proportional to the number of polymerized molecules present. A local conversion gradient will therefore introduce a heterogeneous deformation field. The [200] streaking reveals that two-dimensional interfacial defects parallel to the (200) planes initiate and develop during the phase transformation. These two-dimensional defects are probably the result of heterogeneous deformation across an interface at which there is a local fluctuation in conversion.

To further elucidate the nature of the defects contributing to the diffraction streaking and spot broadening, we imaged the microstructural development during the phase transformation by dynamic dark-field (DF) imaging and HREM. Bend contours arise in DF imaging from heterogeneous deformation within a TEM sample. A dramatic change in the morphology of bend contours occurred during the reaction. Initially, we observed bend contours throughout the monomer crystals. We then found that new bend contours were initiated and moved through the crystal during the transformation. The final polymerized crystals were nearly free of bend contours.

The HREM images of the DCHD lattice during the phase transformation are shown in Fig. 3, A and B. The 1.2-nm (010) lattice image moved from the upper left in Fig. 3A down toward the middle in Fig. 3B. These HREM data confirm that the lattice crystallinity is retained during the phase transformation. The shift in the lattice image reflects the reorientation of the polymerizing crystal domains during the transformation. We expect that this low-dose dynamic HREM technique will provide a powerful means to study phase transformation mechanisms on a molecular level in this and other interesting polymer materials.

REFERENCES AND NOTES

1. G. Wegner, *Z. Naturforsch. Teil B* **24**, 824 (1969).
2. M. Schott and G. Wegner, in *Nonlinear Optical Properties of Organic Molecules and Crystals*, D. S. Chemla and J. Zyss, Eds. (Academic Press, Orlando, FL, 1987), vol. 2, chap. III-1.
3. J. Kaiser, G. Wegner, E. W. Fischer, *Isr. J. Chem.* **10**, 157 (1972).
4. R. H. Baughmann, *J. Appl. Phys.* **43**, 4362 (1972).
5. V. Enkelmann, R. J. Leyrer, G. Schleier, G. Wegner, *J. Mater. Sci.* **15**, 168 (1980).
6. R. H. Baughmann, *J. Polym. Sci. Polym. Phys. Ed.* **12**, 1511 (1974).
7. M. Dudley, J. N. Sherwood, D. Bloor, D. J. Ando, *J. Mater. Sci. Lett.* **1**, 479 (1982).
8. M. Dudley, J. N. Sherwood, D. J. Ando, D. Bloor, *Mol. Cryst. Liq. Cryst.* **93**, 223 (1983).
9. R. Sinclair and M. A. Parker, *Nature* **322**, 531 (1986).
10. R. J. Young and J. Petermann, *J. Polym. Sci. Polym. Phys. Ed.* **20**, 961 (1982).
11. R. T. Read and R. J. Young, *J. Mater. Sci.* **19**, 327 (1984).
12. P. M. Wilson and D. C. Martin, *J. Mater. Res.* **7**, 3150 (1992).
13. We grew the DCHD monomer single crystals (2 mm wide and 2 to 3 cm long) by slowly evaporating chloroform solution saturated with DCHD at room temperature for nearly 1 month. These crystals were embedded in an epoxy matrix and then sectioned (nominally 100 nm thick) perpendicular to the [001] direction with a Reichert Ultracut FC4E microtome. We also prepared samples by evaporating dilute DCHD-chloroform solutions of 0.01% DCHD (by weight) on amorphous carbon films.
14. V. Enkelmann, G. Schleier, G. Wegner, H. Eichele, M. Schwoerer, *Chem. Phys. Lett.* **52**, 314 (1977); P. A. Apgar and K. C. Yee, *Acta Crystallogr. Sect. B* **34**, 957 (1978).
15. D. T. Grubb, *J. Mater. Sci.* **9**, 1715 (1974).
16. R. T. Read and R. J. Young, *ibid.* **16**, 2922 (1981).
17. R. J. Young and P. H. Yeung, *J. Mater. Sci. Lett.* **4**, 1327 (1985).
18. The video digitizer is a Scion Video Image 1000 with a NuBus card installed in a Mac II. The image-processing software is Image from the National Institutes of Health, which is available by anonymous FTP from zippy.nimh.nih.gov. Our version was modified for fast Hartley transforms by A. Reeves of Dartmouth College and further modified by J. Mansfield and D. Crawford of the University of Michigan.

19. The work was supported by the College of Engineering at the University of Michigan, the Michigan Memorial-Phoenix project, and the National Science Foundation (NSF) through grant DMR-9024876. We thank J. S. Moore for help in synthesizing the DCHD monomer and the staff of the Electron Microbeam Analysis Laboratory of the University of Michigan. D.C.M. also thanks the NSF for a National Young Investigator award.

30 November 1992; accepted 1 March 1993

Secondary and Tertiary Structural Effects on Protein NMR Chemical Shifts: An ab Initio Approach

Angel C. de Dios, John G. Pearson, Eric Oldfield*

Recent theoretical developments permit the prediction of ^{1}H, ^{13}C, ^{15}N, and ^{19}F nuclear magnetic resonance chemical shifts in proteins and offer new ways of analyzing secondary and tertiary structure as well as for probing protein electrostatics. For ^{13}C, ϕ,ψ torsion angles dominate shielding for Cα and Cβ, but the addition of hydrogen bonding and electrostatics gives even better accord with experiment. For $^{15}N^{H}$, side chain (χ^{1}) torsion angles are also important, as are nearest neighbor sequence effects, whereas for $^{1}H^{N}$, hydrogen bonding is particularly significant. For ^{19}F, weak or long-range electrostatic fields dominate ^{19}F shielding nonequivalencies. The ability to predict chemical shifts in proteins from known or test structures opens new avenues to structure refinement or determination, especially for condensed systems.

It has been known for more than 20 years (*1*, *2*) that the folding of a protein into its native conformation causes large ranges of nuclear magnetic resonance (NMR) chemical shift nonequivalencies to be introduced—about 10 ppm for ^{13}C (*2*, *3*), 30 ppm for ^{15}N (*4*), and 15 ppm for ^{17}O (*5*) and ^{19}F (*6*). However, surprisingly little progress in computing such chemical shifts from known solid- or liquid-state structures has been made, for without these nonequivalencies modern multidimensional NMR studies of protein structure (*7*) would not be possible. An understanding of the origins of chemical shifts is expected to lead to new ways of determining, or at least refining, protein structure.

In principle, chemical shifts can be computed by using ab initio techniques (*8*–*10*), but full ab initio computations on structures with 1000 atoms or more are not currently feasible. However, it seems unreasonable to suppose that the effects of *all* atoms would need to be incorporated into a chemical shielding calculation because nuclear shielding is fundamentally a local phenomenon. We show that good values for ^{1}H, ^{13}C, ^{15}N, and ^{19}F folding-induced shielding can be obtained from quantum chemical methods in which we separate the total shielding, σ_t, into three parts:

$$\sigma_t = \sigma_s + \sigma_l + \sigma_0 \quad (1)$$

The short-range contribution, σ_s, contains shielding contributions that can best be evaluated through full ab initio calculation. The long-range contributions are divided into electrostatic (σ_l) and magnetic (σ_0) contributions and can be evaluated in several ways. Examples of σ_s would be the dependencies of the shielding on torsion angles, bond lengths, bond angles, and strong hydrogen bonding. Sites in helical or sheet segments normally have characteristic ϕ,ψ torsion angles, and the changes in shielding due to these geometrical parameters are caused by the changes in the electronic wave functions near the site of interest, which necessitates full ab initio calculation. Fortunately, these effects propagate through the bonding framework and are therefore very short range. As a result, one need take into account only the local geometry so that only a small number of atoms require basis functions.

For proteins, the external electrical charge field for evaluating σ_l is generated by

Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801.

*To whom correspondence should be addressed.

multipoles in the vicinity of the nucleus of interest, and weak electrical interactions can be evaluated either by incorporating the charge field at the self-consistent field (SCF) level (*11–13*) or by using the multipole-shielding polarizability (MSP) approach (*14, 15*). In our SCF-level calculations, we use partial atomic charges to reflect the protein's electrostatic field because this is one approach that is readily implemented and that does not add to the number of contracted functions. As shown by calculations on simple model systems (*16, 17*), the charge field approach yields results that are in good agreement with full ab initio calculations and with the MSP approach. The σ_0 term includes bulk susceptibility, ring current, and magnetic anisotropy contributions, such as those arising from nonspherical samples, neighboring aromatic side chains, and nearby carbonyl groups. These effects are mainly important for $^1H\alpha$ and can be treated adequately by using classical methods (*18*). For the heavier elements (^{13}C, ^{15}N, and ^{19}F) of most interest here, only the σ_s and σ_l contributions are considered.

Gauge invariance (*9, 19*) is essential for solving the chemical shift problem in proteins because the molecular fragments investigated are large, making shielding calculations susceptible to gauge errors. In the absence of symmetry, it is impractical to incorporate more than ~400 basis functions because computational time scales up very rapidly with the number of contracted functions. Thus we incorporate the effects of σ_l by representing atoms outside a core fragment by partial atomic charges (*20*) in much the same spirit as Cummins *et al.* have done for computing ^{17}O nuclear quadrupole coupling constants in ice [(*11*); see also (*12, 13*)]. Following Jaman *et al.* (*13*) we denote this method as the charge field perturbation–gauge including atomic orbital (CFP-GIAO) approach. For comparison, we have also used an MSP approach (*14, 15*) to estimate σ_l. Here, the derivatives of the shielding with respect to fields and field gradients are the values computed by Augspurger, Dykstra, and co-workers (*15, 17*), and these shielding polarizability coefficients are combined with the average values of the fields and field gradients determined by using the local reaction field–molecular dynamics (LRF-MD) approach of Lee and Warshel (*21*) to give the electrical contributions to shielding, $\Delta\sigma(\Phi)$. This is essentially the electrostatic field approach outlined previously by Buckingham *et al.* (*22*) and Batchelor (*23*).

We first discuss the validity of the CFP-GIAO and MSP-LRF methods for ^{19}F NMR because ^{19}F is the nucleus most likely to be dominated by σ_l. Next, we consider $^{13}C\alpha$, $^{13}C\beta$ shieldings in *Staphylococcal* nuclease (SNase), where we provide the theoretical basis for the ϕ,ψ correlations with shielding that have been observed in experimental databases. These are then improved upon by adding CFP and hydrogen bonding. Finally, we consider ^{15}N and ^1H shielding in the peptide group and show that in this case, χ^1 torsion angle effects may also be significant (for ^{15}N), although direct hydrogen bonding and σ_l effects are also very important for both nuclei.

We tested our ability to evaluate contributions to σ_l with the D-galactose binding protein (GBP) from *Escherichia coli* that had been labeled at five Trp residues (127, 133, 183, 195, and 284) with [5-F]Trp by Luck and Falke (*24*). Each resonance has been specifically assigned via site-directed mutagenesis, and the x-ray structures of *E. coli* and *Salmonella typhimurium* GBP are known (*25*). Folding causes an ~10 ppm chemical shift nonequivalence, which can plausibly be attributed to electrostatic interactions because σ_s for ^{19}F will be constant from one [5-F]Trp to another and σ_0 will be an order of magnitude too small.

We investigated the ^{19}F shieldings in GBP using the CFP-GIAO method. Shielding results for [5-F]Trp^{127} in *E. coli* GBP (containing Ca^{2+} and glucose) are shown in Fig. 1A. In curve a we show the computed shielding as a function of a cutoff radius with atomic resolution (only atoms within the cutoff are included in the calculation). The large oscillations result from formal charge imbalance. Much smaller oscillations are seen if all atoms within a residue are included (Fig. 1A, curve b), but convergence is not manifest, and the dependence of the shielding at large cutoff radii ($R > 10$ Å) appears to be due to residues with charged side chains. Indeed, the large change in shielding at $R = 10$ to 11 Å coincides with the inclusion of two Lys residues within the cutoff, and in the range $R = 15$ to 18 Å, two Lys, one Arg, and three Asp residues are added. We believe the inclusion of full formal charges for ionized side chains is unrealistic because we have shown elsewhere that chemical conversion of Lys NH_3^+ groups to $-NHCOCH_3$ groups causes essentially no change in ^{19}F shielding for [4-F]Trp- and [4-F]Phe-labeled hen egg white lysozyme (*26*), and ^{19}F chemical shifts are more or less pH-independent. We thus repeated the shielding calculations using neutral charge sets for Lys, Arg, Asp, and Glu and obtained a much more damped response (Fig. 1A, curve c). Results for Trp residues 127, 133, 183, and 195 are shown in Fig. 1B, again by using truncation at the residue level plus use of electroneutral side chains. There is an essentially flat response by ~6 to 7 Å, and the CFP-GIAO shielding patterns for these four sites are in good accord with the experimental chemical shifts of Luck and Falke (*24*). The less good agreement for Trp^{284} (not shown) probably arises because this [5-F]Trp site is highly exposed to solvent water (Fig. 2A), and it is not feasible to adequately describe solvent

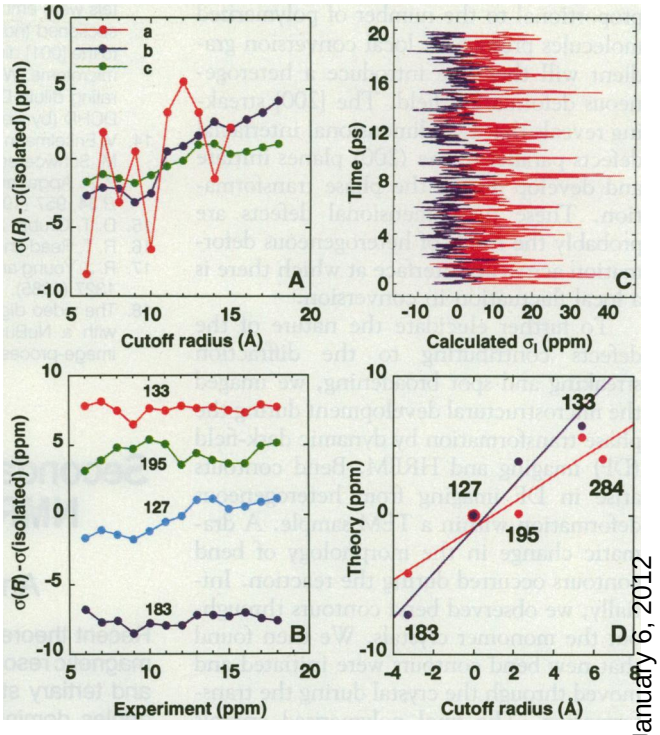


Fig. 1. Fluorine shielding results for the five Trp residues (127, 133, 183, 195, and 284) in the *E. coli* galactose binding protein. (**A**) Ab initio (GIAO) results with the use of a (red), radial cutoff with atomic resolution; b (blue), radial cutoff with whole residue resolution; and c (green), radial cutoff with whole, electroneutral residue resolution. (**B**) GIAO shielding results for Trp residues 127, 133, 183, and 195 with the use of whole electroneutral residue resolution as a function of radial cutoff. (**C**) Results for 20-ps shielding trajectories for [5-F]Trp^{183} (blue) and Trp^{284} (red) residues in GBP. (**D**) Graph showing experimental versus theoretical shielding results in GBP: blue, CFP-GIAO; and red, MSP-LRF method (*27*). Calculations were carried out in our laboratory by using a cluster of RISC computers (International Business Machines (IBM) Corporation, Austin, Texas) equipped with a total of 0.3 gigabytes of RAM and 26 gigabytes of disk space and having a peak theoretical speed of ~0.6 gigaflops, as well as at the University of Illinois at Chicago on an IBM RISC 6000/model 560 computer.

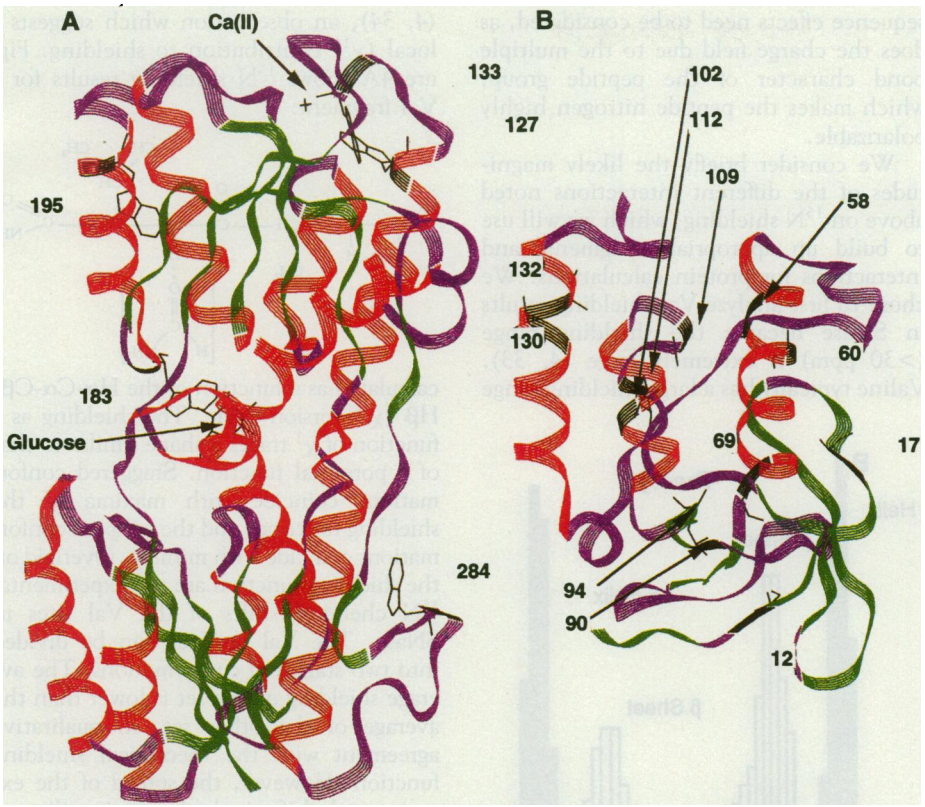
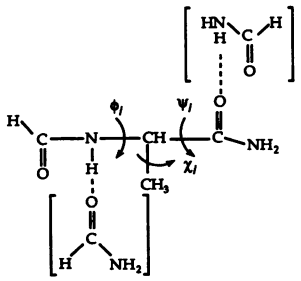


Fig. 2. Ribbon diagrams for *E. coli* galactose binding protein and *Staphylococcal* nuclease, showing positions of residues whose ^{19}F and ^{13}C chemical shieldings were computed. (**A**) GBP, showing [5-F]Trp residues 127, 133, 183, 195, and 284. (**B**) SNase, showing Ala residues 12, 17, 58, 60, 69, 90, 94, 102, 109, 112, 130, and 132. Legend: β sheet, green; α helix, red; random, purple; and computed residue, black.

interaction within a single, static structural picture.

The second approach to determine ^{19}F shielding nonequivalencies in GBP permits incorporation of solvent effects and involves the use of MSPs combined with the fields and field gradients from an LRF-MD approach (*21*) to compute shielding trajectories for each ^{19}F site. Full details are given elsewhere (*27*), but for comparison with the ab initio approach, Fig. 1C shows 20-ps trajectories, $\sigma(E_x, V_{ii})f(\tau)$, where E_x is the field component along the C–F bond axis, V_{ii} are the elements of the electric-field gradient tensor, and τ is the time for [5-F]Trp residues 183 and 284 in GBP (*27*). The shielding of the exposed Trp^{284} fluctuates about twice as much as that of the more buried sites (Trp residues 127, 133, 183, and 195) because of interactions with mobile solvent water molecules. The average of five such shielding trajectories are compared with the CFP-GIAO result in Fig. 1D. Both sets of results are in good general accord with experiment (*24*), with the slope of ~1.6 in the CFP-GIAO method (Fig. 1D) representing an approximate value for the dielectric constant in the vicinity of each F-Trp residue. These results support the idea that ^{19}F chemical shift nonequivalencies in proteins are dominated by σ_I effects that can be computed by quantum chemical methods.

We next considered ^{13}C NMR shielding in proteins because the ability to compute (especially) backbone ^{13}C shieldings could be a useful complement to other distance-based (*7*) or more theoretical approaches to protein folding (*28*), especially in large or more condensed systems. Clear-cut correlations of ^{13}C chemical shifts with structure in polymers (*29*) and proteins (*30*) have only been reported more recently, but as Spera and Bax (*30*) state, "^{13}C chemical shifts remain poorly understood and significant deviations from expected secondary shifts can occur." We have thus investigated Cα, Cβ shieldings of some model fragments containing Ala residues as well as the Ala residues in SNase (Fig. 2B) to see if our understanding of ^{13}C shielding in proteins can be improved. By way of introduction, we show in Fig. 3A histograms of Cα and Cβ shieldings for α-helical and β-sheet residues from the database of Spera and Bax (*30*). There is a clear separation between the shieldings of the two structure types, suggesting an important effect of ϕ,ψ torsion angles on shielding, but the experimental results could be affected by other interactions, such as hydrogen bonding and σ_I effects. To investigate the effects of ϕ and ψ alone, we computed the ^{13}C NMR shieldings for Cα and Cβ in a series of Ala molecular fragments:

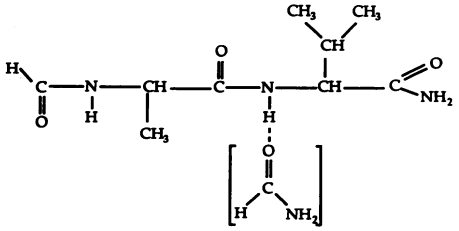


We first used a 6-311G** basis set for all atoms (without the formamide hydrogen-bond partners). The overall widths and ~5 ppm separations between helical and sheet residues found experimentally are reproduced in the calculations (Fig. 3B). Thus, ϕ and ψ torsion angles both appear intimately correlated with Cα and Cβ shieldings. We then computed the theoretical shieldings for the 12 Ala Cα groups in SNase (Fig. 2B) with the fragment shown above and an attenuated basis set following the locally dense scheme proposed by Chesnut (*31*). We used a 6-311++G (2*d*,2*p*) basis for Cα, Cβ, C^0, N, H^N, and Hα (set I) and a 6-31G basis for other atoms in the fragment (set II), together with the known x-ray structure (*32*), energy minimized in the absence of ligands. We obtained a slope of 0.85 and a regression coefficient $r = 0.94$ (a measure of the "goodness of fit")—a respectable agreement between theory and experiment (*33*). We then investigated the effects of hydrogen bonding and the proteins' charge field on Cα shielding with the CFP-GIAO (electroneutral) fragment approach used for ^{19}F, but with the additional incorporation of formamide molecules (to represent hydrogen-bond partners) as pictured above. Results for a small basis set (6-311G**/I/6-31G/II) gave $m = 1.2$ and $r = 0.97$, whereas a larger basis set [6-311++G(2*d*,2*p*)/I/6-31G/II], which included explicit hydrogen bonding and CFP, gave $m = 1.18$ and $r = 0.97$ (Fig. 3C), indicating basis set insensitivity. Generally similar results are obtained for Cβ (Fig. 3D, $m = 1.2$ and $r = 0.93$). Thus, although ϕ,ψ effects are overwhelmingly dominant for ^{13}C, slight improvements are obtained upon incorporation of specific hydrogen-bond partners and the proteins' charge field, and the accord with experiment is quite satisfactory. The helical nature of Ala^{130}, Ala^{132}, or Ala^{60} or the sheet nature of Ala^{17}, Ala^{90}, and Ala^{94} seen in Fig. 2B is clearly observed in the calculations (Fig. 3, C and D), and the root-mean-square deviations from the linear relation are only 0.55 and 0.71 ppm for Cα and Cβ, respectively.

Finally we consider the peptide NH group, both from the standpoint of ^{15}N and ^{1}H NMR, which are routinely determined

in protein structure studies (*7*). Shielding in the NH group is expected to be much more difficult to compute because there are many different interactions to consider and there are no clear-cut helix-sheet shielding correlations, as found with $^{13}C\alpha$, where ϕ,ψ torsional effects were expected to dominate. For ^{15}N, the effects of direct hydrogen bonding can be expected to be very important (and the H atom position may be somewhat uncertain), and, more complex factors involving side chain torsions can be anticipated (since γ-gauche effects may be involved). Also, sequence effects need to be considered, as does the charge field due to the multiple bond character of the peptide group, which makes the peptide nitrogen highly polarizable.

We consider briefly the likely magnitudes of the different interactions noted above on ^{15}N shielding, which we will use to build up appropriate fragments and interactions for protein calculations. We chose to first analyze Val shielding results in SNase because the shielding range (>30 ppm) is extremely large (*4*, *33*). Valine typically has a large shielding range (*4*, *34*), an observation which suggests a local (χ^1) contribution to shielding. Figure 4A shows ^{15}N shielding results for a Val fragment:



calculated as a function of the Hα-Cα-Cβ-Hβ (χ^1) torsion angle. The shielding as a function of χ^1 traces a shape similar to that of a potential function. Staggered conformations coincide with maxima of the shielding function, and the eclipsed conformations coincide with minima. Overlaid on the shielding function are the experimental ^{15}N chemical shifts of the Val sites in SNase. The Val residues can be divided into two staggered conformations. The average shielding of one set is lower than the average of the other set, in qualitative agreement with the theoretical shielding function. However, the spread of the experimental shifts in both sets is still very large (~15 ppm), indicating that there must be other significant contributions to the ^{15}N chemical shifts for the Val sites in SNase, and intuitively this additional effect would seem to be hydrogen bonding. To investigate the magnitude of the effects of hydrogen bonding we carried out calculations involving a Gly-Gly dipeptide model interacting with a formaldehyde molecule. The basis set used in these calculations was [6-311++G(2*d*,2*p*)/I/6-31G/II], where the peptide atoms (CONH) belong to group I and all of the other atoms are in group II (Fig. 4B). Only one orientation was considered—a linear NH···OC hydrogen. Hydrogen bonding induces a deshielding of as much as 13 ppm, about the amount expected for the protein. Close examination of the x-ray structure of SNase reveals that the N–H groups of Val residues 23, 66, 74, 99, and 111 participate in hydrogen bonding (N–O distance <3.0 Å) with the C=O groups of residues 34, 62, 10, 92, and 129, respectively, and the shielding values for sites 23, 74, 99, and 111 (Fig. 4A) are the lowest in the conformer sets they belong to, in general agreement with the theoretical results shown in Fig. 4B.

The last set of model calculations was aimed at extracting the effects of hydrogen bonding of the C=O group of the peptide bond on peptide ^{15}N shielding. The model again consisted of a Gly-Gly peptide, this time with its peptide C=O interacting either with an ammonia molecule or with a point charge distribution to represent the NH_3 molecule. The shapes of both curves

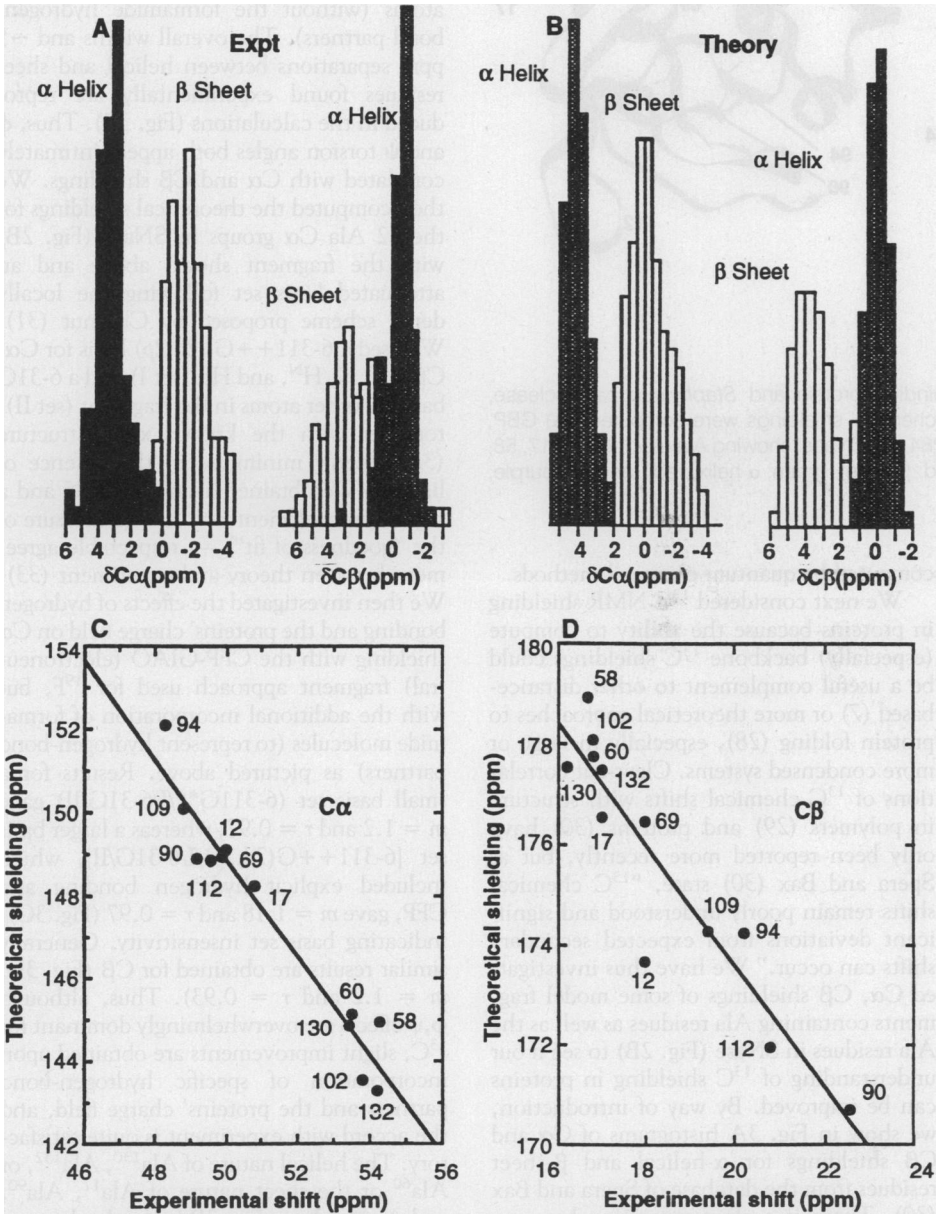


Fig. 3. Carbon-13 NMR shielding results for Cα and Cβ sites in proteins and a model system. (**A**) Histograms showing separation of α-helical and β-sheet chemical shifts for Cα and Cβ sites in proteins [based on data in Spera and Bax (*30*)]. (**B**) Histogram showing Cα and Cβ shifts for model α-helical and β-sheet (Ala) fragments based on ab initio shielding surface calculations. (**C**) Experimental Cα chemical shifts for the 12 Ala sites in SNase versus computed shieldings, $m = 1.2$ and $r = 0.97$. (**D**) Experimental Cβ chemical shifts for the 12 Ala sites in SNase versus computed shieldings, $m = 1.2$ and $r = 0.93$. The x-ray structure (*32*) was energy minimized after "protonation," without ligands, for comparison with the solution NMR results.

are similar (Fig. 4C), and the effects of "secondary" hydrogen bonding in both cases are very small—about 1 or 2 ppm at a 3 Å R_{O-N} distance, to be compared with the >30 ppm experimental range. Because it does not add to the basis functions, and the effect is very small, we thus chose to use the point charge distribution in our SNase calculations.

Thus, for our ^{15}N Val shielding calculations in SNase, we use the Ala-Val fragment shown above, with the coordinates of the heavy atoms derived from the relaxed x-ray structure (*32*) used to extract the short-range contributions to shielding. If the Val N atom participates in hydrogen bonding, a formamide molecule is incorporated as a hydrogen-bond partner. An exception is the N atom of Val^{111}, which is close to the –COOH group of Glu^{129}. Here, the calculation was performed with a formaldehyde molecule placed on the coordinates of the C=O group closest to the Val N atom. Also, site 51 is next to a Gly residue, so in this case only a Gly-Val fragment was used, and the shielding calculations for Val^{51} and Val^{114} were carried out without any additional hydrogen-bond partners (*32*).

The results of our ^{15}N shielding calculations for the nine Val residues in SNase are shown in Fig. 5. Figure 5A shows the correlation obtained by using solely the isolated peptide fragment, Fig. 5B shows the isolated fragment plus a single formamide hydrogen-bond partner, and Fig. 5C shows the isolated fragment plus a single hydrogen-bond partner and the proteins' charge field (m = 0.87 and r = 0.97). Although the general ^{15}N Val shielding features in SNase are seen in the isolated fragment calculation (Fig. 5A) both hydrogen bonding and σ_1 effects need to be incorporated in order to obtain close agreement with the experimental result (Fig. 5C). Equivalent calculations for $^1H^N$ were also carried out and gave m = 1.5 and r = 0.79 for all interactions. Residues 51 and 114 were considerably off the trend because they arise from exposed residues, and as with the ^{19}F result on Trp^{284} in GBP, alternative approaches to investigating solvated groups are needed, such as the MSP method.

The results we have outlined above represent the successful application of quantum chemical methods to analyzing the ^1H, ^{13}C, ^{15}N and ^{19}F chemical shifts (or shieldings) of backbone and side chain atomic sites in proteins. The average regression coefficient between theory and experiment for the heavy elements ($^{13}C\alpha$, $^{13}C\beta$, ^{15}NH, and ^{19}F) is 0.96, without use of any adjustable parameters, whereas the slopes vary between ~0.8 and 1.2. For ^{19}F, the two models agree well with each other and with experiment, suggesting the importance of weak electrical interactions for ^{19}F shielding nonequivalencies in proteins. For ^{13}C NMR, we find excellent agreement between theory and experiment for Cα and Cβ sites in Ala residues in SNase, even without hydrogen bonding and charge-field effects, whereas for ^{15}N in SNase, φ,ψ, χ^1 torsion angles, hydrogen bonding, next neighbors, and electrostatics are all important, and for H^N, hydrogen bonding dominates. The ability to predict ^1H, ^{13}C, ^{15}N, and ^{19}F chemical shifts in proteins from known or test structures using quantum chemical methods could be useful in refining or determining such structures.

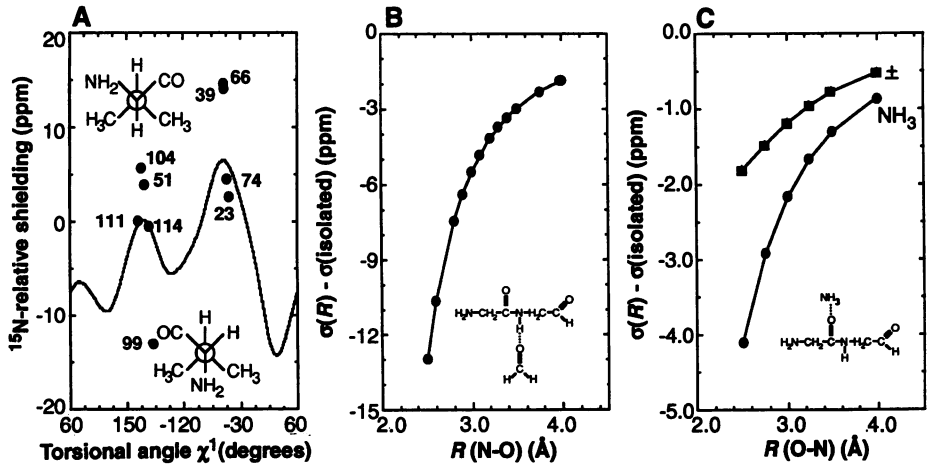


Fig. 4. Nitrogen-15 NMR shielding for model systems. (**A**) Effect of χ^1 torsion angle on ^{15}N shielding in a Val model compound. The experimental shieldings for the Val sites in SNase are indicated. (**B**) Effects of (linear) hydrogen bonding to a carbonyl acceptor (formaldehyde) on the ^{15}N shielding of an NH group in a model peptide. (**C**) Indirect effect of (linear) hydrogen bonding, modeled by interacting the $H_2NCH_2CONHCH_2CHO$ peptide-carbonyl with NH_3 (●) or a point charge model for NH_3 (■).

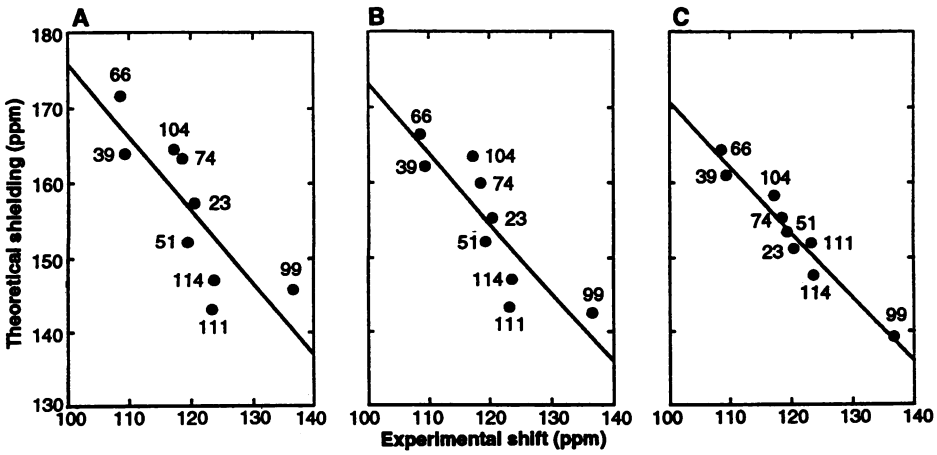


Fig. 5. Nitrogen-15 chemical shieldings for the valine sites in SNase. (**A**) Isolated SNase ^{15}N Val fragments, m = 0.97 and r = 0.81. (**B**) As in (A) but with inclusion of a hydrogen-bond partner, m = 0.92 and r = 0.86. (**C**) As in (B) but with additional inclusion of CFP, m = 0.87 and r = 0.97.

REFERENCES AND NOTES

1. C. C. McDonald and W. D. Phillips, *J. Am. Chem. Soc.* **89**, 6332 (1967).
2. A. Allerhand, R. F. Childers, E. Oldfield, *Biochemistry* **12**, 1335 (1973).
3. L. K. Nicholson *et al.*, *ibid.* **31**, 5253 (1992).
4. D. A. Torchia, S. W. Sparks, A. Bax, *ibid.* **28**, 5509 (1989).
5. K. D. Park *et al.*, *ibid.* **30**, 2333 (1991).
6. E. Oldfield *et al.*, unpublished results.
7. A. Bax, *Annu. Rev. Biochem.* **58**, 223 (1989); G. M. Clore and A. M. Gronenborn, *Science* **252**, 1390 (1991).
8. M. Schindler and W. Kutzelnigg, *J. Am. Chem. Soc.* **105**, 1360 (1983).
9. R. Ditchfield, *J. Chem. Phys.* **56**, 5688 (1972); K. Wolinski, J. F. Hinton, P. Pulay, *J. Am. Chem. Soc.* **112**, 8251 (1990).
10. A. E. Hansen and T. D. Bouman, *J. Chem. Phys.* **82**, 5035 (1985); J. D. Augspurger and C. E. Dykstra, *J. Phys. Chem.* **95**, 9230 (1991).
11. P. L. Cummins, G. B. Bacskay, N. S. Hush, *Mol. Phys.* **61**, 795 (1987).
12. J. A. Tossell, *J. Magn. Reson.* **94**, 301 (1991).
13. A. I. Jaman, T. C. Germann, H. S. Gutowsky, J. D. Augspurger, C. E. Dykstra, *Chem. Phys.* **154**, 281 (1991).
14. J. D. Augspurger, C. E. Dykstra, E. Oldfield, *J. Am. Chem. Soc.* **113**, 2447 (1991).
15. J. Augspurger *et al.*, *J. Magn. Reson.* **100**, 342 (1992); C. E. Dykstra, E. Oldfield, J. P. Aug-

spurger, J. G. Pearson, in *Calculation of NMR Shielding Constants and Their Use in the Determination of the Geometric and Electronic Structures of Molecules and Solids*, J. Tossell, Ed. (Kluwer Academic, Dordrecht, in press).
16. A. C. de Dios and E. Oldfield, *Chem. Phys. Lett.* **205**, 108 (1993).
17. ———, J. D. Augspurger, C. E. Dykstra, unpublished results.
18. K. Ösapay and D. A. Case, *J. Am. Chem. Soc.* **113**, 9436 (1991); M. P. Williamson, T. Asakura, E. Nakamura, M. Demura, *J. Biomol. NMR* **2**, 83 (1992); D. S. Wishart, B. D. Sykes, F. M. Richards, *J. Mol. Biol.* **222**, 311 (1991); *Biochemistry* **31**, 1647 (1992).
19. F. London, *J. Phys. Radium* **8**, 397 (1937); L. Salem, *Molecular Orbital Theory of Conjugated Systems* (Benjamin, New York, 1966), chap. 4.
20. P. K. Weiner and P. A. Kollman, *J. Comput. Chem.* **2**, 287 (1981); B. R. Brooks *et al.*, *ibid.* **4**, 187 (1983).
21. F. S. Lee and A. Warshel, *J. Chem. Phys.* **97**, 3100 (1992).
22. A. D. Buckingham, *Can. J. Chem.* **38**, 300 (1960); ——— and K. P. Lawley, *Mol. Phys.* **3**, 219 (1960).
23. J. G. Batchelor, *J. Am. Chem. Soc.* **97**, 3410 (1975).
24. L. A. Luck and J. J. Falke, *Biochemistry* **30**, 4248 (1991).
25. N. K. Vyas, M. N. Vyas, F. A. Quiocho, *Science* **242**, 1290 (1988); S. L. Mowbray, R. D. Smith, L. B. Cole, *Receptor* **1**, 41 (1990).
26. C. Le and E. Oldfield, unpublished results.
27. J. G. Pearson, E. Oldfield, F. S. Lee, A. Warshel, *J. Am. Chem. Soc.*, in press.
28. R. A. Goldstein, Z. A. Luthey-Schulten, P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9029 (1992).
29. I. Ando, H. Saito, R. Tabeta, A. Shoji, T. Ozaki, *Macromolecules* **17**, 457 (1984).
30. S. Spera and A. Bax, *J. Am. Chem. Soc.* **113**, 5490 (1991).
31. D. B. Chesnut and K. D. Moore, *J. Comput. Chem.* **10**, 648 (1989).
32. P. J. Loll and E. E. Lattman, *Proteins* **5**, 183 (1989).
33. J. Wang, A. P. Hinck, S. N. Loh, D. M. LeMaster, J. L. Markley, *Biochemistry* **31**, 921 (1992).
34. J. Glushka *et al.*, *J. Am. Chem. Soc.* **111**, 7716 (1989).
35. We are grateful to P. Pulay, K. Wolinski, and J. F. Hinton for providing their TEXAS 90 program; to A. Warshel and F. S. Lee for providing a customized version of ENZYMIX; and to C. Jameson and P. G. Wolynes and the above for useful advice. Supported in part by the U.S. Public Health Service (USPHS) [grants HL-19481 and (in part) GM-40426], by the American Heart Association (grant AHA 92-013340), with funds provided in part by the Illinois Heart Association, Inc. and by the Graduate Research Board of the University of Illinois at Urbana-Champaign. J.G.P. was supported by a USPHS postdoctoral fellowship (grant GM-14545).

11 January 1993; accepted 5 April 1993

Poly(phenylcarbyne): A Polymer Precursor to Diamond-Like Carbon

Glenn T. Visscher, David C. Nesting, John V. Badding, Patricia A. Bianconi*

The synthesis of poly(phenylcarbyne), one of a class of carbon-based random network polymers, is reported. The network backbone of this polymer is composed of tetrahedrally hybridized carbon atoms, each bearing one phenyl substituent and linking, by means of three carbon–carbon single bonds, into a three-dimensional random network of fused rings. This atomic-level carbon network backbone confers unusual properties on the polymer, including facile thermal decomposition, which yields diamond or diamond-like carbon phases at atmospheric pressure.

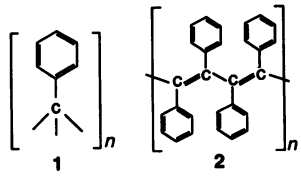
The polyacetylene class of polymers, of stoichiometry $(CR)_n$, has long been a focus of intense research owing to their conductive and electronic properties (*1*). These polymers have linear backbones consisting of alternating C–C and C=C bonds, each carbon bearing one substituent (*2–5*). Recently, inorganic backbone polymers of similar stoichiometry, but different structure, have been synthesized: the polysilynes $(SiR)_n$ and polygermynes $(GeR)_n$ and their copolymers (*6*). These polymers have a continuous random network backbone, each inorganic atom being tetrahedrally hybridized and bound by single bonds to three other inorganic atoms and one substituent. The properties of the network polymers are different from those of linear inorganic backbone polymers because of the network structure. We report an analogous carbon-based network polymer of $(CR)_n$ stoichiometry (*7*). That the linear structure is preferred in $(CR)_n$ polymers is to be expected, given the much greater strength of C=C over Si=Si or Ge=Ge bonds (*8*). However, appropriate synthetic conditions can produce a carbon network-backbone polymer of stoichiometry $(CR)_n$ (R = Ph; Ph, phenyl), which is analogous in structure to the polysilynes and germynes (*9–13*).

The synthesis (*14*) was accomplished by a procedure analogous to those used for the synthesis of the inorganic network polymers (*6*): the reduction of an appropriate $RCCl_3$ monomer with an ultrasonically generated emulsion of Na-K alloy and an ethereal solvent, tetrahydrofuran (THF) (Eq. 1).

$$PhCCl_3 + 3.0\ NaK \xrightarrow[THF]{\text{475-W, 20-kHz ultrasound}} \underset{\mathbf{1}}{(PhC)_n} + 3.0\ Na(K)Cl \quad (1)$$

For R = phenyl, poly(phenylcarbyne) (**1**) is obtained from this reaction as a tan powder that is soluble in common organic solvents. Insoluble crosslinked material and low molecular weight oligomers are the only other products. Chemical analysis (*15*) of **1** was consistent with the empirical formula $(C_6H_5C)_n$, and gel permeation chromatography of the material gave a weight-averaged molecular weight of $\overline{M}_w$ = 4000 and a number-averaged molecular weight of $\overline{M}_n$ = 3077, which indicates that **1** is polymeric (*16*).

Infrared (IR) spectra of **1** show a complete absence of C=C stretching bands, which are characteristic of *cis*-polyacetylenes and are seen in the IR spectrum of poly(diphenylacetylene) (**2**), the linear



1 **2**

polymer whose empirical formula is identical to that of **1** (*17*). The IR spectra also show bands consistent only with monosubstituted phenyl rings (*15*): no di- or trisubstituted phenylene-type aryl groups are present. The ^{13}C nuclear magnetic resonance spectroscopy of **1** exhibits a very broad resonance ($\delta\nu_{1/2}$ = 800 Hz) centered at a chemical shift δ of 51 parts per million (ppm) of the applied field, characteristic of quaternary carbon atoms (*7*, *9–13*), and no resonances other than those of the phenyl rings were detected in the vinylic carbon region, where the resonances of the backbone carbons of phenyl-substituted acetylenes normally appear (*3*, *4*). The resonance at 51 ppm is enhanced when **1** is synthesized with the use of 10 mole percent of α,α,α-trichlorotoluene monomer that has been labeled with ^{13}C in the α position. These data indicate that C=C bonds are not primary structural features of **1** and that this polymer therefore does not adopt the linear polyacetylene structure. The presence of quaternary α-carbons as a primary structural feature and the broadness of the ^{13}C resonances indicate that **1** consists of a randomly constructed rigid network of tetrahedral phenylcarbyne units in which each carbyne carbon forms three C–C bonds to the network and one to the phenyl substituent (Fig. 1). This structure corresponds to the network arrangements found in the analogous silicon and germanium polymers and

Department of Chemistry, The Pennsylvania State University, University Park, PA 16802.

*To whom correspondence should be addressed.