The Measure of Interior Disorder in a Folded Protein and Its Contribution to Stability

Y. Bruce Yu,[†] **Pierre Lavigne,**[†] **Peter L. Privalov,**[‡] **and Robert S. Hodges*,**[†]

Contribution from the Protein Engineering Network of Centres of Excellence, University of Alberta, Edmonton, Alberta T6G 2S2, Canada, and Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218

Received April 20, 1999

Abstract: To investigate whether any parameter other than packing density determines the quality of packing in a folded protein, the contribution of van der Waals interactions between hydrophobic core residues to protein stability at fixed packing density is investigated experimentally in this study. To this end, we employed a novel sequence variation scheme called "residue shuffling", defined as permutation of guest residues at equivalent host sites in a symmetric sequence frame. By comparing the stability of the analogues generated by permutation of hydrophobic core residues in a synthetic two-stranded α-helical coiled-coil scaffold, we conclude that the number of permissible rotamers, based on the avoidance of steric clashes, is another packing parameter. Rotamer number measures the degree of disorder in the hydrophobic core of a folded protein and complements packing density in evaluating packing free energy by gauging packing entropy, the dynamical aspect of packing.

Introduction

The formation of a well-packed interior core is important for the folding of both natural and designed proteins.[1−3] The evaluation of packing quality has so far been framed in terms of packing density, defined as a dimensionless ratio of the actual volume of an object to the volume of space that it occupies.[4] Over the years, however, studies have shown that packing density cannot be the sole parameter of packing. First, native proteins often do not have the highest possible packing densities,[1,5] and second, an increase in packing density can be destabilizing.[6] In this study, we set out to define another parameter of packing.

The criterion for a packing parameter is that it depends on no characteristics other than size and shape of immediate neighboring residues.[1] An unambiguous demonstration of the existence of a packing parameter other than density demands that the following two standards must be satisfied: first, the observed experimental effect has to arise from the packing of interior side chains in the folded state but not other factors, such as hydrophobicity, helix propensity, hydration, or changes in the unfolded state; second, the experiments have to be conducted under fixed packing density. The conventional site-specific substitution approach is not up to this task because each substitution brings with it multiple effects in both the folded and the unfolded states. Here, we adopt a sequence variation scheme which permutes guest residues in a symmetric sequence frame with equivalent host sites. We call this scheme "residue shuffling", the prototype of which was first introduced by us in the study of electrostatic interactions in coiled-coils.[7] Permutation of guest residues at equivalent host sites has the effect that factors other than interactions between the guest residues are conserved among the analogues. Mathematically, symmetry is defined as invariance under transformation,[8] and, for a lattice system such as a polypeptide sequence, the invariance is a translational one, meaning that the host sites have identical neighboring residue pairs. This identity in neighboring residue pairs makes the host sites equivalent. A homopolymer (for example, aaaaa_aaaa_aaaa) and, to a lesser extent, a palindrome (for example, abc_deffed_cba) are exemplary sequences with translational invariance. Translational invariance of the frame sequence necessitates that the interaction to be measured be introduced through a second chain with identical sequence frame. This also makes the unfolded states of the analogues identical, the reason being that differences among the analogues arise predominantly from interchain interactions which are negligible in the unfolded state due to wide separation of the two chains. Consequently, the very difficult issue of the unfolded state can be circumvented.

The two-stranded α-helical coiled-coil conformation[9] provides an ideal system to implement residue shuffling from both sequential and structural considerations. Sequentially, a peptide chain capable of forming a coiled-coil is made of repetitive seven-residue units called heptads.[10,11] This built-in sequential periodicity makes the translational invariance a natural one, because as long as each heptad satisfies the invariance requirement, the whole sequence will automatically satisfy it. However,

* To whom correspondence should be addressed. Tel.: 780-492-6540. Fax: 780-492-1473. E-mail: robert.hodges@ualberta.ca.

† University of Alberta.

‡ The Johns Hopkins University.

(1) Richards, F. M.; Lim, W. A. *Q. Rev. Biophys.* **1994**, *26*, 423−498.

(2) Levitt, M.; Gerstein, M.; Huang, E.; Subbiah, S.; Tsai, J. *Annu. Rev. Biochem.* **1997**, *66*, 549−579.

(3) Dahiyat, B. I.; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10172−10177.

(4) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151−176.

(5) Ponder, J. W.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 775−791.

(6) Karpusas, S.; Baase, W. A.; Matsumura, M.; Matthews, B. W. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8237−8241.

(7) Yu, Y.; Monera, O. D.; Hodges, R. S.; Privalov, P. L. *Biophys. Chem.* **1996**, *59*, 299−314.

(8) Hamermesh, M. *Group Theory and Its Application to Physical Problems*; Addison-Wesley: Reading, MA, 1962.

(9) Crick, F. H. C. *Acta Crystallogr.* **1953**, *6*, 689−697.

(10) Hodges, R. S.; Sodek, J.; Smillie, L. B.; Jurasek, L. *Cold Spring Harbor Symp. Quant. Biol.* **1972**, *37*, 299−310.
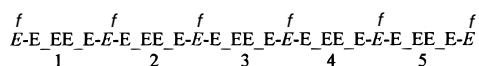
(11) McLachlan, A. D.; Stewart, M. *J. Mol. Biol.* **1975**, *98*, 293−304.

this sequential periodicity also brings with it a complication into the symmetry consideration, as will be illustrated later. Structurally, coiled-coils, in contrast to globular proteins, have two unique features that make them suited for residue shuffling. One is that there are no long-range intrachain or interchain residue contacts.[11] The other is that structural relaxation in coiled-coils is often amplified to a switch in global fold which can be easily tipped from two-stranded to three- or four-stranded by making small changes in the hydrophobic core.[12,13] This fragility in structural specificity provides a selection mechanism for keeping potential structural relaxation local. The extra benefit of employing the coiled-coil motif is that its ubiquity in protein structures[14] means that results obtained from this motif will have direct relevance to many biological problems.[15]

Sequence Design

The various positions in a coiled-coil heptad are conventionally labeled as *a*, *b*, *c*, *d*, *e*, *f*, and *g*, with *a* and *d* occupied largely by nonpolar residues which form the interhelical hydrophobic core.[11,16] In this study, our focus resides on the interactions in the hydrophobic core, which is formed by four residues at positions a_i, a_i', d_{i+3}, and d_{i+3}' (prime refers to residues on the second chain) in a two-stranded coiled-coil. These four residues form an ideal set for probing packing interactions because they form a hydrophobic cluster with extensive van der Waals contacts between the side chains.[17] Here enters a complication, which is that the *a* and *d* sites have different phases in terms of the sequential periodicity (they are $12\pi/7$ radians apart) and, therefore, are nonequivalent, regardless of the frame sequence. Structurally, this inherent asymmetry is reflected in the differences in the orientation of the $C^\alpha-C^\beta$ bond (parallel vs perpendicular)[12] and exposed surface areas of guest residues at *a* and *d*.

The guest residues are chosen as Val and Leu, two of the most common hydrophobic core residues in natural coiled-coils.[16] The length of the frame is five heptads. The sequence with perfect translational invariant symmetry is the following homopolymer, with the *a* and *d* positions in each heptad left blank for guest residues (numbers underneath the sequence indicate the heptad number):

$$\begin{array}{ccccccc} f & f & f & f & f & f \\ E\text{-E\_EE\_E-}E\text{-E\_EE\_E-}E\text{-E\_EE\_E-}E\text{-E\_EE\_E-}E\text{-E\_EE\_E-}E \\ & 1 & 2 & 3 & 4 & 5 \end{array}$$

However, practical considerations of solubility and concentration determination require modification of the sequence (Figure 1A). This is accomplished by replacing Glu by Lys for solubility at all the *f* positions (in *italic* in the above sequence) except the two termini, where Glu is replaced by Tyr for concentration determination by UV absorption spectroscopy. The *f* position is chosen for these replacements because it is equidistant from the two guest positions and, therefore, will cause the least disturbance to symmetry (the resultant frame is palindromic).

Although there are five heptads, the shuffling of guest residues should not take place in all of them for two reasons. First is the end effect, which cannot be neglected in any finite system. Thus,

the terminal heptads should remain unchanged among the peptides. The *a* and *d* positions of the N- and C-terminal heptads are filled as C_aL_d and L_aV_d, respectively (Figure 1A). Cys is introduced so that a disulfide bond can be formed between the two chains to keep them in-register and to maintain any heterodimer as one molecular species. The disulfide bond also simplifies the unfolding reaction by eliminating concentration dependency. End fraying[18] and the asymmetry between the two ends caused by the disulfide bond will be subtracted out when the peptides are compared since they are the same in every peptide. Second, shuffling should be compartmentalized to avoid cross interactions between guest residues, which means that permutation cannot take place in adjacent heptads simultaneously. This compartmentalization is accomplished by filling the *a* and *d* positions of the central heptad by Lys, which is chosen over Leu to ensure that all the coiled-coils are two-stranded.[19] Filling the central heptad with K rather than L creates an asymmetry between the two permuting heptads in the sense that the two *a* (*d*) sites are not equivalent because the hydrophobe preceding a_{10} (succeeding d_{13}) is a Leu (Lys), while that preceding a_{24} (succeeding d_{27}) is a Lys (Leu).

Host−Guest Indexes

The asymmetry between the *a* and *d* sites within one heptad and the asymmetry between the second and fourth heptads impose constraints on permutation. To quantify permutation and symmetry, we introduce a parameter called the host−guest index. A host−guest index is a non-negative integer, defined as the occupancy number of a particular type of host site (such as *a* in the second heptad) by a particular kind of guest residue (such as L). The number of distinct host−guest pairs is the number of indexes a peptide will have. The essence of residue shuffling is to preserve the host−guest indexes among the analogues it generates. Analogues with identical indexes have the same intrinsic properties belonging to a guest residue (such as side chain volume), a host site (such as orientation of the $C^\alpha-C^\beta$ bond), or a host−guest pair (such as exposed surface areas and their hydration). Consequently, differences between these analogues arise solely from interactions between the guest residues.

Four peptides, αα-36, ββ-36, αβ-36, and γδ-36, are generated by shuffling V and L in the second and fourth heptads in identical fashion to double the measurable effect. Their sequences are given in Figure 1A. Table 1 lists the host−guest indexes of the four peptides. Only αβ-36 and γδ-36 have identical indexes. αα-36 and ββ-36 are constructed in such a way that, although individually they have different host−guest indexes, their average, $^1/_2$(αα-36 + ββ-36), has the same host−guest indexes as αβ-36 and γδ-36 (Table 1). Thus, αα-36 and ββ-36 serve as controls to illustrate the consequences of both changing and preserving host−guest indexes.

Materials and Methods

Peptide Synthesis. All peptides were synthesized on MBHA resin using t-Boc chemistry and purified by reversed-phase HPLC, using standard protocols described as before.[18,20] The interchain disulfide bond of each homodimer was formed by air oxidation of the monomers in 100 mM NH_4HCO_3 at pH 8.0, followed by a purification step. For the heterodimers, the dimerization was carried out using the DNTP

(12) Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. *Science* **1993**, *262*, 1401−1407.
(13) Wagschal, K.; Tripet, B.; Hodges, R. S. *J. Mol. Biol.* **1999**, *285*, 785−803.
(14) Cohen, C.; Parry, D. A. D. *Trends Biochem. Sci.* **1986**, *11*, 245−248.
(15) Hodges, R. S. *Biochem. Cell. Biol.* **1996**, *74*, 133−154.
(16) Conway, J. F.; Parry, D. A. D. *Int. J. Biol. Macromol.* **1990**, *12*, 328−334.
(17) O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. *Science* **1991**, *254*, 539−544.
(18) Zhou, N. E.; Kay, C. M.; Hodges, R. S. *Biochemistry* **1992**, *31*, 5739−5746.
(19) Gonzalez, L.; Woofson, D. N.; Alber, T. *Nat. Struct. Biol.* **1996**, *3*, 1011−1018.
(20) Hodges, R. S.; Semchuk, P. D.; Taneja, A. K.; Kay, C. M.; Parker, J. M. R.; Mant, C. T. *Pept. Res.* **1988**, *1*, 19−30.
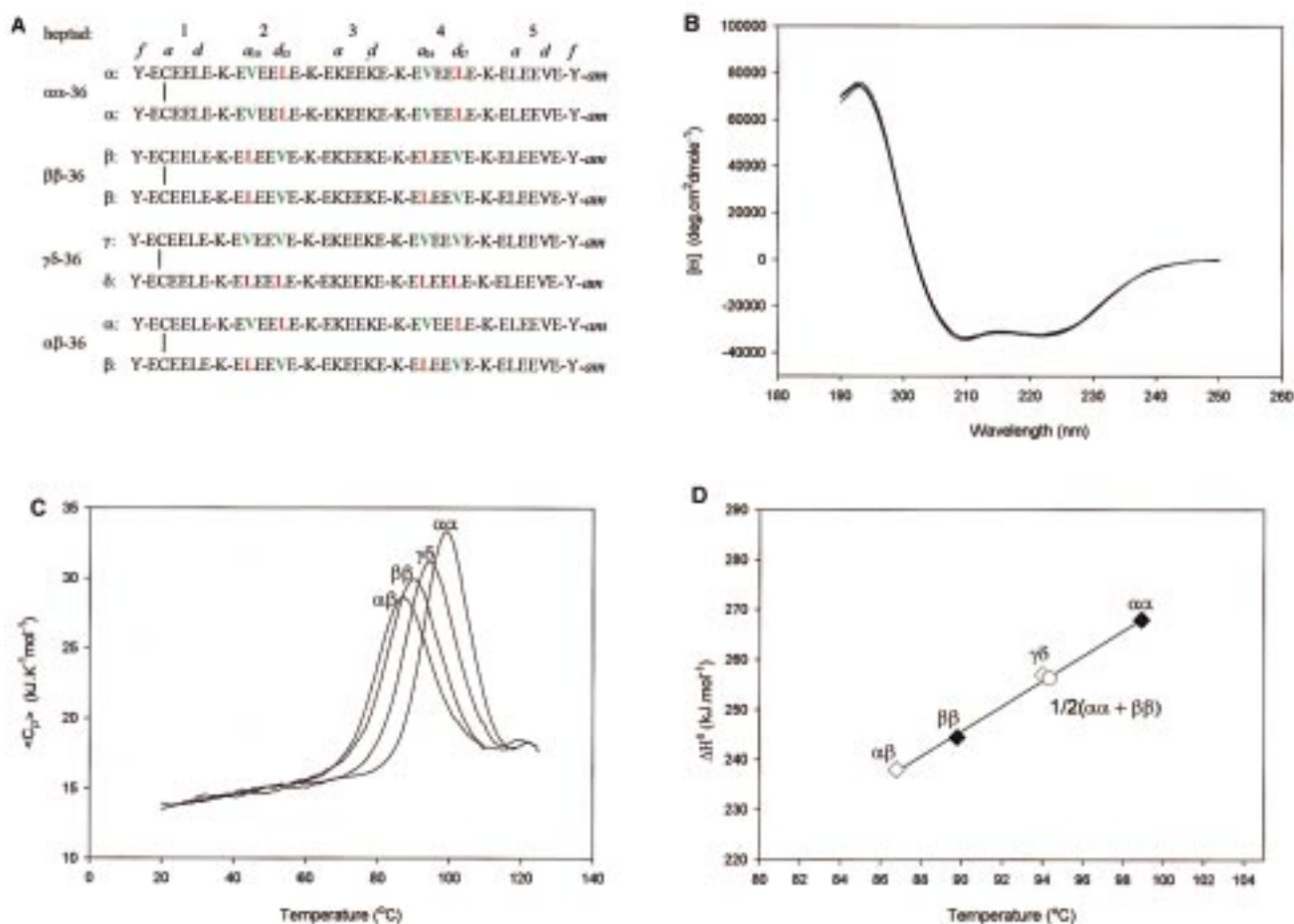
Figure 1. Properties of the coiled-coils. (A) Sequences of peptides generated by residue shuffling. Each peptide is made of two chains of 36 residues, cross-linked by a disulfide bond. There are four different individual chains, denoted as α, β, γ, and δ. The N-terminus of each chain is free, while the C-terminus is amidated. The guest residues at *a* and *d* positions are color-coded, with L in red and V in green. (B) Circular dichroism spectra of the four peptides at 25 °C. They superimpose each other. (C) Partial molar heat capacities of the four peptides. (D) Unfolding enthalpy of the peptides versus transition temperature. Diamonds represent actual experimental peptides, while the circle represents the hypothetical peptide $^1/_2(\alpha\alpha$-36 $+ \beta\beta$-36), which is the average of $\alpha\alpha$-36 and $\beta\beta$-36. Open symbols are for peptides with identical host−guest indexes, and solid symbols are for peptides with differing host−guest indexes.

derivatization procedure.[21] The reaction product was further purified to obtain the heterodimer, the purity of which was verified by analytical HPLC. The retention times of $\alpha\alpha$-36, $\beta\beta$-36, and $\alpha\beta$-36 are 33.2, 34.5, and 35.9 min, respectively, at 25 °C and 1% min^{-1} eluent B (0.05% trifluoroacetic acid in acetonitrile) gradient and a 0.4 mL min^{-1} flow rate, using a ZORBAX 300-SB C$_8$ column (i.d. 2.1 × 150 mm). Therefore, the heterodimer is well separated from the parental homodimers, even though they are isomers. For peptide $\gamma\delta$-36, the heterodimer and the parental homodimers have different molecular weights, and therefore, no possibility arises to confuse the heterodimer with the homodimers. Authenticity of the peptides was verified by mass spectrometry at both the monomer and the dimer stages as well as during synthesis through test cleavages. All final products were within 2 Da of the calculated mass (9096 Da).

Analytical Ultracentrifugation (AU). Sedimentation equilibrium studies were performed at 25 °C on a Beckman XLI analytical ultracentrifuge equipped with Rayleigh interference optics. Each sample was loaded at three different concentrations using a six-sector CFE cell and run at two speeds: 30K and 34K rpm. The radial equilibrium concentrations were analyzed using the program NONLIN.[22] The entire concentration ranges are 0.25−5.14 mg/mL for $\alpha\alpha$-36, 0.22−3.40 mg/mL for $\beta\beta$-36, 0.27−5.78 mg/mL for $\alpha\beta$-36, and 0.20−3.06 mg/mL

for $\gamma\delta$-36. The data were best described by a single-species model with the square root of variance in the vicinity of 0.02. Fitting into self-association models increases the square root of variance.

Circular Dichroism (CD) Spectroscopy. The ellipticity of the peptides was measured on a Jasco 500C spectropolarimeter at 25 °C, using a LAUDA RM6 circulating water bath to maintain temperature. The spectropolarimeter was calibrated using camphorsulfonate-d_{10}. The concentrations of the peptides were in the range of 0.6−0.7 mg/mL.

Differential Scanning Calorimetry (DSC). Calorimetric measurements were carried out on the Nano II microcalorimeter manufactured by Calorimetric Science Corp. (CSC). All measurements were conducted at a scanning rate of 2 °C·min^{-1} under 3.5 atm excess pressure from 0 to 125 °C. The peptide concentrations were about 0.5−5 mg/mL for all peptides except $\beta\beta$-36, which was 0.5−2 mg/mL. Within this concentration range, transition temperature T_t shows no concentration dependency. The partial molar heat capacity of the peptides was determined as described by Privalov and Potekhin,[23] using a value of 0.73 mL·g^{-1} for the partial specific volumes of all peptides, as calculated from their amino acid compositions.[24] The same value of partial specific volume was also used in processing the ultracentrifuge data. Reversibility is over 95% after heating to 125 °C, judged by reappearance of the transition peak upon rescan. Proof of reversibility was also provided by CD spectroscopy, which showed the regaining

(21) Rabanal, F.; DeGrado, W. F.; Dutton, P. L. *Tetrahedron Lett.* **1996**, *37*, 1347−1350.

(22) Johnson, M. L.; Correia, J. J.; Yphantis, D. A.; Halvorson, H. R. *Biophys. J.* **1981**, *36*, 575−588.

(23) Privalov, P. L.; Potekhin, S. A. *Methods Enzymol.* **1986**, *131*, 4−51.
(24) Makhatadze, G. I.; Medvedkin, V. N.; Privalov, P. L. *Biopolymers* **1991**, *30*, 1001−1010.

Table 1. Host−Guest Indexes, Rotamer Numbers, Calculated Rotamer Entropy, and Measured Thermodynamic Quantities of Peptides

| | host−guest indexes[b] | | | | | | | | | $-TR\ln(\omega/\omega_0)$ | $-T\delta\Delta S^0$ | $\delta\Delta H^0$ | $\delta\Delta G^0$ |
peptide[a]	$V_a(2)$	$V_d(2)$	$L_a(2)$	$L_d(2)$	$V_a(4)$	$V_d(4)$	$L_a(4)$	$L_d(4)$	rotamer number[c] (ω)	(calcd)[d]		(measd)[e]	
$\alpha\alpha$-36	2	0	0	2	2	0	0	2	$4 \times 4 = 16$				
$\beta\beta$-36	0	2	2	0	0	2	2	0	$16 \times 16 = 256$				
$^1/_2(\alpha\alpha$-36 $+ \beta\beta$-36)	1	1	1	1	1	1	1	1	$[16 \times 256]^{1/2} = 64$	0	−0.9	0.9	0.1
$\gamma\delta$-36	1	1	1	1	1	1	1	1	$8 \times 8 = 64$	f	f	f	f
$\alpha\beta$-36	1	1	1	1	1	1	1	1	$4 \times 4 = 16$	4.2	5.1	0.1	5.2

^{a} $^1/_2(\alpha\alpha$-36 $+ \beta\beta$-36) refers to the hypothetical peptide whose properties are averages of $\alpha\alpha$-36 and $\beta\beta$-36. In particular, its thermodynamic quantities and host−guest indexes are the arithmetic averages, while its rotamer number is the geometric average of $\alpha\alpha$-36 and $\beta\beta$-36. ^{b} A host−guest index is the occupancy number of a particular type of host site by a particular kind of guest residue. Since there are two types of host positions in each heptad (a and d), two nonequivalent heptads (second and fourth), and two kinds of guest residues being permuted (V and L), there are $2 \times 2 \times 2 = 8$ distinct guest−host pairs. For example, $V_a(2)$ refers to the pair that is formed by V at the a site in the second heptad. Host−guest indexes are used to quantify permutation and symmetry. Note that the hypothetical peptide $^1/_2(\alpha\alpha$-36 $+ \beta\beta$-36) is different from an actual hybrid of $\alpha\alpha$-36 and $\beta\beta$-36, whose second (fourth) heptad is the same as $\alpha\alpha$-36 and whose fourth (second) heptad is the same as $\beta\beta$-36. Such a hybrid will have host−guest indexes (2, 0, 0, 2, 0, 2, 2, 0) or (0, 2, 2, 0, 2, 0, 0, 2), which is the *hybrid* of the indexes of $\alpha\alpha$-36 and $\beta\beta$-36. In contrast, the host−guest indexes of the hypothetical $^1/_2(\alpha\alpha$-36 $+ \beta\beta$-36), (1, 1, 1, 1, 1, 1, 1, 1), is the *average* of the indexes of $\alpha\alpha$-36 and $\beta\beta$-36. This difference between the average and the hybrids of $\alpha\alpha$-36 and $\beta\beta$-36 will disappear if the second and fourth heptads are equivalent because the current eight distinct indexes will degenerate into four. Then, the hypothetical average and the two actual hybrids will have the same host−guest indexes, (2, 2, 2, 2). The asymmetry between the second and fourth heptads is the result of putting Lys, instead of Leu, in the third heptad in order to keep all the coiled-coils two-stranded. This asymmetry splits the degeneracy of host−guest indexes and reduces the number of permutations that preserve these indexes, which is the order of the symmetry group. ^{c} Side-chain conformational flexibility in the folded peptide is addressed by counting the rotamer number, using the rotational isomeric state approximation.[32] Three possibilities are considered for each side-chain χ angle: *gauche$^+$* (g^+), *trans* (t), and *gauche$^-$* (g^-). Therefore, there are three possible rotamers for V and nine for L. Main-chain (including C$^\beta$) overlap check (to avoid steric clashes) reduces rotamers of V at both a and d positions to just one possibility, t. The same overlap check reduces rotamers of L at a position to four possibilities, (t, t), (t, g^+), (t, g^-), and (g^+, t), and that at d position to two possibilities, (g^+, g^-) and (g^+, t). Clashes between guest side chains happen only in peptide $\alpha\beta$-36 between the two Leu residues in each heptad and reduce the rotamers at a to two possibilities, (t, t) and (g^+, t). The rotamer number of each heptad is the product of the rotamer number of each residue minus the number of clashes. The total rotamer number of each peptide, ω, is the product of the rotamer numbers of the two permuted heptads. Only rotamers of the shuffled residues are counted. ^{d} ω ($=64$) refers to the rotamer number of the reference peptide $\gamma\delta$-36. $-TR\ln(\omega/\omega_0)$ gives the entropy effect due to the difference in rotamer numbers between peptide $^1/_2(\alpha\alpha$-36 $+ \beta\beta$-36) or $\alpha\beta$-36 and the reference peptide. Assigning equal probability to different permissible rotamer combinations is, of course, an approximation. At $T = 366$ K (93 °C), where this approximation takes place, the structures are highly fluctuating (between the folded and the unfolded states and between the substates of the folded form), and the energetic differences between different rotamer combinations ($\Delta\epsilon$) are "washed out" by the thermal energy (RT), i.e., $\Delta\epsilon < RT$. At lower temperatures, especially in the crystalline state, it is possible that the different rotamer combinations are not equally probable and that one of them stands out, which is the one observed by X-ray diffraction. In such a scenario, a detailed calculation involving explicit treatment of the van der Waals potential may be necessary.[33] However, such a calculation goes beyond simple geometric consideration. ^{e} The measured thermodynamic quantity, $\delta\Delta X^0$ ($X = S, H, G$) for a peptide ($^1/_2(\alpha\alpha$-36 $+ \beta\beta$-36) or $\alpha\beta$-36) is defined as $\delta\Delta X^0 = \Delta X^0(\gamma\delta$-36$) - \Delta X^0(\text{peptide}) = X^0_{\text{unfolded}}(\gamma\delta$-36$) - X^0_{\text{folded}}(\gamma\delta$-36$) - \{X^0_{\text{unfolded}}(\text{peptide}) - X^0_{\text{folded}}(\text{peptide})\} = X^0_{\text{folded}}(\text{peptide}) - X^0_{\text{folded}}(\gamma\delta$-36$)$. The second equality holds because the unfolded states of all the peptides are identical. Therefore, $\delta\Delta X^0$ measures the difference between the folded states. All thermodynamic values in this table are in kJ·mol^{-1}. ^{f} Reference.

of 100% ellipticity at low temperature (2 °C) by the peptides after being heated to 95 °C. Prolonged exposure to high temperature (such as repeated scans or incubation at 125 °C in the calorimeter cell) reduces reversibility. Apparently, high temperature causes slight degradation of the peptides, such as hydrolysis (as evidenced by the appearance of minor impurity peaks in analytical HPLC profiles after heating) and aggregation (as evidenced by increased optical scattering in the UV absorption spectra after heating). Rescan profiles show that the degradation mainly affects the posttransition heat capacity, particularly in the range of 115−125 °C, while leaving the transition peak (maximum peak position and peak area) basically intact. Consequently, the heat capacity change (ΔC_p) from individual scans is less reliable than that determined from the slope of the enthalpy−temperature plot. It is the ΔC_p value determined from the enthalpy−temperature plot that is used in subsequent calculations.

Solution Conditions and Concentration Determination. The buffer for CD and DSC measurements is H$_3$PO$_4$/H$_2$O, pH 2.0. No salt is added because salt will further stabilize the peptides,[25] the stabilities of which are already reaching the upper limit that can be measured by DSC. The buffer for AU is the same buffer used for CD and DSC measurements with an additional 100 mM NaCl, which is added to prevent nonideality during AU (AU runs with no NaCl have a high nonideality factor and are not amenable to analysis). Peptides are dialyzed extensively (48−60 h) prior to measurements. The concentrations of peptide samples were determined from their UV absorption spectra with light scattering corrected in 5−6 M gaunidinium hydrochloride at pH 6.5. The extinction coefficient used for all peptides is 5945 M^{-1}·cm^{-1} at 275 nm, calculated from their amino acid compositions.[26]

Structural Modeling. The models for the different peptides were generated by replacing the proper residues on the GCN4 core mutant Asn16Lys (1ZIK)[19] with the program InsightII (InsightII 97, MSI). All side-chain rotamer combinations of the guest residues were sampled manually. Steric clashes were monitored using the Biopolymer module of the InsightII suite. A particular rotamer combination is rejected if a steric clash cannot be removed by allowing a deviation of ±12° from the canonical rotamer values (180° for t, −60° for g^+, and 60° for g^-). For all the permissible rotamer combinations, no cavity was detected by a probe with a radius of 1.4 Å using the program GRASP.[27]

Results

The four analogues have, within experimental error, identical CD spectra, indicative of coiled-coil formation (Figure 1B). The stability of these analogues was measured by DSC. Figure 1C shows the partial molar heat capacities of all four peptides. Table 2 lists the thermodynamic parameters of unfolding transition of each peptide, including the transition temperature (T_t), enthalpy (ΔH^0), entropy (ΔS^0) and free energy (ΔG^0). All four analogues unfold in a two-state fashion because the van't Hoff enthalpy (ΔH^0_{vh}) deviates less than 5% from the calorimetric enthalpy (ΔH^0_{cal}).[23] Sedimentation equilibrium experiments show that all the analogues have a molecular weight close to the calculated value of the cross-linked two-stranded monomers (Figure 2 and Table 2), confirming that there is no oligomerization. The monomeric nature of the peptides is also confirmed by the lack of concentration dependency of T_t. The combined

(25) Yu, Y.; Monera, O. D.; Hodges, R. S.; Privalov, P. L. *J. Mol. Biol.* **1996**, *255*, 367−372.

(26) Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, *182*, 319−326.

(27) Nicholls, A.; Sharp, K. A.; Honig, B. *Proteins: Struct., Funct. Genet.* **1991**, *11*, 281−296.
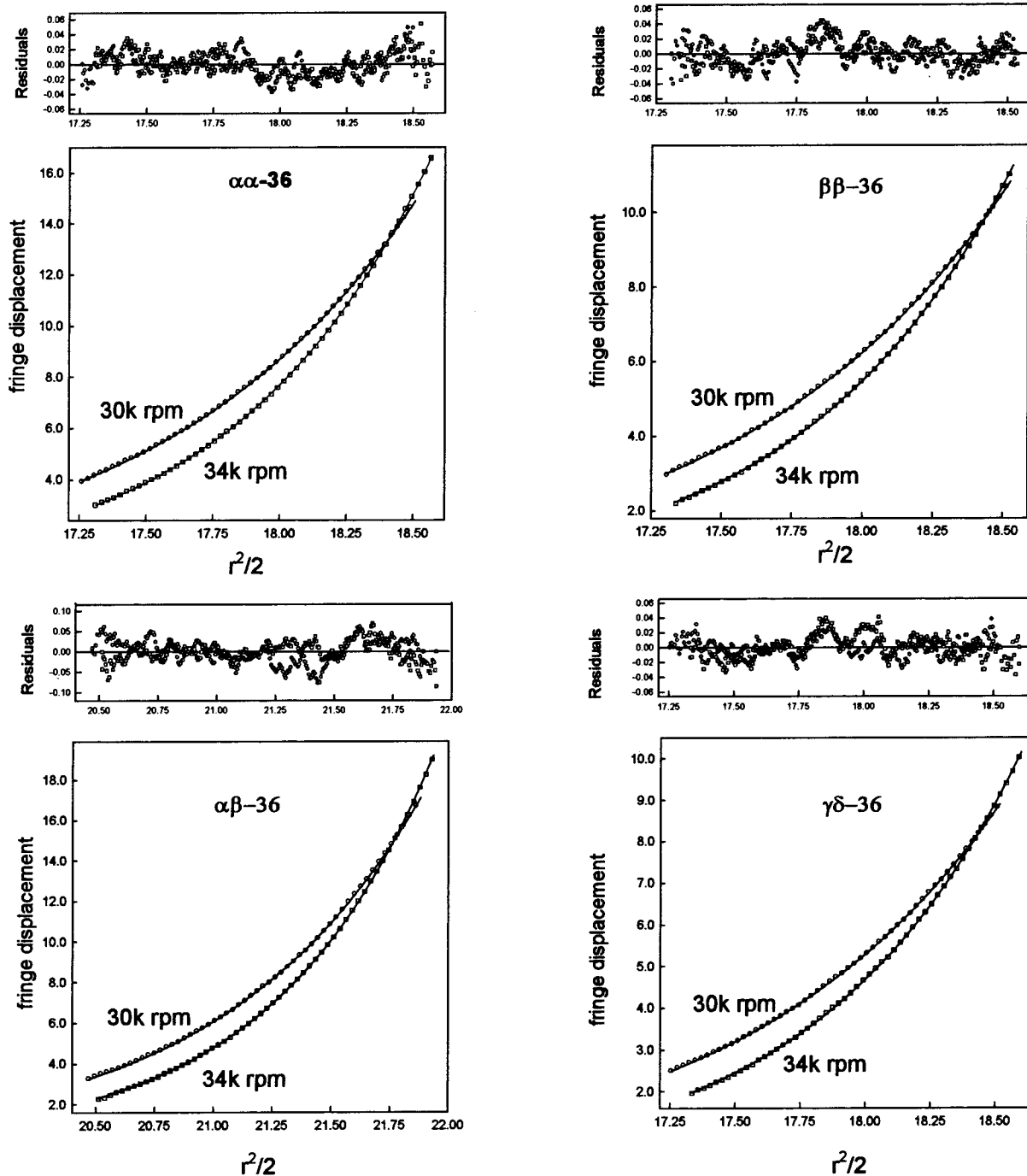
Figure 2. Sedimentation equilibrium analysis of the molecular weights of the peptides. Runs are performed at two speeds: 30K (○) and 34K rpm (□). The solid lines are fitting results of a single-species model.

results of AU, CD spectroscopy, and calorimetry confirm that all four analogues form two-stranded coiled-coils of identical helical content that unfold as single cooperative units. This enables a meaningful comparison of their unfolding energetics. Figure 1D plots ΔH^0 against T_t. The data essentially fall on a straight line, giving a ΔC_p of 2.5 kJ·K^{-1}·mol^{-1}. Thermodynamic quantities at a common temperature, 93 °C, are obtained through extrapolation using this ΔC_p value (Table 2). Here, we assume that the peptides have identical and constant ΔC_p, a reasonable assumption considering that the peptides have identical composition and structure and the extrapolation is within the temperature range where the ΔC_p value is obtained. Also listed in Table 2 is the unfolding thermodynamic quantities of the hypothetical peptide $^1/_2(\alpha\alpha\text{-36} + \beta\beta\text{-36})$ at 93 °C.

Packing density is invariant among all the peptides, as judged

by the total side-chain volumes of the packing core.[5] Among the three peptides with identical host−guest indexes, all the factors other than interactions between the packing core residues are also invariant. Therefore, if density is the only factor determining the energetic effects of packing, then all three of these peptides [$^1/_2(\alpha\alpha\text{-36} + \beta\beta\text{-36})$, $\alpha\beta\text{-36}$, and $\gamma\delta\text{-36}$] should have the same stability. While this is, indeed, the case for $^1/_2\text{-}(\alpha\alpha\text{-36} + \beta\beta\text{-36})$ and $\gamma\delta\text{-36}$, peptide $\alpha\beta\text{-36}$ has a stability much lower than those of the other two (Table 2 and Figure 1C). This lower stability must be caused by the other packing factor we are seeking. The data indicate that the nature of this factor is entirely entropic (Tables 1 and 2). Since all the peptides have identical unfolded states, the fact that $\alpha\beta\text{-36}$ has a larger unfolding entropy suggests that folded $\alpha\beta\text{-36}$ has lower entropy than folded $^1/_2(\alpha\alpha\text{-36} + \beta\beta\text{-36})$ and $\gamma\delta\text{-36}$ (see footnote *e* of

Table 2. Observed Molecular Weights (MW), Ellipticity (Θ_{222}), and Unfolding Thermodynamic Data

peptide	MW[a] (25 °C)	Θ_{222}[b] (25 °C)	T_t[c] (°C)	ΔH_{cal}^0 (T_t)	ΔH_{vh}^0 (T_t)	ΔH^0 (T_t)	ΔS^0 (T_t)	ΔH^0 (93 °C)	$-T\Delta S^0$ (93 °C)	ΔG^0 (93 °C)
$\alpha\alpha$-36	9 800	−32 670	98.9	262	274	268.0	720	253.1	−249.0	4.1
$\gamma\delta$-36	9 300	−32 550	94.0	260	254	257.0	699	253.7	−252.8	0.9
$\beta\beta$-36	9 700	−31 910	89.8	249	240	244.5	674	252.6	−254.8	−2.2
$\alpha\beta$-36	10 700	−32 640	86.8	244	232	238.0	661	253.6	−257.9	−4.3
$^1/_2(\alpha\alpha$-36 + $\beta\beta$-36)[d]			94.3					252.8	−251.9	0.9

[a] The observed molecular weights are obtained from sedimentation equilibrium analysis. The calculated molecular weight for all four peptides is 9100. [b] Θ_{222} is the ellipticity at 222 nm in deg·cm²·dmol⁻¹. [c] T_t is the transition temperature for the thermal unfolding process. It has an error of ±1 °C, judged by scans obtained at different peptide concentrations. ΔH_{cal}^0 is the calorimetric enthalpy change upon unfolding, while ΔH_{vh}^0 is the van't Hoff enthalpy change. ΔH^0 is the average of ΔH_{cal}^0 and ΔH_{vh}^0. This averaging process reduces experimental error in enthalpy. ΔS^0 is the entropy change, and ΔG^0 is the free energy change. $\Delta S^0(T_t) = \Delta H^0(T_t)/T_t$. $\Delta G^0 = \Delta H^0 - T\Delta S^0$. ΔH^0, $T\Delta S^0$, and ΔG^0 of each peptide are extrapolated to 93 °C using a ΔC_p (heat capacity change) value of 2.5 kJ·K⁻¹·mol⁻¹ (Figure 1D). Experimental error in ΔH^0 is about 0.5−1.0 kJ·mol⁻¹, estimated from the standard deviation of ΔH^0(93 °C) of the four synthetic peptides (which gives 0.5 kJ·mol⁻¹) and from difference in $\Delta H^0(T_t)$ from measurements of the same peptide at different concentrations (which gives 1.0 kJ·mol⁻¹). $T\Delta S^0$ has the same magnitude. However, ΔG^0 is hardly affected by the errors in ΔH^0 and $T\Delta S^0$ because they cancel each other (this is true only if the temperature extrapolation is not large, which is the reason why 93 °C, the middle of the transition temperature range, is chosen as the common temperature for comparison). In the vicinity of T_t, $\delta\Delta G^0$ is primarily determined by δT_t and, to a lesser extent, by ΔC_p. Both δT_t and ΔC_p are insensitive to errors in ΔH^0 of each individual scan. Therefore, the error in $\delta\Delta G^0$ should be well under 1.0 kJ·mol⁻¹. ΔH^0, $T\Delta S^0$, and ΔG^0 are in kJ·mol⁻¹, while ΔS^0 is in J·K⁻¹·mol⁻¹. [d] All thermodynamic properties of $^1/_2(\alpha\alpha$-36 and $\beta\beta$-36) are the averages of $\alpha\alpha$-36 and $\beta\beta$-36. See Table 1 for details.
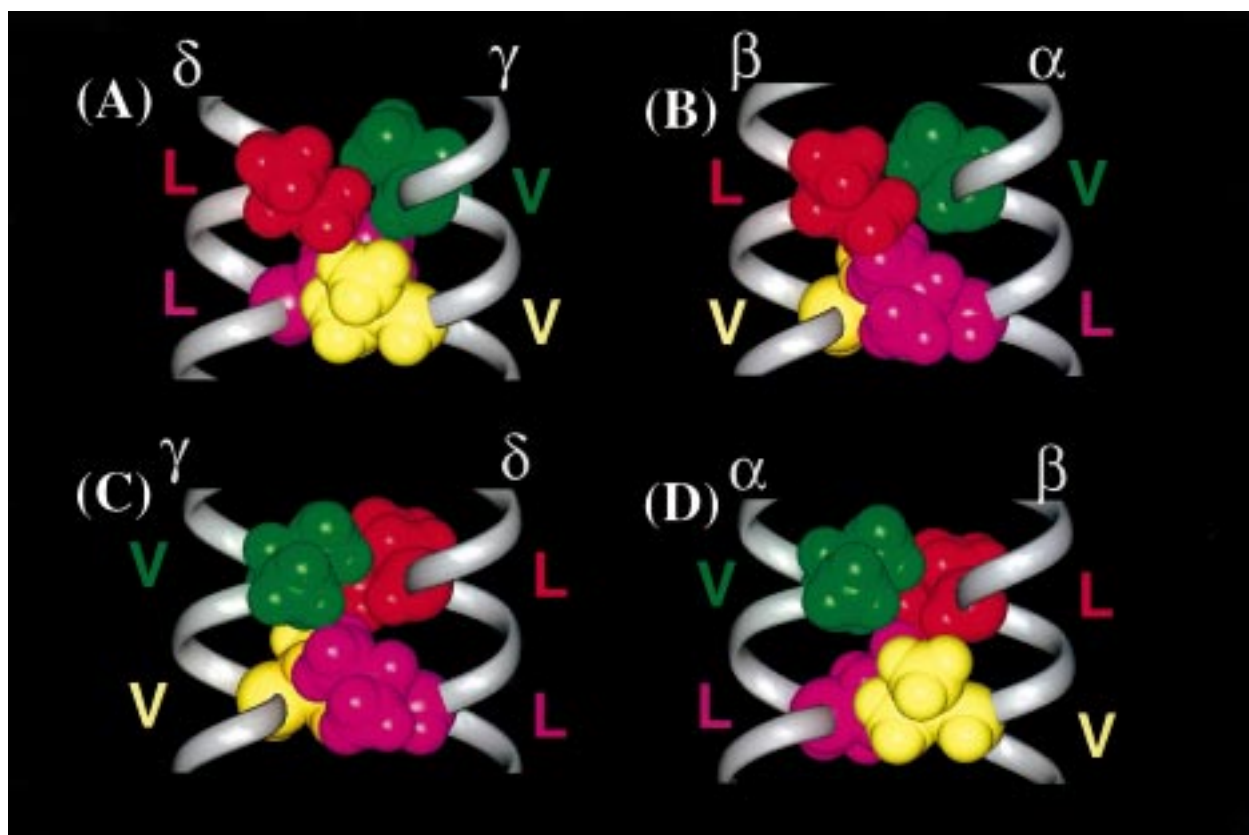

Figure 3. Permissible and forbidden rotamer combination. (A) and (B) are the front views of peptides $\gamma\delta$-36 and $\alpha\beta$-36, respectively, while (C) and (D) are the back views. V is shown in green (*a* site) and yellow (*d* site), and L is shown in red (*a* site) and purple (*d* site). The rotamer conformation of V is *t* at both *a* and *d,* while that for L is (*t, g⁻*) at *a* and (*g⁺, t*) at *d.* The two peptides have the same core residues but different arrangements. As a result, while there is no side-chain clash in $\gamma\delta$-36 (A, C), there is a clash between the two L side chains in $\alpha\beta$-36 (B). Therefore, this particular rotamer combination is permissible in $\gamma\delta$-36 but forbidden in $\alpha\beta$-36. This reduced permissible rotamer number leads to lowered entropy in the folded state of $\alpha\beta$-36 and makes it more ordered but less stable compared to $\gamma\delta$-36.

Table 1). The most plausible explanation is that the arrangement of guest residues in $\alpha\beta$-36 puts more restrictions on the mobility of their side chains. While direct experimental verification of this has to come from side-chain dynamics as measured by NMR spectroscopy,[28,29] counting based on avoidance of steric clashes shows that peptides $^1/_2(\alpha\alpha$-36 + $\beta\beta$-36) and $\gamma\delta$-36 have the

same number of permissible rotamers. On the other hand, $\alpha\beta$-36 has half as many permissible rotamers (Table 1), due to clashes between Leu side chains which are absent in the other two peptides. Figure 3 gives an example of such a rotamer combination that is permissible in $\gamma\delta$-36 but forbidden in $\alpha\beta$-36. Table 1 gives the calculated entropy due to this difference in rotamer numbers. Quantitatively, this rotamer entropy matches the experimentally determined entropy difference between $\alpha\beta$-36 and $\gamma\delta$-36 within 1 kJ·mol⁻¹ at 93 °C.

(28) Akke, M.; Bruschweiler, R.; Palmer, A. G. *J. Am. Chem. Soc.* **1993**, *115*, 9832−9833.

(29) Yang, D.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 369−382.

Discussion

A first principle statistical mechanical calculation of stability based on energy potentials will always produce a free energy containing both enthalpy and entropy. Geometric packing, as a substitute for complex calculations based on van der Waals potentials, must build enthalpy and entropy *a priori*. Packing density catches the enthalpic part of the van der Waals interactions but says nothing about the degeneracy of configurations under the same packing density, which leads to an entropic contribution. Rotamer number, as shown here, reveals this entropic part; it satisfies the criterion of a packing parameter because it depends only on the size, shape, and arrangements of neighboring residues. Packing density (p_d) provides a measure of order, while rotamer number (ω) provides a measure of disorder. The higher stability of peptide $\gamma\delta$-36 compared to that of $\alpha\beta$-36 provides a concrete example of how the global order of a protein, the folded conformation, can be stabilized by local disorder. A complete evaluation of packing quality should include both packing density (p_d) and rotamer number (ω). More precisely (see footnote *d* of Table 1 for a comment),

$$G^0(\text{packing}) = H^0(p_d,T) - TS^0(\omega) = H^0(p_d,T) - RT \ln(\omega) \tag{1}$$

While the actual calculation of G^0(packing) awaits the explicit functional form of $H^0(p_d,T)$, this expanded picture of packing explains the previous observation that, at wild-type packing density, which is lower than the maximum density, there are several permissible rotamer combinations for each sequence.[5] It also explains why maximizing packing density can be destabilizing.[6]

In a broader sense, our conclusion is about side-chain mobility in the folded state and the residual entropy accompanying this mobility. Our results show that mobility of internal side chains in the folded state can contribute significantly to its stability and that the energetic effect of this mobility can be adequately dealt with by counting permissible rotamer numbers. This conclusion essentially provides a way to extract dynamic information out of static structural data as far as stability is concerned and has direct bearing on both structure−energetic parametrization[30,31] and correlation of dynamics with entropy,[28,29] two areas currently under intense pursuit.

The most difficult issue in the study of protein stability is perhaps the context dependency of some of the factors involved and the nonadditivity of energetic parameters caused by this dependency, making the prediction of the stability of the proteins from the properties of individual components impossible. Such context dependency is illustrated here by the stability difference between $\alpha\beta$-36 and $\gamma\delta$-36. The concept of permissible rotamer numbers provides a means to explain and estimate context dependency in the case of packing.

The current work proves that residue shuffling can be a powerful tool in revealing the subtle interactions in protein structures. Crucial to its success is the principle that permutation of guest residues should take place only at equivalent sites created by symmetry, abstracted mathematically as the invariance of the host−guest index. This principle is illustrated perfectly by the dual role played by the control peptides $\alpha\alpha$-36 and $\beta\beta$-36. On the one hand, rotamer number alone fails to explain the stability order of $\alpha\alpha$-36 and $\beta\beta$-36. Permutation from $\alpha\alpha$-36 to $\beta\beta$-36 is not between equivalent host sites; therefore, the difference between them is contaminated by other factors, most likely hydration of apolar surfaces in the folded state.[30] On the other hand, whatever the contaminating factors, they are exactly averaged out in $^1/_2(\alpha\alpha\text{-}36 + \beta\beta\text{-}36)$ as a result of symmetry. Consequently, rotamer number succeeds in explaining quantitatively the stability of $^1/_2(\alpha\alpha\text{-}36 + \beta\beta\text{-}36)$ relative to $\alpha\beta$-36 and $\gamma\delta$-36, all of which have identical host−guest indexes. In fact, the effect of rotamer number is completely masked by the contaminating factors in the case of $\alpha\alpha$-36 *vs* $\beta\beta$-36. Clearly, permutation without symmetry can lead to erroneous conclusions. The symmetry group formed by the index-preserving permutations is a subgroup of the symmetric group made of all possible permutations.

(30) Makhatadze, G. I.; Privalov, P. L. *Adv. Protein Chem.* **1995**, *47*, 307−425.

(31) Luque, I.; Freire, E. *Methods Enzymol.* **1998**, *295*, 100−127.

(32) Flory, P. J. *Statistical mechanics of chain molecules*; John Wiley & Sons: New York, 1969; pp 55−61.

(33) Lee, C.; Subbiah, S. *J. Mol. Biol.* **1991**, *217*, 373−388.