*Proceedings of the National Academy of Sciences*
Vol. 67, No. 1, pp. 239–246, September 1970

# An Approach to Conformational Analysis of Peptides and Proteins in Solution Based on a Combination of Nuclear Magnetic Resonance Spectroscopy and Conformational Energy Calculations

**W. A. Gibbons, George Némethy*, Arnold Stern†, and Lyman C. Craig‡**

THE ROCKEFELLER UNIVERSITY, NEW YORK, N.Y. 10021

*Communicated June 11, 1970*

**Abstract.** Simple criteria, based on the combined use of nmr spectral parameters and potential energy maps, are proposed for the conformational analysis of polypeptides and proteins. Experimentally determined coupling constants $^3J_{NC}$ for the N–$C^\alpha$ bond are consistent with the Karplus-Bystrov relationship. It is proposed therefore that $^3J_{NC}$ can be used to distinguish (*a*) between right- and left-handed $\alpha$-helices, (*b*) between $\alpha$-helical, $\beta$-pleated sheet, and randomly coiled forms of peptides. The average $^3J_{NC}$ for the random coil is predicted. The criteria proposed are valid for both L- and D-amino acids. Correlation between the Karplus-Bystrov relationship for $^3J_{NC}$ and the peptide conformational potential energy map limits the possible values of the N–$C^\alpha$ dihedral angle $\phi$ of each amino acid residue in a polypeptide and protein, and therefore presents a method of conformational analysis in solution superior to the use of either nmr or conformational maps alone. Nmr studies of hydrogen bonding or neighboring-group diamagnetic anisotropy reduce the number of possibilities consistent with the above criteria. A suggestion for evaluating the dihedral angle is presented. These criteria are useful provided the coupling constant is not obscured by line broadening.

---

Simple accurate rules for determining details of the conformation of proteins and peptides would be of great help in understanding their chemistry. X-ray diffraction provides the type of information needed to derive such rules for materials in the crystalline form but cannot be used for those solutes which crystallize poorly or for determining the conformation of molecules in solution. Methods capable of revealing the subtle changes in conformation brought about by a change in solution environment or temperature are also needed.

The requirements obviously demand a combination of techniques which give information not only on the overall size and shape of the molecule but also on local structural features. Nuclear magnetic resonance (nmr) which has the intrinsic potential of revealing something of the spatial arrangement of the immediate neighbors of any magnetic nucleus in the molecule, at present seems to be the most promising spectroscopic technique available.

The difficulty in the use of nmr stems from the numbers of overlapping resonances. Although advances in resolution and sensitivity have been made re-

cently, nmr studies have not yet resulted in the establishment of any side-chain or backbone conformation in a protein. Such information can be obtained for the smaller peptide antibiotics and hormones. These can be considered "mini-proteins" and can serve as excellent models with which to establish rules for the interpretation of nmr and other physical parameters in terms of conformation.

The complex decapeptide gramicidin S-A was the first to be studied in detail by nmr.[1] In that study, nmr criteria were used to establish (*a*) the presence of a $C_2$ axis of symmetry, (*b*) that Val-Orn-Leu tripeptides form an anti-parallel $\beta$-pleated sheet structure linked by two Pro-Phe dipeptides, and (*c*) the relationship of amide exchange rates to their stereochemical environment. The "anomalously high field" chemical shift of the valine amide protons was explained by invoking diamagnetic anisotropic shielding by neighboring carbonyl groups. The results of this study were confirmed by other groups.[2-7]

The approximate approach to conformational analysis of Stern *et al.*[1] has also been applied to oxytocin,[8,9] actinomycin[10] and some synthetic peptides.[11-13] Nmr studies of valinomycin, a depsipeptide, have been reported.[2,14,15]

We propose here that a diagram such as those of Figs. 1*A* and *C* describing the relationship between $^3J_{NC}$, the CH–NH vicinal coupling constant, and the CH–NH dihedral angle $\phi$ can give well-defined structural information when used in conjunction with the ($\phi$, $\psi$) conformational energy maps (Figs. 1*B* and *D*). Possible correlations between nmr parameters and hydrogen bonding are also discussed.

**I. Validity of the Karplus-type Relationship for the CH–NH Dihedral Angle.** The conformations of the peptide units are described according to the new rules, recommended by the IUPAC-IUB Commission on Biochemical Nomenclature, for the description of the conformation of polypeptide chains.[16] These represent a modification of the rules proposed by Edsall *et al.*[17]

A relationship between the CH–NH coupling constant $^3J_{NC}$ and the corresponding dihedral angle for rotation around the C–N bond can be derived on the basis of published studies of model compounds. For *N*-alkyl formamides,[18] three stable rotational isomers can be assumed.

In each isomer the N–H bond is staggered with respect to the substituents on the *C*-atom.[19] We also assume that only two coupling constants have to be considered for each isomer, namely, $^3J_{NC}{}^t$ and $^3J_{NC}{}^g$, with the indices *t* and *g* indicating *trans* and *gauche*, respectively.

For *N*-methyl formamide, $R_1 = R_2 = H$, and the populations of all three rotational isomeric states are equal. The measured value of the coupling constant[18] represents a weighted average over all rotational states:

$$\langle {}^3J_{NC} \rangle = {}^1/_3 \, (2\,{}^3J_{NC}{}^g + {}^3J_{NC}{}^t) = 5.0 \text{ Hz}$$

The measured values are $\langle {}^3J_{NC} \rangle = 6.0$ Hz for *N*-ethyl formamide ($R_1 = CH_3$, $R_2 = H$) and $\langle {}^3J_{NC} \rangle = 7.7$ Hz for *N*-isopropyl formamide ($R_1 = R_2 = CH_3$). These numbers require that the relative population of the *trans* isomer(s) increase progressively with higher substitution. They do not allow the determination of unique values of $^3J_{NC}{}^t$ and $^3J_{NC}{}^g$ (except for the restriction that $^3J_{NC}{}^t > 2\,{}^3J_{NC}{}^g$). The results are consistent with the values (compare below) of $^3J_{NC}{}^g \approx 3$ Hz and $^3J_{NC}{}^t \approx 9$ Hz.

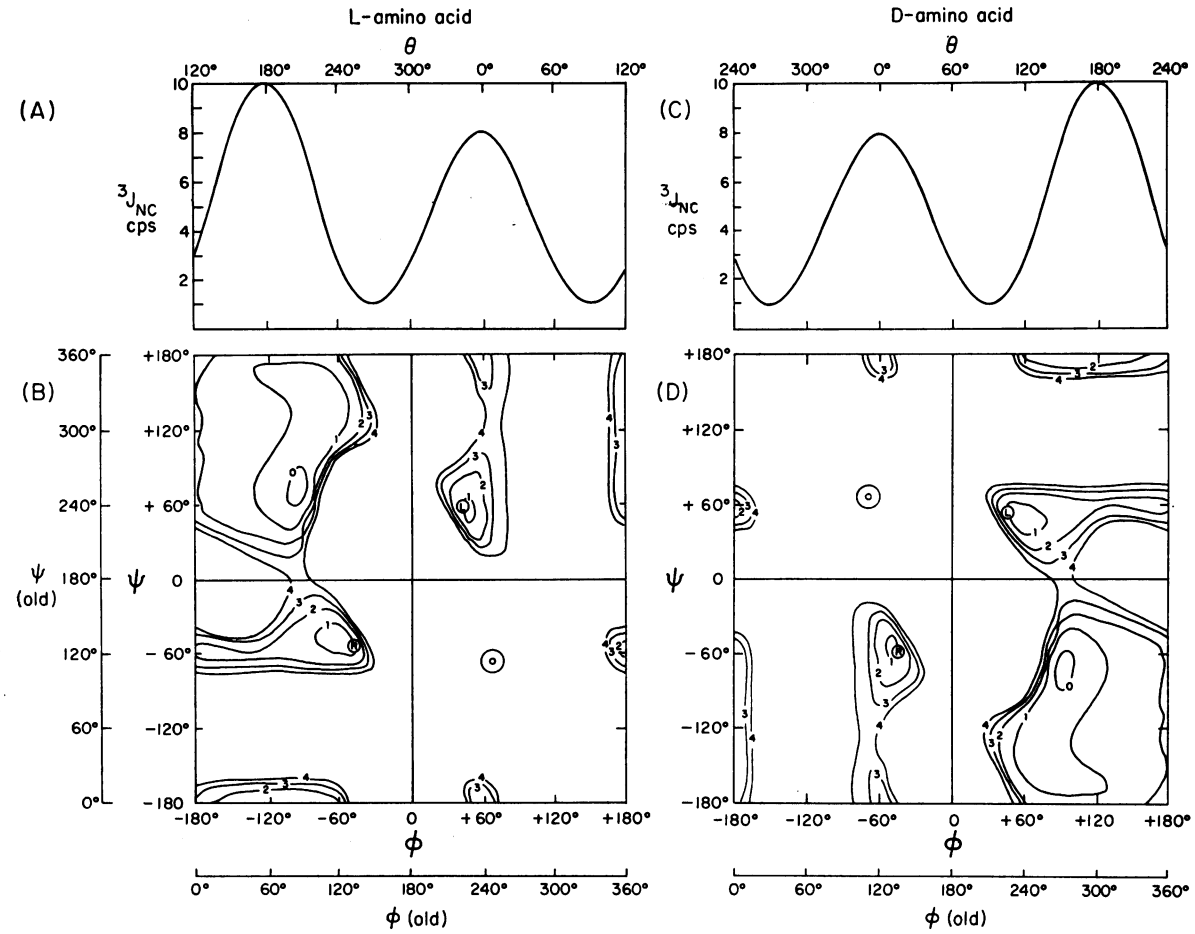



FIG. 1. Correlation of the Karplus-type relationship[21,22] (*A* and *C*) between the constant $^3J_{NC}$ and the N–C dihedral angle $\phi$ with the $(\phi, \psi)$ conformational energy map[19] (*B* and *D*). Scales of $\phi$ and $\psi$ agree with the rules adopted in 1970 by the IUPAC-IUB.[16] [The $\phi$ and $\psi$ in the widely used system of conventions proposed[17] in 1966 differ from the new rules by $\pm 180°$. For comparison with earlier data, the old definitions are indicated on separate scales labeled as $\phi$ (old) and $\psi$ (old).] Parts *B* and *D* have been adopted from refs. 19–21. Energy contours are given at intervals of 1 kcal/mol.

This value of $^3J_{NC}^g$ is also consistent with the data of Kopple[11] on cyclic dipeptides, although no exact comparisons should be made between compounds containing *trans* and *cis* peptide units. Factors such as the geometry of the peptide bond, conformational mobility, and geminal bond angle variations also affect $^3J_{NC}$.

Bystrov *et al.*[21] postulated a relationship between dihedral angle $\phi$ and $^3J_{NC}$, similar to the Karplus relationship.[22] Their arguments were based on the combined infrared and nmr study of *N*-acyl dialanyl methyl esters, with variable degrees of methylation of the amide NH groups.

Further confirmation that a Karplus-type relationship might be applicable to the CH–NH of peptides also comes from the analogy with the work of Bothner-By[23] on alkylethylenes and of Garbisch[24] on vinyl-alkyl coupling relationships.

A good test of the validity of the Karplus-Bystrov relationship between $\phi$ and $^3J_{NC}$ is the nmr study of polypeptides of known conformation. The $\phi$ values in the common conformations are known.[16] If the $^3J_{NC}$ read from the Karplus-Bystrov curve for these angles agrees with the experimentally determined $^3J_{NC}$, it can be assumed that the relation holds for all peptides with *trans* peptide bonds. This could make the curve of Fig. 1 important for determining an unknown $\phi$ in a peptide or protein and thus for determining the local conformation at an amino acid residue, i.e., the conformation of both regular and disordered regions in a polypeptide or protein. Jackman[25] has cautioned against this use of Karplus-type relations, but we feel that this procedure, although approximate now, can eventually be refined, as shown here.

**II. The Correlation Between Coupling Constants and Peptide Backbone Conformation.** On the basis of Figs. 1*A* and *B* for L-amino acids, the right-handed $\alpha$-helix is assigned $^3J_{NC} \approx 2 \pm 1$ Hz, the left-handed $\alpha$-helix $^3J_{NC} \approx 8 \pm 1$ Hz, the antiparallel $\beta$-sheet $^3J_{NC} \approx 9 \pm 1$ Hz, and the parallel $\beta$-sheet $^3J_{NC} \approx 10 \pm 1$ Hz. For D-amino acids the assignments for these conformations, based on Figs. 1*C* and *D*, are $8 \pm 1$ Hz, $2 \pm 1$ Hz, $9 \pm 1$ Hz, and $10 \pm 1$ Hz, respectively. For the random coil $^3J_{NC}$ can be estimated as the average over all accessible conformations $(\phi, \psi)$, by weighting each $^3J_{NC}$ $(\phi, \psi)$ with a Boltzmann factor $\exp[-E(\phi, \psi)/kT]$, where $E(\phi, \psi)$ is the potential energy obtained from conformational energy calculations. Using in this procedure the potential energy map calculated by Scott and Scheraga[19] for an alanyl residue (Fig. 2 of ref. 19), an average $\langle ^3J_{NC} \rangle \approx 7 \pm 1$ Hz is obtained for the random coil. While the numbers cited are all approximate, primarily due to the uncertainty about the detailed shape of the curves in Figure 1, the differences in coupling constants between various conformations seem fairly large. Therefore, if only these common conformations have to be considered, it should be possible to assign the dihedral angles of an amino acid residue on the basis of a measurement of the coupling constant.

It should be noted that the value of the random coil was obtained for an amino acid residue actually possessing the full extent of freedom of internal rotation[26] allowed by the potential energy restrictions of Figs. 1*B* and *D*, and does not apply to an amino acid residue in a nonregular but rigidly fixed "nonhelical" region of a protein. The latter would exhibit a $^3J_{NC}$ corresponding to the particular value of its conformational angle $\phi$.

All these predictions must be regarded today as qualitative. However, they lay the basis for further studies. When more peptide structures, determined crystallographically, become available, together with more extensive information about coupling constants in peptides of known conformation, this approach can be further refined. The applicability of the predictions depends on the further assumption that the coupling constant $^3J_{NC}$ is not obscured by broadening due to the relaxation time of the $^{14}N$ nucleus, exchange processes, or dipolar broadening. However, these are points which can be checked in each case.

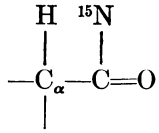
**III. Correlations Between Experiments and Predictions.** Examination of some of the pertinent literature for the helix–coil transition[27,28] in which attention was paid to measuring coupling constants, shows that $^3J_{NC}$ is about 6.5 ± 0.5 Hz for the random coil and about 3 ± 1 Hz for the right-handed $\alpha$-helical form of a given homopolymer. The latter value is very inaccurate since dipolar line broadening can be of the same order as $^3J_{NC}$. There are no authenticated $^3J_{NC}$ values for $\beta$-conformations except perhaps gramicidin S-A. If it is regarded as a genuine antiparallel $\beta$-structure then $^3J_{NC}$ for this conformation should be approximately 8.5 Hz. Thus while few good values seem to be available, the measured coupling constants $^3J_{NC}$ for the various common conformations are consistent with the Karplus-type curve (Fig. 1*A*) and the latter curve can be used with some confidence to distinguish these conformations in solution.

On the other hand, a high $^3J_{NC} \approx 8$ Hz can be predicted from Figure 1*A* for the left-handed $\alpha$-helix. Thus the presence or absence of a large value of $^3J_{NC}$ in a polypeptide known to have $\alpha$-helical structure can be used as a criterion to decide the handedness of the helix. No experimental information is available to confirm this prediction, but we are carrying out tests to confirm it.

Since the prediction for the random coil is $^3J_{NC} \approx 7$ Hz, it should be possible to distinguish it from the right-handed $\alpha$-helix. Distinction between the random coil and the left-handed $\alpha$-helix ought to become possible only when a more exact relationship between $^3J_{NC}$ and $\phi$ becomes available, since $\alpha_L$ is predicted to have a $^3J_{NC}$ slightly higher than the random coil.

A more exact relationship should allow analysis of the occurrence of preferred structures in flexible portions of a polypeptide chain. Since there is a correlation between the magnitude of $^3J_{NC}$ as a function of $\phi$ and the location of low energy conformations of an amino acid residue (Fig. 1), the comparison of experimentally obtained $^3J_{NC}$ values with calculations of weighted averages based on the conformational map could indicate whether a particular amino acid residue experiences further restrictions (beyond those incorporated in the conformational map) due to specific local interactions.

Two groups of workers have used chemical shifts of amide protons to help in conformational analysis. Stern *et al.*[1] postulated that the anomalously high field chemical shift of the valine amide protons is the result of their being in the shielding region of neighboring carbonyl groups. The carbonyl group which dominates this chemical shift is probably that of the phenylalanine located in the Phe-Pro sequence joining the two antiparallel bonded chains formed by the Val-Orn-Leu tripeptides in gramicidin S-A. The amide proton of asparagine in oxytocin[8,9] experiences a similar high field chemical shift because it is in the position analogous to the valine amide protein of gramicidin S-A.

**IV. A More Rigorous Approach to Conformational Analysis Combining nmr and ($\phi$, $\psi$) Energy Maps.** In general four possible $\phi$ angles may correspond to a given experimentally determined $^3J_{NC}$. We propose that the combined use of the Karplus-Bystrov relationship and the ($\phi$, $\psi$) plots (Fig. 1) can reduce this uncertainty[29] in most cases to two values of $\phi$. The approximate values obtained in this manner can serve as the basis for a refined energy minimization for the peptide or protein in question and yield a much more accurate calculation of the most probable conformation in solution. These ideas can be extended to the bond angle $\psi$ through the evaluation of the coupling constant $^3J^{15}_{NH}$ in the fragment  thus permitting conformational analysis of the whole back-bone of the polypeptide or protein.

**V. Hydrogen Bonding and Conformation.** The conformation of the back-bone of a protein or polypeptide can be specified if both dihedral angles $\phi$ and $\psi$ are known for each residue. (Complete specification of the conformation requires knowledge of the side-chain dihedral angles $\chi$ as well.) Since the preceding considerations apply only to $\phi$, they must be supplemented (in the absence of nmr information on $\psi$) by other data related indirectly to the conformation, if the conformation of a polypeptide chain is to be defined accurately. Such auxiliary information in some cases may be obtained from hydrogen exchange measurements or ring current effects between side chains.

The occurrence of many slowly exchanging hydrogens in proteins has been attributed generally to hydrogen bonding or to protection from access by solvent, or both.[30] There have been attempts to relate differences in exchange rates to conformation. For example, the $\alpha$-helical and random coil forms of poly-amino acids may have exchange rates[31] differing by a factor of the order of $10^3$. The usual measurements of exchange rates yield limited information on local stereochemistry, because the exchange rates cannot be assigned to specific protons or to localized conformational features. This problem can be circumvented in part by using nmr to measure hydrogen exchange rates since the particular protons belonging to classes with various exchange rates may be identified. While this in itself would be insufficient for the identification of *both* residues participating in the hydrogen bond, combined use of other information may be sufficient to assign the hydrogen-bonded amino acids. Nmr data on hydrogen exchange were so used in the analysis[1,2] of gramicidin S-A. Kopple and Ohnishi[12] have extended this type of study, although they could not find differences in rates between internally hydrogen-bonded protons and protons hydrogen-bonded with the solvent. They did not, however, study rates in the range $10^{-4}$ to $10^{-1}$ sec. Ohnishi and Urry[2] also studied gramicidin S-A and valinomycin.

The temperature dependence of the chemical shifts is also expected to be indicative of the presence of intramolecular hydrogen bonding. In studies of gramicidin S-A[1,2], synthetic hexapeptides, and valinomycin[2] it has been assumed that a small temperature coefficient indicated intramolecular hydrogen bonding and vice versa. Exchange rates and temperature-dependence of chemical shifts

can be a measure of strengths of hydrogen bonds, protection from solvent, and intramolecular versus intermolecular hydrogen bonds, but generally are not sufficient to establish a particular conformation in the absence of the more specific information on individual dihedral angles.

**VI. Conclusions.** It has been shown that the available proton–proton coupling constants $^3J_{NC}$ for the vicinal dihedral angle $\phi$ in the $\alpha$-helical, $\beta$-pleated sheet, and random-coil forms of peptides are consistent with a Karplus-type relationship between $\phi$ and $^3J_{NC}$. On the basis of this it is suggested that nmr can be used to (*a*) distinguish the right-handed $\alpha$-helix ($\alpha_R$) from the left-handed $\alpha$-helix ($\alpha_L$) and (*b*) distinguish between $\alpha$-helical, $\beta$-pleated sheet, and randomly coiled forms of peptides. It also can provide a means of studying conformational averaging in random-coil structures, and could eventually yield detailed information on protein conformation by providing the values of the dihedral angles determining local conformations of the polypeptide chain. The data on $^3J_{NC}$ for the various common conformations have been available for some time but, to our knowledge, no one has used them to test the validity of the Karplus-type relationship or suggested that $^3J_{NC}$ can offer a means to distinguish these conformations.

The simple approach to conformational analysis of peptides and proteins in which a measured coupling constant is related to dihedral angle through the Karplus-type relationship is ambiguous in that there are in general four possible dihedral angles for every coupling constant. We propose that the combined use of the Karplus plot and the conformational ($\phi$, $\psi$) energy maps (Fig. 1) can reduce the number of possible $\phi$'s significantly, depending on the compound in question. Even greater refinement can be achieved if these latter $\phi$'s are used as a basis for a detailed polypeptide energy minimization.

Another way of resolving, in part, the ambiguity remaining after the use of Figure 1 is the use of supplementary structural information. Exchange rates for amide protons and the temperature dependence of amide proton chemical shifts have been utilized to compensate for lack of knowledge of the dihedral angle $\phi$. These measurements, taken by themselves, do not uniquely specify a conformation. However, the information they provide on hydrogen bonding may permit the resolution of some of the ambiguities in the assignment of $\phi$'s based on Figure 1.

It should be stressed that the main emphasis in this paper is on the method of approach to conformational analysis. The known values of coupling constants are few and often approximate. The Karplus-Bystrov relationship requires more rigorous theoretical and experimental verification. Peptide ($\phi$, $\psi$) conformational maps (whether based on "hard sphere" contact estimates or potential energy calculations) and computations of low-energy conformation of specific polypeptides and proteins contain several approximations, being based on semi-empirically defined interaction parameters. In spite of all of these drawbacks, we feel that even approximate nmr criteria may stimulate rigorous and detailed investigations resulting in refinement of the approach presented here.

* Supported in part by NSF grant GB-8410.
† Supported by NIH postdoctoral fellowship GM 29,420.

‡ Supported in part by NIH grant AM 02493-12 and by NSF grant GB-12278, and grants from the Research Corporation and the Sloan Foundation to a consortium at The Rockefeller University for a 220 MHz nmr facility.

[1] Stern, A., W. A. Gibbons, and L. C. Craig, *Proc. Nat. Acad. Sci. USA*, **61**, 734 (1968).

[2] Ohnishi, M., and D. W. Urry, *Biochem. Biophys. Res. Commun.*, **36**, 194 (1969).

[3] Momany, F. A., G. Vanderkooi, R. W. Tuttle, and H. A. Scheraga, *Biochemistry*, **8**, 744 (1969).

[4] Pysh, E. S., *Science*, **167**, 290 (1970).

[5] Conti, F., *Nature*, **221**, 777 (1969).

[6] Liquori, A. M., and P. De Santis, personal communication.

[7] Schwyzer, R., and U. Ludescher, *Biochemistry*, **7**, 2519 (1968).

[8] Johnson, L. F., I. L. Schwartz, and R. Walter, *Proc. Nat. Acad. Sci. USA*, **64**, 1269 (1969).

[9] Urry, D. W., L. F. Johnson, and R. Walter, personal communication.

[10] Victor, T. A., F. E. Hruska, K. Hikichi, S. S. Danyluk, and C. L. Bell, *Nature*, **223**, 302 (1969).

[11] Kopple, K. D., and D. H. Marr, *J. Amer. Chem. Soc.*, **89**, 6193 (1967).

[12] Kopple, K. D., and M. Ohnishi, *J. Amer. Chem. Soc.*, **91**, 962 (1969).

[13] Kopple, K. D., M. Ohnishi, and A. Go., *J. Amer. Chem. Soc.*, **91**, 4264 (1969).

[14] Ivanov, V. T., I. A. Laine, N. D. Abdulaev, L. B. Senyavina, E. M. Popov, Yu. A. Ovchinnikov, and M. M. Shemyakin, *Biochem. Biophys. Res. Commun.*, **34**, 803 (1969).

[15] Haynes, D. H., A. Kowalsky, and B. C. Pressman, *J. Biol. Chem.*, **244**, 502 (1969).

[16] IUPAC-IUB Commission on Biochemical Nomenclature, Abbreviations and Symbols for Description of Polypeptide Chains, *J. Biol. Chem.*, in press.

[17] Edsall, J. T., P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Némethy, G. N. Ramachandran, and H. A. Scheraga, *J. Biol. Chem.*, **241**, 1004, 4176 (1966); *J. Mol. Biol.*, **15**, 399; **20**, 589 (1966); *Biopolymers*, **4**, 121, 1149 (1966).

[18] La Planche, L., and M. T. Rogers, *J. Amer. Chem. Soc.*, **86**, 337 (1964).

[19] Scott, R. A., and H. A. Scheraga, *J. Chem. Phys.*, **45**, 2091 (1966).

[20] Crippen, G. M., and H. A. Scheraga, *Proc. Nat. Acad. Sci. USA*, **64**, 42 (1969).

[21] Bystrov, V. F., S. L. Portnova, V. I. Tsetlin, V. T. Ivanov, and Yu. A. Ovchinnikov, *Tetrahedron*, **25**, 493 (1969).

[22] Karplus, M., *J. Amer. Chem. Soc.*, **85**, 2870 (1963) and references therein.

[23] Bothner-By, A. A., C. Naar-Colin, and H. Günther, *J. Amer. Chem. Soc.*, **84**, 2748 (1962).

[24] Garbisch, E. W., Jr., *J. Amer. Chem. Soc.*, **86**, 5561 (1964).

[25] Jackman, L. M., *Application of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry* (New York: Pergamon Press, 1969).

[26] Ramachandran, G. N., C. M. Venkatachalam, and S. Krimm, *Biophys. J.*, **6**, 849 (1966).

[27] For review see Roberts, G. C. K. and O. Jardetzky, *Advan. Protein Chem.*, **24**, 447 (1970).

[28] Markley, J. L., D. H. Meadows, and O. Jardetzky, *J. Mol. Biol.*, **27**, 25 (1967).

[29] In some cases, the uncertainty may be reduced even more. Results of a preliminary energy calculation may indicate that additional regions of the $(\phi, \psi)$ map are excluded for specific residues, beyond those shown in Figure 1*B*. This would limit further the ranges of $(\phi, \psi)$ which are consistent with a measured $^3J_{NC}$ value.

[30] Hvidt, A., and S. O. Nielsen, *Advan. Protein Chem.*, **21**, 287 (1966).

[31] Hvidt, A., and S. O. Nielsen, *Advan. Protein Chem.*, **21**, 308 (1966).

[32] Laiken, S L., M. P. Printz, and L. C. Craig, *Biochemistry*, **8**, 519 (1969).