Volume 91, number 2 FEBS LETTERS July 1978

HYDRATION SCHEME OF THE COMPLEMENTARY BASE-PAIRS OF DNA

Amiram GOLDBLUM, David PERAHIA and Alberte PULLMAN
Institute de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique, associé au CNRS,
13, rue P. et M. Curie, 75005 Paris, France

Received 18 May 1978

1. Introduction

As a continuation of our studies on the hydration of the purine and pyrimidine bases [1,2], we present here an investigation on the polyhydration scheme of the complementary base pairs of the nucleic acids: adenine–thymine (A–T) and guanine–cytosine (G–C). Results obtained for the bases indicate that one shell of 'bound' water may be expected to exist around the bases and around the base-pairs. Here we wish to explore the nature of this shell around the base pairs with the particular aim of determining the possible difference in their degree of hydration.

2. Method

The calculations have been carried out with the overlap multipole procedure (OMTP) developed in our Laboratory to compute the electrostatic interaction between molecules [2,3]. Its essential features is the use of a 'polycentric' multipole expansion of the electron density of each of the interacting molecules which serves to compute the electrostatic interaction between them (detailed explanation [2]). A minimization technique has been coupled to the computation in order to determine the optimum positions of the water molecules around the ligands. This method has been shown to reproduce correctly the hydration schemes of molecules obtained by full SCF ab initio supermolecule calculations [2].

The multipole expansions are obtained directly from the SCF ab initio wave functions of the individual molecules (the purine and pyrimidine bases and water). For adenine, thymine and guanine we used the wave functions in [4]. The wave function for cytosine was recomputed in the same basis set for the experimental geometry and the base pairs were constructed by superposing their two components at the appropriate distances [5] without doing an SCF computation of the supersystem. Hence we neglect in the present study the polarizing effect of one base on the electron distribution of the other.

In order to avoid false zones of apparent stability due to the neglect of the exchange repulsion, a distance of closest approach of 2.85 Å has been fixed between the oxygen of water (O(W)) and the first-row atoms, with the exception of the distance between O(W) and pyridine-type nitrogens which was fixed at 2.95 Å, following the indications in [2].

We used two versions of the program computing the electrostatic OMTP interactions:

1. The first version moves one water molecule within a grid defined around the ligand (the O(W) atom being placed at every point of the grid); for every point the best orientation of the water molecule is found by appropriate rotations of the three local axes centered on O(W). This version enables us to locate the favored positions of a single water molecule around the ligand, giving a picture of the monohydration scheme.

2. The second version utilizes a minimization technique which optimizes the interaction energy of a given water molecule with all its surroundings (ligand + the other water molecules), changing the geometrical position of the water molecule con-

sidered. This version enables us to find the polyhydration scheme of the ligand, taking into consideration all the possible interactions between the molecular species of the model.

These two versions have been used in succession, the first one giving the positions of the water molecules around the ligand to be adopted as starting points in the second version. We have used, following the case, different starting positions. The interaction energies of the water molecules have been optimized one by one in succession in the presence of all the other water molecules involved. This procedure was then repeated restarting with the first optimized water molecule and so on until the computed total interaction energy did not change significantly.

The minimization technique used is a version of the simplex method [6] available from the Quantum Chemistry Program Exchange [7].

3. Results

The polyhydration schemes obtained for the A–T and G–C base pairs are presented in fig.1 and fig.2, respectively, which indicate also the interaction energy of each of the bound water molecules with the base pair.

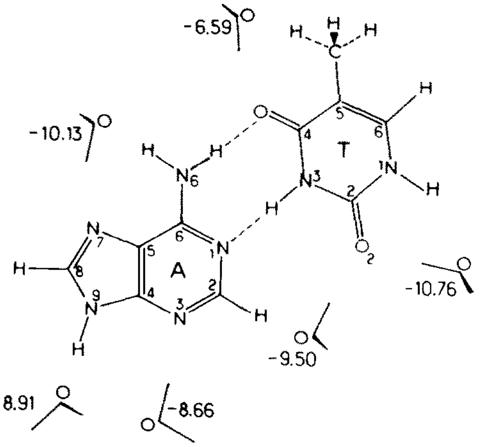

Fig.1. Hydration scheme of A–T base pair. The numbers correspond to the interaction energy (in kcal/mol) of each water molecule with the base pair.

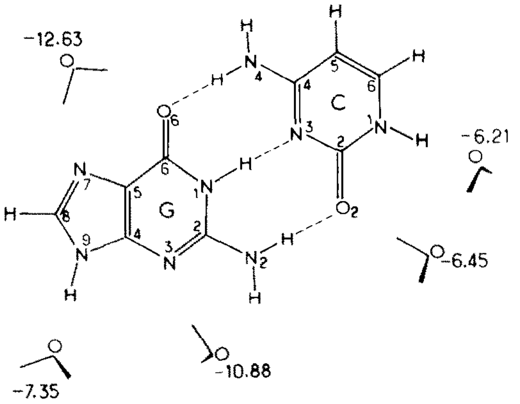

Fig.2. Hydration scheme of G–C base pair. Same comments as in fig.1.

Different arrangements are observed. A water molecule may prefer to bind to two adjacent sites simultaneously (e.g., to N_7 and NH_2 of adenine, O_2 and NH of thymine, N_7 and O_6 or N_3 and NH_2 of guanine) or two water molecules may prefer to bind to two adjacent sites in a dimeric form (e.g., between N_3 and N_9H of adenine or between N_1H and O_2 of cytosine). The most strongly bound water molecule is situated in a bridge position between N_7 and O_6 of guanine, with an interaction energy of –12.6 kcal/mol. A group of strong binding sites are located on guanine, adenine and thymine, with interaction energies comprised between –10.1 and –10.8 kcal/mol. The water molecule bound to O_2 of thymine has its energy of interaction (–9.5 kcal/mol) somewhat enhanced by an incipient interaction with the C_2H bond of adenine. On the other hand, cytosine presents only weak binding sites, of energy exceeding only slightly 6 kcal/mol. The overall hydration preferences of the bases within the base pairs are in the order (kcal/mol):

$$G(-30.86) > A(-27.70) > T(-26.85) > C(-12.66)$$

For the sake of comparison it may be indicated that the energies of water–water interactions, computed within the same systematics amount to –6.7 kcal/mol in the best water dimer and to a mean value of –5.1 kcal/mol in an arrangement in which a central water molecule is surrounded by 4 others. We may use one or the other of these two values as a lower limit for distinguishing the water molecules 'bound' to the base pairs. The use of the higher value leaves 5 water molecules bound to A–T and 3 bound to G–C. The use of the lower values attributes 6 bound water molecules to the A–T pair and 5 to the G–C pair. The overall result points thus to a greater hydration of the A–T pair as compared to the G–C pair, by 1–2 water molecules/base pair.

It may be remarked that we have considered the base pairs themselves and not the base pairs as they occur in the nucleic acids where one NH bond (N_1H in T and C, N_9H in A and G) is replaced by the glycosidic bond to the sugar ring, which blocks the corresponding site for water binding. In this situation the AT pair loses one water molecule (that bound to N_9H of A), the water molecule which bridged N_1H and O_2 in thymine being displaced towards O_2. The GC pair loses two waters of hydration, one bound to N_9H of G, the other to N_1H of C. Hence in the nucleic acids the same conclusion remains valid as to the greater hydration of AT compared to GC.

It is gratifying to observe that these theoretical results are supported and in turn support the conclusions in [8] on base-composition dependence of the net hydration of DNAs which these authors investigated by density gradient ultracentrifugation of DNAs in a series of cesium and lithium salt solutions of different water activities. Their conclusion was that at least 90% of the dependence of buoyant density on base composition can be accounted for on the basis of differential hydration, with a 1 mol A–T pairs binding ~ 2 mol more water than 1 mol G–C pairs in CsCl.

References

[1] Port, G. N. J. and Pullman, A. (1973) FEBS Lett. 31, 70.

[2] Pullman, A. and Perahia, D. (1978) Theoret. Chim. Acta (Berlin) in press.

[3] Dreyfus, M. (1970) Thèse de 3è cycle, University of Paris.

[4] Clementi, E., André, J. M., André, M. Cl., Klint, D. and Hahn, D. (1969) Acta Phys. Acad. Sci. Hung. 27, 493.

[5] Perahia, D. and Pullman, A. (1978) Theoret. Chim. Acta 48, 29.

[6] Melder, J. A. and Mead, R. (1968) Comput. J. 7, 308.

[7] Chandler, J. P. (1969) Simplex Method, QCPE, Dept. Chem. Indiana Univ., Bloomington, IN.

[8] Tunis, M. J. B. and Hearst, J. E. (1968) Biopolymers 6, 1345.