Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 5317–5320, June 1991
Biophysics

α-Helix stabilization by natural and unnatural amino acids with alkyl side chains

(α-helix/protein folding/unnatural amino acids/side chain conformation/β-branching)

PINGCHIANG C. LYU, JOHN C. SHERMAN, AMY CHEN, AND NEVILLE R. KALLENBACH*

Department of Chemistry, New York University, New York, NY 10003

Communicated by Bruno H. Zimm, March 18, 1991 (received for review January 10, 1991)

ABSTRACT **Knowledge of the role of individual side chains in forming different secondary structures such as the α-helix would be useful for prediction of protein structure from sequence or *de novo* protein design. Experimental and theoretical studies on natural and synthetic peptides and proteins indicate that individual side chains differ in their helix-forming potential. Four aliphatic side chains occur in the standard complement of amino acids: alanine and leucine are helix stabilizing, whereas isoleucine and valine are weakly destabilizing. We have synthesized a series of helical peptides containing unnatural aliphatic side chains having two to four carbons to explore some of the factors involved in α-helix stabilization and the basis for selection of the natural set. We find that linear side chains with two, three, or four carbons are as strongly helix stabilizing as the single methyl in alanine and that all linear side chains are stronger helix promoters than leucine. In addition, a *t*-butyl side chain is significantly more helix destabilizing than the *sec*-butyl side chain of isoleucine, the isopropyl side chain of valine, or even the unrestricted side chain of glycine. These results provide experimental evidence that restriction in conformational freedom of a side chain imposed by α-helix formation is a major component of the role of a side chain in stabilizing helical structure.**

The folded structures of native proteins frequently produce characteristic secondary structural patterns, including α-helices, β-sheets, and turns. Statistical surveys of the frequencies with which various amino acids occur in each pattern suggest that each residue might have a distinct propensity for stabilizing these structures (1–3). Recent experiments on synthetic peptides that contain significant helical structure confirm this proposal for the α-helix. Several systems of model peptides containing various amino acid residues as guests in a natural "host" environment have made it possible to determine the propensities of individual amino acids for α-helix formation (4–15). Specific interactions between neighboring side chains such as ion pairing or hydrophobic clusters have been identified that modulate the intrinsic propensities of various amino acids to stabilize or destabilize helical structure. The extent of agreement among recent data from different peptide models indicates that in the absence of these side chain–side chain interactions, the intrinsic propensities are independent of context to a first approximation (10, 12). Thus a consistent quantitative description of the helix-forming potential of a given sequence of amino acids is now available. Attention is now being directed to establishing experimental propensities for β-sheets and turns, as well as to the detailed mechanisms by which a particular side chain affects the stability of the α-helix.

We have described a series of soluble modular peptides in which guest amino acid triplets are substituted in a host sequence consisting of glutamic acid and lysine side chains,

succinyl-Tyr-Ser-Glu_4-Lys_4-(Xaa_3)-Glu_4-Lys_4-NH_2,

to give Xaa_3 peptides, where Xaa denotes the guest amino acid. Tyrosine is included to provide a standard for concentration measurement, and the glutamic acid and lysine residue blocks are positioned at sites *i* and *i* + 4 along the chain allowing them to form a number of potential salt bridges (5, 11, 12, 16). The peptides are C-terminally blocked to eliminate a negative charge interaction with the helix dipole and succinylated at the N terminus to introduce one extra negative charge at this position to further stabilize the helix (17, 18). The results show very different helix-forming propensity among those amino acids with alkyl side chains (alanine > leucine > isoleucine > valine). The difference may come from the conformational constraint of side chains since the chemical properties of these aliphatic side chains are similar. To test this hypothesis, four peptides with unnatural amino acids were designed and synthesized. In addition to those peptides containing natural aliphatic amino acids, three sets of peptides are generated in which the structures of the side chains vary in ordered progressions. Fig. 1 illustrates the structural features of the side chains studied. In set I, the number of unbranched carbons in the side chains increases from one to four. In set II, the β-carbon of the side chain is substituted with one to three methyl groups. Set III consists of a series of isomeric four-carbon side chains with a complete array of alkyl branching.

MATERIALS AND METHODS

Peptide Synthesis and Purification. Peptides were synthesized by fluoren-9-ylmethoxycarbonyl (Fmoc) solid-phase chemistry [Fmoc-Tle (where Tle is *t*-leucine) from Calbiochem; Fmoc-Nva (where Nva is norvaline) and Fmoc-Abu (where Abu is α-aminobutyric acid) from Bachem; and Fmoc-Nle (where Nle is norleucine) from MilliGen/Biosearch] on a Biosearch 9600 peptide synthesizer as described (12), and purified by reverse-phase high performance liquid chromatography with a gradient of 0–30% acetonitrile in 0.1% trifluoroacetic acid (Deltapack 15-m C_{18} column). The correct primary molecular ion of each purified product was verified by fast-atom bombardment mass spectometry (FAB-MS, M-Scan, West Chester, PA).

Circular Dichroism (CD) Spectroscopy. CD spectra were recorded on an Aviv 60DS spectropolarimeter equipped with an HP model 89100A temperature controller. All measurements reported were carried out in 10 mM KF solution, using cuvettes with a 1-mm pathlength. Peptide concentrations were determined from the tyrosine absorbance at 275 nm in 6 M guanidine hydrochloride (19). Concentrations used in CD experiments were 10–45 μM except when the concentration

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Fmoc, fluoren-9-ylmethoxycarbonyl; Tle, *t*-leucine; Nva, norvaline; Abu, α-aminobutyric acid; Nle, norleucine.

*To whom reprint requests should be addressed.

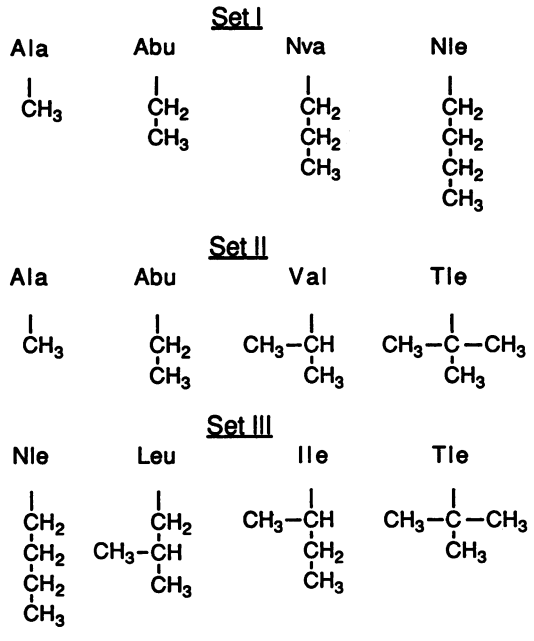

FIG. 1. Code and side chain structure of the amino acid substitutions of this study. Four peptides in the series have been synthesized, with the following residues corresponding to Xaa: norleucine, norvaline, α-aminobutyric acid, and *t*-leucine. In addition to the alanine, leucine, isoleucine, and valine peptides reported (12), these molecules generate the three sets of peptides listed, in which various features of the structure of the side chains vary progressively.

dependence was being determined. The CD spectra are dependent on both temperature and pH.

RESULTS AND DISCUSSION

N-Alkyl Side Chains Are Strongly α-Helix-Stabilizing. The secondary structure of these peptides is most readily characterized by CD spectroscopy; the peptide absorption bands in the UV become optically active on formation of regular secondary structures (20, 21). The helix content of each peptide was monitored by the value of the mean residue ellipticity at 222 nm ($[\theta]_{222}$). As for the other 10 members of the series we have studied, $[\theta]_{222}$ is independent of concentration from 4 to 350 μM peptide as shown in Fig. 2. This is consistent with the structure observed in the Nle_3, Abu_3, Nva_3, and Tle_3 peptides being intramolecular in origin. Fig. 3 contrasts the CD spectra of these peptides. Qualitatively, each of the straight-chain amino acids is very close to alanine in stabilizing helical structure and more stabilizing than leucine or other natural alkyl side chains. The highly branched Tle side chain on the other hand destabilizes helical structure more than glycine. To evaluate the free energy differences among various side chains, it is necessary to introduce a model of the helix–coil transition process (22). By using a multistate description for the helix–coil equilibrium in these peptides based on that of Zimm and Bragg (12), the free-energy differences between the various side chains shown in Table 1 were obtained. The resulting $\Delta\Delta G$ values are model-dependent: the numbers in Table 1 take into account the presence of ion pairs in stabilizing the helix in this system.

The near equivalence in the helix–coil transition energies in set I of these peptides demonstrates that helicity is independent of the chain length for unbranched *n*-alkyl side chains in isolated helices. The similar helical propensities of alanine and α-aminobutyric acid correlate well with the similarity of conformational constraints for these side chains (25). The increase in hydrophobicity with increasing chain length does not correlate with any change in helical propensity for these helices. This is in contrast to computational predictions that the α-helix provides the preferred conformation for hydrophobic residues (26) and the experimental data in which the helicity of alanine, norvaline, and norleucine in lysine copolymers has been examined (27, 28). We believe the explanation is that in both the helix and coil states of this family of peptides the exposure of a guest side chain to solvent is similar. This situation is comparable to sites in a globular protein that are exposed to solvent and not to internal sites where hydrophobic packing plays a crucial role. In contrast to these model peptides, polylysine can be thought of as creating a hydrophobic matrix, especially at high pH, which accommodates the bulkier norleucine and norvaline better than the small alanine side chain.

Effect of β-Branching. The difference in helix–coil transition energies within set II—Ala_3, Abu_3, Val_3, and Tle_3—shows that helicity is extremely sensitive to higher substitution of methyl groups on the β-carbon. Whereas substitution of a single methyl on the β-carbon of alanine to give α-aminobutyric acid causes no loss in helicity, bis substitution at the β-carbon to yield valine results in a substantial loss in helicity. This is consistent with the significant diminution in accessible conformational space calculated for valine relative to alanine (or α-aminobutyric acid) (25). Moreover, the small conformational space available to valine is mostly relegated to regions outside of the α-helix. Substitution of a third

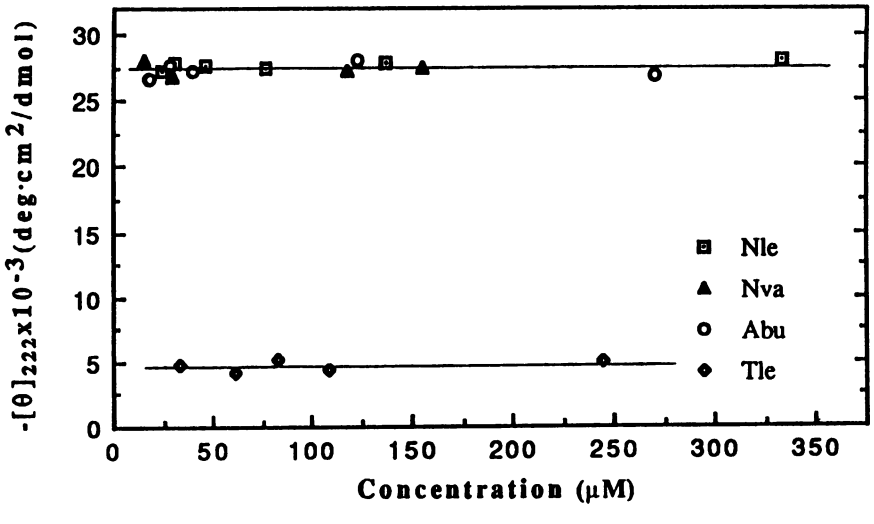

FIG. 2. Concentration dependence of the following peptides: Nle_3, Nva_3, Abu_3, and Tle_3. deg, Degree.

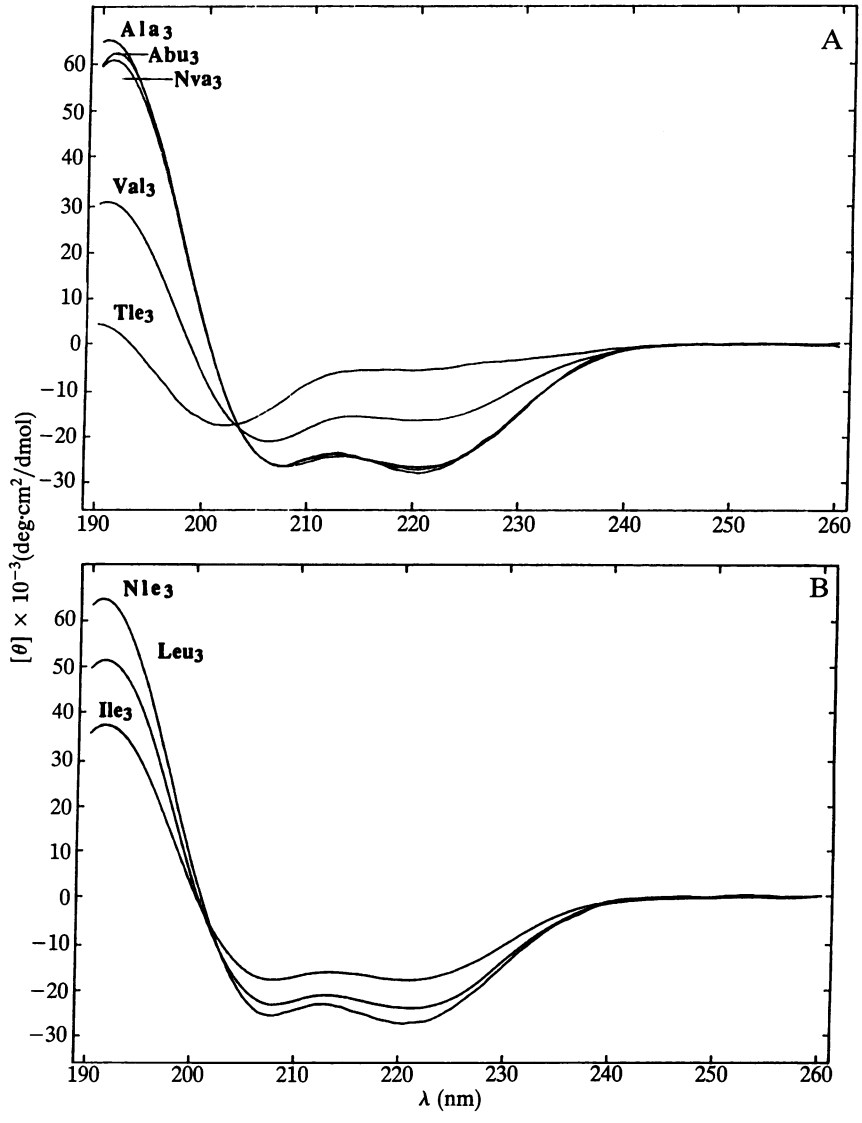

FIG. 3. (*A*) CD spectra of the following five peptides: Ala_3, Abu_3, Nva_3, Val_3, and Tle_3. The spectra shown here were taken at 4°C and neutral pH. (*B*) CD spectra of the following three peptides: Nle_3, Leu_3, and Ile_3.

methyl group on the β-carbon to give *t*-leucine results in nearly total disruption of helix; this peptide has even lower helix content than the glycine-substituted peptide, which has the lowest helix content among peptides containing the 10 side chains studied previously (12). The conformational space for the *t*-butyl side chain is small and constrained to predominantly extended conformations (25, 28).

Table 1. Helix content and free energy of peptides

Peptide	$-[\theta]_{222}$	*f*, %	ΔΔ*G*, kcal/mol
Ala_3	27,300 ± 800	85 ± 3	−0.74 ± 0.07
Abu_3	27,000 ± 900	84 ± 4	−0.70 ± 0.08
Nva_3	27,520 ± 800	86 ± 3	−0.76 ± 0.07
Nle_3	27,370 ± 600	85 ± 2	−0.74 ± 0.06
Leu_3	24,100 ± 700	75 ± 3	−0.55 ± 0.04
Ile_3	17,800 ± 500	56 ± 2	−0.32 ± 0.02
Val_3	16,000 ± 450	50 ± 2	−0.27 ± 0.02
Tle_3	4,800 ± 300	15 ± 4	0.22 ± 0.04

$-[\theta]_{222}$, Mean residue ellipticity (degree·cm²/dmol) of peptides at 222 nm. $f = [\theta]_{obs} - [\theta]_0/[\theta]_{max} - [\theta]_0$ = the fraction of helix. $[\theta]_{obs}$, $[\theta]$ observed from a previous column; $[\theta]_0 = 0 \pm 500$ degree·cm²/dmol, obtained by titrating a series of peptides with the denaturing solvent guanidine hydrochloride (23); $[\theta]_{max} = (n - 4/n)[\theta]_\infty$ = the maximal mean residue ellipticity value for chain length where n = the number of residues and $[\theta]_\infty = -40{,}000$ degree·cm²/dmol (24). ΔΔ*G* = free energy for helix formation of each guest amino acid related to glycine. $\Delta\Delta G_x = \Delta G_x - \Delta G_{Gly}$, where $\Delta G_{Gly} = 0.31$ kcal/mol (12).

The difference in the helix–coil transition energies in the third set of peptides—Nle_3, Leu_3, Ile_3, and Tle_3—demonstrates the effect of branching in the side chain for a series of compositionally identical amino acid residues. The straight-chain norleucine promotes helix formation best in this series. β-Branched isoleucine substantially diminishes helicity and bis-β-branched *t*-leucine severely disrupts the helix due to the restriction and extended locale of its conformational space (25). The γ-branch in leucine creates a more subtle diminution in helicity. Since this position is more removed from the helix backbone, restrictions in the side chain conformations (χ angles) have a dampened effect on the restriction of the Ψ/Φ angles. Interestingly, leucine is found to promote helix slightly better than norleucine in lysine polymers (25). The ε-methyl of norleucine is more mobile than the δ-methyl of leucine (25) and might be more entropically costly in a hydrophobic polylysine matrix.

Analysis of the stereochemistry of amino acid side chains in native protein structures suggests that the nature of the conformational restrictions imposed by helix formation should influence the helix propensities among chemically

similar side chains (29–35). Normally, three favored rotamers occur around the C_α–C_β bond in side chains other than glycine and alanine, corresponding to the 60°($g-$), 180°(t), and 300°($g+$) structures. The $g-$ rotamer is essentially absent in long side chains in the middle of α-helices, because of clashing between C_γ at position i and the carbonyl oxygen from the residue at position $i-3$. Even without β-branching, then, we might expect alanine to stabilize helix slightly more than α-aminobutyric acid, norvaline, or norleucine since it retains the $g-$ rotamer in the helical state. However, Table 1 shows that no significant differences among these are observed. Piela *et al.* (32) have argued that the severe steric restriction imposed by β-branching of the isoleucine side chain effectively destabilizes the α-helix relative to coil, while the γ-branched leucine side chain imposes much less restriction in conformation on helix formation. The results from set III show that increasing β-branching, while holding the side-chain composition constant, decreases helicity. Set II shows that increasing β-branching while concomitantly increasing side chain length likewise decreases helicity. Set I shows that side chain length *per se* does not affect helicity in the absence of branching.

Quantitative Comparisons. Our results allow us to estimate the quantitative effect of β-branching on α-helicity. Comparing the three pairs of peptides, $Abu_3 \rightarrow Val_3$, $Val_3 \rightarrow Tle_3$, and $Nva_3 \rightarrow Ile_3$, should yield the loss of free energy due to adding a β-branching methyl group to a guest residue in each case. The $\Delta\Delta G$ results in Table 1 suggest that adding a β-branched methyl side chain destabilizes helix by 0.45 ± 0.03 kcal/mol of side chain (Abu → Val, 0.43 kcal/mol; Val → Tle, 0.49 kcal/mol; Nva → Ile, 0.44 kcal/mol) (1 cal = 4.184 J). In terms of differences in s values for different side chains, this represents a very significant effect.

Concluding Remarks. From the point of view of stabilizing an α-helix, any of the amino acids from set I could replace alanine as a side chain. However, packing constraints have a major influence on folding of globular proteins (35), and the greater chain entropy associated with longer unbranched side chains could present a problem in creating compact globules. One might wonder why t-leucine is not used in nature to favor extended conformations that require well-defined packing. Glycine can be extended but has great conformational freedom; proline can be extended but often kinks and has cis/trans isomerization problems. One answer may be that nature has the same problem as the chemist: incorporation of t-leucine into a polypeptide is slow and difficult due to its sterically imposing side chain; creation of a t-leucine homopolymer of appreciable length might be formidable synthetically.

We are grateful to Jan Hermans for helpful discussions. This research was supported by research grants (GM 40746 and RR 02497) from the National Institutes of Health to N.R.K., a postdoctoral fellowship from the National Institutes of Health to J.C.S., a Sokol fellowship to P.C.L., and by a fellowship from the undergraduate MOST program at New York University to A.C. during the summer of 1990.

1. Fasman, G. D. (1989) in *Prediction of Protein Structure and the Principles of Protein Conformation*, ed. Fasman, G. D. (Plenum, New York), pp. 193–316.
2. Presta, L. G. & Rose, G. D. (1988) *Science* **240,** 1632–1641.
3. Richardson, J. S. & Richardson, D. C. (1988) *Science* **240,** 1648–1652.
4. Sueki, M., Lee, S., Powers, S. P., Denton, J. B., Konishi, Y. & Scheraga, H. A. (1984) *Macromolecules* **17,** 148–155.
5. Marqusee, S. & Baldwin, R. L. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 8898–8902.
6. Marqusee, S., Robbins, V. H. & Baldwin, R. L. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5286–5290.
7. Padmanabhan, S., Marqusee, S., Ridgeway, T., Laue, T. M. & Baldwin, R. L. (1990) *Nature (London)* **344,** 268–270.
8. Shoemaker, K. R., Fairman, R., Schultz, D. A., Robertson, A., York, E. J., Stewart, J. M. & Baldwin, R. L. (1990) *Biopolymers* **29,** 1–11.
9. Merutka, G. & Stellwagen, E. (1990) *Biochemistry* **29,** 7511–7515.
10. O'Neil, K. T. & DeGrado, W. F. (1990) *Science* **250,** 646–651.
11. Lyu, P. C., Marky, L. A. & Kallenbach, N. R. (1989) *J. Am. Chem. Soc.* **111,** 2733–2734.
12. Lyu, P. C., Liff, M. I., Marky, L. A. & Kallenbach, N. R. (1990) *Science* **250,** 669–673.
13. Liff, M. I., Lyu, P. C. & Kallenbach, N. R. (1991) *J. Am. Chem. Soc.* **113,** 1014–1020.
14. Barskaya, T. V. & Ptitsyn, O. B. (1971) *Biopolymers* **10,** 2181–2197.
15. Finkelstein, A. V., Badretdinov, A. Yu. & Ptitsyn, O. B. (1990) *Nature (London)* **345,** 300.
16. Barlow, D. J. & Thornton, J. M. (1983) *J. Mol. Biol.* **168,** 867–885.
17. Shoemaker, K. R., Kim, P. S., Brems, D. N., Marqusee, S., York, E. J., Chaiken, I. M., Stewart, J. M. & Baldwin, R. L. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 2349–2353.
18. Shoemaker, K. R., Kim, P. S., York, E. J., Stewart, J. M. & Baldwin, R. L. (1987) *Nature (London)* **326,** 563–567.
19. Brandts, J. F. & Kaplan, L. J. (1973) *Biochemistry* **12,** 2011–2024.
20. Woody, R. W. (1985) in *The Peptides* (Academic, New York), Vol. 7, pp. 15–114.
21. Johnson, W. C., Jr. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17,** 145–166.
22. Zimm, B. H. & Bragg, J. K. (1959) *J. Chem. Phys.* **31,** 526–535.
23. Lyu, P. C., Wang, P., Liff, M. I. & Kallenbach, N. R. (1991) *J. Am. Chem. Soc.*, in press.
24. Johnson, W. C., Jr., & Tinoco, I., Jr. (1972) *J. Am. Chem. Soc.* **94,** 4389–4390.
25. Patterson, Y. & Leach, S. J. (1978) *Macromolecules* **11,** 409–415.
26. Lim, V. I. (1978) *FEBS Lett.* **89,** 10–14.
27. Arfmann, H.-A., Labitzke, R. & Wagner, K. G. (1977) *Biopolymers* **16,** 1815–1826.
28. Stokrova, S., Pospisek, J., Sponar, J. & Blaha, K. (1985) *Collect. Czech. Chem. Commun.* **50,** 2925–2936.
29. Janin, J., Wodak, S., Levitt, M. & Maigret, B. (1978) *J. Mol. Biol.* **125,** 357–386.
30. Nemethy, G. & Scheraga, H. A. (1982) *Biopolymers* **21,** 1535–1555.
31. Moult, J. & James, M. N. G. (1986) *Proteins: Struct. Funct. Genet.* **1,** 146–163.
32. Piela, L., Nemethy, G. & Scheraga, H. A. (1987) *Biopolymers* **26,** 1273–1286.
33. Ponder, J. W. & Richards, F. M. (1987) *J. Mol. Biol.* **193,** 775–791.
34. McGregor, M. J., Islam, S. A. & Sternberg, M. J. E. (1987) *J. Mol. Biol.* **198,** 295–310.
35. Chan, H. S. & Dill, K. A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6388–6392.