Carbon-13 Nuclear Magnetic Resonance Chemical Shifts and Polypeptide Structure

Alan E. Tonelli

Contribution from Bell Laboratories, Murray Hill, New Jersey 07974. *Received May 27, 1980*

Abstract: ^{13}C NMR chemical shifts of backbone carbonyl and side chain β carbons in polypeptides provide information about their structure. Utilization of substituent effects on ^{13}C chemical shifts, as applied successfully to synthetic organic polymers, makes it possible to rationalize the relative observed ^{13}C NMR chemical shifts of the backbone carbonyl and side chain β carbons which depend on polypeptide sequence and conformation. As examples, in the polypeptide sequence –Gly-X-L-Ala– the carbonyl carbon of residue X resonates at increasingly higher fields in the series X = L-Ala, L-Pro, Gly. Also the carbon of residue X in the sequence –Gly-X-Gly– resonates downfield from its position in –Gly-X-L-Ala–. These and other sequence-dependent ^{13}C chemical shifts can be understood based on substituent effects (β and γ effects). Furthermore, in homopolypeptides the downfield shift of the side chain β carbon resonance observed when passing through the α-helix to the random-coil conformational transition is consistent with the relative ^{13}C chemical shifts estimated via the γ effect method for the β carbon in the α-helical and the random-coil conformations.

^{13}C NMR chemical shifts observed for the carbon atoms in synthetic organic polymers can be understood in terms of the polymer chain microstructure.[1] The effects on ^{13}C chemical shifts of monomer sequence, configuration, conformation, and defect structures can all be explained based on the substituent effects first deduced for paraffinic hydrocarbons.[2-5]

Relative to an unsubstituted carbon, each carbon substituent in the α or β position produces a downfield shift (deshielding effect) of ca. +10 ppm. On the other hand, a γ carbon substituent shields the observed carbon, resulting in an upfield shift of −2 to −3 ppm. This latter shielding effect (γ effect) has been shown[1,5] to not only require a γ substituent, but the observed and γ carbons must be in a gauche (g) arrangement (see Figure 1). Clearly the γ effect on ^{13}C chemical shifts is sensitive to polymer chain conformation.

The conformational sensitivity of the γ substituent effect has been exploited to understand the ^{13}C NMR chemical shifts and the underlying microstructures of synthetic organic homo- and copolymers.[1] In this report we attempt the same approach to learn something about the microstructure and conformation of polypeptides.

Horsely et al.[6] have derived substituent effects for the ^{13}C NMR chemical shifts observed in amino acids (see ref 7 for a refined version). Agreement between observed and estimated ^{13}C chemical shifts was not found to be as good as that achieved for the paraffinic hydrocarbons or their singly substituted derivatives. These authors suggested that polypeptide sequence effects would be less than a few parts per million based on the relatively small δ and ϵ shift parameters they derived for the ^{13}C chemical shifts in amino acids. However, they did not foresee the possibility that the ^{13}C NMR chemical shifts observed in polypeptides might be sensitive to their conformations via γ substituent effects as described here.

In Figure 2 we present a schematic representation of a polypeptide chain illustrating the backbone and side chain torsional angles which determine its conformation. Among the carbon atoms in each peptide residue only the backbone carbonyl and side chain β carbons have γ substituents whose arrangements depend upon the conformation (ϕ,ψ rotations) about the backbone N–C^α and C^α–C' bonds. Since the γ effect on the ^{13}C chemical shifts in synthetic organic polymers has been shown[1] to depend upon the gauche arrangements of γ substituents with respect to the observed carbon atom, we will focus our attention exclusively on the ^{13}C chemical shifts of the backbone carbonyl and side chain β carbon atoms in an attempt to learn something about the polypeptide backbone microstructure, i.e., sequence of residues and backbone conformations.

Conformational Effects on Polypeptide ^{13}C NMR Chemical Shifts

The backbone C–C bonds in synthetic organic polymers are usually constrained[9] by ~3 kcal/mol inherent rotational barriers to adopt one of the three staggered rotational states $t,g^{\pm}$ (ϕ = 0, ±120°) depicted in Figure 1. In a polypeptide chain, on the other hand, the intrinsic barriers to rotation about the N–C^α and C^α–C' backbone bonds are lower (≤1.0 kcal/mol) and nonstaggered rotational states are more prevalent.[9] However, conformational energy maps of the various residues in randomly coiling polypeptides[10] do indicate that the most probable backbone conformations still occur in the vicinity of ϕ or ψ = 0, ±120°, though deviations from these three rotational states can be rather large. Consequently, in our treatment of the γ effect involving backbone carbonyl and side chain β carbons we will estimate the probabilities[11] that ϕ and ψ rotations adopt values which result in gauche arrangements of C' and C^β with other atoms even if ϕ and $\psi \neq$ 0, ±120°.

In Figure 3 we present a portion of a polypeptide chain containing the sequence –Gly_i-Gly_{i+1}– or –Gly_i-L-Ala_{i+1}– in the planar zigzag ($\phi = \psi = 0°$) conformation. (Ali peptide bonds are assumed to be planar and trans.) The carbonyl carbon of Gly_i is γ to the preceding and succeeding carbonyl carbons and to the β-CH_3 carbon of the L-Ala_{i+1} residue in the –Gly_i-L-Ala_{i+1}– sequence. The arrangements of these γ substituents with C'_{Gly_i} depend on the rotations ϕ_i and ϕ_{i+1} about the N–C^α bonds in residues i and $i+1$.

(1) A. E. Tonelli and F. C. Schilling, *Acc. Chem. Res.*, in press.
(2) H. Spiesecke and W. G. Schneider, *J. Chem. Phys.*, **35**, 722-730 (1961).
(3) D. M. Grant and E. G. Paul, *J. Am. Chem. Soc.*, **86**, 2984-2990 (1964).
(4) L. P. Lindeman and J. Q. Adams, *Anal. Chem.*, **43**, 1245-1252 (1971).
(5) F. A. Bovey, "Proceedings of the International Symposium on Macromolecules", Rio de Janeiro, July 26-31, 1974, E. B. Mano, Ed., Elsevier, Amsterdam, 1975, pp 169-182.
(6) W. Horsely, H. Sternlicht, and J. S. Cohen, *J. Am. Chem. Soc.*, **92**, 680-686 (1970).
(7) J. B. Stothers, "Carbon-13 NMR Spectroscopy", Academic Press, New York, 1972, pp 478-482.
(8) J. J. Edsall, P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Nemethy, G. Ramachandran, and H. A. Scheraga, *Biopolymers*, 4, 121-130 (1966).
(9) P. J. Flory, "Statistical Mechanics of Chain Molecules", Wiley-Interscience, New York, 1969, Chapters III, V, and VII.
(10) Chapter VII in ref 9.
(11) When estimating the probability that ϕ or ψ rotations adopt certain values ϕ' or ψ', $\rho(\phi'$ or $\psi')$, leading to γ gauche interactions involving C' and C^β, we take the following approach

$$\rho(\phi') = \frac{\sum_{\phi=\phi'-30°}^{\phi=\phi'+30°} \sum_{\psi=0°}^{360°} \exp(-[E(\phi,\psi)/RT])}{\sum_{\phi=0°}^{360°} \sum_{\psi=0°}^{360°} \exp(-[E(\phi,\psi)/RT])}$$

as the expression used, for example, to estimate the probability that the N–C^α bond is in rotational state ϕ'. Permitting $\phi = \phi' \pm 30°$ in the numerator takes some account[9,10] of the contribution made by the conformational entropy

0002-7863/80/1502-7635$01.00/0

Table I. Comparison of Predicted and Observed ^{13}C NMR Chemical Shifts of Carbonyl Carbons in Randomly Coiling Polypeptides

polypeptide sequence	residue X	^{13}C chemical shift of C'_X pred[a]	obsd[b]	solvent, ref
-Gly-Gly-X-L-Ala-	L-Ala	L-Ala	175.8 (171.8)[c]	D_2O (Pd = 7.0), 13
	L-Pro	L-Pro	174.1	D_2O (Pd = 6.8), 13
	Gly	Gly	171.8 (168.7)[c]	D_2O (Pd = 6.0), 13
-Gly-Gly-X-Gly-Gly-	L-Ala	L-Ala	176.3	H_2O, 14, 15
	L-Pro	L-Pro	175.8	H_2O, 14, 15
	Gly	Gly	172.8	H_2O, 14, 15
tert-butoxycarbonyl-X-L-Pro	L-Ala	L-Ala	172.6	$CDCl_3$, 17
	Gly	Gly	168.0	$CDCl_3$, 17
poly(L-Pro_1-L-Pro_2-Gly_3)		L-Pro_2	175.2	H_2O, 18
		L-Pro_1	173.4	H_2O, 18
		Gly_3	169.2	H_2O, 18
poly(Gly_1-Gly_2-Pro_3-Gly_4)		Pro_3	176.2	0.15 M sodium acetate (Ph = 4.8), 16
		Gly_1, Gly_4	173.0	0.15 M sodium acetate (Ph = 4.8), 16
		Gly_4, Gly_1	173.0	0.15 M sodium acetate (Ph = 4.8), 16
		Gly_2	170.5	0.15 M sodium acetate (Ph = 4.8), 16

[a] Top to bottom corresponds to increasingly upfield chemical shifts. [b] In ppm downfield from tetramethylsilane (Me_4Si). [c] In ppm downfield from Me_4Si as reported by Grathwohl and Wüthrich[19] for Me_2SO-d_6 solutions.

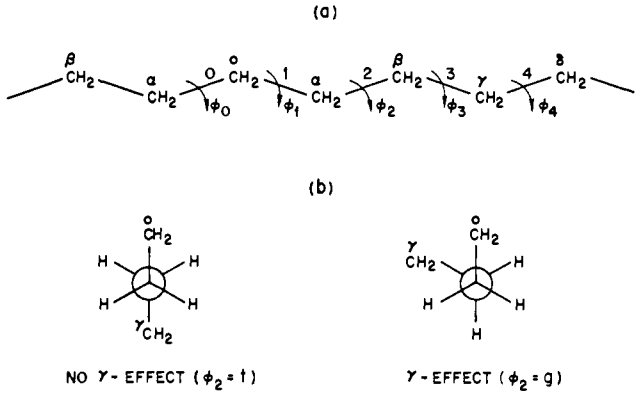

Figure 1. (a) Portion of a paraffinic hydrocarbon chain in the all-trans, planar zigzag conformation. (b) Newman projection along bond 2 in (a) illustrating the γ effect.

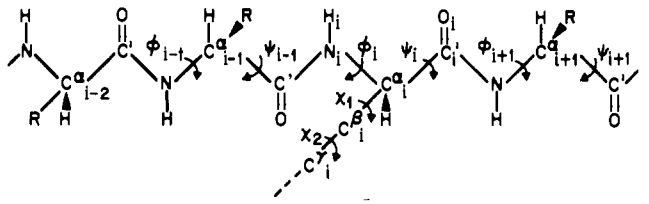

Figure 2. Schematic representation of a polypeptide chain in the planar zigzag conformation where all $(\phi,\psi) = 0°,0°$. We adopt the 1966 definition[8] for polypeptide rotation angles, because it is consistent with the rotation angle convention used for synthetic polymers (see Figure 1 and ref 9).

In a random coil polypeptide containing the $-Gly_i\text{-}Gly_{i+1}-$ and $-Gly_i\text{-}L\text{-}Ala_{i+1}-$ sequences, the conformation (ϕ_i rotation) about the N-C^α bond of the Gly_i residue is independent[10] of whether or not residue $i + 1$ is Gly or L-Ala. However, rotation ϕ_{i+1} about the N-C^α bond of residue $i + 1$ clearly is sensitive[10] to the nature of residue $i + 1$, Gly_{i+1} or L-Ala_{i+1}. From the conformational energy maps[10] appropriate to Gly and L-Ala residues in randomly coiling polypeptide chains, we estimate[11] that C'_i in $-Gly_i\text{-}L\text{-}Ala_{i+1}-$ is gauche to either C'_{i+1} or C^β_{i+1} more than C'_i is gauche to C'_{i+1} in $-Gly_i\text{-}Gly_{i+1}-$.

Assuming that the γ gauche effects (upfield shift) of carbonyl and methyl carbons $i + 1$ upon carbonyl carbon i are comparable,[7,12] then we would expect the C'_i carbon in $-Gly_i\text{-}L\text{-}Ala_{i+1}-$ to resonate upfield from the C'_i carbonyl carbon in $-Gly_i\text{-}Gly_{i+1}-$.

(12) This prediction also holds if the γ effect of C^β_{i+1} on C'_i is larger than the γ effect of C'_{i+1} on C'_i.

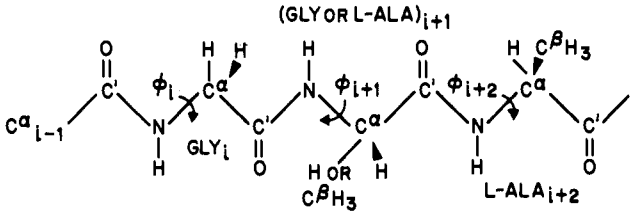

Figure 3. A portion of a polypeptide chain containing the sequence $-Gly_i\text{-}Gly_{i+1}\text{-}L\text{-}Ala_{i+2}-$ or $-Gly_i\text{-}L\text{-}Ala_{i+1}\text{-}L\text{-}Ala_{i+2}-$.

This prediction is borne out by the ^{13}C NMR observations performed on the oligomeric peptides Gly-Gly-X-Ala and Gly-Gly-X-Gly-Gly by Richarz and Wüthrich[13] and Gurd et al.[14,15]

As another example, compare the ^{13}C chemical shifts expected for C'_i in $-Gly_i\text{-}L\text{-}Ala_{i+1}-$ and C'_{i+1} in $-L\text{-}Ala_{i+1}\text{-}L\text{-}Ala_{i+2}-$. Conformational differences between this pair of sequences are the same as in the pair discussed previously. On the basis of γ gauche effects, we might expect C'_{i+1} in L-Ala_{i+1} to resonate upfield from C'_i in Gly_i. The opposite is observed by Richarz and Wüthrich[13] and this is a consequence of the β-CH_3 carbon in the L-Ala_{i+1} residue. β-CH_{3i+1} and C'_{i+1} are β to each other and should result in a substantial downfield shift (deshielding β effect)[7] of C'_{i+1} in L-Ala_{i+1} relative to C'_i in Gly_i. Apparently this downfield β substituent overwhelms the more prevalent upfield γ gauche interactions resulting in the carbonyl carbon C'_{i+1} of $-L\text{-}Ala_{i+1}\text{-}L\text{-}Ala_{i+2}-$ resonating downfield from the Gly_i carbonyl in $-Gly_i\text{-}L\text{-}Ala_{i+1}-$.

Among the residues in the regularly repeatng polypeptide poly(Gly_1-Gly_2-Pro_3-Gly_4), Gly_1 and Gly_4 are both succeeded by glycine residues. Consequently, in the random coil conformation the probabilities that their carbonyl carbons are γ gauche to the carbonyl carbons of the preceding and succeeding residues are the same leading to the prediction of identical $\nu_{C'}$ for Gly_1 and Gly_4.

Gly_2 is succeeded by Pro_3 whose cyclic side group restricts ϕ_{Pro_3} to ca. 120°, resulting in a fixed γ gauche arrangement between C'_{Gly_2} and C'_{Pro_3}. We therefore expect $\nu_{C'_{Gly_2}}$ to be upfield (shielded) relative to $\nu_{C'_{Gly1,4}}$.

The deshielding effect of the β-CH_2 carbon in the Pro_3 side groups should move $\nu_{C'_{Pro_3}}$ downfield from the other carbonyl carbon resonances. With increasing field we would expect to find the following order of carbonyl carbon chemical shifts: $\nu_{C'_{Pro_3}}$, $\nu_{C'_{Gly1,4}}$,

(13) R. Richarz and K. Wüthrich, *Biopolymers*, **17**, 2133-2141 (1978).
(14) P. Keim, R. A. Vigna, R. C. Marshall, and F. R. N. Gurd, *J. Biol. Chem.*, **248**, 6104-4113 (1973).
(15) P. Keim, R. A. Vigna, A. M. Nigen, J. S. Morrow, and F. R. N. Gurd, *J. Biol. Chem.*, **249**, 4149-4156 (1974).

$\nu_{C'_{Gly2}}$. Torchia and Lyerla[16] do observe this order in the random coil polypeptide poly(Gly_1-Gly_2-Pro_3-Gly_4) (see Table I).

Peptide residues with γ substituents in their side chain, such as Val, Leu, Asp, Glu, Met, etc., will generally have carbonyl ^{13}C chemical shifts upfield from the C′ resonances in Ala due to the additional γ effects incurred through χ_1 rotations (see Figure 2) about their C^α–C^β side chain bonds.[13-15] Several more examples of the residue and sequence dependencies of the carbonyl ^{13}C chemical shifts in random coil polypeptides are presented in Table I.

Aside from the possibility of intra-side-chain γ interactions in those residues with long side chains (His, Tyr, Glu, Met, Arg, Orn, Lys, etc.), the side chain β carbon is involved in γ interactions with the C′, O, and N atoms of its own and the succeeding peptide bond. The backbone rotations ϕ and ψ (see Figure 2) determine whether or not the C^β carbon is gauche to the atoms of these peptide bonds. When $\phi_i = 0$, 240° $C^\beta{}_i$ is γ gauche to C'_{i-1}; when $\psi_i = 0$, 120° $C^\beta{}_i$ is γ gauche to N_{i+1}; and when $\psi_i = 180$, 300° $C^\beta{}_i$ is γ gauche to O_i. It is evident that the ^{13}C chemical shift of a residue's side chain β carbon should reflect the residue's backbone conformation (ϕ,ψ) via the γ gauche effect.

As an example, in the conformational transition of a polypeptide from the α-helical to the random-coil state, (ϕ,ψ) change from ≈120°,120° (right-handed α-helix) to all the low-energy values encompassed by the usual (ϕ,ψ) conformational energy map[10] of a randomly coiling peptide residue. From the probabilities[11] that $\psi = 0$, 120, 180, 300° obtained from the conformational energy map appropriate to a randomly coiling peptide residue with a β-CH_2 group in its side chain,[10] it is possible to estimate[20] that the expected difference in the chemical shift of the side chain β carbon in the α-helix and the random-coil states is $\Delta\nu_{C^\beta}(\alpha\text{-helix} - \text{random coil}) \approx 0.6(\gamma_{C^\beta,N} - \gamma_{C^\beta,O})$.

From ^{13}C NMR studies of carbon- and nitrogen-substituted alkanes, as described by Stothers,[7] it is apparent that $\gamma_{C,N}$ is significantly larger than $\gamma_{C,O}$. Clearly then we would expect a downfield shift of the C^β side chain resonance on passing from the α-helical to the random-coil conformation.

Such a downfield shift in the ^{13}C NMR resonance of C^β is observed[21-23] upon disrupting the α-helical polypeptide conformation and passing to the random coil, regardless of whether or not temperature, pH, or salt concentration is the perturbing influence which unwinds the α-helix. The carbonyl carbon ^{13}C NMR chemical shift is not a suitable indicator of the peptide residue conformation in the α-helix to random-coil transition, because the state of the carbonyl oxygen (hydrogen bonded, solvated, or not) also strongly effects the chemical shift of this carbon.

When an L-Pro residue[24] succeeds an amino acid residue with a β-CH_2 or CH_3 group, such as L-Ala, the conformation about the C^α–C′ bond in the residue preceding L-Pro is restricted[10] to $\psi \approx 300°$, while ϕ rotations are unimpeded by the succeeding L-Pro residue. Consequently, the difference in chemical shifts expected at $C^\beta{}_{L\text{-Ala}}$ in the randomly coiling polypeptide sequences –L-Ala-X– and –L-Ala-L-Pro–, where X is not L-Pro, is $\Delta\nu_{C^\beta L\text{-Ala}}$ (–L-Ala-L-Pro---L-Ala-X–) ≈ $0.4(\gamma_{C^\beta,N} - \gamma_{C^\beta,O})$. Since $\gamma_{C^\beta,N} > \gamma_{C^\beta,O}$ we expect $\nu_{C^\beta L\text{-Ala}}$ in –L-Ala-L-Pro– to come upfield from the resonance position in –L-Ala-X–. This expectation is confirmed by the ^{13}C chemical shifts reported in ppm downfield from Me_4Si for $C^\beta{}_{L\text{-Ala}}$ in *tert*-butoxycarbonyl-L-Ala-L-Pro,[17] 16.95 (Me_2SO) and 15.55 (D_2O), and in –Gly-Gly-L-Ala-L-Ala,[13,19] 18.2 (Me_2SO) and 17.7 (D_2O).

Based on the examples discussed in this report, it seems reasonable to conclude that substituent effects (principally γ substituents) on the ^{13}C NMR chemical shifts of the backbone carbonyl and side chain β carbons can be utilized to understand the microstructure of polypeptides. Even though the quantitative details remain to be established, the relative ^{13}C chemical shifts of backbone carbonyl and side chain β carbons already provide us with a means to determine the residue sequence and conformations of polypeptide chains.

(16) D. A. Torchia and J. R. Lyerla, Jr., *Biopolymers*, **13**, 97–114 (1974).
(17) W. Voelter and O. Oster, *Org. Magn. Reson.*, **5**, 547–548 (1973).
(18) R. Di Blasi and A. S. Verdini, *Biopolymers*, **18**, 735–738 (1979).
(19) C. Grathwohl and K. Wüthrich, *J. Magn. Reson.*, **13**, 217–225 (1974).
(20) This estimate of the difference in the chemical shift of C^β in the α-helix and the random-coil conformation ignores the γ effect of C'_{i+1} on $C^\beta{}_i$ when $\phi_i = 0$ or 240°, because the probabilities for these two rotational states are relatively small and because[7] $\gamma_{C^\beta,C'} \ll \gamma_{C^\beta,O}; \gamma_{C^\beta,N}$.
(21) F. A. Bovey, *J. Polym. Sci., Macro. Rev.*, **9**, 1–81 (1975), and references cited therein.
(22) H. J. Lader, R. A. Komoroski, and L. Mandelkern, *Biopolymers*, **16**, 895–905 (1977).
(23) H. Saito, T. Ohki, M. Kodama, and C. Nagata, *Biopolymers*, **17**, 2587–2599 (1978).
(24) We are assuming that the L-Ala·L-Pro and L-Ala·X– peptide bonds are both trans in this discussion.

Determination of the Absolute Rates of Decay of Arylcarbenes in Various Low Temperature Matrices by Electron Spin Resonance Spectroscopy

V. P. Senthilnathan[1] and Matthew S. Platz*

Contribution from the Department of Chemistry, The Ohio State University, Columbus, Ohio 43210. *Received April 11, 1980*

Abstract: The absolute decay rates of diphenylcarbene and fluorenylidene have been measured by ESR. The decay is pseudo first order and arises from reaction of the carbene with the glassy or crystalline host. The kinetics are sensitive to the chemical nature of the matrix, the viscosity of the matrix, the concentration of the diazo precursor, and the history of the sample with respect to photolysis. The signal decay is nonexponential due to site problems in the matrix. The decay can be fitted to either a $t^{1/2}$ or $t^{1/3}$ vs. log I dependence. The predominant carbene decay pathway is by hydrogen atom tunneling through a small barrier. This is indicated by very low Arrhenius parameters and anomalous isotope effects. The kinetic study explains the predominance of hydrogen atom abstraction–recombination products observed by other workers.

The chemistry and spectroscopy of arylcarbenes have been exhaustively studied.[2] The solution chemistry of these species is best interpreted by two reactive states, a very reactive stereoselective singlet and the less reactive, less stereoselective triplet.[3]

0002-7863/80/1502-7637$01.00/0