Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 5869–5873, August 1988
Biochemistry

Structure of the carboxyl terminus of the *RAS* gene-encoded P21 proteins

(conformational energy/membrane binding/α-helix)

PAUL W. BRANDT-RAUF*, ROBERT P. CARTY†, JAMES CHEN‡, MATTHEW AVITABLE§, JACK LUBOWSKY§, AND MATTHEW R. PINCUS¶||

*Division of Environmental Sciences and Department of Medicine, Columbia–Presbyterian Medical Center, 60 Haven Avenue, New York, NY 10032; †Department of Biochemistry, and §Scientific Academic Computing Center, State University of New York Downstate Medical Center, 450 Clarkson Avenue, Brooklyn, NY 11203; ‡Department of Chemistry, New York University, 4 Washington Place, New York, NY 10003; and ¶Department of Pathology, New York University Medical Center, 550 First Avenue, New York, NY 10016

Communicated by H. A. Scheraga, May 5, 1988

ABSTRACT **The three-dimensional structures of the carboxyl-terminal regions of the P21 protein products of the human Harvey (Ha), Kirsten (Ki_A and Ki_B), and neuroblastoma (N) *RAS* oncogenes and various mutants have been determined by using conformational energy analysis. The carboxyl-terminal region of P21 has been strongly implicated in the binding of the protein to the inner surface of the plasma membrane without which the protein is inactive. The only invariant residue in this region is Cys-186, which is necessary for the post-translational addition of palmitic acid. The surrounding sequences of the active native proteins differ considerably. Nevertheless, certain amino acid substitutions in this region are known to eliminate membrane binding and protein activity, suggesting that there is a conserved common structural feature in this region in the native proteins that is disrupted in the mutant proteins. Conformational energy analysis shows that the four native P21 proteins have a common structure in the form of an α-helix for the terminal pentapeptide. A mutant, pBW277, that fails to bind to the membrane and is inactive cannot adopt an α-helical structure in this region because of a proline at position 188. Another mutant, pBW766, that retains membrane binding and activity, on the other hand, retains the preference for an α-helical conformation in the terminal pentapeptide. These findings suggest that, despite various amino acid sequences in this region, the carboxyl-terminal pentapeptides of the P21 proteins form a distinctive structural domain that must have an α-helical structure for membrane binding and intracellular activity.**

It has been well-documented that the *ras* gene-encoded protein product (P21) must be membrane-bound to function intracellularly (1, 2). Naturally occurring single amino acid substitutions at positions 12, 13, 59, 61, and 63 in this protein result in cell transformation (1, 2). P21 proteins with substitutions at position 12 and also with Cys-186 deleted or substituted with another amino acid, such as serine, were not active in cell transformation and failed to bind to the inner cell membrane (3–7). These cytosolic proteins nonetheless exhibited typical P21 protein activity—i.e., the naturally occurring P21 proteins and the substituted P21 proteins had the same binding affinities for GDP (3–7).

Cys-186 is the site at which palmitic acid is covalently linked as a thioester to the free thio group of this residue post-translationally (4, 5, 8, 9). It is this event that presumably results in the attachment of the protein to the inner cell membrane (4, 5, 8, 9). Amino acid sequences for many *ras* gene-encoded proteins from diverse eukaryotic cells, including yeast, reveal that this residue is conserved (4, 5). The surrounding sequences, however, vary considerably among the *ras* gene-encoded proteins (4, 5). For example, human Harvey (Ha), Kirsten (Ki_A and Ki_B), and neuroblastoma (N) *RAS* gene-encoded P21 proteins [P21(Ha), P21(Ki_A), P21(Ki_B), and P21(N), respectively] have strongly homologous sequences in other regions but differ substantially in the last 10 amino acid residues from position 180 to position 189 (10–13). The sequences of these proteins are as follows (10–13):

```
180                 185
Gly-Cys-Met-Ser-Cys-Lys-Cys-Val-Leu-Ser  H
Cys-Val-Lys-Ile-Lys-Lys-Cys-Ile-Ile-Met  Ki_A
Lys-Lys-Ser-Lys-Thr-Lys-Cys-Val-Ile-Met  Ki_B
Gly-Cys-Met-Gly-Leu-Pro-Cys-Val-Val-Met  N.
```

The full-length proteins are fully active and transform NIH 3T3 cells in culture if substituted, for example, at position 12.

The carboxyl-terminal decapeptide may not interact with the remainder of the protein. It has been shown in genetic splicing experiments that if residues from position 164 to position 183 are deleted from P21(Ha) (and, because of frame-shifting, the sequence Pro-Asp-Gln is inserted), the protein retains full transforming activity (3–7). Conversely, changes further into the sequence from position 180 to position 189 sequence that, nonetheless, leave Cys-186 intact can significantly lower the transforming activity of the protein (3–7). For example, if the sequence to residue 187 is left intact and the sequence Thr-Pro replaces the carboxyl-terminal Val-Leu-Ser sequence of P21(Ha) (the pBW277 mutant), then the resulting protein does not bind to the cell membrane (4, 5).

It, therefore, appears that despite differences in sequences among the four membrane-binding decapeptides of P21(Ha), P21(K_i), and P21(N) there may be a common structural feature that is conserved and that may be destroyed by the substitutions in the pBW277 mutant. A truncated form of P21 containing residues 1–171—i.e., without the carboxyl-terminal end—was analyzed by x-ray crystallography at 2.7-Å resolution (14). The carboxyl terminus of this protein contains an α-helix from about residue 150 to residue 171, the last residue. Further toward the carboxyl-terminal end of the complete molecule, sequences, such as Pro-Pro (residues 173 and 174) and Pro-Gly (residues 179 and 180), occur that may terminate this helix or introduce "kinks" in a propagating helix past residue 171. We have computed (15, 16) that P21(Ha) and P21(N) contain helical segments for residues

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: P21(Ha), P21(Ki_A), P21(Ki_B), and P21(N), Harvey, Kirsten A and B, and neuroblastoma *RAS*-encoded P21 proteins, respectively.

||To whom reprint requests should be addressed.

Table 1. Low-energy conformations for the decapeptide *N*-acetyl-Gly-Cys-Met-Ser-Cys-Lys-Cys-Val-Leu-Ser-$NHCH_3$, the carboxyl-terminal sequence for P21(Ha)

	Conformational state*										
Conformer	Gly	Cys	Met	Ser	Cys	Lys	Cys	Val	Leu	Ser	Energy, kcal/mol†
1	C‡	A	A	A	A	A	A	A	A	A	0.0
2	C‡	A	A	A	A	A	A	A	A	G	1.4
3	C‡	A	A	A	A	A	A	A	A	C	1.7
4	C‡	A	A	A	A	A	A	A	A	D	1.8

*These are defined in ref. 23. The single letter conformation code is defined by ranges of dihedral angles as follows: A, $-110° \leq \Phi < -40°$, $-90° \leq \Psi < -10°$; C, $-110° \leq \Phi < -40°$, $50° \leq \Psi < 130°$; D, $-180° \leq \Phi < -110°$, $20° \leq \Psi < 110°$; E, $-180° \leq \Phi < -140°$, $-180° \leq \Psi < -140°$ or $130° \leq \Psi < 180°$; F, $-110° \leq \Phi < -40°$, $-180° \leq \Psi < -140°$ or $130° \leq \Psi < 180°$; G, $-180° \leq \Phi < -110°$, $-90° \leq \Psi < -40°$.

†Energies are expressed relative to the energy of conformer 1. All conformers with relative energies within a cutoff of 2 kcal/mol of the energy of the global minimum are shown.

‡States were obtained by multiplying the corresponding single letter state by −1 and reversing the inequality.

180–189 and may in fact form amphipathic helices in this segment.

In this communication by using conformational analysis, we explore the possible conformations for the carboxyl-terminal decapeptide sequences in P21(Ha), P21(Ki), and P21(N) to determine the common structures. We also examine the preferred conformations for mutant carboxyl-terminal sequences that do not bind to cell membranes or bind with lower affinities than do the parent proteins. Conformational analysis has been used to explore the structural effects of amino acid substitutions in the amino-terminal and central transforming regions of P21 proteins and to determine structural commonalities in various other biologically active peptides (17–21).

METHODS

Conformational analysis was performed on six carboxyl-terminal sequences of the following P21 proteins: P21(Ha), P21(Ki_A), P21(Ki_B), P21(N), the pBW277 mutant, and the pBW766 mutant. As noted, the pBW277 mutant has the same sequence as P21(Ha) except Thr-Pro replaces the terminal Val-Leu-Ser (4, 5); the pBW766 mutant has the same sequence as P21(Ha) except there is a deletion from position 166 to position 183 and Pro-Asp-Gln is inserted before the sequence from position 184 to position 189 (6).

The general methods used are based on the program ECEPP (Empirical Conformational Energy for Polypeptides and Proteins) (22). This approach has been used to compute the allowed conformations of the naturally occurring amino acids (23), short oligopeptides (24), and long constrained oligopeptides (25, 26), as well as long unconstrained polypeptides including other regions of the P21 proteins (15–20, 27, 28).

In this study, all of the peptides were constructed by the successive addition of single amino acid residues from the amino- to the carboxyl-terminal end. Thus, the single residue minima for the first two amino acids were combined and subjected to energy minimization. The conformers whose energies lay within a cutoff energy of that of the global minimum (5.0 kcal/mol; 1 cal = 4.184 J) were retained and then combined with the single residue minima for the next residue in the chain. This process was repeated until all residues of the peptide were added. At all stages of the procedure, the amino and carboxyl termini contained *N*-acetyl and $NHCH_3$ groups, respectively, to include end effects of neighboring residues not included in a particular peptide.

RESULTS AND DISCUSSION

The low-energy conformers for the carboxyl-terminal peptides of P21(Ha), P21(Ki_A), P21(Ki_B), P21(N), and pBW277 are presented in Tables 1–5, respectively. Examination of the conformations of the decapeptides for the four naturally occurring P21 proteins [P21(Ha), P21(Ki_A), P21(Ki_B), and P21(N)] reveals them to be relatively inflexible well-defined structures with relatively few conformers within 2 kcal/mol of the global minimum; in fact, for P21(Ki_A), there were no other conformations than the α-helix in Table 2 whose

Table 2. Low-energy conformations for the decapeptide *N*-acetyl-Cys-Val-Lys-Ile-Lys-Lys-Cys-Ile-Ile-Met-$NHCH_3$, the carboxyl-terminal sequence for P21(Ki_A)

	Conformational state										
Conformer	Cys	Val	Lys	Ile	Lys	Lys	Cys	Ile	Ile	Met	Energy, kcal/mol
1	A	A	A	A	A	A	A	A	A	A	0.0

Conformational states and energy are as described in Table 1.

Table 3. Low-energy conformations for the decapeptide *N*-acetyl-Lys-Lys-Ser-Lys-Thr-Lys-Cys-Val-Ile-Met-$NHCH_3$, the carboxyl-terminal sequence for P21(Ki_B)

	Conformational state										
Conformer	Lys	Lys	Ser	Lys	Thr	Lys	Cys	Val	Ile	Met	Energy, kcal/mol
1	A	A	A	A	A	A	A	A	A	C	0.0
2	A	A	A	A	A	A	A	A	A	A	0.5
3	A	A	A	A	A	A	A	A	A	E	0.8
4	F	A	A	A	A	A	A	A	A	C	1.2
5	C	A	A	A	A	A	A	A	A	C	1.9

Conformational states and energies are as described in Table 1.

Table 4. Low-energy conformations for the decapeptide *N*-acetyl-Gly-Cys-Met-Gly-Leu-Pro-Cys-Val-Val-Met-$NHCH_3$, the carboxyl-terminal sequence for the P21(N)

	Conformational state										
Conformer	Gly	Cys	Met	Gly	Leu	Pro	Cys	Val	Val	Met	Energy, kcal/mol
1	C	A	C	D	D	A	A	A	C	C	0.0
2	C	A	C	D	D	A	A	A	A	A	0.6
3	C	A	C	D	D	A	A	A	C	A	0.6
4	A	A	C	D	D	A	A	A	C	C	0.7
5	C	A	C	D	D	A	A	A	A	A	0.8

Conformational states and energies are as described in Table 1.

Table 5. Low-energy conformations for the P21 protein nonapeptide *N*-acetyl-Gly-Cys-Met-Ser-Cys-Lys-Cys-Thr-Pro-$NHCH_3$, the carboxyl-terminal sequence for the pBW277 mutant protein

	Conformational state									
Conformer	Gly	Cys	Met	Ser	Cys	Lys	Cys	Thr	Pro	Energy, kcal/mol
1	C	A	A	A	A	A	A	D	C	0.0
2	C	A	A	A	A	A	A	D	A	0.7
3	C	A	A	A	A	A	A	D	F	1.1

Conformational states and energies are as described in Table 1.

energies were within 3 kcal/mol of the global minimum. Furthermore, variations in the low-energy structures of each of the peptides tend to occur principally at the amino- or carboxyl-terminal residues, whereas the central eight residues tend to be invariant. None of the global minima for these four peptides are identical. However, overall they do have conformational similarities, especially in the region around the invariant residue Cys-186. In particular there is only one long sequence that is common in structure for all these proteins and that is the all α-helical terminal pentapeptide. This occurs as the global minimum for P21(Ha) and P21(Ki_A) and as the second lowest energy conformation (only 0.5 kcal/mol above the energy of the global minimum) in the P21(Ki_B) and as the second and fifth lowest energy conformations (only 0.6 and 0.8 kcal/mol, respectively, above the energy of the global minimum) in P21(N). The global minimum conformation for the P21(Ha) decapeptide is presented in stereoview in Fig. 1, and the α-helical conformation of the terminal pentapeptide for P21(N) is presented in stereoview in Fig. 2 for comparison.

All of the four naturally occurring P21 proteins behave similarly with regard to membrane binding and transformation activity. As noted, however, the pBW277 mutant fails to bind to the membrane and transform cells, despite the fact

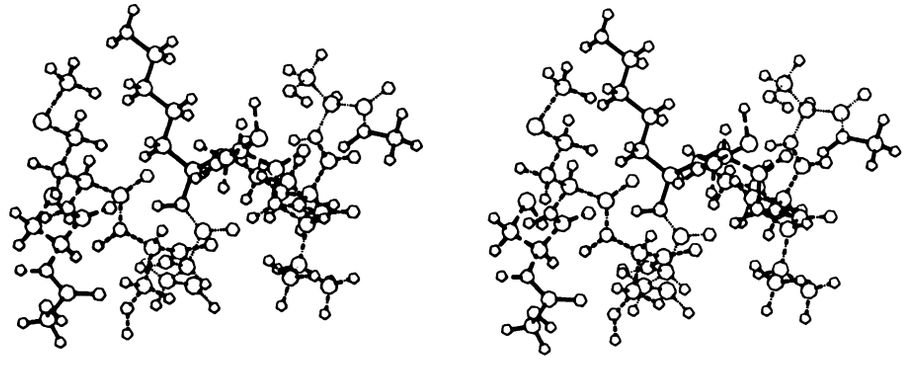

FIG. 1. Stereoview of the global minimum energy conformation of the terminal decapeptide of P21(Ha). The carboxyl terminus is to the right of the figure.

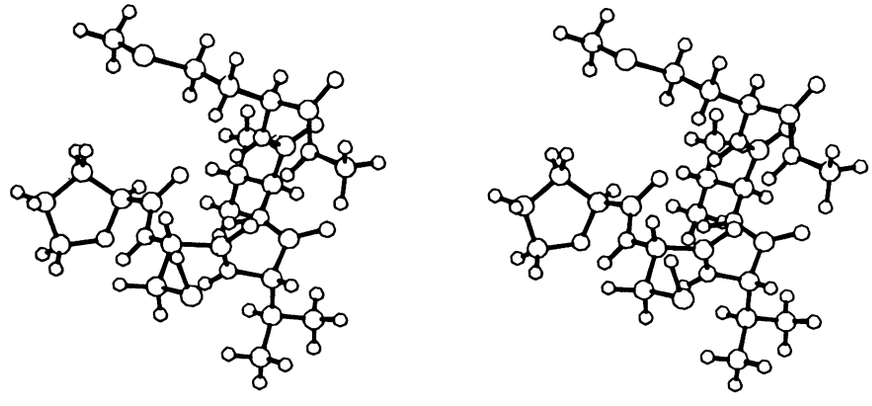

FIG. 2. Stereoview of the all α-helical conformation of the terminal pentapeptide of P21(N). The carboxyl terminus is to the right of the figure.

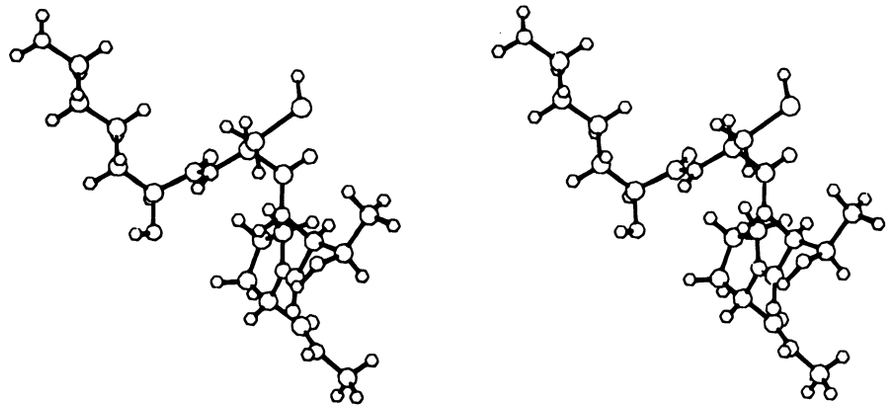

FIG. 3. Stereoview of the global minimum energy conformation of the terminal tetrapeptide of the pBW277 mutant. The carboxyl terminus is to the bottom of the figure.

that the Cys-186 residue is preserved. As can be seen in Table 5, the low-energy conformations for the terminal segment of this protein are different from those in P21(Ha), P21(Ki_A), P21(Ki_B), and P21(N), particularly at the carboxyl terminus. In the global minimum conformation, the terminal α-helix is disrupted by a DC conformation at positions 187 and 188. In fact, this is not unexpected due to the substitution of proline. It has been shown that proline is very likely to disrupt α-helices and other types of regular structures because the amino acid residue preceding proline makes unfavorable backbone–backbone contacts with proline if it adopts the A (α-helical) state (27, 28). Thus, the substitution of proline at position 188 forces the preceding threonine in this case to adopt a nonhelical conformation. This is true for all the low-energy conformers of the pBW277 mutant (Table 5). Therefore, this protein shares no low-energy conformations in common with the active naturally occurring proteins in the region between positions 185 and 188. The global minimum conformation for the terminal tetrapeptide of pBW277 is presented in stereoview in Fig. 3.

On the other hand, the pBW766 mutant, which is substituted at positions 181–183, and similar mutants, such as pBW1092 (6), which has Cys-184 repeated and Pro-Asp-Gln inserted between the two cysteine residues (and thus would have the same amino acid sequence for the nine C-terminal residues as the pBW766 mutant), retain membrane-binding ability and transforming activity (6). In all of the low-energy conformations for this nonapeptide, the terminal pentapeptide segment (data not shown) was found to be α-helical and close to the common structure for the four naturally occurring proteins for the corresponding last five residues.

Conclusions. These findings tend to support the conclusion that the carboxyl terminus of the P21 proteins, in particular the terminal five amino acids, form a distinctive structural domain despite various amino acid sequences in this region. It is likely that this domain is in the form of an α-helix, which with the invariant Cys-186 for the post-translational palmitation, is necessary for membrane binding and intracellular activity. Any amino acid substitutions in this region that would disrupt this conformation would be expected to alter membrane binding. Thus, we would predict, for example, that the substitution of proline at any position from position 186 to position 189 would break the α-helix and cause loss of membrane binding and transforming activity.

It is interesting to speculate on possible conformations for the segment of the protein from position 150 to position 189. As noted above, x-ray crystallographic analysis shows (14) that the entire segment from residue 150 to residue 171 (the last residue of the crystallized protein) is an α-helix. Since in P21(Ha) the sequences Pro-Pro and Pro-Gly occur several residues toward the carboxyl-terminal end, it is likely that these residues would cause hairpin turns. If indeed the carboxyl-terminal decapeptides are α-helical as computed in this paper, it may be that the entire carboxyl-terminal segment from residue 150 to residue 189 exists in a helical hairpin conformation. Such conformations have been proposed for the membrane-binding domains of membrane-bound proteins (29). This consideration would also apply to the carboxyl-terminal end of the P21 protein since this is the region that is membrane-bound.

This work was supported in part by a grant from the National Cancer Institute (CA-42500).

1. Lowy, D. R. & Willumsen, B. M. (1986) *Cancer Surv.* **5**, 275–289.
2. Shih, T. Y. & Weeks, M. O. (1984) *Cancer Invest.* **2**, 109–123.
3. Willumsen, B. M., Christensen, A., Hubbert, N. L., Papageorge, A. G. & Lowy, D. R. (1984) *Nature (London)* **310**, 583–586.
4. Willumsen, B. M. & Lowy, D. R. (1985) *Dev. Oncol.* **28**, 25–40.
5. Willumsen, B. M., Norris, K., Papageorge, A. G., Hubbert, N. L. & Lowy, D. R. (1984) *EMBO J.* **3**, 2581–2585.
6. Willumsen, B. M., Papageorge, A. G., Hubbert, N., Bekesi, E., Kung, H.-F. & Lowy, D. R. (1985) *EMBO J.* **4**, 2893–2896.
7. Willumsen, B. M., Papageorge, A. G., Kung, H.-F., Bekesi, E., Robins, T., Johnsen, M., Vass, W. C. & Lowy, D. R. (1986) *Mol. Cell. Biol.* **6**, 2646–2654.
8. Sefton, B. M., Trowbridge, I. S., Cooper, J. A. & Scolnick, E. M. (1982) *Cell* **31**, 465–474.
9. Chen, Z.-Q., Ulsh, L. S., DuBois, G. & Shih, T. Y. (1985) *J. Virol.* **56**, 607–612.
10. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, D. H. & Goeddel, D. V. (1983) *Nature (London)* **302**, 33–37.
11. Taparowsky, E., Shimizu, K., Goldfarb, M. & Wigler, M. (1983) *Cell* **34**, 581–586.
12. Shimizu, K., Birnbaum, D., Ruley, M. A., Fasano, O., Suard, Y., Edlund, L., Taparowsky, E., Goldfarb, M. & Wigler, M. (1983) *Nature (London)* **304**, 497–500.
13. McGrath, J. P., Capon, D. J., Smith, D. H., Chen, E. Y., Seeburg, D. H., Goeddel, D. V. & Levinson, A. D. (1983) *Nature (London)* **304**, 501–506.
14. de Vos, A. M., Tong, L., Milburn, M. V., Matias, P. M., Jancarik, J., Noguchi, S., Nishimura, S., Miura, K., Ohtsuka, E. & Kim, S.-H. (1988) *Science* **239**, 888–893.
15. Pincus, M. R. & Brandt-Rauf, P. W. (1986) *Cancer Invest.* **4**, 185–195.
16. Pincus, M. R., Carty, R. P., Brandt-Rauf, P. W., Lubowsky, J. & Avitable, M. (1986) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **45**, 1921.
17. Pincus, M. R., van Renswoude, J., Harford, J. B., Chang, E. H., Carty, R. P. & Klausner, R. D. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5253–5257.
18. Pincus, M. R. & Brandt-Rauf, P. W. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3596–3600.

19. Pincus, M. R., Brandt-Rauf, P. W., Carty, R. P., Lubowsky, J. & Avitable, M. (1985) *J. Protein Chem.* **4,** 345–352.
20. Pincus, M. R., Brandt-Rauf, P. W., Carty, R. P., Lubowsky, J., Avitable, M., Gibson, K. D. & Scheraga, H. A. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 8375–8379.
21. Pincus, M. R., Carty, R. P., Chen, J., Lubowsky, J., Avitable, M., Shah, D., Scheraga, H. A. & Murphy, R. B. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4821–4825.
22. Scheraga, H. A. (1984) *Carlsberg Res. Commun.* **49,** 1–9.
23. Zimmerman, S. S., Pottle, M. S., Nemethy, G. & Scheraga, H. A. (1977) *Macromolecules* **10,** 1–9.
24. Simon, I., Nemethy, G. & Scheraga, H. A. (1980) *Macromolecules* **11,** 797–804.
25. Dygert, M., Go, N. & Scheraga, H. A. (1975) *Macromolecules* **8,** 750–761.
26. Miller, M. H. & Scheraga, H. A. (1976) *J. Polym. Sci. Polym. Symp.* **54,** 171–200.
27. Pincus, M. R., Klausner, R. D. & Scheraga, H. A. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5107–5110.
28. Pincus, M. R. & Klausner, R. D. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 3413–3417.
29. Engelman, D. M. & Steitz, T. A. (1981) *Cell* **23,** 411–422.