(16) K. Suzuki and I. Kumashiro, *Chem. Abstr.*, **70**, 58231n (1969).
(17) The usefulness of this structural test was pointed out by a referee.
(18) C. G. Skinner, W. Shive, R. G. Ham, D. C. Fitzgerald, Jr., and R. E. Eakin, *J. Am. Chem. Soc.*, **78**, 5097 (1956).
(19) (a) A. G. Beaman and R. K. Robins, *J. Am. Chem. Soc.*, **83**, 4038 (1961); (b) G. B. Elion, E. Burgi, and G. H. Hitchings, *J. Am. Chem. Soc.*, **74**, 411 (1952).
(20) This tricyclic structure was assumed for compounds IVa and IVb in view of the similarities of their preparation and of their spectral properties with those of the 9-ethyl analogue, for which this structure has been unambiguously established.[21b]
(21) (a) S. H. Burstein and H. J. Ringold, *Can. J. Chem.*, **40**, 561 (1962); (b) T. P. Johnston, A. L. Fikes, and J. A. Montgomery, *J. Org. Chem.*, **27**, 973 (1962).
(22) T. Itaya, F. Tanaka, and T. Fujii, *Tetrahedron*, **28**, 535 (1972).
(23) T. C. Myers and L. Zeleznick, *J. Org. Chem.*, **28**, 2087 (1963).
(24) N. J. Leonard, K. L. Carraway, and J. P. Helgeson, *J. Heterocycl. Chem.*, **2**, 291 (1965).
(25) J. W. Jones and R. K. Robins, *J. Am. Chem. Soc.*, **84**, 1914 (1962).
(26) (a) N. J. Leonard and T. Fujii, *J. Am. Chem. Soc.*, **85**, 3719 (1963); (b) Z. Neiman and F. Bergmann, *Isr. J. Chem.*, **6**, 9 (1968).
(27) R. G. Bates in "Determination of pH", Interscience, New York, N.Y., 1973, p 375.
(28) D. J. Brown and S. F. Mason, *J. Chem. Soc.*, 682 (1957).
(29) S. F. Mason, *J. Chem. Soc.*, 1281 (1959).
(30) S. F. Mason, *J. Chem. Soc.*, 3619 (1958).
(31) J. J. Elliott and S. F. Mason, *J. Chem. Soc.*, 1275 (1959).
(32) Cf., for example, C. N. R. Rao in "Chemical Applications of Infrared Spectroscopy", Academic Press, New York, N.Y., 1963, pp 577–583, and references therein.
(33) Cf. pp 20–25 of ref 7a.
(34) S. F. Mason, *J. Chem. Soc.*, 2071 (1954).
(35) E. C. Taylor and P. K. Loeffler, *J. Am. Chem. Soc.*, **82**, 3147 (1960).
(36) The contribution of the benzyl groups to the spectrum of 1,7-dibenzyladenine is considered negligible in this wavelength range.
(37) (a) R. J. Pugmire and D. M. Grant, *J. Am. Chem. Soc.*, **93**, 1880 (1971); (b) R. J. Pugmire, D. M. Grant, L. B. Townsend, and R. K. Robins, *ibid.*, **95**, 2791 (1973); (c) M. T. Chenon, R. J. Pugmire, D. M. Grant, R. P. Panzika, and L. B. Townsend, *Ibid.*, **97**, 4627, 4636 (1975).
(38) D. E. Dorman and J. D. Roberts, *Proc. Natl. Acad. Sci. U.S.A.*, **65**, 19 (1970).
(39) This should be compared with the observation that the alkylation of neutral 1-alkyladenines leads only to 1,9-disubstituted adenines (cf. ref. 14a and A. D. Broom, L. B. Townsend, J. W. Jones, and R. K. Robins, *Biochemistry*, **3**, 494 (1964)).
(40) N. J. Leonard, S. Achmatowicz, R. N. Loeppky, K. L. Carraway, W. A. H. Grimm, A. Szweykowska, H. Q. Hamzi, and F. Skoog, *Proc. Natl. Acad. Sci.*, **56**, 709 (1966).
(41) Cf. P. Beak, F. S. Fry, Jr., J. Lee, and F. Steele, *J. Am. Chem. Soc.*, **98**, 171 (1976), and references therein.
(42) For further examples, cf.: (a) J. Powling and H. J. Bernstein, *J. Am. Chem. Soc.*, **73**, 4353 (1951); (b) M. T. Rogers and J. L. Burdett, *Can. J. Chem.*, **43**, 1516 (1965).
(43) CNDO calculations have been found to be a reliable tool for estimating the dipole moments of this type of compounds; see, for example, ref 3 and 10.
(44) See, e.g., A. G. Moritz, *Spectrochim. Acta*, **18**, 671 (1962).
(45) B. C. Pal and C. A. Horton, *J. Chem. Soc.*, 400 (1964).
(46) See, e.g., M. Eigen, *Angew. Chem., Int. Ed. Engl.*, **3**, 1 (1964).
(47) H. Beens, K. H. Grellmann, M. Gurr, and A. H. Weller, *Discuss. Faraday Soc.*, **39**, 183 (1965).
(48) In fact, we have observed[11] that, in tautomeric α-aminopyridines structurally related to the (VIIIa $\rightleftharpoons$ VIIIb) system, the "direct" nondissociative proton transfer mechanism contributes to the interconversion rate by a factor 10^3–10^4 s^{-1}. However, this is an order of magnitude smaller than the expected contribution from acid catalysis by water in the VIIIa $\rightleftharpoons$ VIIIb interconversion.
(49) Rather conflicting data have been reported for the p*K* value of 1-methyladenosine (cf. D. M. G. Martin and C. B. Reese, *J. Chem. Soc. C*, 1731 (1968), and ref 50). Therefore, we reestimated this value by potentiometric titration at 25 °C and zero ionic strength (Table I). The volume of HCl needed to neutralize a weighted amount of substance was within 3% of theory, indicating that 1-methyladenosine had not rearranged significantly to *N*-methyladenosine during measurements.
(50) R. V. Wolfenden, *J. Mol. Biol.*, **40**, 307 (1969).
(51) C. D. Poulter and G. D. Frederick, *Tetrahedron Lett.*, **26**, 2171 (1975).
(52) See G. E. Baas and L. J. Schaad, *J. Am. Chem. Soc.*, **93**, 4585 (1971), and references therein.
(53) M. D. Topal and J. R. Fresco, *Nature (London)*, **263**, 285 (1976).
(54) See, e.g., A. L. Lehninger, "Biochemistry, the Molecular Basis of Cell Structure and Function", Worth Publishers, 1972, Chapter 29.
(55) M. C. Gefter, *Prog. Nucleic Acid Res. Mol. Biol.*, **14**, 101 (1974).

Hydrogen Bonding between Nucleic Acid Bases and Carboxylic Acids

G. Lancelot

Contribution from the Centre de Biophysique Moléculaire, 45045 Orleans Cédex, France. *Received March 9, 1977*

Abstract: High-resolution proton magnetic resonance (^{1}H NMR) studies on adenine derivatives in chloroform reveal a restricted rotation of the amino group in 9-ethyladenine and in 9-methyl-6-methylaminopurine. ^{1}H NMR studies of butyric acid–6-ethyladenine mixtures provide evidence for H-bonding interactions and for double proton magnetization transfer between amino and hydroxylic groups. Measurements of equilibrium constants (K) by ^{1}H NMR and absorbance experiments in $CDCl_3$ at 303 K for association of nucleic acid bases with butyric acid (used as a model of side chain for glutamic or aspartic acid) lead to the following preferential order of associations: 2-dimethylamino-9-methylguanine (K = 660) > 1-cyclohexylcytosine (K = 270) > 9-ethyladenine (K = 160) > 1-cyclohexyluracil (K = 80). Comparative studies of interactions with methylamino derivatives of adenine allow the computation of association constants for the two types of 1:1 complexes with two hydrogen bonds formed between 9-ethyladenine and butyric acid, K_1/K_7 = 2.8 (1 and 7 refer to the position of the nitrogen atom which is bound to the OH group). The results are discussed with respect to the contribution of charge transfer, new repartition of electronic density, and new geometry of monomers in the complexed state superimposed on direct Coulombic interactions.

Among the fundamental mechanisms which can contribute to the specificity of recognition of nucleic acids by proteins and enzymes, direct interactions through hydrogen bonds might be of great importance. Some work has already been devoted to such studies in chloroform[1,2a,b] and cyclohexane.[2a,b] Association constants between amino acid side chains and 9-ethyladenine or 1-cyclohexyluracil have been found in the range 1–22 000 in cyclohexane at 10 °C. Although studies in chloroform have shown that competition with solvent is important, the order of specificity is respected: 1-cyclohexylcytosine > 9-ethyladenine > 1-cyclohexyluracil for association with butyric acid. Since several 1:1 complexes can be made between nucleic acid bases and butyric acid, the constants reported in these studies were the sum of the association constants corresponding to the different 1:1 complexes. Sites of H-bonding interactions cannot be altogether available during the recognition of nucleic acids by proteins and enzymes. It is therefore important to know the association constants of the different 1:1 complexes. The present paper is concerned with an investigation of the interactions of the four nucleic acid bases with butyric acid used as a model for the side chain of glutamic or aspartic acid. Among the number of techniques used to study hydrogen-bond formation, proton magnetic resonance and infrared measurements have been the most

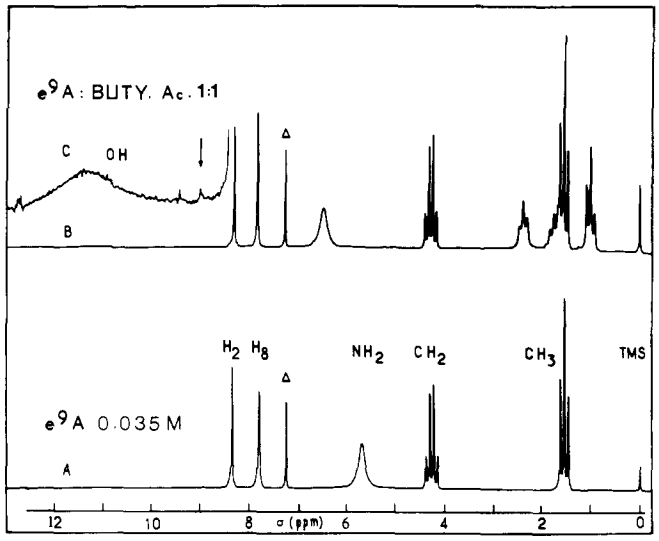

Figure 1. ^1H NMR spectra of 3.5×10^{-2} M e^9A alone (A) and in the presence of 3.5×10^{-2} M butyric acid (B and C) in $CDCl_3$ at 303 K. Δ shows the resonance line of $CHCl_3$ and the arrow shows the position of the OH resonance line for 3.5×10^{-2} M butyric acid in the absence of e^9A. The intensities of resonance lines are 16 time expended in spectrum C.

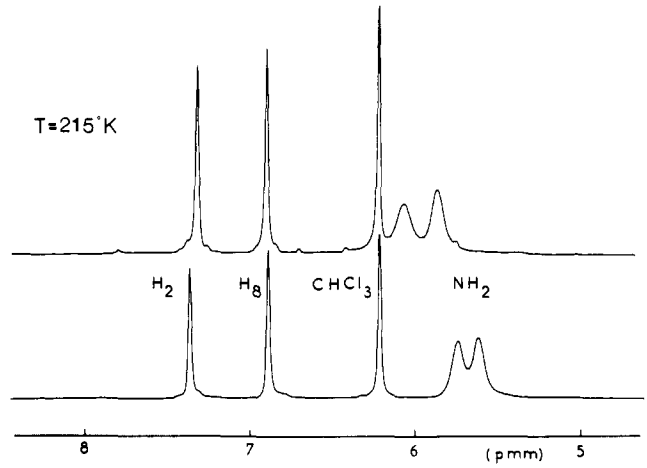

Figure 2. ^1H NMR spectra of 2×10^{-2} M e^9A in the absence (bottom spectra) and the presence (top spectrum) of 2×10^{-2} M butyric acid at 215 K. We can see the two resonance lines of amino protons due to restricted rotation.

powerful because they allow the direct observation of the linked protons. We have investigated hydrogen-bond formation between adenine, uracil, cytosine, guanine derivatives, and butyric acid by proton magnetic resonance and by UV absorption.

Experimental Section

9-Ethyladenine (e^9A), 6-methylamino-9-methylpurine (m^6A), 6-dimethylamino-9-ethylpurine (m_2^6A), 1-cyclohexyluracil ($(chx)^1U$) and 2-dimethylamino-6-hydroxy-9-methylpurine (m_2^2G) were purchased from Cyclo Chemical. 1-Cyclohexylcytosine ($(chx)^1C$) was prepared by thiation of 1-cyclohexyluracil and amination of sulfur compound.[3] 2′,2′,5′-Triacetylguanosine was purchased from Aldrich. Butyric acid was fractionally distilled and after gas chromatography analysis only the end of the distilled product was used and kept on appropriate molecular sieves. During this work, great care was taken to avoid the presence of any water or polar impurities in deuterated chloroform (C.E.A.). Chloroform was purified on an alumina column and kept on appropriate molecular sieves, NMR spectra were obtained with a Bruker WH 90 Fourier transform spectrometer. For all experiments the spectrometer magnetic field was locked on an internal deuterium reference (deuterated chloroform). The temperature was regulated at ±0.5 °C and known at ±3 °C. The positions are given in parts per million (ppm) downfield from tetramethylsilane (Me_4Si = 0.0) at 303 K. For other temperatures, the position of the Me_4Si resonance line was corrected with respect to its position at 303 K. The resonances at H_2 and H_8 protons of the adenine analogues were assigned by deuterium replacement of the slightly acidic H_8 proton, using the method of Chan et al.[4] The other NMR spectral assignments of nucleic acid bases were previously reported by Ts'o et al.[5]

Difference absorption spectra were recorded on a Cary 14 spectrophotometer. All spectra were obtained in 1-cm path length cells. Both reference and sample cells contained the investigated base and were thermostated at 303 ± 0.5 K. The difference spectra were recorded on 0–0.2 scale absorbance after addition of amino acid derivatives in the sample cell. The ratio between the added volume and the total initial volume was always less than 0.004, so that base dilution was neglected.

Results

e^9A–Butyric Acid Associations. Evidence for Complex Formation. Double Proton Transfer. Figure 1 shows the ^1H NMR spectra of e^9A and e^9A + butyric acid in $CDCl_3$ at 303 K. The binding of butyric acid to e^9A leads to a downfield shift of amino and H_8 protons and an upfield shift of the H_2 proton of e^9A. The downfield shift of proton signal resonances is a general characteristic of protons participating in hydrogen bond formation.[6] Only one signal resonance of adenine NH_2 protons was observed either in the free or complexed state. Generally the slow time scale of NMR spectroscopy prevents observation of resonances for both associated and free states of two magnetically inequivalent protons. By lowering the temperature, however, two distinct peaks of amino protons appear below 230 K (Figure 2) and the splitting between these two peaks increases until the freezing point of $CDCl_3$ is reached, showing a restricted rotation of the amino group. By adding butyric acid the two resonance lines of the amino group are shifted downfield and their splitting increases (Figure 2). This increase of the splitting could be due to a stabilization of the restricted rotation of NH_2 upon complex formation or to a different shift for each proton if two different 1:1 complexes are formed. The effect of saturating the spin levels by narrow-band irradiation of the acid hydroxylic proton leads to a vanishing of e^9A amino proton resonances. Conversely the irradiation of amino proton of e^9A decreases the intensity of the hydroxylic proton resonance of butyric acid. This experiment provides evidence for transfer of magnetization between protons of NH_2 (e^9A) and OH (acid). Since these chemical groups are not directly linked by an H bond, this must result from a double proton transfer. The amino group of e^9A can be deuterated by dissolving e^9A in D_2O followed by lyophilization. ^1H NMR spectra show 90% effective deuteration. Upon adding butyric acid, exchange of protons takes place between NH_2 and OH groups. The integral of the ^1H NMR spectra shows the same loss of proton from acid than gain of proton by e^9A. This also is evidence for double proton transfer.

Determination of Association Constants. Binding of e^9A and butyric acid could lead to the formation of two 1:1 complexes with two hydrogen bonds and one 1:1 complex with only one hydrogen bond (Figure 3). The equilibrium:

$$A + B \rightleftharpoons (AB)_i$$

leads to

$$[AB]_i = K_i[A][B] \quad (i = 1, 3, 7) \qquad (1)$$

where K_i represents the association constants of the different complexes, the subscript indicating the N atom involved in H bonding. Let δ_0^j and δ_{1i}^j represent the chemical shifts of proton j in the free molecule and in complex i, respectively. The observed chemical shift will be:

$$\delta^j = \delta_0^j \frac{[A]}{[A_0]} + \sum_i \delta_{1i}^j \frac{[AB]_i}{[A_0]} \qquad (2)$$

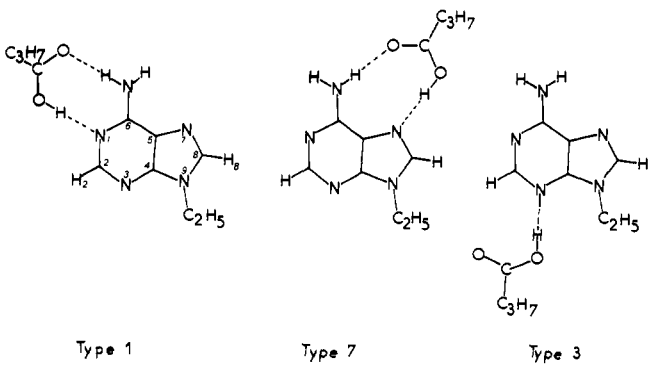

Figure 3. Geometry of the three types of H-bonded complexes formed between e^9A and butyric acid. The geometry of e^9A was taken from x-ray crystallographic data on 9-methyladenine[18] and geometry of butyric acid was obtained from data reported for formic acid.[19]

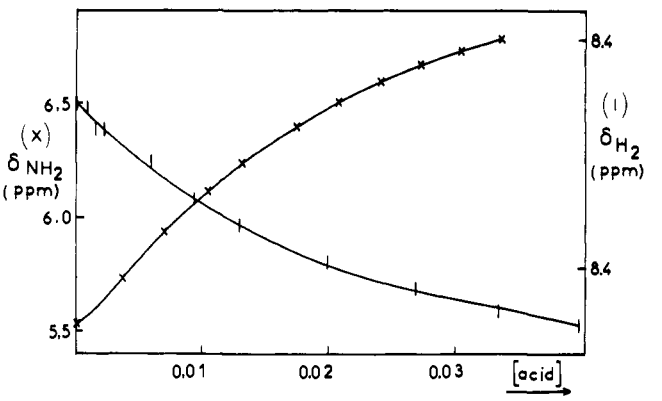

Figure 4. Concentration dependence of e^9A proton resonances (2×10^{-2} M) in the presence of butyric acid in $CDCl_3$ at 303 K. Experimental shifts (error bars for H_2) are fitted by a least-squares program (full lines). The parameters computed are listed in Table I. The small slope of $\delta(NH_2)$ for small concentrations in acid shows the effect of e^9A self-association.

where $[A_0]$ is the total concentration in molecule A. Equations 1 and 2 lead to a variation of chemical shift:

$$\Delta\delta = \delta^j - \delta_0{}^j = \sum_i (\delta_{1i}{}^j - \delta_0{}^j) \frac{K_i}{K} \frac{[AB]}{[A_0]} \qquad (3)$$

with

$$[AB] = \sum_i [AB]_i$$
$$K = \sum_i K_i \qquad (4)$$

A fit of the variation of $\Delta\delta$ vs. concentration in A or B will give:

$$\sum_i (\delta_{1i}{}^j - \delta_0{}^j) \frac{K_i}{K} \qquad (5)$$

and the concentration in AB or, in other words, $K = \Sigma_i K_i$.

The ratio of changes in chemical shifts of H_2, H_8, and NH_2 protons of e^9A (1.4×10^{-2} M) is constant when mixed with butyric acid for concentrations in acid less than 3×10^{-2} M, showing the existence of only 1:1 complexes. By use of a least-squares program, we have computed the association constant $K = K_1 + K_3 + K_7 = 160$ (Figure 4) taking into account the self-association constant of butyric acid as well as that of e^9A (which have been computed by a fit of the concentration dependence of chemical shifts of OH or NH_2 proton resonances). All results are given in Table I.

Restricted Rotation of the Amino Group of m^6A. Engel and Von Hippel[7] have demonstrated the restricted rotation of the amino group of 6-methylamino-9-methylpurine. By measuring the intensities of the two singlet peaks of the methylamino group at low temperature, they have computed an equilibrium population ratio of 96:4 (syn:anti) in CD_3OD. By the same procedure we have computed an equilibrium population ratio of 95:5 in $CDCl_3$ which does not change between 220 and 250 K. Since the restricted rotation data on m^6A (Figure 6) can be interpreted in terms of an equilibrium preference for the rotamer in position syn, viz. CH_3 on the N_1 side,[7] the determination of the association constant of m^6A with butyric acid can allow us to compute the ratio of the association constants for the two 1:1 hydrogen-bonded complexes between e^9A and butyric acid.

m^6A, $m_2{}^6A$, and Butyric Acid. Interaction between m^6A and butyric acid leads to the formation of two 1:1 complexes, whose probability of formation is governed by the ratio K_1/K_7, and the possibility of finding the methylamino group in position syn or anti. Adding butyric acid to a solution of m^6A in $CDCl_3$ at 303 K shifts the amino and H_8 resonance lines downfield, whereas the H_2 resonance line is shifted slightly upfield. The computed association constant is $K = 54$ (Table I).

The value of association constant K_3 can be estimated by investigating complex formation between $m_2{}^6A$ and butyric acid. As a matter of fact three complexes with only one H bond could exist in this system. The computed association constant of the three complexes is $K = 15$ (Table I). By assuming that the probability of formation is equal and by taking into account the population ratio of the two rotamers of m^6A we computed (see Discussion): $K_1 = 114 \pm 20$ M^{-1}, $K_3 = 5 \pm 0.5$ M^{-1}, and $K_7 = 41 \pm 7$ M^{-1}, i.e. $K_1/K_7 = 2.8 \pm 1$.

1-Cyclohexyluracil and Butyric Acid. Interaction of butyric acid (self-associated for concentrations used in the range 10^{-2} to 9×10^{-2} M) with $(chx)^1U$ leads to an upfield shift of the hydroxylic resonance line of acid and the vanishing of the NH

Table I[a]

	K, M^{-1}	$\Delta\delta$, ppm	δ_0, ppm	K (base self-association)	Proton used for fit
e^9A	160 ± 15	−2.37 ± 0.20	5.344	2.28	NH_2
	140 ± 35	0.17 ± 0.02	8.372	2.28	H_2
	150 ± 50	−0.052 ± 0.005	7.806	2.28	H_8
m^6A	55 ± 5	−4.14 ± 0.34	5.514	1.1	NH
	53 ± 13	−0.11 ± 0.01	7.695	1.1	H_8
$m_2{}^6A$	15 ± 2	−0.24 ± 0.02	7.721	0	H_8
$(chx)^1U$[b]	80 ± 15	$\Delta\epsilon_{288} = 2400 \pm 200$		0	
$(chx)^1C$	270 ± 44	−5.09 ± 0.18	2.176	0	OH
$m_2{}^2G$[c]	660 ± 60	$\Delta\epsilon_{298} = 1200 \pm 100$		0	

[a] Equilibrium constants and chemical shifts computed for 1:1 complex formation between nucleic acid bases and butyric acid at 303 K in $CDCl_3$. The dimerization constant for butyric acid ($K = 80 \pm 15$ M^{-1}) was computed by a least-squares program using the chemical shift of OH proton vs. concentration in acid at 303 K in $CDCl_3$. [b] K value of 60 ± 20 M^{-1} was reported by Sellini et al.[1] from infrared measurements at 308 K. [c] K was computed from UV absorption measurements in $CHCl_3$.

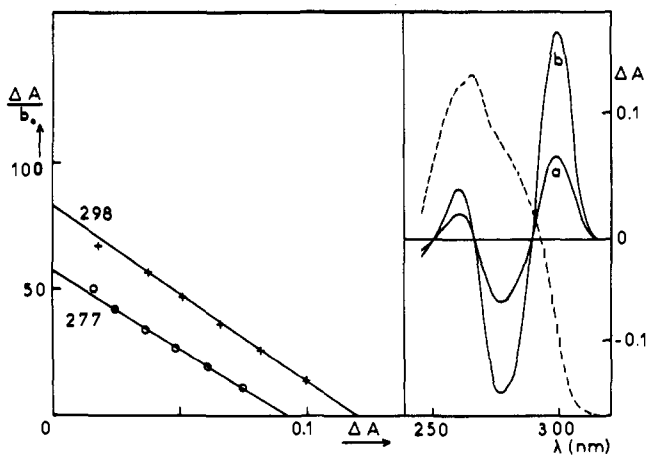

Figure 5. Difference absorption spectra of m_2^2G (10^{-4} M) in the presence of butyric acid in $CHCl_3$ at 303 K (l = 1 cm). (Right): (a) [acid] = 6.32 × 10^{-4} M; (b) [acid] = 4.73 × 10^{-3} M. The broken line shows the absorption spectrum of m_2^2G 10^{-4} M (scale must be multiplied by 5). (Left): Plot of $\Delta A/b_0$ vs. ΔA at 298 and 277 nm. We computed at 298 nm K = 691 M^{-1} and $\Delta\epsilon$ = 1200 and at 277 nm K = 625 M^{-1} and $\Delta\epsilon$ = 930. Here $b_0 \gg a_0$ and $\Delta A/b$ is computed by taking into account the self-association of butyric acid (K_D = 80).

resonance line of (chx)^{1}U. The integrated ^{1}H NMR spectra show that OH and NH resonance lines are superimposed. By taking into account the concentration in OH and NH protons, we can show that the OH resonance line is shifted upfield by interaction with (chx)^{1}U. This can be attributed to a downfield shift greater for self-associated acid species than for (chx)^{1}U-butyric acid complexes, the net result being an upfield shift. This superposition of resonance lines as well as the small changes of chemical shifts observed for non-H-bonded protons of the (chx)^{1}U aromatic ring did not allow us to compute the association constant. Table I gives the value reported by Sellini et al.,[1] $K = 60 \pm 20$ M^{-1}, in $CHCl_3$ at 308 K obtained from IR measurements as well as the value computed from UV absorbance results, $K = 80 \pm 15$ M^{-1}, in $CHCl_3$ at 303 K (Lancelot[2a]; in that paper we reported $K = 170 \pm 30$ M^{-1} in $CHCl_3$ at 283 K).

1-Cyclohexylcytosine and Butyric Acid. As for (chx)^{1}U, the broadening of the NH_2 resonance, as well as the small changes in chemical shifts of H_5 and H_6 of (chx)^{1}C in the presence of butyric acid, did not allow us to compute the association constant. We have only fitted the experimental chemical shifts of the hydroxylic proton of butyric acid to compute the association constant K = 270 M^{-1} (Table I).

Interaction between Guanine Derivatives and Butyric Acid. Guanine and its derivatives are known to have very low solubility in nonpolar solvents. Accumulation of 40 000 ^{1}H NMR spectra of 2′,3′,5′-triacetylguanosine (10^{-4} M) made it possible to observe resonance lines of ring and sugar protons but did not allow us to detect the most important resonance lines for H-bonding studies, namely NH and NH_2. Moreover, photochemical reactions preclude using UV difference spectra in the presence of butyric acid. We have investigated the interactions between butyric acid and 2-dimethylamino-6-hydroxy-9-ethylpurine (m_2^2G) which is soluble enough in $CDCl_3$ for ^{1}H NMR and absorption studies. ^{1}H NMR spectra of m_2^2G (7 × 10^{-3} M) in the presence of butyric acid show a broadening of the N_1H resonance line of m_2^2G as well as a mixing with resonance line of butyric acid protons (OH). We have therefore used difference absorbance spectra to compute the association constant. The absorption spectrum of m_2^2G (10^{-4} M) in $CHCl_3$ at 303 K is very sensitive to the presence of small amounts of butyric acid (Figure 5). The difference spectra show a red shift of the first absorption band. This is typical for hydrogen-bonding effects on the first $\pi\pi^*$ absorption band of a molecule whose chromophore acts as a proton donor.[6] For butyric acid concentrations (b_0) greater than that of m_2^2G (a_0), complex concentration can be calculated according to:

$$c = a_0 b_0 K/(1 + b_0 K) \quad (6)$$

where K refers to the association constant in the equilibrium:

$$A + B \rightleftharpoons C \quad (7)$$

The absorbance measured at a wavelength where B does not absorb is:

$$A_2 = [(a_0 - c)\epsilon_A + c\epsilon_C]l \quad (8)$$

where δ_A and δ_C are the extinction coefficients of A and C, respectively, and l is the path length. For $b_0 = 0$, the absorbance is:

$$A_1 = a_0 \epsilon_A l \quad (9)$$

The absorbance difference measured is:

$$\Delta A = A_2 - A_1 = c\Delta\epsilon l \quad (10)$$

where

$$\Delta\epsilon = \epsilon_c - \epsilon_a \quad (11)$$

One can then write as done in ref 8 and 9:

$$\Delta A/b_0 = -K\Delta A + \Delta\epsilon a_0 l \quad (12)$$

As Person[10] has previously noted, by plotting $\Delta A/b_0$ vs. ΔA, we are able to obtain good values of both K (slope) and $\Delta\epsilon$ from spectrophotometric measurements if:

$$0.1K_1^{-1} \leqslant b_0 \leqslant 9K^{-1} \quad (13)$$

Such a linear relation is shown in Figure 5 from which we determined $K = 660 \pm 60$ M^{-1}. The dimerization of butyric acid was taken into account (K_D = 80 M^{-1}) and b_0 was taken as the concentration in acid monomers.

Discussion

Equilibrium constants for the association of butyric acid with e^9A, m^6A, and m_2^6A are the sum of the association constants of several 1:1 complexes. Methylation of the amino group of adenine leads to the appearance of two new complexes with only one hydrogen bond (K_1' and K_7'). Since $K(m_2^6A) = K_1' + K_7' + K_3 = 15$ is small as compared to $K(e^9A)$ and $K(m^6A)$, the exact values of K_1' and K_7' are not very important in determining the ratio K_1/K_7 as shown by the following calculations using different assumptions: (i) if $K_1' = K_7' = 0$—$K_1 = 109 \pm 20$ M^{-1}, $K_3 = 15 \pm 2$ M^{-1}, $K_7 = 36 \pm 6$ M^{-1}, $K_1/K_7 = 3 \pm 1$; (ii) if $K_1' = K_7' = K_3$—$K_1 = 114 \pm 20$ M^{-1}, $K_3 = 5 \pm 1$ M^{-1}, $K_7 = 41 \pm 7$ M^{-1}, $K_1/K_7 = 2.8 \pm 1$; (iii) if $K_3 = 0$ and $K_1' = K_7'$—$K_1 = 116 \pm 20$ M^{-1}, $K_3 = 0$, $K_7 = 44 \pm 7$ M^{-1}, $K_1/K_7 = 2.6 \pm 1$.

For each of these assumptions K_1/K_7 is in the range 2.6-3. From the radius of van der Waals and the possibility of rotation around the O-H bond (viz. O side to C_2 or C_8), the steric effect of the NCH_3 group or H atom on the oxygen atom is not sufficiently important to prevent the formation of complexes with one hydrogen bond on N_1 and N_7. Therefore, we conclude that: $K_1' \simeq K_7' \simeq K_3 = 5$ and $K_1/K_7 = 2.8 \pm 1$. Linear H bonds are stronger than bent H bonds. The geometry of complexes (Figure 3) is such that O-H···N and O···H-N are linear for a complex of type I and slightly bent for type 7. This could explain that K_1 is about three times greater than K_7.

We have been able to measure the equilibrium constant for association of m_2^2G with butyric acid in $CHCl_3$ at 303 K. Methylation of the amino group of guanine prevents the formation of a complex with two hydrogen bonds involving NH_2

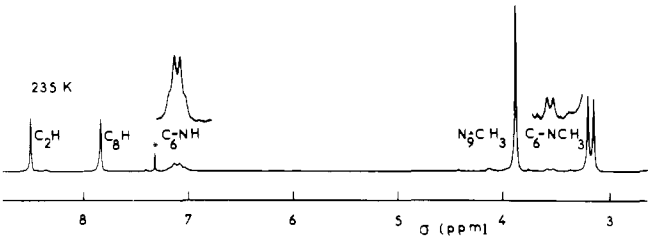

Figure 6. ^1H NMR spectrum of m^6A in $CDCl_3$ at 235 K showing the restricted rotation of the $NHCH_3$ group. Two resonance lines are observed for the NCH_3 group (3.56 and 3.18 ppm) corresponding to the two rotamers. The expended spectrum has been multiplied by 8.0f the two NH resonances only one could be detected at 7.10 ppm. The peak marked by an asterisk is residual $CHCl_3$.

Table II[a]

Adenine	N_1, −0.30944	NH_2, H side to N_1	0.32803
	N_7, −0.26798	H side to N_7	0.37406
Cytosine	N_3, −0.33509	NH_2, H side to N_3	0.33598
Thymine	O_2, −0.40615	N_3H, H	0.36648
	O_4, −0.36348		
Guanine	O, −0.36836	N_1H, H	0.35188

[a] Net atomic charges on atoms participating in H bonds for nucleic acid bases. All electron SCF-LCAO-MO computations were reported by Clementi et al.[23]

and N_3. The poor solubility of guanine in nonpolar solvents has led several investigators to use a polar solvent such as Me_2SO which gives rise to strong competition for hydrogen bond formation. For example, the equilibrium constant for G-C association is larger than 10^4 M^{-1} in $CHCl_3$,[11] 3.7 M^{-1} in Me_2SO,[12] and $K = 33$ M^{-1} in Me_2SO-d_6-benzene-d_6 (1:1). It is not possible then to compare equilibrium constants for e^9A, $(chx)^1U$, or $(chx)^1C$ in $CHCl_3$ with that of guanine determined in Me_2SO. On the other hand, comparative studies in Me_2SO could be strongly perturbed by important errors generally encountered in the determination of small association constants. Nevertheless, we can estimate an order of magnitude of the association constant for interaction of butyric acid with guanosine in position N_3. The H bonds O-H···N_3 and O···H-N_4 are equivalent in nature with H bonds made with adenine or cytosine (K in the range 40-300). Moreover, using the geometry given by x-ray crystallographic data we can predict that one of these hydrogen bonds will be slightly bent (about 10°). With this information it is not unreasonable to predict that interaction with butyric acid in position N_3 gives a smaller constant than interaction in position N_1.

It should be noted that the equilibrium constants for the association of butyric acid with nucleic bases decrease in the order $m_2^2G > (chx)^1C > e^9A$ (in position 1) $> (chx)^1U > e^9A$ (in position 7). There is an eightfold difference between m_2^2G and U although in both cases the same H bonds are formed (O-H···O and N-H···O) and have approximately the same geometry (Figures 3 and 6). The association constant for C is two- to threefold larger than that for A in position 1 although the same H bonds with the same geometry are formed (N-H···O, O-H···N). The lowest of the association constants is obtained for A in position 7, sevenfold smaller than for $(chx)^1C$, but the difference can be partly explained by the presence of bent H bonds (see above). Though net atomic charges of atoms participating in H bonds are slightly different (Table II) in these compounds, this cannot explain the difference in free energy of these complexes. For a small dimer such as H_2O···H_2O Van Duijneveldt-Van de Rijdt and Van Duijneveldt[13] have computed stabilization energies of −7.2 kcal mol^{-1} (electrostatic), −1.0 kcal mol^{-1} (dispersion), and −1.1 kcal mol^{-1} (delocalization), and a destabilization energy

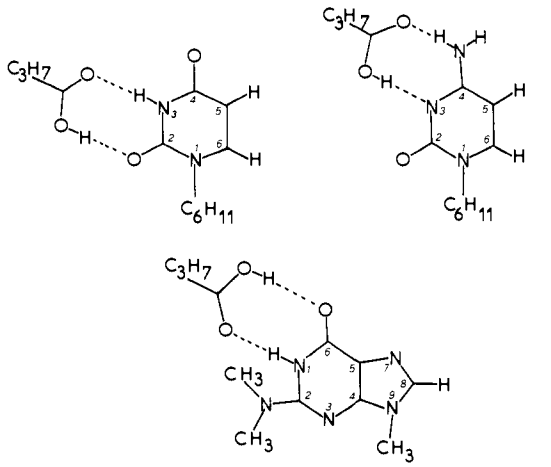

Figure 7. Geometry of 1:1 complexes with two hydrogen bonds between $(chx)^1U$, $(chx)^1C$, m_2^2G, and butyric acid. Hydrogen bonding of butyric acid with guanine leads to a new complex involving NH_2 and N_3. The geometry of butyric acid was determined by published data reported on formic acid.[19] Geometries of $(chx)^1U$, $(chx)^1C$, and m_2^2G were determined by using neutron diffusion crystallographic data on 1-methylthymine[20] or x-ray crystallographic data on cytosine and on guanosine dihydrate.[21,22]

(exchange) of +5 kcal mol^{-1}. This gives a net stabilization energy of −4.3 kcal mol^{-1}, a value comparable with experimental data (−5.2 kcal mol^{-1}).[14] Due to the complexity of the problem even for small molecules, it is not reasonable to argue that a difference in total atomic charges will explain energy differences in our systems. On the other hand, Del Bene and Kochenour[15] have recently computed the stabilization energy for dimer formation by formic acid assuming either identical geometries for formic acid monomer in the free and complexed states (constrained dimer) or a dimer geometry optimized to obtain the best stabilization energy (fully optimized dimer) by STO-3G computations.[16] For a variation of about 3% of angles and bond lengths (except for O···H bonds, 1.65 and 1.53 Å), they have computed a difference of stabilization energy of 2.7 kcal mol^{-1}. For complexes of butyric acid with $(chx)^1U$ or m_2^2G the lengths of H bonds on simple geometric models are in disagreement with experimentally reported average values. Assuming the usual length of the N-H···O bond (1.9 Å), the O-H···O bond length is 2.5 Å for guanine and 2.25 Å for uracil, values greater than generally reported: 1.7 Å.[6] The variation of energy due to a fully optimized geometry can be of importance in explaining differences in H-bond energy formation. Redistribution of charges on different atoms for the four bases in interaction with butyric acid could also be of great importance.

The results reported above concern carboxylic acids in their un-ionized state. However, it should be pointed out that many amino acid side chains have abnormal pK values in proteins.[17] Therefore, un-ionized carboxylic acids might be involved in hydrogen bonding interaction under biological conditions. Formation of double hydrogen bonds allows an amino acid side chain (carboxylic acid) to discriminate between different nucleic acid bases. There is a 15-fold difference in association constants between guanine (N_1H and O_6) and adenine (N_7 and N_6H). Although results obtained with the carboxylic group (Glu, Asp) cannot be simply extended to amide groups (Gln, Asn), both side chains can give rise to the same types of H bonds. Moreover, it should be noted that in double-stranded nucleic acids many of the H-bonding possibilities described here (Figures 3 and 7) are eliminated. Only two sites remain: adenine (N_7 and N_6H) in the large groove and guanine (N_3 and N_2H) in the small groove. Double hydrogen bond formation might be one of the interactions involved in selective recognition of nucleic acid bases by proteins.

Acknowledgments. We wish to thank Dr. C. Helene for helpful discussions.

References and Notes

(1) M. Sellini, J. C. Maurizot, J. L. Dimicoli, and C. Helene, *FEBS Lett.*, **30**, 219 (1973).
(2) (a) G. Lancelot, *Biophys. J.*, **17**, 243–254 (1976); (b) G. Lancelot, submitted for publication.
(3) J. J. Fox, D. V. Van Pragg, I. Wempen, I. L. Doerr, L. Cheong, J. E. Knoll, M. L. Eidinoff, A. Bendich, and G. B. Brown, *J. Am. Chem. Soc.*, **81**, 178 (1959).
(4) S. I. Chan, M. P. Schweizer, P. O. P. Ts'o, and G. K. Helmkamp, *J. Am. Chem. Soc.*, **86**, 4182 (1964).
(5) P. O. P. Ts'o, N. S. Kondo, M. P. Schweizer, and D. P. Hollin, *Biochemistry*, **8**, 997 (1969).
(6) M. D. Joesten and L. J. Schaad, "Hydrogen Bonding", Marcel Dekker, New York, N.Y., 1974.
(7) J. D. Engel and P. H. Von Hippel, *Biochemistry*, **13**, 4143 (1974).
(8) H. A. Benesi and J. H. Hildebrand, *J. Am. Chem. Soc.*, **91**, 2703 (1969).
(9) R. L. Scott, *Recl. Trav. Chim. Pays-Bas*, **75**, 787 (1956).
(10) W. B. Person, *J. Am. Chem. Soc.*, **87**, 167 (1965).
(11) J. Pitha and R. N. Jones, *Can. J. Chem.*, **44**, 1045 (1966).
(12) R. A. Newmark and C. R. Cantor, *J. Am. Chem. Soc.*, **90**, 5010 (1968).
(13) J. G. C. M. Van Duijneveldt-Van de Rijdt and F. B. Van Duijneveldt, *Theor. Chim. Acta*, **19**, 83 (1972).
(14) H. A. Gielbic, W. J. Burroughs, J. Chamberlain, J. E. Harries, and R. G. Jones, *Nature (London)*, **221**, 143 (1969).
(15) J. E. Del Bene and W. L. Kochenour, *J. Am. Chem, Soc.*, **98**, 2041 (1976).
(16) J. A. Pople, in "Aspect de la Chimie Quantique Contemporaine", edited by CNRS, Paris, 1971, p 17.
(17) D. M. Blow and T. A. Steiz, *Annu. Rev. Biochem.*, **39**, 63 (1970).
(18) R. F. Stewart and L. H. Jensen, *J. Chem. Phys.*, **40**, 2071 (1964).
(19) R. G. Lener, J. P. Friend, and B. P. Dailey, *J. Chem. Phys.*, **23**, 201 (1955).
(20) A. Kvich, T. F. Koetzle, and R. Thomas, *J. Chem. Phys.*, **61**, 2711 (1974).
(21) V. Thewalt, C. E. Bugg, and R. E. Marsh, *Acta Crystallogr., Sect. B*, **26**, 1089 (1970).
(22) D. L. Barker and R. E. Marsh, *Acta Crystallogr.*, **17**, 1581 (1975).
(23) E. Clementi, J. M. Andre, M-Cl. Andre, D. Klint, and D. Hahn, *Acta Phys. Acad. Sci. Hung.*, **27**, 493 (1969).

Theoretical Studies of Environmental Effects on Protein Conformation. 1. Flexibility of the Peptide Bond

Steve Scheiner[1] and C. W. Kern*

Contribution from the Department of Chemistry, Ohio State University, Columbus, Ohio 43210. *Received April 20, 1977*

Abstract: The effects of hydrogen bonding on peptide bonds are studied by quantum mechanical methods. The peptide unit is modeled by *trans-N*-methylacetamide (NMA) which is allowed to interact with various hydrogen bonding species that are similar to those typically found in the environment of a peptide within a protein molecule. These species include waters of hydration, other peptide units, and the side chains of amino acid residues. The effects of these species on the flexibility and electronic charge distribution of NMA can be interpreted in terms of resonance structures and atomic orbital overlap. All species are found to have a restricting influence on the conformation of a peptide bond. Nonplanar deformations of the peptide unit require more energy in the presence of these species. The effects of partial as well as full hydration of the peptide are considered. It is found that many of the effects of a full hydration shell can be simulated by the interaction with two water molecules. A correlation is found between the increased rigidity of the peptide and several parameters of charge redistribution. Interpeptide hydrogen bonding is found to have much the same effects as hydration. Interaction of a peptide unit with the electrically charged side chains of several residues is also predicted to result in a significantly more rigid peptide.

Hydrogen bonding is one of the most important determinants of three-dimensional protein structure. The peptide units of the various residues normally engage in extensive hydrogen bonding with one another. They also interact with waters of hydration which are found both within the protein and along its periphery. These effects on the peptide linkage must be understood before the principles underlying protein conformation can be ensured.

Quantum mechanical methods are useful for studying these interactions on a molecular level. The interaction between amide units, which serve as models for the peptide units within proteins, has been studied by both ab initio[2-7] and semiempirical[8,9] theory. The hydration of an amide has also been examined by ab initio techniques.[5,6,10-13] Although such studies provide useful information about the nature of the interactions, they have dealt mostly with amide units in their fully planar conformations. It has become increasingly more evident from x-ray crystal-structure determinations of proteins,[14] as well as open and cyclic polypeptides and other amides,[15-20] that the peptide units within proteins deviate significantly from planarity. Distortions of "isolated" peptide units from planar form have been examined by both semiempirical[9,15,16,21-24] and ab initio[21,25-28] techniques which predict[15,16,21] that the preferred conformation is the fully planar one but that relatively little energy is required for significant deviations from planarity to occur.

A major limitation of previous studies is that the amide unit was modeled in vacuo, much unlike the true environment of a protein. In the present paper, we examine the effects that various hydrogen bonding species (HBS) produce on a model peptide unit, particularly the role they play in altering its flexibility, and the possible implications on the secondary protein structure. The molecule chosen to model the peptide unit is *trans-N*-methylacetamide (NMA). The trans arrangement is chosen because the overwhelming majority of peptide units in proteins is found in this configuration. Of more than 20 globular proteins whose crystal structures have been determined, the cis configuration is observed only rarely and is usually associated with the presence of a proline residue.[29]

Planar Systems

The primary computational method used in our study is partial retention of diatomic differential overlap (PRDDO) which has been proven[30] to simulate accurately and rapidly ab initio minimum Slater basis-set results. Optimization of the geometry of NMA (Figure 1) with PRDDO yielded a planar molecule with the bond lengths and bond angles shown in Table