the unpaired electron with the oxygen's x orbitals contributes to the magnitude of g^{yy} while the oxygen's y orbitals similarly contribute to g^{xx}. This direction dependency of the g factors has been of great assistance in the interpretation[27] of semiquinone-solvent interactions[28] and the anomalous saturation behavior[29] exhibited by semiquinones in rigid media.

The g factors calculated both from perturbation theory (eq 4) using INDO wave functions and energy levels and from the modified form of Stone's semiempirical equation (eq 6) give results in close agreement with the experimentally determined g factors. On the other hand, both perturbation theory and the modified Stone's equation yield g^{YY} noticeably smaller than the experimental g^{yy}. The reason for the difference is not known. One possible explanation, however, is that the oxygen's $2p_x$ orbitals are significantly perturbed (possibly by the solvent[27]) to cause a mixing of these orbitals with the $2p_y$ orbitals on oxygen. The result of this mixing would be a relative increase in the value of g^{YY} and possibly g^{XX}. Such a prediction is consistent with the above results.

Acknowledgment. This research was supported by a Cottrell Research Grant from Research Corporation.

References and Notes

(1) For a recent review of the file, see H. M. Swartz, J. R. Bolton, and D. C. Borg, Ed., "Biological Applications of Electron Spin Resonance", Wiley, New York, N.Y., 1972.
(2) P. Jost, A. S. Waggoner, and O. H. Griffith, "Structure and Function of Biological Membranes", L. I. Rothfield, Ed., Academic Press, New York, N.Y., 1971, pp 83–144.
(3) H. F. DeLuca and J. W. Suttie, Ed., "The Fat-Soluble Vitamins", University of Wisconsin Press, Madison, Wisconsin, 1969.
(4) R. E. Olson, "Perspectives in Biological Chemistry", R. E. Olson, Ed., Marcel Dekker, New York, N.Y., 1970, p 83.
(5) W. A. Pryor, *Chem. Eng. News*, **49**, 34 (1971).
(6) R. Bentley and L. M. Campbell, "Chemistry of the Quinonoid Compounds", S. Patai, Ed., Wiley, New York, N.Y., 1974, pp 683–736.
(7) P. A. Loach and R. L. Hall, *Proc. Natl. Acad. Sci. U.S.*, **69**, 786 (1972).
(8) G. Feher, M. Y. Okamura, and J. D. McElroy, *Biochim. Biophys. Acta*, 267, 222 (1972).
(9) A. F. Esser, *Photochem. Photobiol.*, **20**, 167 (1974).
(10) J. T. Warden and J. R. Bolton, *Photochem. Photobiol.*, **20**, 245 (1974).
(11) M. S. Blois, H. W. Brown, R. M. Lemmon, R. O. Lindblom, and M. Weissbluth, Ed., "Free Radicals in Biological Systems", Academic Press, New York, N.Y., 1961.
(12) J. R. Bolton and G. K. Fraenkel, *J. Chem. Phys.*, **40**, 3307 (1964).
(13) T. E. Gough and M. C. Symons, *Trans. Faraday Soc.*, **62**, 269 (1966).
(14) E. Charney and E. Becker, *J. Chem. Phys.*, **42**, 910 (1970).
(15) T. Yonezawa, T. Kamamura, M. Ushio, and Y. Nakao, *Bull. Chem. Soc. Jpn.*, **43**, 1022 (1970).
(16) E. W. Stone and A. H. Maki, *J. Chem. Phys.*, **36**, 1944 (1962).
(17) J. Gendell, J. H. Freed, and G. K. Fraenkel, *J. Chem. Phys.*, **37**, 2836 (1962).
(18) B. G. Segal, M. Kaplan, and G. K. Fraenkel, *J. Chem. Phys.*, **43**, 4191 (1965).
(19) G. Vincow, *J. Chem. Phys.*, **38**, 917 (1963).
(20) D. T. Wilkinson and H. R. Crane, *Phys. Rev.* [*Sect.*] *A*, 852 (1963).
(21) H. M. McConnell and R. E. Robertson, *J. Phys. Chem.*, **61**, 1018 (1957).
(22) A. J. Stone, *Proc. R. Soc. London, Ser. A*, **271**, 424 (1963).
(23) A. J. Stone, *Mol. Phys.*, **6**, 509 (1973).
(24) A. J. Stone, *Mol. Phys.*, **7**, 311 (1964).
(25) D. S. McClure, *J. Chem. Phys.*, **17**, 905 (1949).
(26) H. M. McConnell and J. Strathdee, *Mol. Phys.*, **2**, 129 (1959).
(27) B. J. Hales, in preparation.
(28) J. J. Harbour and G. Tollin, *Proc. Natl. Acad. Sci. U.S.*, **69**, 2066 (1972).
(29) B. J. Hales and J. R. Bolton, *J. Am. Chem. Soc.*, **94**, 3314 (1972).

Carbon-13 Magnetic Resonance Study of Structural and Dynamical Features in Carbamylated Insulins

Jens J. Led,*[1a] David M. Grant,*[1b] W. James Horton,[1b] F. Sundby,[1c] and K. Vilhelmsen[1c]

Contribution from the Department of Chemical Physics, University of Copenhagen, H. C. Ørsted Institutet Universitetsparken 5 Copenhagen O, Denmark, Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, and NOVO Research Institute, Copenhagen, Denmark. *Received September 9, 1974*

Abstract: Carbon-13 NMR Fourier transform spectra were recorded and assigned and nuclear Overhauser enhancement parameters and spin-lattice relaxation times, T_1, were measured for 90% ^{13}C enriched carbamyl groups attached to the A1 glycine, B1 phenylalanine, and the B29 lysine in zinc insulin and nickel insulin. Similar measurements were made for metal-free insulin in which only the two N-terminal amino acids were carbamylated. All measurements were made with the derivatives in water solution at pH 7.8. The results indicate that the derivatives contain the A1 glycine in two different environments suggesting that the insulin dimers have an imperfect twofold axis. The correlation times for the overall tumbling of the derivatives were estimated which led to evaluation of the state of aggregation of the derivatives. Likewise, information was obtained as to the relative mobility of the location in the insulin molecules to which the carbamyl groups were attached, via an estimation of the correlation times for the internal rotation of the carbamyl groups. Finally, the effective distance between the Ni^{2+} ion in B1 carbamylated nickel insulin and the ^{13}C nucleus in the carbamyl group was estimated to be 10.5 Å or less.

Natural abundance carbon-13 magnetic resonance (^{13}C NMR) spectra of biopolymers have recently been shown[2-9] to contain valuable information on both the structural conformation and the segmental motion present in these molecules. However, the detectability of ^{13}C signals in natural abundance for such large molecules, even with Fourier transform techniques, and the complexity of the spectra resulting from the overlapping of a large number of ^{13}C resonance lines combine to limit an extensive study of these structural and dynamical characteristics.

These difficulties can be overcome in part through ^{13}C enrichment of specific positions of the molecules, thereby increasing the signal to noise level of the labeled position beyond the background signals due to the unenriched parts of the molecule which contain ^{13}C at the 1.1% natural abundance. Reliable data for the enriched position at biologically significant concentrations have been obtained in this manner as recently illustrated[10] with ^{13}C enriched histidine residues incorporated, in vivo, into a native enzyme. This interesting approach in general is expensive and time consuming, and at this stage a more practicable way in which ^{13}C isotopes can be incorporated is through highly

specific ^{13}C enriched reagents that react with specific groups in the biomolecule.[11,12] This method provides experience with isotopically labeled samples at lower concentrations and explores the biological importance of data which can be obtained at these levels. The principal disadvantage of this method is that the modifications may alter structural features of the molecule important to its biological function and thereby reduce or destroy its activity. Conversely, if it can be established that the chemical modification connected with labeling produces no decrease in the biological activity, it can be argued that the important structural and dynamic characteristics of the biomolecule may not be greatly altered from their native state.

In order to explore this possibility a ^{13}C NMR study was initiated on insulins modified with highly ^{13}C enriched (90%) potassium cyanate at the two α-amino groups, A1 and B1, and the ϵ-amino group of the lysine residue, B29. It has been demonstrated[13] recently that such changes in the A1 amino group do not affect the biological potency of insulin significantly, while the modification of the B1 amino group results in a decrease in the immunological potency.

Theoretical Considerations

The spin-lattice relaxation times, T_1, and the spin-spin relaxation times, T_2, of ^{13}C nuclei may involve several mechanisms. These include dipole-dipole interaction between ^{13}C and its nearest protons, chemical shift anisotropy, spin-rotation as well as scalar, dipole, and quadrupole interaction with spin other than protons.[14] In addition, it has been demonstrated[15] that even nonprotonated carbons have a purely dipolar relaxation mechanism when they are incorporated in relatively large asymmetric molecules as sucrose, cholesteryl chloride, and AMP. Similar results have been obtained for larger biomolecules.[7,10]

Under these circumstances and considering only intramolecular dipolar interactions between ^{13}C and its adjacent protons, the T_1 and T_2 relaxation times for the ^{13}C nuclei are given by the following expressions, provided rotational diffusion of the molecules is isotropic and all protons are saturated:[16]

$$\frac{1}{T_{1D}} = \frac{1}{20}\hbar^2\gamma_I^2\gamma_S^2[J_0(\omega_S - \omega_I) + 3J_1(\omega_I) + 6J_2(\omega_S + \omega_I)] \quad (1)$$

$$\frac{1}{T_{2D}} = \frac{1}{40}\hbar^2\gamma_I^2\gamma_S^2[4\tau_r + J_0(\omega_S - \omega_I) + 3J_1(\omega_I) + 6J_1(\omega_S) + 6J_2(\omega_S + \omega_I)] \quad (2)$$

Here $J_k(\omega)$, the spectral densities in the case of an isotropic tumbler, is given by[16,17]

$$J_k(\omega) = \Sigma_i r_i^{-6}\frac{2\tau_r}{(1 + \omega^2\tau_r^2)} \quad (3)$$

where the subscripts I and S refer to ^{13}C and ^1H, respectively, r_i is the distance between the ^{13}C nucleus and its i'th neighboring proton, and τ_r is the correlation time for the ^{13}C-^1H dipolar coupling assuming isotropic rotational reorientation. These equations hold even if $\omega^2\tau_r^2 \gg 1$.[18]

Using the same assumptions the differential nuclear Overhauser enhancement (NOE) parameter is given by[19]

$$\eta = \frac{I_z - I_0}{I_0} = \rho_{CH}S_H\frac{\gamma_H}{\gamma_C} \quad (4)$$

where I_z and I_0 are the intensities of the signal with and without enhancement and S_H is the degree of saturation of the proton absorption signal defined by

$$S_H = 1 - \frac{1}{1 + \gamma_H^2H_2^2(T_1T_2)_H} \quad (5)$$

while ρ_{CH} is a measure of the effectiveness of the coupling between the magnetic dipoles of the two types of nuclei. When it is due to intramolecular dipolar coupling modulated by rotational diffusion, it takes the form[20]

$$\rho_{CH} = \frac{6J_2(\omega_S + \omega_I) - J_0(\omega_S - \omega_I)}{J_0(\omega_S - \omega_I) + 3J_1(\omega_I) + 6J_2(\omega_S + \omega_I)} \quad (6)$$

where $J_k(\omega)$ has the same meaning as above. From eq 6 it appears that ρ_{CH} and thereby the NOE is independent of the distances r_i for a purely dipolar relaxation mechanism. From eq 3 to 6 it follows that η can assume a maximum value of 1.988 if the overall reorientation of the molecule is sufficiently fast, i.e., $\omega^2\tau_r^2 \ll 1$. However, for large biomolecules it is found more often that $\omega^2\tau_r^2 \approx 1$. Thus, by nanosecond polarization spectroscopy it has been found[21] for a number of biopolymers in their native states that the overall rotational correlation times vary between 8×10^{-9} and 40×10^{-9} sec when the molecular weight range is between 17,000 and 66,000. For $\omega_I = 1.58 \times 10^8$ rad/sec this gives $1.6 \leq \omega_I^2\tau_r^2 \leq 40$. Recently correlation times have been found[7,10] by ^{13}C NMR studies of similar magnitude. Under these conditions η becomes smaller than its maximum value and varies with τ_r as shown in Figure 1 for $\omega_I = 1.58 \times 10^8$.

The simple isotropic model is only applicable to rigid biomolecules for which the only important motion is the overall reorientation of the entire molecule. If internal segmental motion must be considered this description is inadequate, and the theoretical treatment becomes more complex. Freely moving side chains or groups, internal rotations, etc., all could contribute to relaxation processes when their effective correlation times are in the proper range. While eq 1, 2, 4, 5, and 6 are still correct, eq 3 now becomes invalid. Woessner has described the case in which the internal motion of a given group consists of a random reorientation about an axis which has a fixed orientation relative to the rigid frame of the molecule undergoing reorientation.[22] For a carbon located in such a group the spectral densities are given by:

$$J_k(\omega) = \Sigma_i 2r_i^{-6}\left[\frac{A_i\tau_r}{1 + \omega^2\tau_r^2} + \frac{B_i\tau_b}{1 + \omega^2\tau_b^2} + \frac{C_i\tau_c}{1 + \omega^2\tau_c^2}\right] \quad (7)$$

where for a stochastic diffusion process

$$\tau_b^{-1} = \tau_r^{-1} + (6\tau_{int})^{-1}$$
$$\tau_c^{-1} = \tau_r^{-1} + 2(3\tau_{int})^{-1}$$
$$\tau_r^{-1} = 6R$$
$$\tau_{int}^{-1} = 6R_{int} \quad (8)$$

where R and R_{int} are the overall rotational diffusion constant and internal jump rates, respectively. The geometric constants in the Woessner equation are

$$A = \tfrac{1}{4}(3\cos^2\theta - 1)^2$$
$$B = 3\sin^2\theta\cos^2\theta$$
$$C = \tfrac{3}{4}\sin^4\theta \quad (9)$$

Here θ is the angle between the axes of internal motion and the vector between the ^{13}C and ^1H giving rise to dipolar relaxation.

Using eq 7 Schaefer and Natusch[23] pointed out that the NOE is almost independent of the internal motion whenever $\tau_{int} \ll \tau_r$ and $\omega\tau_r$ is in the range 0.0-0.3 since the first

term in eq 7 dominates under these circumstances. As $\omega\tau_r$ continues to increase beyond this range the overall motion of the molecule again loses efficiency in the relaxation processes and the NOE now is affected by the internal motion. This has been demonstrated in considerable detail by Doddrell, Glushko, and Allerhand[24] in the form of plots of NOE and T_1 vs. τ_r and τ_{int}.

If the observed molecules contain a paramagnetic ion (e.g., Zn^{2+} ion in zinc insulin is substituted by the paramagnetic Ni^{2+} ion), the relaxation times and the NOE's of the ^{13}C in the molecule may be affected due to the large magnetic moment of the unpaired electrons whenever the observed nuclei is sufficiently close to the paramagnetic ion.

The contributions (T_{1M} and T_{2M}) to the overall T_1 and T_2 resulting from the paramagnetic ion include an electron-nuclear dipolar term with equations similar to eq 1 and 2 and a term for the scalar coupling between the electrons and the ^{13}C nuclei. When $\tau_s \ll \tau_r$ where τ_s is the relaxation time of the paramagnetic electrons, the expressions for T_{1M} and T_{2M} can to a good approximation be written as

$$\frac{1}{T_{1M}} = \frac{2}{15}\frac{\mu_{eff}^2\beta^2\gamma_I^2}{r^6}\left[3\tau_s + \frac{7\tau_s}{1+\omega_s^2\tau_s^2}\right] + \frac{2}{3}\frac{S(S+1)A^2}{\hbar^2}\left[\frac{\tau_s}{1+\omega_s^2\tau_s^2}\right] \quad (10)$$

$$\frac{1}{T_{2M}} = \frac{1}{15}\frac{\mu_{eff}^2\beta^2\gamma_I^2}{r^6}\left[7\tau_s + \frac{13\tau_s}{1+\omega_s^2\tau_s^2}\right] + \frac{1}{3}\frac{S(S+1)A^2}{\hbar^2}\left[\tau_s + \frac{\tau_s}{1+\omega_s^2\tau_s^2}\right] \quad (11)$$

Here μ_{eff} is the effective magnetic moment of the paramagnetic ion while S is its spin and A is the scalar coupling constant between the paramagnetic ion and the observed nuclei. The Bohr magnetron is given by β.

The η parameter in the presence of the paramagnetic species is given by[25]

$$\eta = \frac{I_z - I_0}{I_0} = \rho_{CH}S_H\frac{\gamma_H}{\gamma_C}\frac{1}{1+T_{1D}k[E]} \quad (12)$$

whenever the paramagnetic species relaxes the ^{13}C nuclei by a direct mechanism and not through the molecular protons. In eq 12 it is furthermore assumed that no mechanisms contribute to T_1 other than the paramagnetic contribution and the ^{13}C-^1H dipolar contribution given by T_{1D}. The efficiency of the direct mechanism is represented by $k[E]$ and in the case of fast chemical interaction between the paramagnetic species and the observed nuclei is given by eq 10. This assumes that intermolecular interaction with paramagnetic ions can be neglected.

Experimental Section

Chemicals. The pork insulin used in this work was provided in high purity by the NOVO Research Institute. The zinc insulin, purified by column chromatography, yielded only one component when subjected to polyacrylamide gel disk electrophoresis and gel chromatography. These analytical methods can detect the presence of less than 0.5% of noninsulin components. The sample contains 0.49% zinc, corresponding to 2.7 zinc atoms per hexamer, and 14.64% nitrogen. The biological potency was 27.2 IU/mg (24.8–29.9 IU/mg) as determined by the mouse convulsion test. The figures in the brackets give the $P = 0.95$ confidence intervals.

The $KO^{13}CN$ was prepared from 90% ^{13}C enriched KCN[26] purchased from Prochem. All other chemicals obtained commercially were of the highest purity available.

N-Carbamylglycine and *N*-carbamyl-L-phenylalanine were prepared according to known methods.[27,28] The melting points were 173.5–174.5°C (reported[27] 169–170°C) and 195.5°C (reported[28] 195–196°C), respectively, and the yields were 52 and 83%.

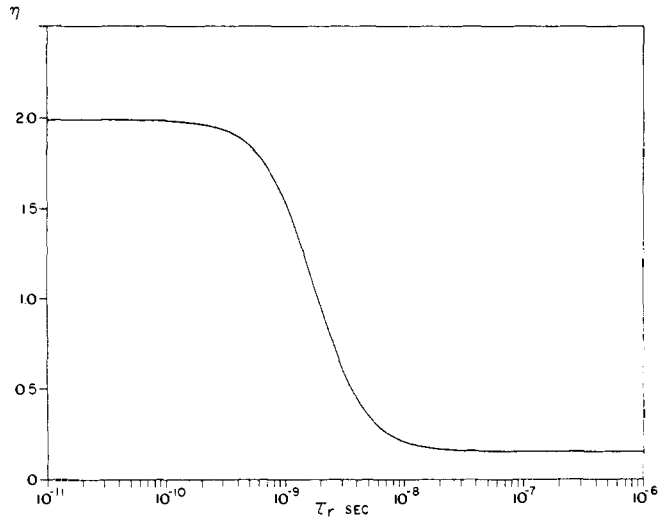


Figure 1. Semilog plot of the differential nuclear Overhauser enhancement parameter, η, for a proton-decoupled ^{13}C nucleus as a function of the correlation time for assumed isotropic ^1H–^{13}C dipolar interaction. The curve correspond to a magnetic field of 23.5 kG.

***N*-ε-Carbamyl-L-lysine.** Two grams (7.2 μmol) of *N*-α-carbobenzoxy-L-lysine (Sigma Chemical Co.) was dissolved in 20 ml of warm water and cooled to room temperature. After addition of 0.874 g of potassium cyanate the solution was warmed on the water bath for 30 min with magnetic stirring and then evaporated to a viscous syrup. On acidification with dilute (1:2) hydrochloric acid and rubbing with a rod, the material crystallized. After addition of 3 ml of water, chilling, and filtration, 2.28 g (7 μmol, 97%) of colorless crystals was obtained. The *N*-α-carbobenzoxy-*N*-ε-carbamyl-L-lysine melted at 115–120°C, turbid melt, with gas evolution at 127°.

The above was dissolved in 60 ml of anhydrous methanol, combined with 0.5 g of 5% palladium–charcoal, and shaken under hydrogen at 60 lb pressure for 3 hr. After addition of 10 ml of water, the catalyst was removed by filtration through Celite and the filtrate was taken to dryness in vacuo. The residue was dissolved in 7 ml of water and 50 ml of methanol and the product precipitated by addition of 100 ml of ether. After standing overnight in the ice box, 1.14 g (86%) of colorless product was obtained which melted at 180°C (in at 160°C) with gas evolution at 181°C (reported melting point 211–212°C).[29] The compound is insoluble (50 mg/l ml) in chloroform, acetonitrile, dioxane, dimethylformamide, and dimethyl sulfoxide.

Preparation of Metal-Free Insulin and Nickel Insulin. The metal-free insulin was prepared from recrystallized pork insulin by a salting-out procedure previously described.[30]

The nickel insulin was prepared from the metal-free insulin by crystallization from a medium containing a surplus of Ni^{2+} ions. The isolated crystals were then recrystallized once from a medium of the same composition as that used from the first crystallization except that no Ni^{2+} ions were added (see ref 30). The Ni^{2+} content of the final crystals was 0.33% corresponding to 2.2 atoms per hexamer.

Preparation of Carbamyl Derivatives. A solution of 220 mg of zinc insulin (37 μmol) and ca. 75 mg, 90% ^{13}C enriched potassium cyanate (925 μmol) in 10 ml water was kept at pH 8.4 and room temperature for 24 hr. The reaction mixture was extensively dialyzed against a 0.05 *M* phosphate buffer, pH 7.8, and concentrated to ca. 4 ml on Amicon PM-10 ultrafilter. To ascertain a maximum carbamylation of the α-NH_2 groups the reaction was repeated as just described except for the pH value, which was kept in the range 7.2–7.4 during another 24-hr period. The final sample was a 3-ml solution of carbamylated insulin in 0.05 *M* phosphate buffer at pH 7.8. The concentration of the protein was ca. 37 mg/ml, according to spectrophotometric measurements.

The nickel insulin derivatives were prepared following the same procedure as described for the zinc insulin, while the carbamylation of the metal-free insulin (sample 1) was performed at pH 7.2–7.4 only. A second sample of carbamylated metal-free insulin (sample 11) was obtained by reacting 300 mg of metal-free insulin

(50 μmol) with 27 mg (333 μmol) potassium cyanate in 15 ml of 0.1 *M* phosphate buffer at pH 7.2. The reaction mixture was kept at 32–35°C for 18 hr and then desalted on a column (40 × 2.5 cm) of Sephadex G-25 with 0.05 *M* ammonium carbonate, pH 8.5, as eluent. For ^{13}C NMR measurements the desalted protein was dissolved in 0.05 *M* phosphate buffer, pH 7.8, to a concentration of 40 mg/ml.

^{13}C NMR Measurements. High-resolution ^{1}H noise-decoupled ^{13}C NMR spectra were obtained at 25.2 MHz using a Varian XL-100-15 spectrometer equipped with Varian Fourier transformation gear and a Varian 620f computer. The width of the 90° pulse was 100 μsec for a carrier frequency which differs 500 Hz from the resonance line. All spectra were recorded using 8K data points and a spectral width of 2500 Hz. This ensures the best possible signal to noise level for a spectrum of sufficiently high resolution. The 100 MHz proton decoupling field was produced by a Hewlett-Packard 5105A frequency synthesizer combined with a Hewlett-Packard 5110B frequency synthesizer driver and amplified with a 10 W ENI Model 310 L Wide Band radiofrequency amplifier connected with a 100 MHz filter. The probe was adjusted to receive the decoupling radiofrequency with practically no reflection. Noise modulation of the H_2 field was achieved by phase modulating the frequency synthesizer with a $(\sin x)/x$ noise band. During all experiments the temperature was kept at 30 ± 1°C.

The spin–lattice T_1 relaxation times were measured employing the inversion recovery technique.[31] Since all the observed resonance signals are within a relatively narrow frequency range (~30 Hz vide infra) the T_1 relaxation times of all lines could be determined simultaneously even though a relative weak rf pulse (H_1 = 2500 Hz) was used. In all experiments a delay time between pulse trains in excess of four times the longest T_1's was used to ensure equilibration of the ^{13}C nuclei.

The η parameters were determined by comparing the intensities of the spectra with and without proton decoupling while the intensities were determined by integrating the signals using a planimeter. The width of the noise band was approximately 2500 Hz with the center of the band placed between the NH and NH_2 proton signals arising from the carbamyl groups. At the magnetic field used here these signals fall within a 60 Hz band. It was found that the η factors were unchanged within experimental error even though the center of the noise band was varied by 300 Hz. The same result was obtained under constant decoupling frequency with a bandwidth of 1500 or 5000 Hz. These results would indicate complete saturation of the protons (S_H = 1 in eq 5).

Results

Amino Acid and Electrophoretic Analysis. The extent of carbamylation in the various insulins was determined by difference in amino acid analysis. The relative decrease of the amino acid residues in the derivatives compared to the unmodified insulins specifies the fraction of amino acids which are carbamylated. The results thus obtained after a 20 hr hydrolysis of the protein in 6 *M* HCl at 110°C are given in Table I for carbamylated lysine, N-terminal glycine, and N-terminal phenylalanine residues. As indicated in the table all three residues have been substantially labeled in the zinc and nickel insulins. Conversely, only a negligible fraction of the lysine residues were labeled in the metal-free sample while the two *N*-terminal residues are still labeled to a large extent. This result holds, of course, also when proper correction is made for a 17–30% recovery of the lysine residue[32] from its carbamylated derivative during the hydrolysis of the protein and might well have been expected[33,34] if one considers the pH and the low concentration of carbamylation reagent used in the preparation of the metal-free sample. To confirm that the lysine residue was relatively unaffected during carbamylation an amino acid analysis of the second carbamylated metal-free insulin (sample II) was undertaken in which the amount of carbamylated lysine was determined both directly, as homocitrulline, and by difference. The uncarbamylated form of lysine was found as before to be in a 20-fold excess. Furthermore, by performing amino acid analyses after 24, 48, and 74 hr of hydrolysis, respectively, it is found as shown in Table I that a significant amount of glycine and phenylalanine are recovered from their carbamyl derivatives during the hydrolysis of the protein, in the same way as reported[32] for lysine.

Table I. Number of Carbamylated Amino Acid Residues per Monomer Insulin According to Amino Acid Analysis

	Lys	N-terminal Gly	N-terminal Phe
Metal-free carbamylinsulin I[a]	0.05 ± 0.05	0.6 ± 0.2	0.4 ± 0.1
Zn^{2+} carbamylinsulin[a]	0.22 ± 0.05	0.6 ± 0.2	0.2 ± 0.1
Ni^{2+} carbamylinsulin[a]	0.24 ± 0.05	0.8 ± 0.2	0.5 ± 0.1
Metal-free carbamylinsulin II duration of hydrolysis, hr			
24	0.03 ± 0.03[b]	0.5 ± 0.1	0.45 ± 0.1
48	0.03 ± 0.03[b]	0.3 ± 0.1	0.3 ± 0.1
74	0.03 ± 0.03[b]	0.15 ± 0.1	0.2 ± 0.1

[a] 20 hr of hydrolysis. [b] Determined both as homocitrulline and by difference (see text).

A polyacrylamide gel electrophoresis was also performed on sample II with the carbamylated metal-free insulin dissolved in 7 *M* urea to a concentration of 1–2 g/l and at pH 8.7. These results show only one strong band (more than 95% of the total amount of protein), migrating slightly faster than expected for uncarbamylated insulin, but definitely slower than expected for insulin carbamylated at the B29 lysine residue, which at pH 8.7 is still primarily protonated. Furthermore, a weak band corresponding to no more than 5% of the total sample was found in the region expected for the lysine carbamylated insulin. It, like the main band, migrated slightly faster than expected for an insulin less one positive charge. We have attributed these minor discrepancies in migration rates to the partial elimination with carbamylation of the *protonated* glycine and phenylalanine residues, which at pH 8.7 will still be present at about 1/20th of the concentration of the unchanged glycine and phenylalanine residues. Thus, both the electrophoretic and amino acid analysis indicates that the carbamylation of the lysine residue is limited to something less than 5%. A similar electrophoresis, performed at pH 4.5, showed two strong bands in a ratio of about 1:2 corresponding to a mono- and a dicarbamylated derivative, respectively. A third band, corresponding to uncarbamylated insulin, indicated that this component was present in only minor amounts. The electrophoretic results agree with the amino acid analysis when the above mentioned recovery of the amino acid residues from the carbamyl derivative during the hydrolysis of the protein is taken into account. Having excluded the possibility of significant lysine carbamylation both on the basis of preparative conditions[33,34] and from direct electrophoretic and amino acid analysis, it may be concluded that mono- and dicarbamylated insulin gives electrophoretic bands that are restricted primarily to labels on the A1 glycine and B1 phenylalanine residues. Since the conditions under which the electrophoresis was performed favor the monomeric form of carbamylated insulin[34] as well as unmodified insulin[35] the electrophoretic results indicate a 0.82 [$1(\frac{2}{3}) + 0.5(\frac{1}{3})$] carbamylation factor on the average of glycine and phenylalanine in good agreement with the results from the amino acid analysis when extrapolating these data to $t = 0$.

Analysis of the ^{13}C NMR Spectra. ^{13}C NMR Fourier transform spectra were obtained on 3–4% solutions of the carbamylated derivatives of metal-free, zinc, and nickel insulin dissolved in phosphate buffer at pH 7.8. In all three

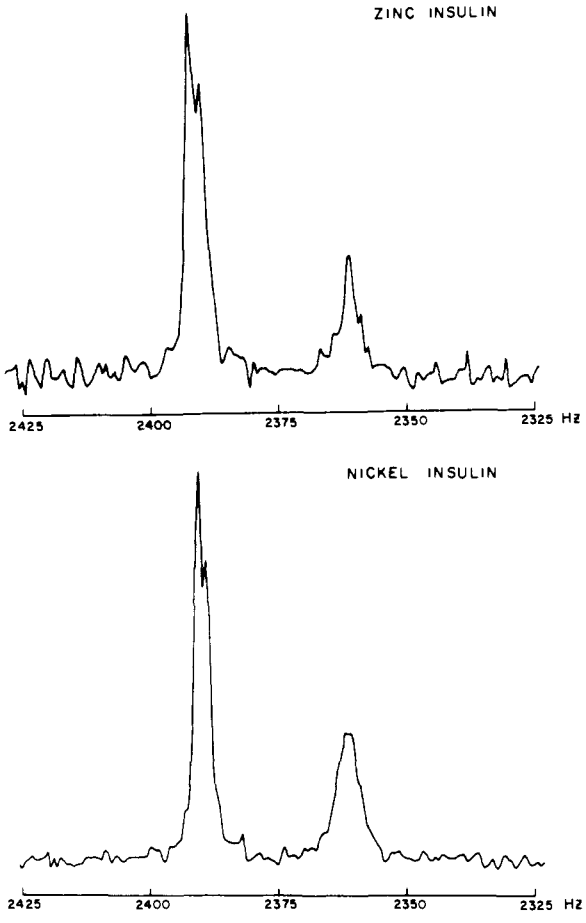


Figure 2. Proton-decoupled ^{13}C NMR Fourier transform spectra of zinc and nickel insulins labeled with 90% ^{13}C enriched carbamyl groups and in about 5×10^{-3} *M* water solutions at pH 7.8 and 30°C. The two observable bands are I and II from higher to lower frequencies. The frequencies indicated in the figure are downfield from *p*-dioxane as external standard.

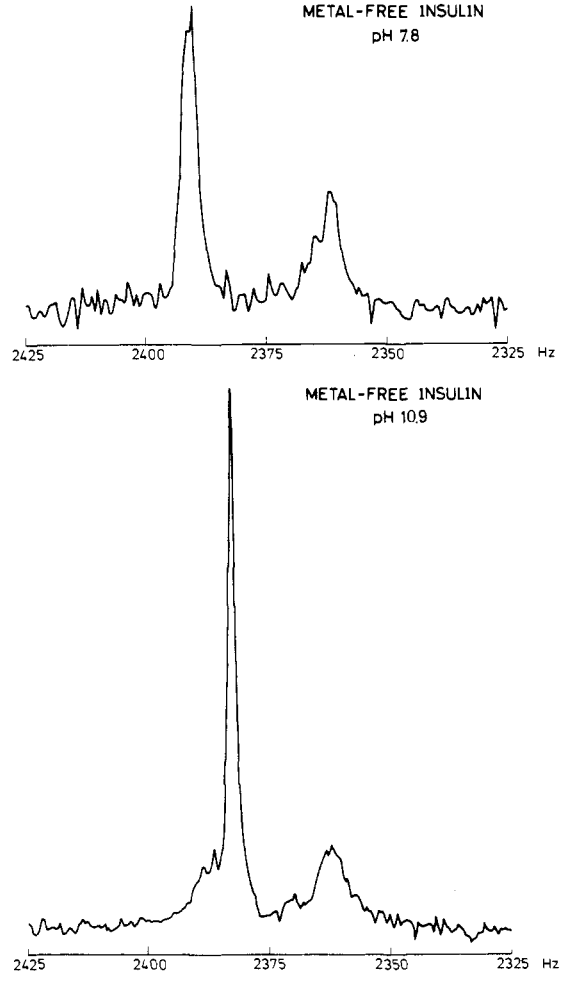


Figure 3. Proton-decoupled ^{13}C NMR Fourier transform spectra of metal-free insulin labeled with 90% ^{13}C enriched carbamyl groups. For further explanation see caption of Figure 2.

cases at least two distinct bands were observed with line structure in the low-field band. Typical spectra are shown in Figures 2 and 3 and the chemical shift values for the three derivatives are presented in Table II.

As indicated in Figures 2 and 3 the ^{13}C enrichment is sufficiently high that interference from the naturally occurring ^{13}C isotope is insignificant.

In order to assign the chemical shifts a comparison was made with the corresponding ^{13}C signals in the carbamyl derivatives of the free amino acids, lysine, glycine, and phenylalanine, and these values are given in Table II. On the basis of these data and the results of the amino acid analysis at least three lines are expected in the ^{13}C NMR spectra of the carbamylated zinc and nickel insulin corresponding to the three different labeled amino acids. Two of the lines due to carbamylated glycine and lysine are expected to be closely spaced from the data taken on the carbamylated free amino acids while a third line arising from the carbamylated phenylalanine is well separated from the other two. As the lysine is carbamylated only to a negligible extent in the metal-free insulin only two separated lines were expected in the ^{13}C NMR spectrum of this derivative. Indeed the ^{13}C NMR spectra of the two metal insulin derivatives are in agreement with the anticipated results, and band I at low field, containing two closely spaced lines, is assigned to the carbamyl carbons on the A1 glycine and B29 lysine residues. Furthermore, the band II at higher field is assigned to the carbamyl carbon on the B1 phenylalanine residues. The ^{13}C NMR spectrum of the metal-free insulin at pH 7.8 (upper spectrum in Figure 3) exhibited an unusual feature in that it also manifests two very closely spaced lines of ap-

Table II. Carbon-13 Chemical Shift Values[a] at 30°C and 25.2 MHz of the Carbamyl Carbons in Carbamylinsulins and Special Monomeric Derivatives

Compd[b]	Concentration, mg/ml	Line Ia	Line Ib	Band II
Insulins				
Metal-free carbamylinsulin	40	2391.1	2390.2	2361.8
Zn^{2+} carbamylinsulin	37	2392.1	2390.2	2362.1
Ni^{2+} carbamylinsulin	33	2390.9	2389.6	2362.7
Amino acids				
ε-Carbamyllysine	133[c]	2392.3		
Carbamylglycine	82[c]	2387.0		
Carbamylphenylalanine[d]	145[c]			2366.1

[a] The chemical shift values are in Hz, downfield relative to dioxane as external standard. The values in ppm can be obtained by dividing by 25.1800. The frequencies were reproducible within 1 Hz. The uncertainties of the relative positions of the lines within the same spectrum are less than 0.5 Hz. [b] These compounds were dissolved in 0.05 *M* phosphate buffer at pH 7.8. [c] Corresponding to a 0.7 *M* solution. [d] Due to the low solubility of the Phe derivative at pH 7.8, the pH was raised to 12 in order to obtain a 0.7 *M* solution. To check if the increase in pH has any effect on the chemical shift value, a spectrum of the glycine derivative was recorded under similar conditions. The measured chemical shift value was identical with the one indicated above, within the experimental error.

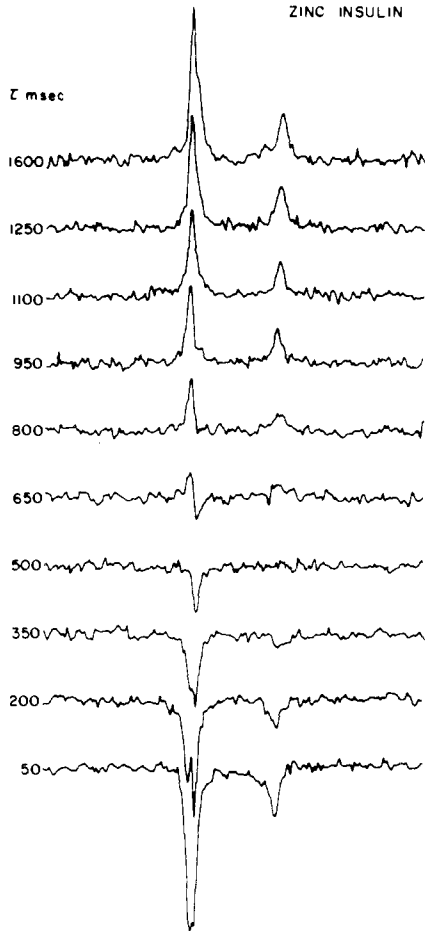


Figure 4. Proton-decoupled ^{13}C NMR partially relaxed Fourier transform spectra of carbamylated zinc insulin. See also caption of Figure 2. The number of transients per spectrum was 2000. The repetition time of the 180°-τ-90° pulse sequence was 4.4 sec. Delay times, τ, between pulses are given in the figure.

Figure 5. Proton-decoupled ^{13}C NMR partially relaxed Fourier transform spectra of carbamylated nickel insulin. For further details see the captions of Figures 2 and 4.

parent equal intensity in the lower field band. The band at higher field arising from the carbamyl carbon on the B1 phenylalanine group showed no line structure although it was somewhat broad. As both of the two closely spaced lines in the lower field band can only arise from the A1 glycine carbamyl group, these results constitute evidence for the presence of the glycine residues in two slightly different chemical environments. It should be emphasized that none of the ^{13}C spectra of the carbamyl derivatives of the three free amino acids glycine, phenylalanine, and lysine show any splitting in the ^{13}C carbamyl lines. Furthermore, the equal intensity of the two lines rules out the correspondence to the 2:1 ratio of di- and monocarbamylated species found in the electrophoretic results. Likewise, the equal intensity feature tends to mitigate against all similar proposals which explain the difference in chemical environments for the two types of glycine on the basis of two different insulin derivatives. Conversely, these intensity data support the presence of a dimeric species in which the two A1 glycine residues are in slightly different molecular environments, when the insulin is dissolved in 0.05 *M* phosphate buffer at pH 7.8. On the other hand, when the same derivative is dissolved in 0.05 *M* phosphate buffer at pH 10.9 the two glycine lines coalesce to a single sharp line, shifted about 0.4 ppm (10 Hz) more upfield, while the phenylalanine line is unaffected, as shown in the lower spectrum of Figure 3. Exactly the same change was observed for a sample at pH 7.8 after the sample had been heated to ca. 60°C for 12 hr. Following this treatment a spectrum recorded at room temperature and 90.5 MHz using a Bruker HX 360 spectrometer showed the same pattern with a relative sharp and a relative broad line, now shifted 72 Hz apart.

In the case of carbamylated zinc and nickel insulin the labeled lysine is present in about one-third of the amount of labeled glycine according to the amino acid analysis. One would therefore expect a carbamylated lysine peak. Assuming the continued presence of two glycine peaks in these cases, the lysine peak could be expected to significantly overlap one of the two glycine peaks. As line Ia is generally broader and more intense than line Ib it is concluded that the lysine signal contributes to line Ia. This assignment is further supported by the facts that the signal from the carbamyl group in carbamylated free lysine is at lower field than in the case of carbamylated free glycine and that the separation between the line Ia and Ib is larger in the lysine carbamylated derivatives of zinc and nickel insulin than in metal-free insulin where the lysine label will be negligible. Using these assignments the agreement between the intensities of lines Ia and Ib and the extent of carbamylation obtained from the amino acid analysis is good. The glycine labeling is assumed of course to be equally distributed between the two separate signals.

Relaxation Parameters (T_1 and NOE) for Insulin Derivatives. Values for T_1 and $S(\infty)$ were extracted from the partially relaxed Fourier transform spectra by a two-parameter least-squares fit of eq 13. In this equation

$$S(t) = S(\infty)\left(1 - K\exp\left(-\frac{\tau}{T_1}\right)\right) \qquad (13)$$

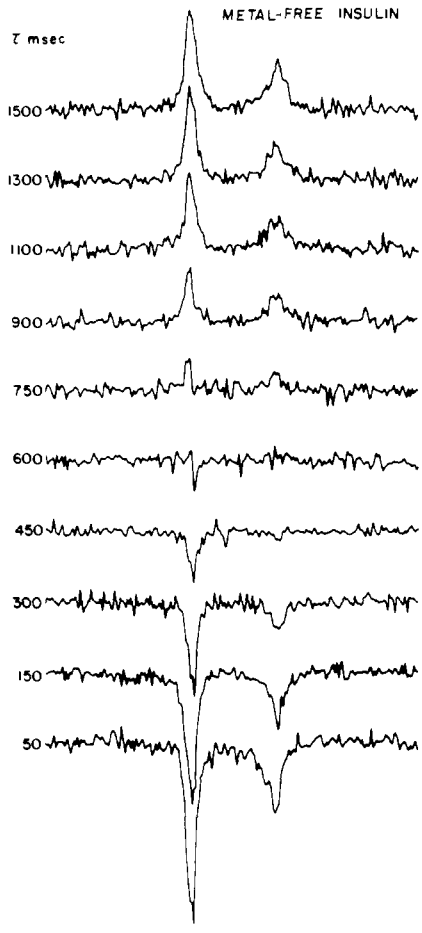


Figure 6. Proton-decoupled ^{13}C NMR partially relaxed Fourier transform spectra of carbamylated metal-free insulin. For further details see captions of Figures 2 and 4.

$K = 2$ for a perfect 180°-τ-90° pulse sequence. In the present experiments, however, since the applied H_1 field is equivalent only to 2500 Hz, the residual static magnetic field H_0 in the rotating frame corresponding to the 500 Hz off-resonance position of the carrier frequency is not negligible, thereby making H_{1eff} slighly unequal to H_1. This causes $|S(0)| < S(\infty)$. A simple calculation using the above mentioned magnitude of H_1 and H_0 indicates that $K = 1.92$ and this value was used in eq 13 to obtain the two-parameter fit of T_1 and $S(\infty)$. Typical plots of partially relaxed spectra are given in Figures 4, 5, and 6, while the success of the exponential fits is shown in Figure 7. Even though the two lines in band I significantly overlap especially in the case of metal-free insulin (Figure 6), it is still possible to follow the differential relaxation of line Ia and Ib separately when the partially relaxed spectra are carefully frequency aligned as in Figures 4-6. It is readily apparent from the spectra in these figures that the relaxation times of the two lines in band I are clearly different as one line passes through the null point before the other. For this intermediate delay time, the two lines are actually oriented in opposite directions and the presence of two different transitions is confirmed. These data also unequivocally rule out the possibility of the splitting arising from a scalar coupling. Table III contains the T_1 data of the carbamyl ^{13}C in the three insulin derivatives together with the 1 σ confidence limits. The NOE data obtained by methods outlined in the experimental section are given in Table IV along with estimated uncertainties in the determination. It is worth noting that no significant difference in the η parameter was observed for the two lines in band I. Also it was found that the

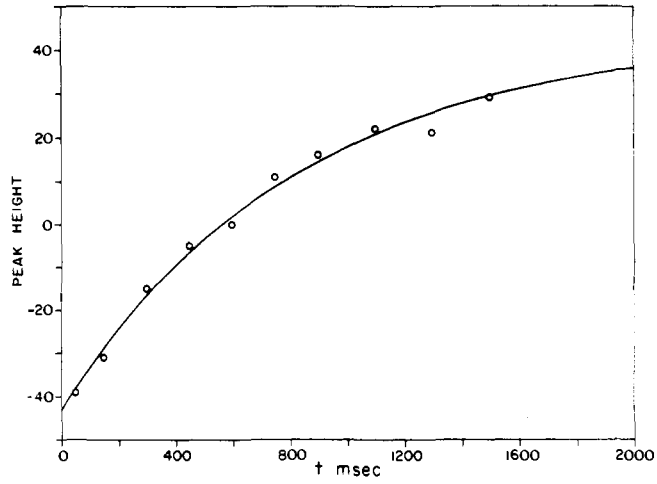


Figure 7. Peak height of band II in the proton-decoupled ^{13}C NMR partially relaxed Fourier transform spectra of metal-free insulin as a function of the delay time between the 180 and 90° pulses. Peak heights are in arbitrary units. The points are the experimental values while the curve is the two-parameter least-squares fit.

Table III. Proton-Decoupled T_1-Relaxation Times for ^{13}C in Carbamyl Groups in Carbamylinsulin at 30°C and 25.2 MHz[a]

	Line Ia in msec	Line Ib in msec	Band II in msec
Metal-free carbamylinsulin	943 ± 16	1210 ± 49	874 ± 17
Zn^{2+} carbamylinsulin	765 ± 13	1155 ± 19	758 ± 18
Ni^{2+} carbamylinsulin	722 ± 22	1164 ± 9	536 ± 16

[a] The indicated uncertainties are the 66% confidence limits.

Table IV. Carbamyl ^{13}C Differential Nuclear Overhauser Enhancement in Carbamylinsulins at 30°C and 25.2 MHz[a]

	Band I	Band II
Metal-free carbamylinsulin	0.8 ± 0.1	0.6 ± 0.1
Zn^{2+} carbamylinsulin	0.6 ± 0.1	0.5 ± 0.1
Ni^{2+} carbamylinsulin	0.6 ± 0.1	0.4 ± 0.1

[a] The numbers are in arbitrary units.

η parameters obtained for a sample of carbamylated metal-free insulin were unaffected within the experimental error by the addition of 1.2 mg/ml of EDTA·H_2O. This would indicate that one can exclude contributions to the measured η parameters as well as to the obtained T_1 values due to paramagnetic ions that, unintentionally, could have been introduced into the system by the phosphate buffer or the sample tubes.

Discussion

Conformation and Aggregation of Insulins. The crystal structure of zinc insulin consists of a hexamer unit cell comprised of three dimers oriented about a three-fold axis defined by the two Zn^{2+} ions. In the individual dimers the monomeric units are tied together by four hydrogen bonds in an antiparallel arrangement, relating the two monomers by an imperfect twofold axis.[36-38] The deviation from twofold symmetry is quite significant in the area of the A1 glycine and the A19 tyrosine residues.[38] In one of the two monomers the A19 tyrosine hydroxyl group appears to be linked through a water molecule to the carbonyl group of the A1 glycine while in the other, approximately twofold related monomer the A19 tyrosine hydroxyl group is more

closely linked with the A5 glutamine residue. These conformational differences may be associated with an adjustment of the side chains along the twofold axis, necessary for establishing the hydrogen bonds between the two monomers.[37] A second interpretation is that the distortions are caused by the specific crystal packing of the hexamers.[37] The effect has also been attributed to sequential zinc binding during the hexamer formation.[39] When the concentration is above 2 mg/ml in neutral or moderately alkaline solutions, zinc insulin is known to form a monodisperse system with a molecular weight corresponding to the hexamer form.[40,41] It is reasonable to assume that this hexamer is essentially the same as that found in the crystal phase,[37] although some changes in the conformation of the dimers might be expected if the deviation from twofold symmetry is caused by the crystal packing. For zinc-free insulin the dimer formation prevails at pH 8,[37] even though further aggregation undoubtedly occurs at high insulin concentrations. Thus, metal-free insulin in 0.03 *M* phosphate buffer at pH 7.02 and 25°C is reported to form polydispersed solutions with an average molecular weight ranging from 9,500 to 55,000 for insulin concentrations in the range 0.1–8.0 mg/mg.[41]

An important question in this study deals with the influence of carbamylation of the insulins on the state of aggregation. As mentioned earlier it has been found[13] that carbamylation of the two end amino groups has no significant effect on the biological potency, while the carbamylation of the B1 phenylalanine group causes a decrease in the immunological potency. Likewise it has been reported[37] that insulin modified at the A1 amino group by carbamyl and other smaller groups give the normal rhombohedral zinc insulin crystals, while insulin acetylated at the B1 amino group cannot be crystallized in the rhombohedral space group.[33] These results indicate that the A1 derivatives form hexamers and may have a similar structure to insulin, while a modification of B1 phenylanine can impair the hexamer formation in agreement with the fact that this residue plays an important role in the formation of the insulin hexamer in the crystal phase.[36]

The zinc insulin samples used here contain 2.7 zinc atoms per hexamer and it is known[42,43] that the presence of more than two zinc atoms per hexamer in solution can cause aggregation beyond the hexamer state. However, it has been found[42] that the binding of the excess zinc is relatively loose and does not occur in phosphate buffer due to stronger complexation of this zinc with phosphate. Likewise the extra zinc can be removed by dialyses at neutral pH. No extra aggregation caused by the surplus of zinc atoms is therefore expected in the present study considering the experimental conditions (vide supra). In order to clarify to what extent the relative high concentration of insulin derivatives used here can cause further aggregation the chemical shift and T_1 values were determined for the carbamyl carbons using samples with twice as high a concentration of zinc insulin and nickel insulin as in the samples mentioned in Table II. However, it was found that the increase in the concentrations causes no changes in the values of the above mentioned parameter within experimental error.

Interpretation of the Results. Perhaps the most unusual feature of the experimental results is the existence of the two closely spaced carbamyl glycine lines in the spectrum of the 4% solution of the metal-free derivative in phosphate buffer at pH 7.8. This indicates that the glycine residue under these conditions exists in two different chemical environments. Since the predominant form of aggregation for metal-free insulin under similar experimental conditions is the dimer, these results constitute evidence that the imperfect twofold axis found in the solid is preserved in the solution. Further evidence for this conclusion is obtained from the spectrum of carbamylated metal-free insulin at pH 10.9 (Figure 3) as well as of the preheated sample at pH 7.8. Thus, the coalescence of the two carbamyl glycine lines observed in these cases strongly suggests that a transition from dimers with two A1 glycines in slightly different environments to identical monomers has occurred. The change in environments of the A1 carbamyl groups that follows this dissociation is, furthermore, indicated by the change in chemical shift of the glycine carbamyl signal. That the chemical shift of the B1 phenylalanine carbamyl signal is unaffected by the increased pH value is also compatible with a dissociation of the dimers since the B1 residues, in contrast to the A1 residues, are far removed from the contact surface between the two monomers in the dimers. Finally, such a dissociation would be expected for the sample at pH 10.9 on the basis of the observation[43] that metal-free insulin in alkaline solution dissociates in proportion to its negative charge, the molecular weight being minimal above pH 9. Since the charge of the insulin molecules in the alkaline region hardly can be affected by carbamylation of the two N-terminal amino groups, it seems reasonable to assume a similar behavior of the carbamylated derivatives in alkaline solution. In contrast to this a difference in aggregation between native and carbamylated insulin should be expected at low pH values, since here the carbamylation of the amino groups reduces the positive net charge that causes dissociation in this pH range.[44,45]

The difference in the relaxation times of the two A1 glycine carbamyl ^{13}C in the sample at pH 7.8 furthermore exhibits that local mobility of these nuclei are different in the two environments. This would be expected if the A1 carbamyl group is tied to the A19 tyrosine hydroxyl group in one monomer but not in the other as suggested[38] by the crystal structure of zinc insulin and discussed briefly above.

The data presented here therefore strongly suggest that the insulin dimers in water solution at pH 7.8 and in the crystalline phase exhibit the same structural features. If this is the case, then the structural details such as the imperfect twofold symmetry found from crystallographic measurements in the dimers are due to structural adjustments required by the monomer–monomer interactions and it is not necessary to propose distortions from either crystal packing or sequential zinc binding with hexamer formation.

In order to assess the degree of motional freedom of the carbamyl ^{13}C in the different positions, the T_1 relaxation time and the η parameter for the carbamyl ^{13}C have been calculated as a function of the correlation time, τ_r, for the overall tumbling of an isotropic tumbler using different values of the correlation time, τ_{int}, for internal rotation about the N-C_α bond. Equations 1 and 3–9 are used to study the effect which these two correlation times have on the NOE and T_1 values. The calculations are made with the reasonable assumption that all three C–H distances are identical and for a Larmor frequency corresponding to a magnetic field strength of 23.5 kG. The geometrical parameters in eq 9 are calculated from the neutron diffraction structure determination of urea.[46] The calculations have been performed with internal motion for two different conformations of the carbamyl group as shown in Figure 8 and the results are presented graphically in Figures 9 and 10.

In this simplified model only dipole–dipole contributions from the three protons in the carbamyl group have been taken into account. Other possible mechanisms like scalar relaxation of the first kind[14] due to coupling to the exchangeable protons at the adjacent nitrogens or scalar relaxation of the second kind[14] arising from interactions with the quadrupolar nitrogen-14 nucleus are not expected to play any significant role in the relaxation of the carbamyl

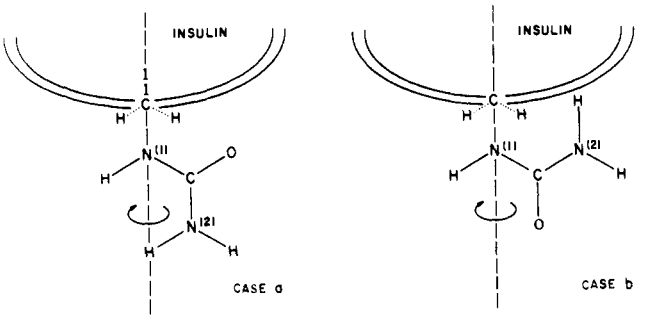


Figure 8. Conformation of the carbamyl group relative to the insulin molecule. The bond distances and angles are taken from the structure of urea.[46] In case a the carbamyl oxygen and the hydrogen on N(1) are in the trans conformation while in case b they are in the cis conformation. The barrier to internal rotation around the N(1)-CO bond is large compared with the barrier around the N(1)-C_α bond. The rotation around this bond is indicated by the arrow. For further details see text.

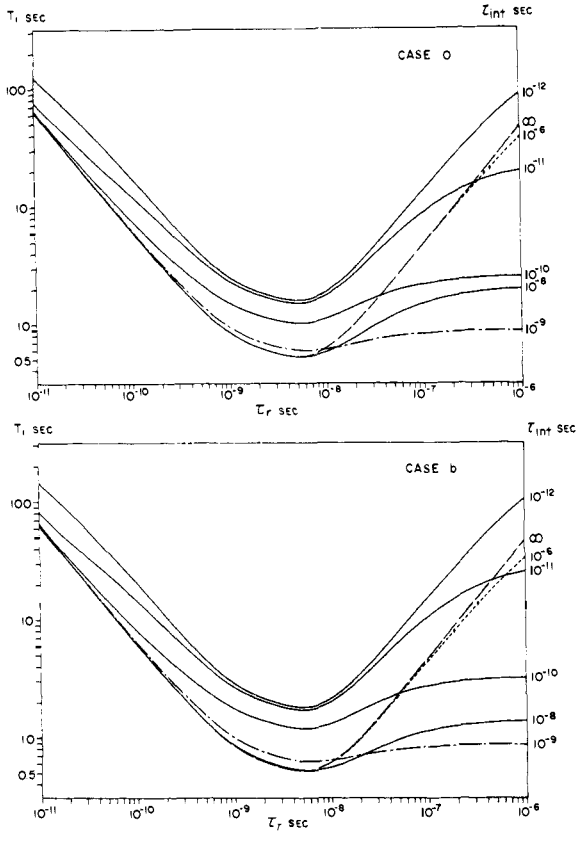


Figure 9. Log-log plot of T_1 for ^{13}C in a carbamyl group attached to an isotropic tumbler as a function of τ_r, the correlation time for the overall tumbling of the molecule, and for different values of τ_{int}, the correlation time for the internal rotation around the N(1)-C_α bond. Cases a and b correspond to the two conformations in Figure 8. The calculations are made assuming complete proton decoupling and for $\omega_C = 1.58 \times 10^{-8}$ rad/sec. Only intramolecular dipolar interactions between the three nitrogen-bound protons and the ^{13}C nucleus are considered.

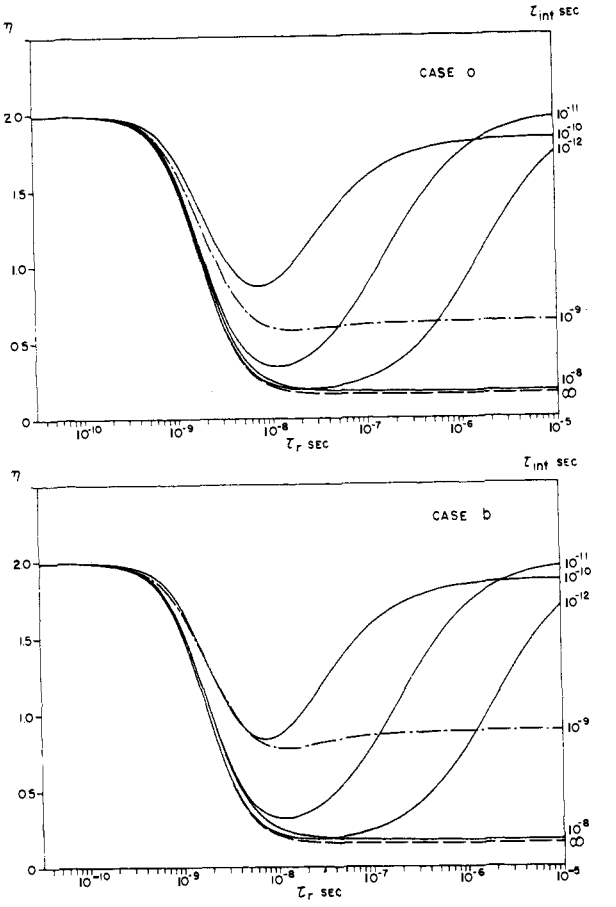


Figure 10. Semilog plot of the η parameter for ^{13}C in a carbamyl group attached to an isotropic tumbler as a function of τ_r and τ_{int}. For further details see caption of Figure 9.

^{13}C in the insulin derivative since no such efficient mechanisms were found in the relaxation of the same carbon in the carbamyl derivatives of the free amino acids where the relaxation times are of the order of 10 sec. This conclusion is based on the reasonable assumption that the correlation time governing proton exchange for the scalar relaxation processes would be the same in the insulin derivatives and the derivatives of the free amino acids. Furthermore, the relative contribution due to ^{14}N quadrupolar interactions should be the same in both instances because quadrupolar relaxation is governed by the reorientational correlation time important in dipole-dipole processes. There may be some dipolar contributions to the relaxation of the carbamyl ^{13}C from other protons such as those hydrogen bonded to the carbonyl oxygen, etc. However, the effect on T_1 of these extra contributions would be to decrease the calculated T_1 and thereby lower the curves in Figures 9 and 10. Thus, the calculated values for the dipolar contribution to the relaxation time of the carbamyl ^{13}C constitute an upper limit.

The rotation about the bonds between the carbamyl carbon and the two nitrogens is negligible compared to the rotation around the N-C_α bond because of the partial double bond character of the two first bonds. For example the barrier to internal rotation about the C-N bond in formamide has been found to be ~19 kcal,[47] while the barrier around the N-C_α bond can hardly exceed a few kilocalories. It is reasonable, therefore, to describe the internal motion in the present case as a rotation around the N-C_α bond of a rigid planar carbamyl group in which the carbamyl oxygen and the hydrogen on N(1) can be either in trans or cis position to each other (see Figure 8).

The plots presented in Figure 10 support the well-known fact that η is almost totally insensitive to the resonance frequency and internal rotation when $0 < \omega_I\tau_r < 0.1$. Conversely, η becomes very sensitive to internal rotation as $\omega_I\tau_r$ increases above 0.3[23] and moves out of the extreme narrowing region. However, it is interesting to observe from Figures 10a and 10b that the η parameter is rather insensitive to the geometrical differences in the two cases exhibiting almost identical dependence on the correlation time for internal rotation. Similar patterns are also found in both models for T_1 as may be observed in Figures 9a and 9b. These re-

sults are interesting when it is realized that the contributions to the total power spectrum, see eq 7, from the individual protons given by the geometrically dependent constants A, B, and C in eq 9 vary considerably and often with opposite signs for different proton–carbon orientations relative to the axis of internal rotation. As more and more protons contribute to the power spectrum, however, they will tend to occupy a greater variety of positions relative to the observed ^{13}C giving an overall average effect which in the case of the two models used here led to very similar results. For the T_1 relaxation times this is only true as long as the distances between the protons and the observed ^{13}C remain unchanged by the variations in geometry, since the magnitude of T_1 depends strongly on the various dipole–dipole distances.

By comparing the experimental T_1 relaxation times and η parameters with the calculated values in Figures 9 and 10, it is possible to estimate the correlation times for the overall tumbling and the internal motion. In the metal-free insulin derivative the relaxation time of the carbamyl ^{13}C on the B1 phenylalanine is 874 msec. If there was no internal motion, this T_1 value would correspond to an overall correlation time, τ_r, of either 1×10^{-9} or 1.4×10^{-8} for either model given in Figure 9. It should be mentioned that if there are dipolar contributions other than those taken into account here, or if the experimental T_1 values contain nondipolar contributions, the effect would be to increase the difference in these two possible values. Comparison of these two τ_r values with the η curves in Figure 10 requires that internal rotation must be involved in order to rationalize the 0.6 η value obtained experimentally. Otherwise a $\eta \geq 1.5$ would be required for $\tau_r \leq 1 \times 10^{-9}$ and for $\tau_r \geq 1.4 \times 10^{-8}$ the η would equal 0.2. When internal rotation is taken into account, the two solutions to τ_r for $T_1 = 874$ msec will, according to either model in Figure 9, both increase for $10^{-8} \geq \tau_{int} \geq 10^{-9}$. According to Figure 10, however, internal rotation has almost no effect on the η parameter at and about $\tau_r \approx 1 \times 10^{-9}$, while in the region $\tau_r \geq 1.4 \times 10^{-8}$, η depends strongly upon τ_{int} and has a value of 0.6 at $\tau_{int} \approx 1 \times 10^{-9}$ in case a and at τ_{int} slightly smaller than 10^{-9} in case b. While η is almost independent of τ_r for this general value of τ_{int} and $\tau_r \geq 10^{-8}$, the T_1 curve indicates a value for τ_r of about 6×10^{-8} in both cases a and b. It should be mentioned that for $\tau_r \approx 6 \times 10^{-8}$ a η parameter of 0.6 is also obtained for $\tau_{int} \approx 10^{-11}$. However, this value for τ_{int} is not compatible with a T_1 value of 874 msec, and can therefore be excluded. Thus, both T_1 and NOE values may be used to obtain the reasonable estimates for both τ_r and τ_{int} of 60 and 1 nsec, respectively.

The relaxation times of the two lines due to the glycine carbamyl groups in the metal-free insulin derivative are somewhat longer than for the ^{13}C in the phenylalanine carbamyl group. Assuming that a single τ_r exists in the overall protein reorientation for labels at both residues, these increases in the T_1 relaxation times correspond to more rapid internal motion according to either model in Figure 9 and thereby to a larger η value in accordance with Figure 10 and the data in Table IV. For $\tau_r \approx 6 \times 10^{-8}$, the T_1 data give $\tau_{int} \approx 0.6$ nsec for line Ia and $\tau_{int} \approx 0.3$ nsec for line Ib. Thus, the carbamyl ^{13}C on B1 phenylalanine has a slightly longer τ_{int} than the carbamyl ^{13}C on both of the two A1 glycines although the B1 residue appears from the crystal structure to be on the surface of the dimers and a relative high degree of motional freedom might be expected as long as the dimers are not a part of a well-defined hexamer where it is known[37] that two B1 phenylalanines are interlaced. These experimental results therefore support the suggestion[37] that on dissociation to dimers, the B1, B2, and B3 residues would move to pack more closely against the body of the monomer.

For the zinc insulin derivatives the experimental values for the T_1 relaxation times and the η parameters are slightly smaller than in the case of metal-free insulin. For the carbamyl ^{13}C at the B1 phenylalanine it is found by similar arguments that $\tau_r \approx 4 \times 10^{-8}$ corresponding to slightly less molecular aggregation than would be found in the metal-free insulins sample. A value of τ_{int} of 1–2 nsec, however, indicates that the local mobility of the carbamylated B1 phenylalanine group is even more restricted in the zinc insulin derivative than in the metal-free compound. This restricted mobility may indicate that the phenylalanine group is playing a more important role in the formation of molecular aggregates in the derivatives of zinc vs. metal-free insulins.

The T_1 and η data for the two lines at 2392.1 and 2390.2 in the zinc insulin derivative cannot be easily analyzed since they, according to the assignments, contain the signals from the carbamyl groups in the B29 lysine and on the two different A1 glycines. According to the data in Table I the contribution from the lysine carbamyl group is about 30% of the total intensity of the two lines, and will contribute accordingly to the relaxation parameters. However, except for the T_1 value for the line at 2390.2 Hz, that stays unchanged, the trend compared with the metal-free insulin derivative is the same as for the carbamyl ^{13}C at the B1 phenylalanine, in that both the relaxation time of the line at 2392.1 Hz gets shorter and the combined η parameter smaller. Because we are on the nonextreme narrowing side of the T_1 minima, this corresponds to a smaller molecular aggregation and a decrease in the internal mobility of the labeled positions.

Like the zinc compound, nickel insulin is known to crystallize in the rhombohedral space group[48] with a minimum nickel content of 2 atoms per hexamer.[49] It seems therefore reasonable to anticipate that nickel insulin and its carbamyl derivatives in solution will behave similar to the corresponding zinc compound with respect to aggregation and peptide chain conformation. As the nickel ion is paramagnetic it can alter the relaxation processes unless the metal is sufficiently far away from the labeled positions that it will have no significant effect on either the relaxation time or the NOE. The data in Tables III and IV show that the experimental T_1 relaxation times for lines Ia and Ib in band I in both the nickel and zinc insulin derivatives are the same in the two cases within the experimental error. The same holds for the combined η parameter for these lines. Since, furthermore, the labeling is the same for the zinc and the nickel insulin according to the amino acid analysis this agreement between the two sets of data therefore indicates that in solution of this species the paramagnetic nickel ion is sufficiently far away from the ^{13}C labels on the A1 glycine and B29 lysine to be unimportant in the relaxation of these nuclei.

For the carbamyl ^{13}C on the B1 phenylalanine, however, there is a significant decrease in the T_1 relaxation time when going from the zinc to nickel insulin derivative. This result is compatible with a simultaneous decrease in the experimental η values. Using the experimental T_1 relaxation times for the B1 carbamyl ^{13}C in both the nickel and the zinc insulin derivatives, and the experimental η parameter for the zinc compound, one may estimate the corresponding η parameter for the nickel insulin with eq 12 to be 0.35. The experimental value of 0.4 ± 0.1 is in good agreement. Assuming that the nonparamagnetic contribution to the relaxation time is the same as in the zinc compound the contribution due to the presence of the paramagnetic ion can be calculated as:

$$\frac{1}{T_{1P}} = \frac{1}{T_{1,Ni}} - \frac{1}{T_{1,Zn}} \qquad (14)$$

which for the carbamyl ^{13}C at the B1 phenylalanine gives $T_{1P} = 1.8$ sec. Furthermore, if the number of insulin molecules that can be complexed to the nickel ions is small compared with the total number of insulin molecules T_{1P} can be written as:[50]

$$\frac{1}{T_{1P}} = \frac{qp}{T_{1M} + \tau_M} \quad (15)$$

where q is the number of insulin molecules that can be bound to a single nickel ion at one time, while p is the number of nickel ion per insulin molecule and τ_M is the average time the insulin molecules spend in the first coordination sphere of the nickel ions. T_{1M} is the pure paramagnetic contribution to the relaxation time given by eq 10. If the exchange of the insulin molecules in and out of the coordination sphere is fast compared with the paramagnetic relaxation rate, that is $(1/\tau_M) > (1/T_{1M})$, T_{1M} will assume from eq 15 a maximum value given by $T_{1M} = qpT_{1P}$.

This preferential effect of the nickel ion on the phenylalanine carbamyl groups over that of the glycine suggests that the interaction sites are closer to the first mentioned residues. The T_{1M} value calculated above and eq 10 may be used, at least in principle, to estimate a maximum value of the average distance, r, between the paramagnetic ion and the carbamyl ^{13}C at the B1 phenylalanine. The lack of an observable isotropic contact shift indicates that the scalar coupling between this ^{13}C and the nickel ion is negligible and therefore only the dipolar term in eq 10 has to be taken into account. In the present case, however, at least two possibilities exist for estimating the paramagnetic effect on the carbamyl relaxation in phenylalanine. First, the T_{1P} contribution to the relaxation time may be caused by a surplus of 0.2 nickel ion per six insulin molecules and p in eq 15 is $0.2/6 = 0.033$. For $q = 3$, the value found in crystalline hexameric zinc insulin, $\mu_{eff} = 2.8\ \mu_B$, the magnetic moment of a free Ni^{2+} ion, and assuming a spin-lattice relaxation time for the electrons[51] of 2×10^{-10} sec the dipolar term in eq 10 gives an $r = 7.2$ Å in the case of fast exchange. For $q < 3$ as well as in the case of slower exchange r would be less than 7.2 Å. Second, the T_{1P} contribution in nickel insulin may also be caused by a general breakdown of the normal crystal phase hexamer and formation of a different type of complex in which the carbamyl ^{13}C at the B1 phenylalanine is closer to the metal ion. Then in the extreme case $p \approx 2.2/6$, and for $q = 3$ one has $pq \approx 1$. This would allow all insulin molecules to be in the first coordination sphere of the nickel ion at the same time. Therefore, by setting $T_{1M} = T_{1P}$ and using eq 15 an average distance between the paramagnetic ion and the carbamyl ^{13}C at the B1 phenylalanine is found to be $r = 10.5$ Å. If $q < 3$ and T_{1P} is affected by an exchange of the insulin molecules in and out of the coordination sphere, then $T_{1M} < T_{1P}$ and thereby $r < 10.5$ Å.

Both of these estimates for the metal-carbon distance use a reasonable but not precisely known value of the electron relaxation time, and therefore only relative structural details can be given for the nickel insulin complex. Failure to observe paramagnetic contributions in the T_1 of the carbamyl group at glycine and lysine indicates that these metal-residue distances are significantly longer than in the phenylalanine case. This result is in contrast with the following metal-amino nitrogen distances obtained from crystallographic work: $Zn\text{-}N_{Gly}$, 19 ± 0.5 Å; $Zn\text{-}N_{Lys}$, 20 ± 2 Å; $Zn\text{-}N_{Phe}$, 22 ± 0.5 and 21 ± 0.5 Å.[52] This discrepancy can be explained by additional complexation in the vicinity of the phenylalanine group with nickel ions arising from a slight surplus or from a breakdown of the hexamer structure as proposed above.

Acknowledgment. This work was supported in part by the U.S. Public Health Services through the National Institutes of Health under Awards GM 08521-14 and RR-00574-04. The Atomic Energy Commission has also recently provided this laboratory with resources under Grant AT(11-1)-2451 for isotopically labeling studies of this type and their support is also acknowledged. One of the authors (J.J.L.) also wants to thank the Danish Scientific Research Board and the NATO Science Fellowship Program for financial support during this work. We wish to thank Professor S. F. Velick for his assistance in preparing samples used in the exploratory stages of this work, and the Carlsberg Laboratory for their assistance with the amino acid analysis. An exponential curve-fitting program of Dr. P. E. Fagerness was used to obtain the T_1 values.

References and Notes

(1) (a) University of Copenhagen; on leave at the University of Utah 1971–1973; (b) University of Utah; (c) NOVO Research Institute.
(2) P. G. Lauterbur, *Appl. Spectrosc.*, **24**, 450 (1970).
(3) A. Allerhand, D. W. Cochran, and D. Doddrell, *Proc. Natl. Acad. Sci. U.S.*, **67**, 1093 (1970).
(4) J. C. W. Chien and J. F. Brandts, *Nature (London), New Biol.*, **230**, 209 (1971).
(5) V. Glushko, P. J. Lawson, and F. R. N. Gurd, *J. Biol. Chem.*, **247**, 3176 (1972).
(6) F. Conti and M. Paci, *FEBS Lett.*, **17**, 149 (1971).
(7) A. Allerhand, D. Doddrell, V. Glushko, D. W. Cochran, E. Wenkert, P. J. Lawson, and F. R. N. Gurd, *J. Am. Chem. Soc.*, **93**, 544 (1971).
(8) R. A. Komoroski and A. Allerhand, *Proc. Natl. Acad. Sci. U.S.*, **69**, 1804 (1972).
(9) R. B. Moon and J. H. Richards, *Proc. Natl. Acad. Sci. U.S.*, **69**, 2193 (1972).
(10) D. T. Browne, G. L. Kenyon, E. L. Packer, H. Sternlicht, and D. M. Wilson, *J. Am. Chem. Soc.*, **95**, 1316 (1973).
(11) D. J. Saunders and R. E. Offord, *FEBS Lett.*, **26**, 286 (1972).
(12) A. M. Nigen, P. Keim, R. C. Marshall, J. S. Morrow, and F. R. N. Gurd, *J. Biol. Chem.*, **247**, 4100 (1972).
(13) D. G. Lindsay, O. Loge, W. Losert, and S. Shall, *Biochim. Biophys. Acta*, **263**, 658 (1972).
(14) J. R. Lyerla, Jr., and D. M. Grant, *MTP Int. Rev. Sci.: Phys. Chem., Ser. One, 1972*, **4**, 155 (1972).
(15) A. Allerhand, D. Doddrell, and R. Komorosk, *J. Chem. Phys.*, **55**, 189 (1971).
(16) I. Salomon, *Phys. Rev.*, **99**, 559 (1955).
(17) K. F. Kuhlmann, D. M. Grant, and R. K. Harris, *J. Chem. Phys.*, **52**, 3439 (1970).
(18) K. F. Kuhlmann, private communication.
(19) D. F. S. Natusch, R. E. Richards, and D. Taylor, *Mol. Phys.*, **11**, 421 (1966).
(20) A. Abragam, "Principles of Nuclear Magnetism", Claredon Press, Oxford, England, 1961, Chapter 8.
(21) J. Ygerabide, H. F. Epstein, and L. Stryer, *J. Mol. Biol.*, **51**, 573 (1970).
(22) D. E. Woessner, *J. Chem. Phys.*, **36**, 1 (1962).
(23) J. Schaefer and D. F. S. Natusch, *Macromolecules*, **5**, 416 (1972).
(24) D. Doddrell, V. Glushko, and A. Allerhand, *J. Chem. Phys.*, **56**, 3683 (1972).
(25) H. S. Gutowsky and D. F. S. Natusch, *J. Chem. Phys.*, **57**, 1203 (1972).
(26) A. A. Plentl and W. T. Kelly, *Biochem. Prep.*, **12**, 103 (1968).
(27) T. L. McMeekin, E. J. Cohn, and J. H. Weare, *J. Am. Chem. Soc.*, **57**, 627 (1935).
(28) H. D. Dakin and H. W. Dudley, *J. Biol. Chem.*, **17**, 34 (1914).
(29) A. C. Kurtz, *J. Biol. Chem.*, **180**, 1253 (1949).
(30) J. Schlichtkrull, *Acta Chem. Scand.*, **10**, 1455 (1956).
(31) R. Freeman and H. D. W. Hill, *J. Chem. Phys.*, **53**, 4103 (1970).
(32) G. R. Stark, *Methods Enzymol.*, **11**, 125 (1967).
(33) D. G. Lindsay and S. Shall, *Biochem. J.*, **121**, 737 (1971).
(34) D. E. Massey and D. G. Smyth, *Eur. J. Biochem.*, **31**, 470 (1972).
(35) D. W. Kupke, *C. R. Trav. Lab. Carlsberg*, **32**, 107 (1961).
(36) M. J. Adams, T. L. Blundell, E. J. Dodson, G. G. Dodson, M. Vijayan, E. M. Baker, M. M. Harding, D. C. Hodgkin, B. Rimmer, and S. Sheat, *Nature (London)*, **224**, 491 (1969).
(37) T. Blundell, G. G. Dodson, D. C. Hodgkin, and D. Mercola, *Adv. Protein Chem.*, **26**, 279 (1972), and references therein.
(38) T. L. Blundell, J. F. Cutfield, S. M. Cutfield, E. J. Dodson, G. G. Dodson, D. C. Hodgkin, D. A. Mercola, and M. Vijayan, *Nature (London)*, **231**, 506 (1971).
(39) A. S. Brill and J. H. Venable, *J. Mol. Biol.*, **36**, 343 (1968).
(40) B. Sjögren and T. Svendberg, *J. Am. Chem. Soc.*, **52**, 2657 (1931).
(41) B. H. Frank and A. J. Veros, 154th National Meeting of the American Chemical Society, Chicago, Ill., Sept 1967, Abstract No. 248.
(42) L. W. Cunningham, R. L. Rischer, and C. S. Vertling, *J. Am. Chem. Soc.*, **77**, 5703 (1955).
(43) E. F. Fredericq, *Arch. Biochem. Biophys.*, **65**, 218 (1956).
(44) P. D. Jeffrey and J. H. Coates, *Biochemistry*, **5**, 489 (1966).
(45) P. D. Jeffrey and J. H. Coates, *Biochemistry*, **5**, 3820 (1966).
(46) J. E. Worsham, Jr., H. A. Levy, and S. W. Peterson, *Acta Crystallogr.*,

10, 319 (1957).
(47) H. Kamei, *Bull. Chem. Soc. Jpn.*, **41**, 2269 (1968).
(48) D. A. Scott and A. M. Fischer, *J. Pharmacol. Exp. Ther.*, **55**, 206 (1935).
(49) J. Schlichtkrull, "Insulin Crystals", Munksgaard, Copenhagen, 1958.
(50) Z. Luz and S. Meiboom, *J. Chem. Phys.*, **40**, 2686 (1964).
(51) I. D. Campbell, J. P. Carber, R. A. Dwek, A. F. Nummelin, and R. E. Richards, *Mol. Phys.*, **20**, 913 (1971).
(52) G. G. Dodson, private communication.

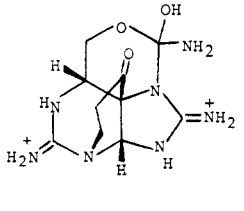
The Structure of a Crystalline Derivative of Saxitoxin. The Structure of Saxitoxin[1]

Jon Bordner,[2a] William E. Thiessen,[2b] Hans A. Bates,[2c] and Henry Rapoport*[2c]

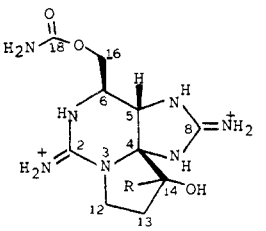
Contribution from the Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27607, the Chemistry Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, and the Department of Chemistry, University of California, Berkeley, California 94720. *Received February 27, 1975*

Abstract: The crystalline ethyl hemiketal of saxitoxin, the paralytic shellfish poison, was obtained, and its structure was deduced as **2b** by X-ray crystallography. Carbon-13 magnetic resonance studies demonstrate that saxitoxin in solution exists primarily as a carbamate rather than a cyclol, and that in aqueous solution, saxitoxin exists as the hydrate **2a**.

Saxitoxin, the paralytic shellfish poison, is one of the most toxic known compounds of low molecular weight. The toxin accumulates in normally edible bivalves such as *Saxidomus giganteus* and *Mytilus californianus* after they injest the marine dinoflagellate *Gonyaulax catenella*.[3] Oxidative degradation showed that saxitoxin hydrate dihydrochloride ($C_{10}H_{15}N_7O_3 \cdot 2HCl \cdot H_2O$) is a tetrahydropurine, and structure **1** was proposed on the basis of ^1H NMR and chemical evidence.[4]



1



2a; R = OH
2b; R = OEt
2c; R = D

Saxitoxin dihydrochloride is an amorphous, hygroscopic material which has never been crystallized.[5] Despite numerous diverse efforts in the past, no crystalline derivative of saxitoxin was ever obtained.[6] We have now obtained a crystalline ethyl hemiketal dihydrochloride, the structure of which was deduced as **2b** by X-ray crystallography. On the basis of this evidence and ^{13}C NMR studies, we now assign structure **2a** to the hydrate of saxitoxin.

X-Ray Diffraction Analysis

Crystals of saxitoxin ethyl hemiketal dihydrochloride (**2b**) suitable for an X-ray analysis were mounted in glass capillary tubes to prevent the loss of solvate molecules. These crystals were surveyed and a 1 Å data set (maximum $\sin\theta/\lambda = 0.5$) was collected on a Syntex $P\bar{1}$ diffractometer equipped with a graphite monochromator and copper radiation (λ 1.5418 Å). All diffraction data were collected at room temperature. Details of the crystal survey and data collection parameters are summarized in Table I.

Atomic scattering factors for C, N, O, and Cl were taken from The International Tables for X-ray Crystallography.[7] Anomalous dispersion corrections for Cl were included.[8] The scattering factor for H is that given by Stewart, Davidson, and Simpson.[9] No corrections were made for absorption (μ = 32.2 cm^{-1}). Routine crystallographic calculations were facilitated by the CRYM crystallographic computer system.[10]

The weights used throughout the refinement were set equal to $1/\sigma^2(F_o^2)$. $\sigma^2(F_o^2)$ was based on the variance of the intensity calculated by the formula: $\sigma^2(I) = S + \alpha^2(B_1 + B_2) + (dS)^2$ where S is the total counts collected during the scan, B_1 and B_2 are the numbers of counts collected for each background, α is the scan time to total background time ratio, and d is an empirical constant of 0.02.

The phase problem was solved using a computer program developed at Oak Ridge National Laboratory.[11] This program facilitated a rather conventional phasing using direct methods, and the portion of the program utilizing transforms of a known molecular fragment was not used. After the intensities were reduced to normalized structure factor magnitudes, $|E|$, triples were selected which met these conditions: (a) their indices add to zero; (b) each had $|E| \geq 1.5$; and (c) the magnitude of the product[12]

$$A(\mathbf{h},\mathbf{k}) = \frac{^2\sigma_3}{\sigma_2^{3/2}} |E(-\mathbf{h})E(\mathbf{k})E(\mathbf{h}-\mathbf{k})|$$

was greater than 1.3.

For each of these triples the quantity[13]

$$D(\mathbf{h},\mathbf{k}) = \langle (|E(\mathbf{h}+\mathbf{l})|^2 - 1) || E(\mathbf{l})| \geq 1.5; |E(\mathbf{k}+\mathbf{l})| \geq 1.5 \rangle_{\mathbf{l}}$$

was computed. Only those triples with $D > 0$ were retained for manipulation. Phases were obtained by a multisolution approach which utilized $D(\mathbf{h},\mathbf{k})$ as the weight assigned to each triple in convergence mapping.[14] Phase extension and refinement were accomplished by a weighted tangent formula[15]

$$\tan\phi(h) = \frac{S}{C} = \frac{\langle w(\mathbf{h},\mathbf{k}) \sin[\phi(\mathbf{k}) + \phi(\mathbf{h}-\mathbf{k})] \rangle_{\mathbf{k}}}{\langle w(\mathbf{h},\mathbf{k}) \cos[\phi(\mathbf{k}) + \phi(\mathbf{h}-\mathbf{k})] \rangle_{\mathbf{k}}}$$

where $w(\mathbf{h},\mathbf{k}) = D(\mathbf{h},\mathbf{k})t(\mathbf{k})t(\mathbf{h}-\mathbf{k})$ and $t(\mathbf{h}) = ((S^2 + C^2)^{1/2}| \sum_{\mathbf{k}} w(\mathbf{h},\mathbf{k})$ is the consistency of the phase for reflection **h** as determined on the previous cycle.

Four phase sets were produced; an E map based on the set with the highest overall self-consistency ($\sum_{\mathbf{h}}|E(\mathbf{h})|t(\mathbf{h})/$