HYDROGEN BONDING OF AMINO ACID SIDE CHAINS TO NUCLEIC ACID BASES

H. SELLINI, J. C. MAURIZOT, J. L. DIMICOLI and C. HELENE
Centre de Biophysique Moléculaire,
45 045 Orleans Cedex, France

Received 19 December 1972

1. Introduction

The origin of the specificity of interactions between nucleic acids and proteins is a fundamental problem of molecular biology. Many studies have been made on the electrostatic interactions between basic polypeptides and nucleic acids [1, 2]. More recently, results on the formation of stacked complexes between aromatic amino acids and polynuclectide bases have been reported [3–6]. Another type of force might be of importance in nucleic acid–protein interactions, namely hydrogen bonding. Study of this type of force, on model compounds, is very difficult in aqueous solutions due to competition of hydrogen bonding to the solvent. However, nucleic acid derivatives have been recently shown to form hydrogen bonds in concentrated aqueous solutions [7]. Preliminary investigation on the existence of such interactions can be made in organic solvents. This method has been successfully applied to the study of specific pairing of nucleic acid bases either by infrared (IR) spectroscopy [8–11] or by proton magnetic resonance (PMR) [12–15].

We wish to report preliminary results on the interactions between derivatives of adenine or uracil, and compounds bearing chemical functions identical to several amino acid side chains. The experiments are performed in chloroform solutions and hydrogen-bond association is investigated by IR and PMR spectroscopies.

2. Experimental

9-Ethyladenine (A) and 1-cyclohexyluracil (U) were obtained from Cyclochemical Co., Los Angeles, and used without further purification. Chloroform was distilled after being dried over $CaCl_2$, and then passed through an alumina gel column 20 cm long. *p*-Cresol, indole, and imidazole were purified by sublimation. Butyric acid was distilled over $CaCl_2$. Methylguanidine was prepared from its hydrochloride by addition of NaOH in ethanol and dried under vacuum with P_2O_5.

Infra-red spectra were recorded with a Beckman IR 12 double beam spectrophotometer, using AgCl cells. Proton magnetic resonance spectra were obtained with a Brüker HFX 90 MHz spectrometer. TMS was used as an internal reference.

3. Results

3.1. *9-Ethyladenine and p-cresol*

In the wavenumber range 3200–3650 cm^{-1} the IR spectrum of A (0.02 M) in chloroform, is composed of essentially two bands at 3526 and 3417 cm^{-1}, corresponding to the antisymmetric and symmetric stretching vibrations of the free amino groups [8]. Shoulders at 3490 and 3350 cm^{-1} are assigned to the vibration of the bound amino groups arising from self association of A. The band corresponding to the free hydroxyl vibration of *p*-cresol is located at 3604 cm^{-1} whereas the bonded hydroxyl vibration due to self association appears as a broad band around 3400 cm^{-1}. A comparison of the spectrum of the equimolar mixture

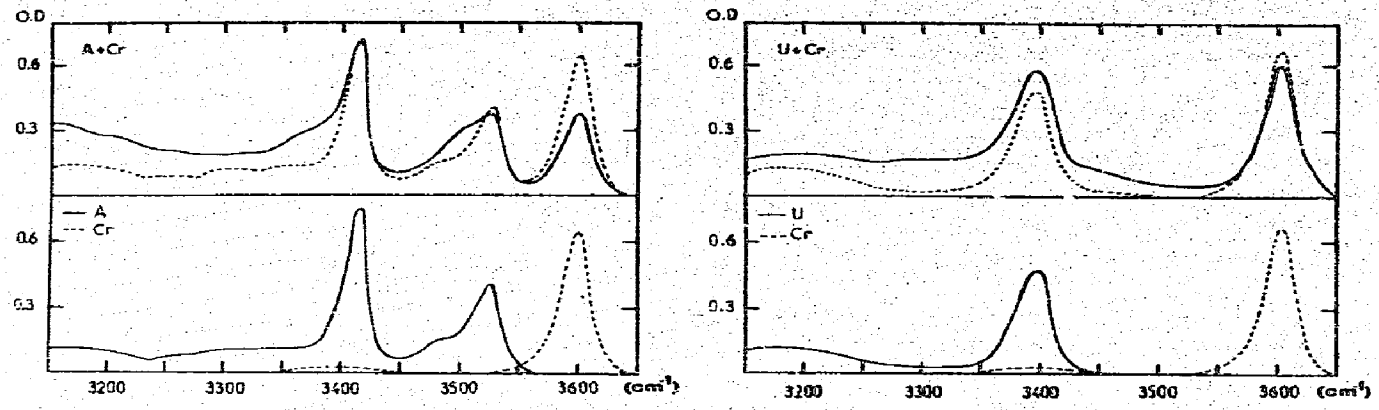

Fig. 1. Infra-red spectra of 9-ethyladenine (A), 1-cyclohexyluracil (U), *p*-cresol (Cr) and equimolar mixtures of A and U with Cr at 277° K in $CHCl_3$. Concentration: 2×10^{-2} M. In the upper spectra the full line is the observed spectrum of the mixture and the broken line is the sum of the spectra of separated components.

of A and *p*-cresol with that of the non interacting compounds leads to the following observations (fig. 1): the intensity of the *p*-cresol free hydroxyl group band is strongly decreased; the intensity at 3490 cm^{-1} is markedly increased and a new band appears at 3390 cm^{-1} on the low energy side of the free NH_2 symmetric stretching band; the absorption in the region 3190–3390 cm^{-1} increases due to bound species. The intensity of the bands at 1631 and 1589 cm^{-1}, characteristic of the NH_2 scissor vibrations and ring stretching modes of A [8], decreases while there is an increase of intensity on the high frequency side of the band at 1631 cm^{-1}, characteristic of NH_2 bonded groups.

From all these results, it can be concluded that the hydroxyl group of *p*-cresol, as well as the amino group of A, are involved in complex formation between A and *p*-cresol. No interaction is detected between A and *p*-cresol-methyl ester.

Proton magnetic resonance (PMR) can be used to study hydrogen-bond formation because the proton in hydrogen bonding is usually strongly deshielded due to the electric field of the acceptor molecule and its resonance is shifted downfield [16]. The magnetic anisotropy of the acceptor group (or ring) may also contribute (either downfield or upfield) to the observed shift. The association process is usually controlled by diffusion and unless the association constants are unusually large (for small organic molecules), the exchange between bound and free species is fast enough for only the weighted average chemical shift to be observed [17].

In the mixtures of 9-ethyladenine and *p*-cresol, the cresol OH resonance is strongly shifted downfield (fig. 2) and broadened as compared to free cresol. The resonance of the adenine amino protons is also shifted downfield although to a much lesser extent that the OH resonance. These observations are in quite good agreement with the IR results presented above and indicate that both cresol OH and adenine NH_2 groups are involved in complex formation. It must also be noticed that small shifts (either upfield or downfield)

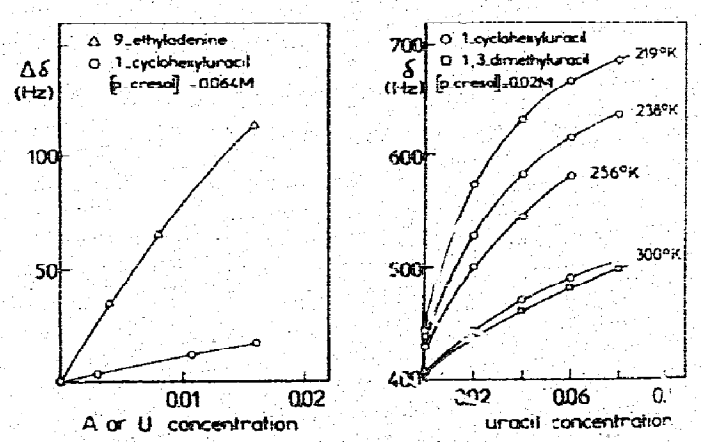

Fig. 2. *Left*: Change in chemical shift of *p*-cresol OH resonance in $CDCl_3$ at 300° K versus 1-cyclohexyluracil or 9-ethyladenine concentration. *p*-Cresol concentration is kept constant (0.064 M). *Right*: Chemical shift of *p*-cresol OH resonance in $CDCl_3$ at different temperatures in the presence of 1-cyclohexyluracil and 1, 3-dimethyluracil.

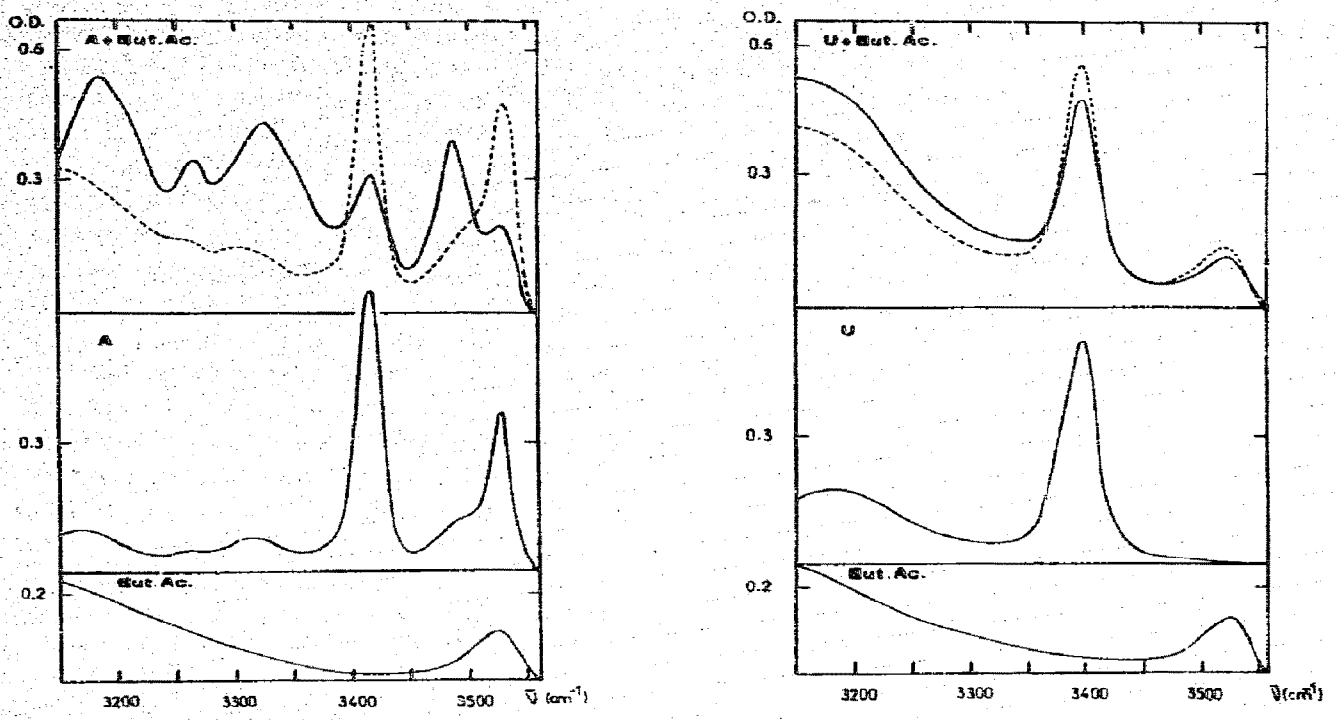

Fig. 3. *Left*: Infra-red spectra of butyric acid (4×10^{-2} M), 9-ethyladenine (4×10^{-2} M) and their equimolar mixture (4×10^{-2} M each) at 308°K in $CHCl_3$. *Right*: Infra-red spectra of butyric acid (5.9×10^{-2} M), 1-cyclohexyluracil (5.9×10^{-2} M) and their equimolar mixture (5.9×10^{-2} M each). In the upper parts, the full line is the observed spectrum of the mixture and the broken line is the sum of the individual spectra.

of the aromatic ring protons of both cresol and adenine are observed. They can be attributed to the redistribution of electronic charges in the hydrogen bonded complex. The effect of cresol upon the C_8H and C_2H resonances of adenine is very similar to that observed when uracil binds to adenine [14] (downfield shifts of C_8H and C_2H, the shift of the C_8H resonance being about twice that of C_2H). Also the shifts of the cresol aromatic protons are consistent with the charge redistribution due to the electrostatic effect upon electrons in the hydrogen bond (ortho protons are shifted downfield and meta protons upfield).

There are three nitrogen atoms on the adenine ring whose basicities are in the order $N_1 > N_7 > N_3$. The involvement of the adenine NH_2 group complex formation suggests that cyclic dimers are formed. At low concentration of adenine as compared to cresol, the amount of free cresol measured by IR is less than that which would be expected if all adenene molecules had been involved in a 1:1 complex. Thus the stoichiometry is probably 2:1 at high (cresol)/(adenine) ratios. This is supported by an IR study using the method of continuous variation: maximum complex formation occurs for a ratio close to 2 cresol per adenine. However the fact that adenine self-associates does limit the validity of the measurements at high adenine concentration.

3.2. *9-Ethyladenine and butyric acid*

Fig. 3 shows the IR spectra of butyric acid (B), 9-ethyladenine (A) and of their equimolar mixture. Drastic changes are observed upon mixing the two compounds. There is a decrease in intensity of the bands due to free NH_2 and OH groups (3530–3417 and also 1630–1590 cm^{-1}) and strong bands characteristic of bound species appear at 3490, 3325, 3265, 3180 and 1600 cm^{-1}. The method of continuous variations indicates the formation of a 1:1 complex. In order to calculate the association constant between the two compounds (K_{AB}) one needs to take into account the self association of butyric acid ($K_{BB} \simeq 58\ M^{-1}$) but one can neglect the self association

of A which is small (K_{AA} = 2.3 M^{-1} at 306°K [8]). At 308°K we find K_{AB} = 130 ± 30 M^{-1}. The enthalpy of reaction (ΔH) and the entropy change (ΔS) were obtained from a Van't Hoff plot using the values of K_{AB} determined at different temperatures: $\Delta H = -7.4 \pm 0.5$ kcal·mole^{-1}; $\Delta S = -13.3 \pm 2$ e u. The value of ΔH seems to be too large to be attributed to a single hydrogen bond and does indicate the existence of a cyclic dimer between A and butyric acid. It must be pointed out that the values we have found for ΔH and ΔS are very close to the values found by Kyogoku et al. [8] for the formation of cyclic dimers between A and U ($\Delta H_{AU} = -6.2$ kcal·mole^{-1} and $\Delta S_{AU} = -1.1$ e u).

3.3. Cyclohexyl uracil and p-cresol

The IR spectrum of U is composed of one band at 3400 cm^{-1} characteristic of the free NH group [8]. At 1710 cm^{-1} there is a band due to the C(2) carbonyl group vibration and at 1687 cm^{-1} another band due to the coupling between C(4) carbonyl stretching and NH bonding modes. Upon mixing U and *p*-cresol there is a decrease of the intensity of the free hydroxyl band of *p*-cresol, and an increase in intensity around 3400 cm^{-1} (fig. 1) due to bound hydroxyl groups.

The binding of 1-cyclohexyluracil to cresol leads to a downfield shift and a broadening of the cresol OH resonance. Using 1,3 dimethyluracil instead of 1-cyclohexyluracil, quite similar downfield shifts of the OH resonance are observed, indicating that hydrogen bonding probably involves only the carbonyl group(s) of uracil. As already observed in the presence of adenine, the resonances of protons in the ortho position of the OH group are slightly shifted downfield in the presence of uracil while the meta protons are shifted upfield.

The broadening of the OH resonance of *p*-cresol upon binding to adenine or uracil derivatives could possibly be ascribed to hydrogen exchange between OH (cresol) and NH_2 (adenine) or NH (uracil groups). This chemical exchange can be easily demonstrated by IR studies if one of these groups is deuterated prior to mixing with the other non-deuterated compound. (OD exchanges with NH_2 and NH.) However, the rate of exchange is too high to be measured by the IR technique. The absence of broadening of the cresol OH group in the presence of 1,3-dimethyluracil as well as the narrowing observed when the temperature of the mixtures with A or U is decreased are consistent with chemical exchange as a source of broadening.

The chemical shifts of the OH resonance at constant cresol concentration (0.02 M) have been measured at different temperatures in the presence of various uracil concentrations (fig. 2). Self-association of cresol at 0.02 M is weak, even at 219 K (see fig. 2). At low [cresol]/[uracil] ratios, it is likely that only 1:1 complexes are formed (uracil self-associates to form cyclic dimers; but each uracil molecule still has a free carbonyl group available for hydrogen bonding). The determination of the association constant requires the knowledge of the chemical shift (δ_c) in the hydrogen-bonded cresol molecule. An estimation of δ_c can be made on the basis of the experimental results shown in fig. 2. Using this estimated δ_c value, the association constant K can be calculated for every uracil concentration. The value of δ_c is then changed until the values of K are independent of uracil concentration. At 219 K, the best fit is obtained when using $\Delta\delta = \delta_c - \delta_o = 3.3$ ppm (where δ_o is the chemical shift of free cresol) and the K value is 69 ± 4 M^{-1}. The IR frequency shift of the OH vibration ($\Delta\nu_{OH}$) and the change in chemical shift of the OH resonance ($\Delta\delta$) which have been observed upon binding of phenol to various hydrogen-bond acceptors have been linearly correlated to the enthalpy of complex formation [16]. Using the relationships reported by Eyman and Drago [16] and the change in chemical shift ($\Delta\delta$ = 3.3 ppm) corrected for the magnetic anisotropy of the carbonyl group (≃ 1.1 ppm), one can calculate $\Delta\nu_{OH}$ = 150 cm^{-1} and $\Delta H = -3.0$ kcal·mole^{-1}. The association constant has been calculated at different temperatures assuming that δ_c is independent of temperature. A plot of log K versus $1/T$ is linear and the calculated thermodynamic parameters are $\Delta H = -3.4$ kcal·mole^{-1}

$\Delta S = -7.5$ cal·mole^{-1}·K^{-1}

The ΔH value obtained from the temperature dependence of the association constant is thus quite similar to that determined above from the change in chemical shift.

At 300°K, the association constant for the binding of cresol to dimethyluracil is determined to be K = 6.3 ± 0.2 M^{-1} with the same value of δ_c (3.3 ppm). Thus the associations of 1,3 dimethyluracil and 1-cyclohexyluracil with cresol are quite similar (6.3 and 7.4 M^{-1}, respectively, at 300°K).

3.4. *Cyclohexyluracil and butyric acid*

The IR spectra of U, butyric acid and their equimolar mixture are shown in fig. 3. These spectra indicate that new bound species appear (vibration between 3150 and 3350 cm^{-1}) and that the NH group of U is involved in the association. However modifications of the spectra are less important than in the case of A. The association constant is calculated to be K = $60 \pm 20\ M^{-1}$ at 308°K.

3.5. *Other compounds*

We could not detect any change in the IR spectra when A or U were mixed with ethanol, indole, imidazole, acetamide and *N*-methylacetamide (equimolar 0.02 M mixtures). In the cases of *N*-methyl guanidine and 5-hydroxyindole interactions have been observed. *N*-Methyl guanidine is not soluble enough in chloroform to record its IR spectrum, however its solubility is strongly increased in the presence of U but not in the presence of A. This indicates that complexes are formed between methyl guanidine and U. Mixing U and 5-hydroxyindole does not induce any change in the IR spectrum, but mixing A and 5-hydroxy indole gives a precipitate with 1:1 stoichiometry indicating a strong interaction between these two compounds.

4. Conclusion

The results which are reported above demonstrate that hydrogen bonded complexes between bases of nucleic acids and compounds bearing the same chemical functions as side chains of amino acids are formed in chloroform. Although the overall environment of the complex in a non polar solvent differs from that in biologically active systems, the more immediate environment of a hydrogen bond may be similar. Using chloroform does not allow the study of such compounds as ionized compounds, e.g. anions of carboxylic acids. Extension of these IR studies to other solvents than chloroform should help overcome this difficulty. However, it must be pointed out that many residues have abnormal pK values in proteins, so that the study of complex formation between nucleic acid bases and unionized carboxylic acids or amines could be relevant to the problem of protein–nucleic acid interactions in biological conditions.

It is surprising that ethanol does not form hydrogen bonds with A and U under our experimental conditions (0.02 M, in chloroform). Interactions involving OH groups have been demonstrated in crystal complexes of ribonuclease and cytidine 3'-monophosphate between threonine 45, serine 123 and the pyrimidine ring [18]. As A and U (as well as ethanol) self associate in chloroform, this might only indicate that complex formation with ethanol is weaker than self association. Aromatic hydroxyl groups are able to form hydrogen bonds with nucleic acid bases (see results obtained with *p*-cresol). This is probably due to the more acidic character of cresol as compared to non aromatic alcohols, but it may also reflect the energy due to direct interactions between the aromatic rings themselves in the hydrogen bonded complexes.

We do not observe any interaction between acetamide or *N*-methyl acetamide and A or U despite the fact that it is possible to build a model with double hydrogen bond formation in a way similar to base pairing of nucleic acids. Kyogoku et al. [11] have shown that compounds bearing the –CO–NH– group like barbiturates could form hydrogen bonds with adenine, but valerolactam fails to form such a complex. This was attributed to the very poor proton donating power of this compound associated to a high pK. A similar explanation could be put forward in the case of acetamide and *N*-methylacetamide.

Our results suggest that among the different amino acid residues, tyrosine and carboxylic acids (aspartic and glutamic acids) could participate in the formation of protein–nucleic acid complexes by hydrogen bonding to the bases of DNA. Experiments with guanine and cytidine derivatives are required before any conclusion can be drawn about the specificity of hydrogen bond interactions between nucleic acid bases and amino acid residues. Work along this line is currently in progress in our laboratory.

References

[1] J.T. Shapiro, M. Leng and G. Felsenfeld, Biochemistry 8 (1969) 3219.
[2] O.E. Olins, A. Olins and P.H. Von Hippel, J. Mol. Biol. 24 (1967) 157.
[3] C. Helene, Nature New Biology 234 (1971) 120.
[4] C. Helene, T. Montenay-Garestier and J. L. Dimicoli, Biochim. Biophys. Acta 254 (1971) 349.
[5] C. Helene, J.L. Dimicoli and F. Brun, Biochemistry 10 (1971) 3802.

[6] C.Helene, J.L. Dimicoli, FEBS Letters 26 (1972) 6.
[7] M. Rasyka and N.O. Kaplan, Proc. Natl. Acad. Sci. U.S. 69 (1972) 2025.
[8] Y. Kyogoku, R.C. Lord and A. Rich, J. Am. Chem. Soc. 89 (1967) 496.
[9] J. Pitha, R N. Jones and P. Pithova, Can. J. Chem. 44 (1966) 1045.
[10] Y. Kyogoku, R.C. Lord and A. Rich, Proc. Natl. Acad. Sci. U.S. 57 (1967) 250.
[11] Y. Kyogoku, R.C. Lord and A. Rich, Nature 218 (1968) 69.
[12] R.R. Shoup, H.T. Miles and E.D. Becker, Biochem. Biophys. Res. Commun. 23 (1966) 194.
[13] R.A. Newmark and C.R. Cantor. J. Am. Chem. Soc. 90 (1968) 5010
[14] L. Katz, J. Mol. Biol. 44 (1969) 279.
[15] S.I. Chan, G.C.Y. Lee, C.F. Schmidt and G.P. Kreishman, Biochem. Biophys. Res. Commun. 46 (1972) 1536.
[16] D P. Eyman and R.S. Drago, J. Am. Chem. Soc. 88 (1966) 1617, and references cited therein.
[17] L.C.M. De Maeyer, Israel J. Chem. 9 (1971) 351.
[18] F.M. Richards, H.M. Wyckoff and N.M. Allewell, in Neurosciences, ed. F.O. Schmitt, 1969.