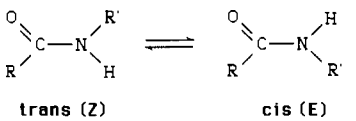
Blackburn, P., & Moore, S. (1982) *Enzymes (3rd Ed.) 15*, 317-433.
Blank, A., & Dekker, C. A. (1981) *Biochemistry 20*, 2261-2267.
Chang, J.-Y. (1983) *Methods Enzymol. 91*, 455-466.
Cranston, J. W., Perini, F., Crisp, E. R., & Hixson, C. V. (1980) *Biochim. Biophys. Acta 616*, 239-258.
Crestfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem. 238*, 622-627.
Dayhoff, M. O., Ed. (1972) in *Atlas of Protein Sequence and Structure*, Vol. 5, p D6, National Biomedical Research Foundation, Washington, DC.
Delaney, R. (1963) *Biochemistry 2*, 438-444.
Dickerson, R. E., & Geis, I. (1969) *The Structure and Action of Proteins*, Benjamin, Menlo Park, CA.
Gaastra, W., Welling, G. W., & Beintema, J. J. (1978) *Eur. J. Biochem. 86*, 209-217.
Gleich, G. J., Loegering, D. A., Bell, M. P., Checkel, J. L., Ackerman, S. J., & McKean, D. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3146-3150.
Hartley, B. S. (1970) *Biochem. J. 119*, 805-822.
Hewick, R. M., Hunkapiller, M. W., Hood, L. E., & Dreyer, W. J. (1981) *J. Biol. Chem. 256*, 7990-7997.
Hopp, T. P., & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3824-3828.
Hunkapiller, M. W., & Hood, L. E. (1983) *Methods Enzymol. 91*, 486-493.
Iwahashi, K., Nakamura, K., Mitsui, Y., Ohgi, K., & Irie, M. (1981) *J. Biochem. (Tokyo) 90*, 1685-1690.
Iwama, M., Kunihiro, M., Ohgi, K., & Irie, M. (1981) *J. Biochem. (Tokyo) 89*, 1005-1016.
Katoh, H., Yoshinaga, M., Yanagita, T., Ohgi, K., Irie, M., Beintema, J. J., & Meinsma, D. (1986) *Biochim. Biophys. Acta 873*, 367-371.
Morita, T., Niwata, Y., Ohgi, K., Ogawa, M., & Irie, M. (1986) *J. Biochem. (Tokyo) 99*, 17-25.
Morrissey, J. H. (1981) *Anal. Biochem. 117*, 307-310.
Niwata, Y., Ohgi, K., Sanda, A., Takizawa, Y., & Irie, M. (1985) *J. Biochem. (Tokyo) 97*, 923-934.
Ohta, T., Ogawa, M., Kurihara, M., Kitahara, T., & Kosaki, G. (1982) *Clin. Chim. Acta 124*, 51-62.
Palmer, K. A., Scheraga, H. A., Riordan, J. F., & Vallee, B. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1965-1969.
Shapiro, R., Fett, J. W., Strydom, D. J., & Vallee, B. L. (1986) *Biochemistry 25*, 7255-7264.
Sierakowska, H., & Shugar, D. (1977) *Prog. Nucleic Acid Res. Mol. Biol. 20*, 59-130.
Slifman, N. R., Loegering, D. A., McKean, D. J., & Gleich, G. J. (1986) *J. Immunol. 137*, 2913-2917.
Strydom, D. J., Fett, J. W., Lobb, R. R., Alderman, E. M., Bethune, J. L., Riordan, J. F., & Vallee, B. L. (1985) *Biochemistry 24*, 5486-5494.
Sugiyama, R. H. (1981) Thesis, University of California, Berkeley.
Sugiyama, R. H., Blank, A., & Dekker, C. A. (1981) *Biochemistry 20*, 2268-2274.
Vereijken, J. M., De Vlieg, J., & Beintema, J. J. (1984) *Biochim. Biophys. Acta 788*, 298-305.
Yamanaka, M., Akagi, K., Murai, K., Hirano, N., Fujimi, S., & Omae, T. (1977) *Clin. Chim. Acta 78*, 191-201.

Influences of Solvent Water on Protein Folding: Free Energies of Solvation of Cis and Trans Peptides Are Nearly Identical†

Anna Radzicka, Lee Pedersen, and Richard Wolfenden*
Departments of Biochemistry and Chemistry, University of North Carolina, Chapel Hill, North Carolina 27514
Received October 23, 1987; Revised Manuscript Received February 10, 1988

ABSTRACT: Peptide bonds interact so strongly with water that even a modest difference between the free energies of solvation of their cis and trans isomers could have a significant bearing on protein structure. However, proton magnetic resonance studies at high dilution in deuteriated solvents show that *N*-methylformamide exists as the cis isomer to the extent of 8% in water, 10.3% in chloroform, 8.8% in benzene, and 9.2% in cyclohexane. Integrated intensities of proton and carbon resonances show that *N*-methylacetamide exists as the cis isomer to the extent of only 1.5% in water, not changing much in nonpolar solvents. Quantum mechanical calculations using the 6-31G** basis set reproduce these relative abundances with reasonable accuracy and show that there is little difference between the dipole moments of the cis and trans isomers, for either amide. The remarkable insensitivity of cis/trans equilibria to the solvent environment and the heavy preponderance of trans isomers regardless of the polarity of the surroundings (ca. 98.5% for *N*-methylacetamide, whose properties may resemble those of a typical peptide bond) accord with the overwhelming preference of peptide bonds for the trans configuration that is consistently observed in the three-dimensional structures of globular proteins.

The peptide bond prefers to adopt either of two planar configurations (Corey & Donohue, 1950), stabilized by resonance to such an extent that their interconversion is slow on the NMR time scale at room temperature (Drakenberg et al., 1972). Both configurations have been observed in proteins,



but the trans configuration appears to be much more common. Among the rare exceptions are three non-proline cis peptide bonds in carboxypeptidase A (Rees et al., 1981). Steric difficulties would arise if peptide bonds were to adopt the cis

† This work was supported by NSF Research Grant PCM-7823016 (A.R. and R.W.) and by NIH Research Grants HL-27995 and HL-26309 (L.P.).

0006-2960/88/0427-4538$01.50/0

configuration at many adjacent positions (Ramachandran & Mitra, 1976), but it has not been clear why isolated cis peptide bonds should be so rare.

Earlier work has shown that free energies of interaction of amino acid side chains with solvent water are closely related to their relative tendencies to appear at the surfaces of proteins (Radzicka & Wolfenden, 1988), and solvation is one of several factors that might contribute to the difference in stability between cis and trans peptides. Dynamic vapor pressure measurements have shown that the peptide bond is among the most strongly solvated of uncharged functional groups that are found in biological systems (Wolfenden, 1976). The negative free energy of solvation of the peptide bond is so great that even a modest difference between cis and trans isomers could have a significant effect on protein structure. For example, if the trans configuration were associated with a longer and stronger dipole than the cis configuration, then the trans configuration would be expected to be relatively favored in water; conversely, the cis configuration might be relatively favored in less polar surroundings such as the interior of a globular protein. As a further consequence, other steric constraints, that might tend to force any particular peptide bond to adopt a cis or a trans configuration, might be expected to affect its thermodynamic tendency to seek a solvent-accessible position in the overall structure of a protein.

This possibility could be tested by examining the positions of cis/trans equilibrium of model peptides in a variety of solvents. Earlier efforts to measure the relative abundances of isomers, under various conditions, did not resolve this issue clearly [for a review, see Stewart and Liddall (1970)]. Liler (1972) concluded from proton magnetic resonance measurements that the rare cis isomer of *N*-methylacetamide was 3-fold more abundant in water than in chloroform, whereas the rare cis isomer of *N*-methylformamide was substantially less abundant in water (7.7%) than in chloroform (11%). Using natural abundance ^{15}N spectroscoy, Nakanishi and Roberts (1981) detected 8% of the cis isomer in neat *N*-methylformamide but did not observe detectable amounts of the cis isomer in neat *N*-methylacetamide.

With the availability of instruments of greater sensitivity, it seemed desirable to extend these equilibrium measurements to solvents less polar than chloroform and to include solvents that are not capable of acting as hydrogen-bond donors. The results described below indicate that the position of cis/trans equilibrium of *N*-methylformamide is almost completely insensitive to the solvent environment. Thus, the abundance of the cis isomer does not increase even when the polarity of the surroundings is reduced severely, as in cyclohexane. Quantum mechanical calculations, using the 6-31G** basis set, reproduce the observed abundances of the cis and trans isomers in the vapor phase with resonable accuracy and indicate that there is not much difference between their dipole moments.

Materials and Methods

Materials. *N*-Methylformamide and *N*-methylacetamide were purchased from Aldrich Chemical Co. *N*-Methylformamide was dried with molecular sieves, distilled, and fractionally frozen. *N*-Methylacetamide, after drying and redistillation, was purified by zone melting (Knecht, 1971). [^{13}C]-*N*-Methylacetamide was synthesized by treating [^{13}C]methylamine hydrochloride with acetic anhydride; after removal of substances boiling below 130 °C, the product was purified by ion-exchange chromatography on Bio-Rex MSZ-501D gel from Bio-Rad Laboratories, eluting with water. Each of these solutes was free of impurities that could be detected by proton or carbon magnetic resonance spectroscopy (see below). D_2O, $CDCl_3$, and C_6D_6 were purchased from Aldrich Chemical Co. C_6D_{12} was purchased from MSD Isotopes.

Methods. Proton and ^{13}C magnetic resonance spectra were recorded with a Varian XL-400FT spectrometer operating at 20 °C. Proton chemical shifts, referred to solvent peaks, and ^{13}C chemical shifts, referred to added 1,4-dioxan, are shown in Table I. The relative concentrations of the cis and trans isomers were calculated by comparing the integrated intensities of signals from the methylamine and carboxylic acid moieties as indicated in Table I. In each case, signals from the two isomers could be resolved completely with this instrument.

Table I: Chemical Shifts and Abundances of the Cis Isomers of *N*-Methylformamide and *N*-Methylacetamide in Various Solvents at 20 °C

nucleus	solvent	solute (M)	δ (trans)	δ (cis)	% cis[a]
		N-Methylformamide			
$HCONHCH_3$	D_2O	0.050	2.74 (s)	2.88 (s)	8.0 ± 0.2
$HCONHCH_3$	D_2O	0.050	8.04 (s)	7.95 (s)	8.95 ± 0.5
$HCONHCH_3$	$CDCl_3$	0.025	2.75 (d)	2.91 (d)	10.5 ± 0.1
$HCONHCH_3$	$CDCl_3$	0.025	8.18 (s)	8.02 (d)	10.1 ± 0.1
$HCONHCH_3$	C_6D_6	0.025	2.37 (d)	1.96 (d)	8.7 ± 0.1
$HCONHCH_3$	C_6D_6	0.025	7.85 (s)	7.54 (d)	8.8 ± 0.3
$HCONHCH_3$	C_6D_{12}	0.025	2.71 (d)	2.74 (d)	9.25 ± 0.2
$HCONHCH_3$	C_6D_{12}	0.025	8.02 (s)	7.87 (d)	7.85 ± 2.4
		N-Methylacetamide			
$CH_3CONH-CH_3$	D_2O	0.025	2.00 (s)	2.06 (s)	1.39 ± 0.4
$CH_3CONH-CH_3$	D_2O	0.025	2.69 (s)	2.74 (s)	1.88 ± 0.4
$CH_3CONH-{}^{13}CH_3$	D_2O	0.10	17.4 (s)	2.07 (s)	1.46 ± 0.09
$CH_3CONH-CH_3$	$CDCl_3$	0.025	1.96 (s)	2.01 (s)	2.65 ± 0.03
$CH_3CONH-CH_3$	$CDCl_3$	0.025	2.78 (d)	2.86 (d)	2.87 ± 0.04
$CH_3CONH-CH_3$	C_6D_{12}	0.020	1.75 (s)	*b*	<2.0
$CH_3CONH-CH_3$	C_6D_{12}	0.020	2.67 (d)	*b*	<2.0

[a] Abundance of the cis isomer, relative to the total abundance of both isomers, estimated from integrated signal intensities at the resonances indicated. [b] At this concentration, near the solubility limit in water, no signal could be observed from the cis isomer. The signal-to-noise ratio of these samples would have allowed the cis isomer to pass undetected if its relative abundance had been less than approximately 2%, so that this can be considered an upper limit for its abundance.

Results

Concentrations of solutions, chemical shifts for the more abundant trans isomer, and the ratio of integrated intensities observed for the cis and trans isomers are given in Table I. At the concentrations indicated, there were no signs of aggregatation of these solutes in any of the solvents examined. However, when concentrations were elevated above 0.4 M in cyclohexane solution, new peaks began to appear, in accord with the known tendencies of these compounds to aggregate at high concentrations (Graham & Chang, 1971). In water, chloroform, and cyclohexane, proton spectra of *N*-methylformamide showed methyl resonances for the cis isomer that were shifted downfield from those for the trans isomer, whereas the resonance of the formyl proton of the cis isomer was situated slightly upfield from that of the trans isomer, (Table I). In benzene, these differences were larger, and in this exceptional solvent, resonances of the cis methyl protons were shifted upfield from those of the trans isomer.

With changes in solvent polarity, there was a striking absence of significant variation in the cis/trans equilibrium of *N*-methylformamide. Values ranged from 8% in water to approximately 9% in cyclohexane. In the case of *N*-methylacetamide, the rarity of the cis isomer limited the accuracy with which its abundance could be determined. The abun-

dance we observed by ^1H NMR for the cis isomer in water, 1.5%, was much lower than that reported by Liler (1972), 7%. We do not understand this discrepancy but believe that it may have arisen from the difficulty of removing small amounts of contaminating acetic acid from *N*-methylacetamide by simple vacuum distillation, the method of purification that was apparently used by Liler. Our material was further purified by zone melting and mixed-bed ion-exchange chromatography (see Methods) to remove the last traces of methylamine and acetic acid and contained no peak that might have corresponded to the substantial amounts of cis isomer reported by Liler. In the earlier work, overlapping of resonances from this material and the amide may have led to the overestimation of the cis isomer.

Whereas our proton magnetic resonance results for *N*-methylacetamide showed appreciable experimental error, we were able to obtain a more definitive result by carbon magnetic resonance spectroscopy, using material 99% enriched with ^{13}C at the *N*-methyl group (Table I). This material was found to contain 1.46 ± 0.09% of the cis isomer in water, in good agreement with the average value obtained by proton magnetic resonance. In chloroform, the amount of cis isomer increased to 2.8%. No cis resonance could be observed in cyclohexane, but the signal-to-noise ratio placed limits on our ability to detect the cis isomer in this solvent, allowing us only to place an upper limit of approximately 2% on the amount of cis isomer that could have been present. Thus, the results for *N*-methylacetamide suggest greater variability in abundance of the cis isomer than was observed for *N*-methylformamide but that there should be no systematic increase or decrease with changing solvent polarity.

Our quantum mechanical calculations used the 6-31G** basis set of the Gaussian 82 ab initio program (Binkley et al., 1986) on a Cray XMP computer. This basis set places polarization functions on all atoms, whereas a previous study (Fogarasi et al., 1979) used a less extensive ab initio basis set without polarization. Full geometry optimizations were performed for both *N*-methylformamide and *N*-methylacetamide in a number of starting configurations. The trans form was found to be lower in energy, by 1.073 kcal/mol for *N*-methylformamide and by 2.459 kcal/mol for *N*-methylacetamide. In the absence of entropy effects, these energy separations would suggest that the cis isomer should be present to the extent of 13.7% in *N*-methylformamide and 1.45% in *N*-methylacetamide. These values are not very different from the experimental values we observe in several solvents (Table I) [Fogarasi et al. (1979) reported a trans preference of 2.0 kcal/mol for *N*-methylformamide]. The resonable agreement observed here suggests that entropic contributions to the relative free energies of cis and trans isomers are minor. To evaluate these entropy contributions would require a full quantum mechanical calculation in solvent water, a task that cannot be undertaken reliably at present.

Our computations indicate that in the low-energy forms of *N*-methylformamide, the *N*-methyl hydrogen atoms are oriented gauche to the amide hydrogen. In the low energy forms of *N*-methylacetamide, the *N*-methyl group is oriented in such a way that one of its hydrogen atoms is eclipsed by the amide hydrogen atom, the other methyl group being eclipsed by the C=O bond. Apart from the methyl hydrogen atoms, both molecules were essentially planar. In the cis isomer of *N*-methylacetamide, there appears to be less steric hindrance than in the trans isomer, but this difference is somewhat less in *N*-methylformamide. Thus, the geometry changes in going from cis to trans were larger in the case of *N*-methylacetamide than in the case of *N*-methylformamide, as indicated by the calculated bond angles in Table II.

Table II: Bond Angles[a] for *N*-Methylamides

R	bond angle (deg) trans <C–N–R′	bond angle (deg) cis <C–N–R′
CH_3	121.37	127.33
H	123.28	125.22
R	bond angle (deg) trans <C–N–H	bond angle (deg) cis <C–N–H
CH_3	119.43	113.92
H	117.75	115.52

[a] Calculated for atoms as described in the text, with the chemical groups denoted in the structures shown in the text.

Calculated dipole moments of minimum energy forms of *N*-methylformamide were 4.296 D for the cis isomer and 4.030 D for the trans isomer. Dipole moments of minimum energy forms of *N*-methylacetamide were 4.213 D for the cis isomer and 4.032 D for the trans isomer. These small differences are in reasonable accord with the observed lack of a substantial solvent effect on the cis/trans equilibrium. In each case, the calculated dipole vector was oriented roughly parallel to the C → O vector.

A personal communication from Drs. W. L. Jorgensen and J. Gao of Purdue University indicates that they have arrived at similar conclusions from ab initio molecular orbital calculations and Monte Carlo statistical mechanics simulations.

Discussion

The observed and calculated positions of peptide cis/trans equilibria suggest that the strong preponderance of the trans configuration in the backbones of proteins is an intrinsic property of the peptide bond, independent of the polarity of the environment. This preference should be reflected not only in the structures of proteins but also in those of much smaller compounds with biological activities, such as neuropeptides. Thus, even in molecules as simple as dipeptides, 98–99% of the population can be expected to adopt the trans configuration.

The small difference between the calculated dipole moments of the isomers agrees with our experimental finding that the cis/trans equilibrium of *N*-methylformamide is almost completely insensitive to the solvent environment, changing by a factor of only 15% when *N*-methylformamide is transferred from water (dielectric constant 80.1) to cyclohexane (dielectric constant 2.0). Thus, although the total free energy of solvation of the peptide bond is approximately −10.0 kcal/mol (Wolfenden, 1976, 1978), free energies of solvation of the cis and trans isomers appear to differ by less than 0.1 kcal/mol. As a consequence, the transfer of a peptide bond from the interior of a protein to an exposed position on its surface would not, in itself, be expected to change its intrinsic preference for the cis or trans configuration. Conversely, other factors that might force such a bond to adopt the cis or trans configuration would not be expected to influence its tendency to seek a buried or exposed condition. Similarly, natural and synthetic peptides should continue to exhibit a strong preference for the trans configuration when, as may often be the case, they are bound by receptors in a local environment less polar than water.

These conclusions apply only to equilibria, in the alternative environments represented by dilute solution in water and in the nonpolar solvents of the kinds considered here. In a solvent or microenvironment capable of acting only as a donor, or only as an acceptor, of hydrogen bonds to the solute, the positions of cis/trans equilibrium might change. Nor is the *rate* of

peptide isomerization expected to be insensitive to the polarity of the environment. Recently, a "prolyl isomerase" has been found to produce modest enhancements in the rate of cis–trans isomerization of proline imidic peptide bonds in oligopeptides and proteins (Lang et al., 1987). In the transition state for peptide isomerization, there is presumably a major loss of resonance stabilization and susceptibility to polar interactions. Indeed, Drakenberg et al. (1972) have demonstrated that peptide isomerization becomes much more rapid in nonpolar environments than in water. It seems likely that the catalytic effects of prolyl isomerase might be produced, at least in part, by withdrawal of the susceptible peptide bond from water into an active site of lower polarity.

ACKNOWLEDGMENTS

We acknowledge helpful discussions with Drs. David Deerfield, III, Jan Hermans, David Richardson, and Jane Richardson and thank Dr. Jorgensen for sending us a manuscript in advance of publication. L.P. also thanks the NIEHS (DBRA/OD), Research Triangle Park, NC, for a summer appointment during which part of this work was done, and the NCI, Frederick, MD, for access (through NIEHS) to the Cray XMP facility.

Registry No. $HCONHCH_3$, 123-39-7; $CH_3CONHCH_3$, 79-16-3.

REFERENCES

Binkley, J. S., Frisch, M. J., DeFrees, D. J., Rahavachari, K., Whiteside, R. A., Schlegel, H. B., Fulder, E. M., & Pople, J. A. (1986) copyright, Carnegie Mellon University, Pittsburgh, PA.

Corey, R. B., & Donohue, J. (1950) *J. Am. Chem. Soc. 72*, 2899.

Drakenberg, T., Dahlqvist, K.-I., & Forsén, S. (1972) *J. Phys. Chem. 76*, 2178.

Fogarasi, G., Pulay, P., Torok, F., & Boggs, J. E. (1979) *J. Mol. Struct. 57*, 259.

Graham, L. L., & Chang, C. Y. (1971) *J. Phys. Chem. 75*, 776.

Jorgensen, W., & Swenson, C. J. (1985) *J. Am. Chem. Soc. 107*, 1489.

Knecht, L. A. (1971) *Pure Appl. Chem. 27(1–2)*, 281.

Liler, M. (1972) *J. Chem. Soc., Perkin Trans. 2*, 720.

Nakanishi, H., & Roberts, J. D. (1981) *Org. Magn. Reson. 15*, 7.

Radzicka, A., & Wolfenden, R. (1988) *Biochemistry 27*, 1664–1670.

Ramachandran, G. N., & Mitra, A. K. (1976) *J. Mol. Biol. 107*, 85.

Rees, D. C., Lewis, M., Honzatko, R. B., Lipscomb, W. N., & Hardman, K. D. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3048.

Stewart, W. E., & Siddall, T. H., III (1970) *Chem. Rev. 70*, 517.

Wolfenden, R. (1976) *J. Am. Chem. Soc. 98*, 1987.

Wolfenden, R. (1978) *Biochemistry 17*, 201.

Solubility of Different Folding Conformers of Bovine Growth Hormone

David N. Brems

Control Division, The Upjohn Company, Kalamazoo, Michigan 49001

Received December 28, 1987; Revised Manuscript Received February 22, 1988

ABSTRACT: Previous folding studies have shown that equilibrium denaturation of bovine growth hormone (bGH) is a multistate process with stable intermediates. The native and unfolded species are monomeric, but intermediates are both monomeric and associated. In this study, the relative insolubility of an associated intermediate is used to distinguish its presence in equilibrium denaturation and during kinetic refolding. To study the role of the associated intermediate in the refolding pathway, a two-step procedure for its detection was developed. The first step of this procedure is used to populate the associated intermediate, and the second step involves dilution to solvent conditions in which only the associated intermediate precipitates. The amount of precipitate is quantitated either directly by formation of turbidity or indirectly by quantitation of the remaining soluble protein. The results show that an associated species is transiently populated during folding, it is incorrectly folded, and it occurs due to specific interactions of monomeric folding intermediates at moderate to high protein concentrations. This association of intermediates is a competing reaction that decreases the folding rate. The location of this competing reaction in the refolding pathway occurs after the formation of an early framework-type intermediate that contains considerable secondary structure but prior to the rate-limiting formation of the native tertiary structure. When refolding occurs in solutions that solubilize the associated intermediate, then native protein is obtained quantitatively. However, if refolding occurs in solutions that do not solubilize the associated intermediate, then most of the product results in an insoluble protein aggregate. The formation of precipitate that occurs upon refolding is inhibited by addition of fragments 96–133 or 109–133 that are derived from bGH. It is suggested that these fragments prevent precipitation by binding to the framework-type intermediate in a manner that prevents it from participating in the association reaction. The relationship of these results to general pathways of protein precipitation is discussed.

The solubility of proteins in aqueous solutions is a physical property of fundamental biological importance. Solubility occurs when solute (protein) to solvent (water) interactions are more favorable than solute to solute interactions. Insolubility occurs when the solute to solute interactions are more favorable than the solvent interactions. Protein insolubility at the molecular level is poorly understood and frequently causes experimental difficulties for protein studies in general.

0006-2960/88/0427-4541$01.50/0