Krantz, B. A., Dothager, R. S., and Sosnick, T. R. (2004) Discerning the structure and energy of multiple transition states in protein folding using ψ-analysis, *J. Mol. Biol. 337*, 463−475.
29. Went, H. M., and Jackson, S. E. (2004) Ubiquitin folds through a highly polarized transition state, *Protein Eng. 18*, 239−246.
30. Crespo, M. D., Platt, G. W., Bofill, R., and Searle, M. S. (2004) Context-dependent effects of proline residues on the stability and folding of ubiquitin, *Eur. J. Biochem. 271*, 4474−4484.
31. Mariamayagam, N., and Jackson, S. E. (2004) The folding pathway of ubiquitin from all-atom molecular dynamics simulations, *Biophys. Chem. 111*, 159−171.
32. Fersht, A. R., Itzahki, L. S., ElMasry, N. F., Matthews, J. M., and Otzen, D. E. (1994) Single versus parallel pathways of protein folding and fractional formation of structure in the transition state, *Proc. Natl. Acad. Sci. U.S.A. 91*, 10426−10429.
33. Fersht, A. R., and Sato, S. (2004) Φ-value analysis and the nature of protein-folding transition states, *Proc. Natl. Acad. Sci. U.S.A. 101*, 7976−7981.
34. Gianni, S., Guydosh, N. R., Khan, F., Caldas, T. D., Mayor, U., White, G. W. N., DeMarco, M. L., Daggett, V., and Fersht, A. R. (2003) Unifying features in protein-folding mechanisms, *Proc. Natl. Acad. Sci. U.S.A. 100*, 13286−13291.
35. Bofill, R., Simpson, E. R., Platt, G. W., Crespo, M. D., and Searle, M. S. (2005) Extending the folding nucleus of ubiquitin with an independently folding β-hairpin finger: Hurdles to rapid folding arising from the stabilisation of local interactions, *J. Mol. Biol. 349*, 205−221.
36. Northey, J. G. B., Maxwell, K. L., and Davidson, A. R. (2002) Protein folding kinetics beyond the Φ value: Using multiple amino acid substitutions to investigate the structure of the SH3 domain folding transition state, *J. Mol. Biol. 320*, 389−402.
37. Simpson, E. R., Meldrum, J. K., Bofill, R., Crespo, M. D., Holmes, E., and Searle, M. S. (2005) Engineering enhanced protein stability through β-turn optimisation: Insights for the design of stable peptide β-hairpin systems, *Angew. Chem., Int. Ed. 44*, 4939−4944.
38. Thiruv, B., Quon, G., Sadanha, S. A., and Steipe, B. (2005) Nh3D: A reference dataset of non-homologous protein structures, *BMC Struct. Biol. 5*, 12.
39. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) The Protein Data Bank: A computer-based archival file for macromolecular structures, *J. Mol. Biol. 112*, 523−542.
40. Frishman, D., and Argos, P. (1995) Knowledge-based protein secondary structure assignment, *Proteins: Struct., Funct., Genet. 23*, 566−579.
41. Kraulis, P. J. (1989) ANSIG: A program for the assignment of protein ^{1}H 2D NMR spectra by interactive computer graphics, *J. Magn. Reson. 84*, 627−633.
42. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures, *J. Mol. Graphics 14*, 51−55.
43. Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D., Vierstra, R. D., Hatfield, P. M., and Cook, W. J. (1987) Comparison of the three-dimensional structures of human, yeast and oat ubiquitin, *J. Biol. Chem. 262*, 6396−6399.
44. Pace, N. C., and Scholtz, J. M. (1997) in *Protein Structure: A Practical Approach* (Creighton, T. E., Ed.) 2nd ed., IRL Press, New York.
45. Ferguson, N., Capaldi, A. P., James, R., Kleanthous, C., and Radford, S. E. (1999) Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9, *J. Mol. Biol. 286*, 1597−1608.
46. Bachman, A., and Kiefhaber, T. (2005) Kinetic mechanisms for protein folding, in *Protein Folding Handbook, Part I* (Buchner, J., and Kiefhaber, T., Eds.) pp 279−410, Wiley-VCH, Weinheim, Germany.
47. Main, E. R. G., Fulton, K. F., and Jackson, S. E. (1999) Folding of FKBP12: Pathway of folding and characterisation of the transition state, *J. Mol. Biol. 291*, 429−444.
48. Griffiths-Jones, S. R., Maynard, A. J., and Searle, M. S. (1999) Dissecting the stability of a β-hairpin peptide that folds in water: NMR and molecular dynamics analysis of the β-turn and β-strand contributions to folding, *J. Mol. Biol. 292*, 1051−1069.
49. Ramirez-Alvarado, M., Blanco, F. J., Niemann, H., and Serrano, L. (1997) Role of β-turn residues in β-hairpin formation and stability in designed peptides, *J. Mol. Biol. 273*, 898−912.
50. Khorasanizadeh, S., Peters, I. D., and Roder, H. (1993) Folding and stability of a tryptophan mutant of ubiquitin, *Biochemistry 32*, 7054−7063.
51. Khorasanizadeh, S., Peters, I. D., and Roder, H. (1996) Evidence for a three state-model of protein folding from kinetic analysis of ubiquitin variants with altered core residues, *Nat. Struct. Biol. 3*, 193−205.
52. Krantz, B. A., and Sosnick, T. R. (2000) Distinguishing between two-state and three-state models for ubiquitin folding, *Biochemistry 39*, 11696−11701.
53. Went, H. M., Benitez-Cardoza, C. G., and Jackson, S. E. (2004) Is an intermediate state populated on the folding pathway of ubiquitin? *FEBS Lett. 567*, 333−338.
54. Maynard, A. J., Sharman, G. J., and Searle, M. S. (1998) Origin of β-hairpin stability in solution: Structural and thermodynamic analysis of the folding of a model peptide supports hydrophobic stabilisation in water, *J. Am. Chem. Soc. 120*, 1996−2007.
55. Griffiths-Jones, S. R., and Searle, M. S. (2000) Structure, folding and energetics of co-operative interactions between the β-strands

of a *de novo* designed three-stranded antiparallel β-sheet peptide, *J. Am. Chem. Soc. 122*, 8350−8356.
56. Leffler, J. E. (1953) Parameters for the description of transition states, *Science 117*, 340−341.
57. Kiefhaber, T., Sanchez, I. E., and Bachmann, A. (2005) Characterisation of protein folding barriers with rate-equilibrium free energy relationships, in *Protein Folding Handbook, Part I* (Buchner, J., and Kiefhaber, T., Eds.) pp 411−444, Wiley-VCH, Weinheim, Germany.
58. Fersht, A. R. (2004) Relationship of Leffler (Brønsted) α values and protein folding ϕ values to position of transition-state structures on reaction cordinates, *Proc. Natl. Acad. Sci. U.S.A. 101*, 14338−14342.
59. Sanchez, I. E., and Kiefhaber, T. (2003) Origin of unusual ϕ-values in protein folding: Evidence against specific nucleation sites, *J. Mol. Biol. 334*, 1077−1085.
60. Zhou, H. X., Hoess, R. H., and DeGrado, W. F. (1996) *In vitro* selection of thermodynamically stable turns, *Nat. Struct. Biol. 3*, 447−451.
61. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L. J., Dobson, C. M., and Schwalbe, H. (2002) Long-range interactions within a nonnative protein, *Science 295*, 1719−1722.
62. Baldwin, R. L. (2002) Making a network of hydrophobic clusters, *Science 295*, 1657−1658.
63. Friel, C. T., Beddard, G. S., and Radford, S. E. (2004) Switching two-state to three-state kinetics in the helical protein Im9 via the optimisation of stabilising non-native interactions by design, *J. Mol. Biol. 342*, 261−273.
64. Capaldi, A. P., Kleanthous, C., and Radford, S. E. (2002) Im7 folding mechanism: Mis-folding on a path to the native state, *Nat. Struct. Biol. 9*, 209−216.
65. Viguera, A.-R., and Serrano, L. (2001) Bergerac-SH3: "Frustration" induced by stabilising the folding nucleus, *J. Mol. Biol. 311*, 357−371.
66. Viguera, A.-R., and Serrano, L. (2003) Hydrogen-exchange stability analysis of Bergerac-Src homology 3 variants allows the characterisation of a folding intermediate in equilibrium, *Proc. Natl. Acad. Sci. U.S.A. 100*, 5730−5735.
67. Bofill, R., and Searle, M. S. (2005) Engineering stabilising β-sheet interactions into a conformationally flexible region of the folding transition state of ubiquitin, *J. Mol. Biol. 353*, 373−384.
68. Gu, H., Kim, D., and Baker, D. (1997) Contrasting roles for symmetrically disposed β-turns in the folding of a small protein, *J. Mol. Biol. 274*, 588−596.
69. Dyer, R. B., Maness, S. J., Peterson, E. S., Franzen, S., Fesinmeyer, R. M., and Andersen, N. H. (2004) The mechanism of β-hairpin formation, *Biochemistry 43*, 11560−11566.
70. Du, D., Zhu, Y., Huang, C.-Y., and Gai, F. (2004) Understanding the key factors that control the rate of β-hairpin folding, *Proc. Natl. Acad. Sci. U.S.A. 101*, 15915−15920.
71. Chen, R. P., Huang, J. J., Chen, H.-L., Jan, H., Velusamy, M., Lee, C., Fann, W., Larsen, R. W., and Chan, S. I. (2004) Measuring the refolding of β-sheets with different turn sequences on a nanosecond time scale, *Proc. Natl. Acad. Sci. U.S.A. 101*, 7305−7310.
72. Munoz, V., Thompson, P. A., Hofrichter, J., and Eaton, W. A. (1997) Folding dynamics and mechanism of β-hairpin formation, *Nature 390*, 196−199.
73. Du, D., Tucker, M. J., and Gai, F. (2006) Understanding the mechanism of β-hairpin folding via ϕ-value analysis, *Biochemistry 45*, 2668−2678.
74. Griffiths-Jones, S. R., Maynard, A. J., and Searle, M. S. (1999) Dissecting the stability of a β-hairpin peptide that folds in water: NMR and molecular dynamics analysis of the β-turn and β-strand contributions to folding, *J. Mol. Biol. 292*, 1051−1069.
75. Jourdan, M., and Searle, M. S. (2000) Cooperative assembly of a nativelike ubiquitin structure through peptide fragment complexation: Energetics of peptide association and folding, *Biochemistry 39*, 12355−12364.
76. Briggs, M. S., and Roder, H. (1992) Early hydrogen bonding events in the folding reaction of ubiquitin, *Proc. Natl. Acad. Sci. U.S.A. 89*, 2017−2021.
77. Gladwin, S. T., and Evans, P. A. (1996) Structure of the very early protein folding intermediates: New insights through a variant of hydrogen exchange labelling, *Folding Des. 1*, 407−417.

BI052495G