Model studies of interactions between nucleic acids and proteins: hydrogen bonding of amides with nucleic acid bases

Gérard Lancelot and Claude Hélène
Centre de Biophysique Moléculaire, 45045 Orléans Cédex, France

Received 13 December 1978

Summary
The formation of hydrogen bonded complexes between nucleic acid bases and acetamide has been studied by nuclear magnetic resonance in CDCI₃ at different temperatures. Pairs of hydrogen bonds are formed when acetamide binds to nucleic acid bases. Thermodynamic parameters have been computed and compared to those obtained for the association of carboxylic acids with nucleic acid bases. The role of hydrogen bonded complexes in the association of proteins with nucleic acids is discussed.

Among the fundamental mechanisms which may 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 cyclohexane, chloroform and dimethylsulfoxide in the presence of water. Amino acids and nucleic acid bases possess different polar groups which should be able to form hydrogen bonds. Every polar amino acid side chain is able to form hydrogen bonds with all four bases. This peculiarity has led Seeman et al. and Hélène to conclude that a single hydrogen bond is inadequate for uniquely identifying any particular base or base pair. However, model building studies suggest that fidelity of base pair recognition may be achieved if a pair of hydrogen bonds is formed. Binding of carboxylic acids or acetamide groups to nucleic acid bases can lead to such complexes involving two hydrogen bonds. Association constants between carboxylic acids and bases have been found to be in the range 80 - 660 M⁻¹ in chloroform at 303 K. Amides can form the same type of 1:1 complexes with two hydrogen bonds. We report here a study of the asso-
ciation of amides with purine and pyrimidine bases in chloroform. This should provide a basis for the understanding of interactions between nucleic acid bases and glutamine or asparagine.

Experimental

9-ethyladenine (e9A), 1-cyclohexyluracil ((chx)1U) and 2-dimethylamino, 9-methylguanine (m9m2G) were obtained from Cyclo Chemical (New York). Acetamide 15N (95%) was purchased from CEA (France). Great care was taken to avoid the presence of any water or polar impurities in deuterated chloroform (CEA, France). NMR spectra were recorded 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 to +0.5°C. Chemical shifts are given in parts per million (ppm) downfield from tetramethylsilane (TMS = 0.0).

Results

In order to avoid the broadening of the NH2 resonances of acetamide due to the 14N quadrupole, and the overlap with the resonances of nucleic acid bases, 15N acetamide was used throughout this study. At 273 K the NH2 resonance of 15N acetamide appears as two groups of several resonance lines (figure 1). Each of these two groups is composed of a doublet (J = 3.5 Hz) and two quadruplets separated by 3.5 Hz. The splitting of 3.5 Hz represents the coupling between the two protons of the NH2 group while the quadruplet structure is due to the coupling of one of the two NH protons with the methyl group in the trans position. Coupling of the methyl group with the NH proton in the cis position is too small to be observed here. The splitting of 89 Hz between the two groups of resonance lines represents the 15N-1H spin-spin coupling constant. The NMR spectrum of 15N-acetamide depends on the concentration, indicating self association of this molecule in CDCl3. The structure of the hydrogen bonded dimer shown in figure 2 indicates that self association should lead to a downfield shift of the NH proton in the trans position with respect to the methyl group. As a matter of fact, increasing the concentration of
Figure 1 Proton magnetic resonance of acetamide $^{15}N$ (0.025 M) (lower spectrum, full line), $e^9A$ (0.025 M) (lower spectrum, dotted line) and the mixture acetamide $^{15}N + e^9A$ (0.25 M each) (upper spectrum) in CDCl$_3$ at 273 K. The arrows show the variation of the chemical shifts of NH protons.

$^{15}N$-acetamide induces a larger separation between quadruplets and doublet lines. The association constant for dimerization of acetamide was calculated to be 5.5 M$^{-1}$ at 273 K.

**Determination of Association Constants**

Association of nucleic acid bases and acetamide may lead to the formation of several 1:1 complexes involving one or two hydrogen bonds (figure 2). For each complex $i$, the equilibrium equation (1) leads to relationship (2)

$$A + B \rightleftharpoons (AB)_i \quad (1)$$

$$[AB]_i = K_i [A] [B] \quad (2)$$

where $K_i$ represents the association constant of complex $i$. The change in chemical shift of proton $j$ will be given by equation (3)

$$\Delta \delta = \delta^j - \delta^j_o = \Sigma_i \left( \delta^j_i - \delta^j_o \right) \frac{K_i [AB]}{K [A]_o} \quad (3)$$
Figure 2. Structure of complexes formed between acetamide and nucleic acid bases and dimer of acetamide. The structures of adenine and uracil were taken from data reported for 9-methyladenine\(^{13}\) and 1-methylthymine\(^{14}\).

\[
[AB] = \left( \frac{K_i}{K} \right) [AB]_i
\]

\(\delta_o^j\) and \(\delta_i^j\) represent the chemical shifts of proton \(j\) in the free molecule A and in complex i, respectively. A fit of the variation of \(\Delta \delta\) versus A or B concentration will give

\[
\frac{\delta^j}{\delta_o^j} = \frac{K_i}{K}
\]

The concentration dependence of the chemical shifts of the amino protons of nucleic acid bases and acetamide provides evidence for self association of the two compounds.

**Complex formation between nucleic acid bases and acetamide**

Interaction between acetamide and \(e^9\)A, \((\text{chx})^1\)U or \(m^2m^9\)G leads to a downfield shift of the NH resonances of each compound. This downfield shift is a general characteristic of protons participating in hydrogen bond formation. The high resolution NMR spectra of \(e^9\)A, acetamide and their equimolar mixture are shown in figure 1. Only one resonance is observed for the NH\(_2\) protons of \(e^9\)A whereas several resonances...
are observed for acetamide either in the free or in the complexed state. In CDC13 the two protons of the NH2 group of e9A could not be observed separately above 230 K (4).

As shown in figure 2, only one proton of the NH2 group of acetamide can be involved in the formation of a pair of hydrogen bonds with nucleic acid bases. This is proton Hb in the trans position with respect to the methyl group. Therefore this proton gives rise to the highest downfield shift. Consequently the chemical shift difference between the doublet and quadruplets of 15N-acetamide is increased as shown in figure 1.

Association constants for dimerization of acetamide, e9A and (chx)1U, on one hand, and for association of acetamide with e9A or (chx)1U, on the other hand, are given in tables 1 and 2. Table 1 gives the average downfield shifts computed for different protons in the complexes. The experimental procedure used to compute these parameters was based on the measurements of the chemical shifts of different protons for a series of solutions whose total concentration ([acetamide] + [base]) was kept constant (see figure 3).

Guanine and its derivatives are known to have a very low solubility in non-polar solvents. We have investigated the interactions between acetamide and 2-dimethylamino, 6-hydroxy, 9-methylpurine (m9m2G)

<p>| Table 1 |
| Association constants (average of the K values calculated for the different protons indicated in the corresponding row), chemical shifts and changes in chemical shifts (Δδc = δc - δ0) upon complex formation between e9A, (chx)1U and acetamide in CDCl3 at 273 K. |
| K (M⁻¹) | N(3)H [U] | NH2 (A) | NH (acetamide) |
| δ0 | Δδc | δ0 | Δδc | δ0 | Δδc | δ0 | Δδc |
| (chx)1U-(chx)1U | 10.5 | 7.915 | 3.702 | - | - | - | - | - | - |
| e9A - e9A | 5.5 | - | - | 5.513 | 2.173 | - | - | - | - |
| acetamide-acetamide | 2.0 | - | - | - | - | 5.921 | 1.008 | 5.780 | 2.846 |
| (chx)1U-acetamide | 9.0 | 7.915 | 4.762 | - | - | 5.921 | 0.989 | 5.780 | 2.456 |
| e9A-acetamide | 6.3 | - | - | 5.513 | 1.727 | 5.921 | 1.140 | 5.780 | 2.728 |</p>
<table>
<thead>
<tr>
<th></th>
<th>263 K</th>
<th>273 K</th>
<th>293 K</th>
<th>303 K</th>
<th>313 K</th>
<th>ΔH</th>
<th>ΔS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(chx)[U] - (chx) U</td>
<td>12.5</td>
<td>10.5</td>
<td>5.5</td>
<td>4.0</td>
<td>-</td>
<td>- 5.3</td>
<td>- 14.7</td>
</tr>
<tr>
<td>e^[A] - e^[A]</td>
<td>7.4</td>
<td>5.5</td>
<td>3.2</td>
<td>2.6</td>
<td>2.0</td>
<td>- 4.3</td>
<td>- 12.2</td>
</tr>
<tr>
<td>acetamide-acetamide</td>
<td>2.5</td>
<td>2.0</td>
<td>1.0</td>
<td>0.8</td>
<td>0.6</td>
<td>- 5.2</td>
<td>- 17.5</td>
</tr>
<tr>
<td>(chx)[U] + acetamide</td>
<td>12.0</td>
<td>9.0</td>
<td>5.0</td>
<td>4.0</td>
<td>-</td>
<td>- 4.6</td>
<td>- 12.6</td>
</tr>
<tr>
<td>e^[A] + acetamide</td>
<td>8.0</td>
<td>6.3</td>
<td>4.1</td>
<td>3.4</td>
<td>2.9</td>
<td>- 3.4</td>
<td>- 8.7</td>
</tr>
</tbody>
</table>

which is soluble enough in CDCl₃ for ¹H NMR studies. The ¹H NMR spectrum of m²G in presence of acetamide shows a broadening of the N_H resonance of m²G as well as a downfield shift proportional to acetamide concentration (2.6 ppm/M). This linear relationship shows that the reciprocal of the association constant is much greater than the concentration of m²G (4.7 x 10⁻⁵ M) and acetamide (5 x 10⁻⁵ - 10⁻¹ M). Under such conditions the concentration of complex AB and the variation of the chemical shift are given by equation (4) and (5), respectively

AB = \frac{K \cdot A \cdot B}{1 + K \cdot B} \quad (4)

Δδ = \frac{K \cdot B}{1 + K \cdot B} \cdot Δδ_c \quad (5)

Since A < B < K⁻¹, one gets Δδ ∝ K \cdot B \cdot Δδ_c.

Usually this type of complex exhibits a value of Δδ_c for N_H in the range 2 to 4 ppm (3). This leads to a value of K in the range 0.5 to 1 M⁻¹ at 293 K.

**Discussion**

Sugars (such as ribose or deoxyribose), phosphate groups as well as purine and pyrimidine bases are able to bind polar side chains of amino acids. To avoid such a competition in molecular complex formation, we have used nucleic acid bases substituted by non polar groups at the same positions as the glycosidic bond (N₂ for purines and N₁ for pyrimidines).
Figure 3  Concentration dependence of the chemical shifts of acetamide $^{15}$N and (chx)$^1$ U or $^{9}$A. Open symbols correspond to the chemical shifts of the investigated compound alone in solution. Filled symbols correspond to the chemical shifts in the presence of the associating compound, the sum of the concentrations being kept constant (0.1 M for acetamide $^{15}$N + (chx)$^1$ U, upper curves; 0.05 M for acetamide $^{15}$N + $^9$A, lower curves).

Acetamide was investigated here as a model of the amide side chains of glutamine or asparagine. $^{15}$N enriched acetamide was used because the $^{14}$N nucleus possesses a quadrupolar moment. This quadrupolar moment interacts with an electric field gradient and spreads the amino proton resonances (for example the half width is 90 Hz for NH$_2$ in $^{14}$N acetamide). This did not allow us to measure the chemical shifts of amino protons with precision. Moreover this broad resonance line overlaps the amino reso-
nance of adenine. In the case of uracil the N\textsubscript{1}H proton resonance line can be used to follow the association with \textsuperscript{14}N and \textsuperscript{15}N acetamide. No difference was detected.

The restricted rotation of the amino group observed in \textsuperscript{15}N acetamide (Figure 1) allowed us to discriminate between protons participating and non-participating in the formation of acetamide dimer (Figure 3). The largest variation of chemical shift was observed for the same NH proton resonance in the presence or in the absence of nucleic acid bases. It was therefore necessary to take into account the self association of acetamide in the determination of the association constant and free energy. The association constants for the binding of acetamide to nucleic acid bases decrease in the order (chx)\textsuperscript{1}U (4.0 M\textsuperscript{-1}) > e\textsuperscript{9}A (3.4 M\textsuperscript{-1}) > m\textsuperscript{9}m\textsubscript{2}\textsuperscript{G} (0.5 - 1 M\textsuperscript{-1}) (values in parentheses are those obtained at 303 K).

From the change in association constants with temperature the thermodynamic parameters AH and AS could be determined (table 2). They fall in the range expected for the type of hydrogen bonded complexes shown in figure 2.\textsuperscript{(11, 15)}

Carboxylic acids (unionized form) are able to form the same type of complexes with nucleic acid bases as amides. However previous results obtained by NMR in CDCl\textsubscript{3} have shown that the association constants are much higher in the case of carboxylic acids\textsuperscript{(4)}. Moreover the strength of interaction decreases in the reverse order m\textsuperscript{9}m\textsubscript{2}\textsuperscript{G} (660 M\textsuperscript{-1}) > (chx)\textsuperscript{1}C (270 M\textsuperscript{-1}) > e\textsuperscript{9}A (150 M\textsuperscript{-1}) > (chx)\textsuperscript{1}U (80 M\textsuperscript{-1}). In the case of adenine a study of the monomethyl and dimethylamino derivatives allowed us to show that two complexes involving two hydrogen bonds are formed whose respective association constants were \sim 110 M\textsuperscript{-1} \left[ \text{NH(6) and N(1)} \right] and \sim 40 M\textsuperscript{-1} \left[ \text{NH(6) and N(7)} \right]

These results show that even though the same kinds of structures can be written for the two families of amino acid side chains (carboxylic acids and amides), there is a strong difference in the binding constants: a factor of about 10\textsuperscript{3} for G, 50 for A and 20 for U. These results agree with previous results showing that amides form weaker hydrogen bonds than carboxylic acids\textsuperscript{(10)}.

Under the same experimental conditions in CDCl\textsubscript{3} the associa-
tion constants for base pair formation are of the order of \(10^2\) M\(^{-1}\) and \(10^4\) M\(^{-1}\) (G-C). Carboxylic acids form pairs of hydrogen bonds with nucleic acid bases which have about the same energy as that of the two hydrogen bonds between A and U. The energy of complex formation between amides and nucleic acid bases is one to two orders of magnitude smaller.

Carboxylic (unionized) and amide side chains of proteins may form pairs of hydrogen bonds with all four nucleic acid bases. However in double stranded nucleic acids such complexes are restricted to A in A-T base pairs and to G in G-C base pairs. The first interaction takes place in the large groove while the second one is characteristic of the small groove. The only specific interaction described until now concerns carboxylate anions which form hydrogen bonded complexes only with guanine.

Acknowledgements

This work has been supported in part by the D. G. R. S. T. (contract MRM 77-7-1741).

References

(1) G. Lancelot, Biochimie (1977) 59, 587-596
(3) G. Lancelot, Biophys. J. (1976) 17, 243-254
(7) C. Hélène, Studia Biophysica (1976) 57, 211-222
(8) C. Hélène, FEBS Letters (1977) 74, 10-13