

Hydration Scheme of Uracil and Cytosine

A Comparison between Electrostatic and Complete Supermolecule Computations

Alberte Pullman and David Perahia

Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique associé au C.N.R.S.,
13, rue P. et M. Curie, F-75005 Paris, France

The comparison of a pure electrostatic approximation and complete supermolecule SCF *ab initio* computations on the hydration scheme of uracil and cytosine shows that the electrostatic procedure is capable to reproduce the general aspects of the results of the supermolecule treatment provided that different distances of shortest approach be adopted for the distances between the oxygen of water and the nitrogen of NH₂ or NH groups or the oxygen of C=O groups on the one hand and the oxygen of water and pyridine-type nitrogens on the other hand.

Key words: Uracil, hydration scheme of ~ – Cytosine, hydration scheme of ~ – Electrostatic vs. supermolecule computation

1. Introduction

We have investigated in a series of papers (for reviews see [1, 2]) the hydration scheme of a large number of fundamental biomolecules using the SCF *ab initio* method within the “supermolecule” procedure. Our studies involved, among others, the main components of the essential biopolymers, proteins and nucleic acids. The direct extension of these studies to larger units within the same method is, however, prohibitively expensive.

These studies having shown that the electrostatic part of the interaction energy plays a major role in the determination of the hydration scheme, we wish to investigate in this paper the possibility of establishing the conditions of an appropriate pure electrostatic approach which would compare satisfactorily with results of the supermolecule computations and which could then be applied with relatively moderate effort and cost to the study of the hydration scheme of larger biosystems. We have chosen for this comparative study uracil and cytosine, two of the fundamental pyrimidine bases of the nucleic acids.

2. Methods

The supermolecule computations are SCF *ab initio* calculations using a contracted $7s\ 3p/3s$ basis set of Gaussian functions with the exponents and coefficients of Clementi *et al.* [3]. Whenever necessary, the components of the supermolecule SCF interaction energies (Coulombic, ΔE_C , exchange ΔE_{Ex} , delocalization, ΔE_{Del}) have been computed following the procedure of Dreyfus and Pullman [4–6].

The electrostatic approximation tested in the present work is the overlap multipole procedure (OMTP) developed in our laboratory [6, 7]. Its essential distinct feature is the use of a “polycentric” multipole expansion of the electron density of each of the molecules in interaction constructed as a superposition of the individual multipole expansions of all overlap distributions. Specifically, for a molecular wave function expressed as a Slater determinant built on a set of doubly occupied molecular orbitals φ_i written as linear combinations of atomic orbitals χ_r ,

$$\varphi_i = \sum_r c_{ir} \chi_r \quad (1)$$

the electron density distribution:

$$\rho = \sum_{rs} P_{rs} \chi_r \chi_s, \quad (2)$$

appears as a summation of one-center (r,s on the same atom including r,r), and two-center (r,s on different atoms) overlap products weighted by the density matrix elements P_{rs} . In our expansion, *each weighted overlap product* $P_{rs} \chi_r \chi_s$ is approximated by a multipole expansion centered at an appropriate point: for a one-center product, the point is the corresponding nucleus, for a two-center product it is the middle of the segment joining the nuclei which carry χ_r and χ_s respectively. A molecule with n atoms is thus represented by $n(n+1)/2$ centers carrying a set of appropriate multipoles. In the version used here the expansion is limited to the quadrupoles. The nuclear charges remain on each atom.

The interaction energy between the two molecules considered is calculated as the total interaction between the two sets of multipoles, using the classical electrostatic expressions of the charge–charge, charge–dipole, charge–quadrupole, dipole–dipole, dipole–quadrupole and quadrupole–quadrupole interactions [8].

The program is built in such a way as to rotate easily a water molecule around the solute, rotating also the water molecule about three local axes centered on its oxygen atom, so as to find its most favorable orientation.

In order to avoid false zones of apparent stability due to the neglect of the exchange repulsion, important at small distances, a distance of closest approach of the atoms of water to the atoms of the large molecule must be maintained during the rotations.

The wave functions used to obtain the multipole expansions are the SCF *ab initio* wave functions of the individual ligands computed in the same fashion as the global supermolecule.

3. Results

3.1. Uracil

Figure 1 indicates the preferred hydration sites for the fixation of a water molecule upon uracil computed by the OMTP procedure. In this case the oxygen of water was maintained in the plane of the base. As in Ref. [7], a uniform distance of 2.85 Å between the oxygen of the water molecule, O(W), and the first-row atoms of

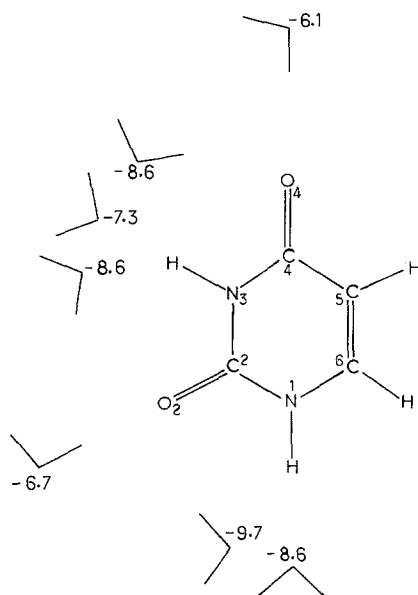


Fig. 1. OMTP electrostatic interaction energies (kcal/mole) of water with uracil. The shortest distance of approach of the oxygen of water to all the first-row atoms of the ligand was uniformly fixed at 2.85 Å

the ligand was adopted as the closest limit of approach. It is seen that the preferred sites obtained in this fashion are those situated between adjacent CO and NH groups, the most stable hydration site occurring between N₁H and the C₂O₂ carbonyl oxygen. For water binding to individual groups of the substrate, the NH groups seem to be preferred to the carbonyl oxygens.

The SCF *ab initio* results obtained by a supermolecule computation of the interaction of water with uracil in the corresponding positions are indicated in Fig. 2. There is an extremely close agreement with the results of Fig. 1 as concerns the location and the preferences of the main binding sites. As in the OMTP results, the preferred site is again that occurring between N₁H and C₂O₂. The supermolecule results also agree with the OMTP ones in the fact that, individually considered, the NH groups of uracil bind water more tightly than the CO groups. The overall agreement between the OMTP and supermolecule results appears to indicate that in the interaction of water with uracil, the electrostatic component commands the interaction.

Concerning the numerical values of the binding energies, it is observed that the supermolecule results are smaller than those of the OMTP calculations. This is due

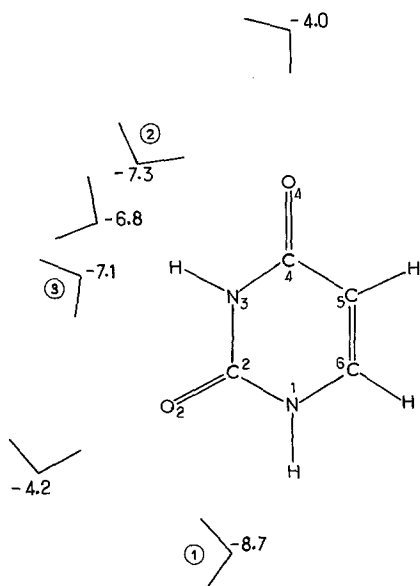


Fig. 2. Supermolecule interaction energies (kcal/mole) of water with uracil. The distances $O \cdots O(W)$ and $N \cdots O(W)$ (optimized for the non-bridged positions) are 2.95 and 2.80 Å respectively

mainly to the exchange repulsions present in the first and omitted in the second; moreover, the fact that the distance of closest approach chosen for the OMTP search is somewhat smaller than the equilibrium distance, overemphasizes the electrostatic energy itself (particularly for $O \cdots O_W$).

In view of the question raised sometimes on the possibility of water binding to the π -electron cloud of conjugated aromatics or hetero-aromatics, we have performed complementary supermolecule computations for some positions of water above the plane of uracil.

Figure 3 indicates the explored sites of interaction. In the first position one OH bond of water points directly above N_3 , the other being oriented in the plane containing N_3C_6 ; in the second position one OH bond of water points directly above O_2 , the other OH pointing outside the ring, *trans* to C_2O_2 ; in the third position the water molecule is placed above the C_5-C_6 double bond in a bridge conformation. The details of these interactions are indicated in Table 1. It is seen that the interaction is either repulsive, being dominated by the (repulsive) exchange component, or slightly attractive. The Coulomb component in these out-of-plane

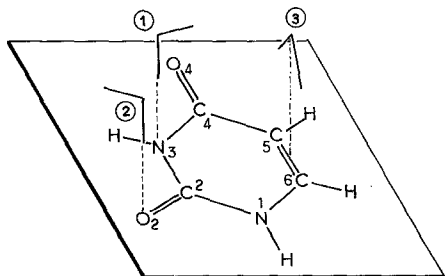


Fig. 3. Sites of out-of-plane interaction of water with uracil

Table 1. Supermolecule binding energies of uracil to water placed above its plane

Site	Distance (Å)	Binding energy (kcal/mole)			
		total	Coulombic	exchange	delocalization
1	$N_3 \dots O(W) = 2.95$	4.6	-2.1	8.4	-1.7
2	$O_2 \dots O(W) = 2.75$	0.6	-8.2	12.8	-4.0
	$O_2 \dots O(W) = 2.90$	-1.0	-5.6	7.4	-2.8
	$O_2 \dots O(W) = 3.10$	-1.8	-3.5	3.4	-1.7
3	$C_5 \dots O(W) = 2.96$	5.7	-2.9	9.7	-1.1
	$C_6 \dots O(W) = 2.96$				

interactions with water is small as could be inferred from the shape of the molecular electrostatic potential of uracil [9] and the out-of-plane delocalization component does not take on particularly large values.

3.2. Cytosine

The essential results of the investigation of the hydration sites of this base by the multipole procedure utilized as before are indicated in Fig. 4 and those obtained by the supermolecule computation in Fig. 5. The positions of water in Fig. 5 are kept the same as in Fig. 4. The *ab initio* results appear here appreciably different from the electrostatic ones, even with respect to the preferred binding site. Thus while this site is found in the vicinity of N_3 in the electrostatic computations it is located between N_1H and C_2O_2 in the *ab initio* ones.

The different components of the binding energies of water to cytosine for each position shown in Fig. 5 are indicated in Table 2. It may be seen in this table that for the 2' position of water facing the N_3 atom, the exchange component of the interaction energy is relatively large (11.1 kcal/mole) at the distance $N_3 \dots O(W) = 2.85 \text{ \AA}$ as compared to the corresponding values observed for the other positions

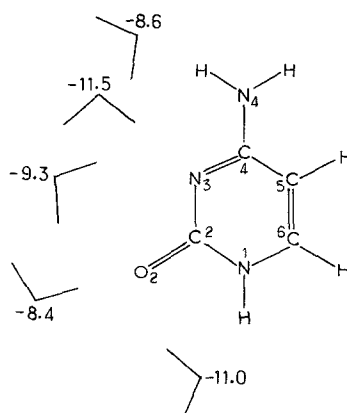


Fig. 4. OMTP electrostatic interaction energies (kcal/mole) of water with cytosine. The shortest distance of approach of the oxygen of water to the heavy atoms of the base is fixed at 2.85 \AA

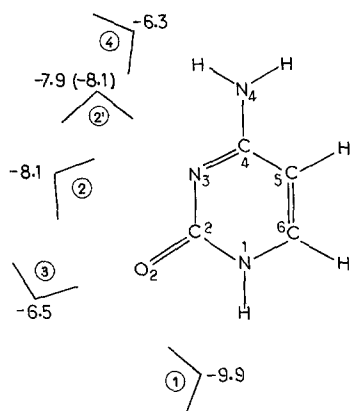


Fig. 5. *Ab initio* interaction energies (kcal/mole) of water with cytosine (-8.1 is the optimal value at position 2', see Table 2)

Table 2. Binding energies of water to cytosine

Site	Distance (Å)	Binding energy (kcal/mole)			
		total	Coulombic	exchange	delocalization
1	$N_1 \cdots O(W) = 3.20$	-9.9	-13.1	7.7	-4.5
	$O_2 \cdots O(W) = 2.85$				
2	$N_3 \cdots O(W) = 3.23$	-8.1	-10.5	4.5	-2.1
	$O_2 \cdots O(W) = 2.85$				
2'	$N_3 \cdots O(W) = 2.85$	-7.9	-15.1	11.1	-3.9
	$N_3 \cdots O(W) = 2.95$	-8.1	-12.5	7.6	-3.2
	$N_3 \cdots O(W) = 3.05$	-7.8	-10.5	5.2	-2.5
3	$O_2 \cdots O(W) = 2.85$	-6.5	-9.7	6.1	-2.9
	$N_4 \cdots O(W) = 2.89$	-6.3	-10.3	6.7	-2.7

of water. When the distance between N_3 and $O(W)$ is increased from 2.85 \AA to 2.95 \AA (which corresponds to the equilibrium distance), the value of the exchange component decreases substantially, from 11.1 to 7.6 kcal/mole, whereas the Coulombic component decreases from -15.1 to -12.5 kcal/mole. In the OMTP electrostatic calculations of Fig. 4, the binding site facing N_3 of cytosine corresponds to $N_3 \cdots O(W) = 2.85 \text{ \AA}$ and corresponds therefore to a too large electrostatic attraction. In the light of the above discussion, a more correct hydration scheme may be expected to be obtained by OMTP computations if the closest distance of approach between the pyrimidine-type nitrogen and the oxygen of water is fixed at 2.95 or 3 \AA , while the other closest distances remain fixed at 2.85 \AA . The OMTP results obtained with the above conditions are presented in Fig. 6. The results are now comparable with the *ab initio* ones of Fig. 5, in particular the same overall order of preferences for the binding of water to the different available sites is found provided that the slight energy differences (up to 0.2 kcal/mole) between some sites be not considered as significant in both methods.

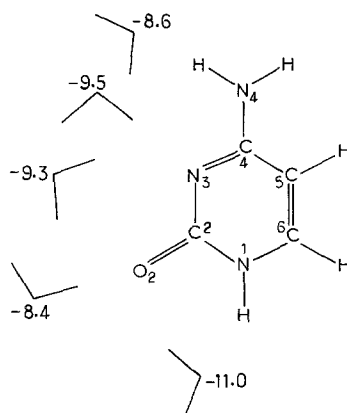


Fig. 6. OMTP electrostatic interaction energies (kcal/mole) of water with cytosine. The distance of shortest approach between $O(W)$ and the NH_2 or NH nitrogens and the carbonyl oxygen is fixed at 2.85 Å, that between $O(W)$ and N_3 at 3 Å.

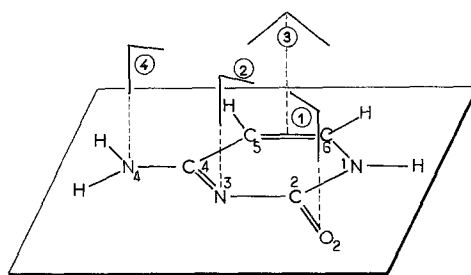


Fig. 7. Sites of out-of-plane interaction of water with cytosine. The non-hydrogen-bonded proton of water is in the plane $O(W) N_3 O_2$ for site 1, in the plane $O(W) O_2 N_1$ for site 2 and in the plane $O(W) N_4 N_3$ for site 4. Position (3) is a symmetrical bridge

Table 3. Out-of-plane binding energies of water to cytosine

Site	Distance (Å)	Binding energy (kcal/mole)			
		Total	Coulombic	exchange	delocalization
1	$N \cdots O(W) = 2.85$	1.0	-8.1	12.6	-3.5
	2.95	-0.5	-6.4	8.6	-2.7
	3.05	-1.4	-5.1	5.9	-2.2
	3.15	-1.9	-4.2	4.0	-1.7
	3.23 ^a	-2.2	—	—	—
2	$O \cdots O(W) = 2.85$	-1.2	-6.8	9.1	-3.5
	2.95	-2.0	-5.4	6.2	-2.8
	3.05	-2.3	-4.4	4.2	-2.1
	3.10 ^a	-2.5	—	—	—
3	$C \cdots O(W) = 2.85$	6.6	-3.2	11.1	-1.3
	3.05	3.4	-1.1	5.3	-0.8
4	$N \cdots O(W) = 2.85$	1.9	-8.0	12.3	-2.4

^a Extrapolated minimum.

The above conditions may then be considered satisfactory for the study by the OMTP procedure of the general features of the hydration scheme of related substrates presenting the same type of atoms as cytosine.

Similarly to the case of uracil we have also studied the possibilities of out-of-plane binding of water to cytosine. The results are presented in Fig. 7 and Table 3. The values obtained indicate that the out-of-plane binding is unfavorable compared to the in-plane binding.

4. Conclusion

The above results show that a correct picture of the hydration scheme of a substrate (here the nucleic acid bases), mimicking SCF *ab initio* results, may be obtained with the help of an electrostatic approximation alone provided that appropriate precautions are taken. The supermolecule results show that at a near equilibrium distance of 2.85 Å between the oxygen of water and the first-row atoms of the substrate (nitrogens of NH₂ and NH groups and the carbonyl oxygens), involved in hydrogen bonding, the exchange repulsions have comparable magnitudes. For the pyridine-type nitrogens the exchange repulsions are on the other hand relatively large at this same distance due to the fact that the atomic orbitals of a nitrogen atom extend further out than those of an oxygen atom [10, 11]. The adoption of a larger distance of closest approach (2.95 to 3 Å) between O(W) and this last type of nitrogen in the electrostatic calculations, leads to results in general agreement with the *ab initio* ones.

Acknowledgment. The authors wish to express their thanks to Dr. A. S. Kolaskar for his help in the achievement of some calculations.

References

1. Pullman, A., Pullman, B.: *Ann. Rev. Biophys.* **7**, 505 (1975)
2. Pullman, A., in: *Environmental effects on molecular structure and properties*, Proceedings of the 8th Jerusalem symposium in quantum chemistry and biochemistry, Pullman, B., ed., pp. 1-15. Dordrecht, Holland: Reidel Publishing Co.
3. Clementi, E., André, J. M., André, M. Cl., Klint, D., Hahn, D.: *Acta Phys. Acad. Sci. Hungaricae* **27**, 493 (1969)
4. Dreyfus, M., Pullman, A.: *Comptes Rendus Acad. Paris* **271**, 457 (1970)
5. Dreyfus, M., Pullman, A.: *Theoret. Chim. Acta (Berl.)* **19**, 20 (1970)
6. Dreyfus, M.: Thèse de 3^e cycle, University of Paris (1970)
7. Port, G. N. J., Pullman, A.: *FEBS Letters* **31**, 70 (1973)
8. See, for instance: Rein, R.: *Advan. Quantum Chem.* **7**, 335 (1973)
9. Perahia, D., Pullman, A., Pullman, B.: *Theoret. Chim. Acta (Berl.)* **42**, 23 (1976)
10. Pullman, A. in: *Mécanismes d'altération et de réparation du DNA*, colloque n° 256 du C.N.R.S., p. 103 (1976)
11. Perahia, D., Pullman, A., Pullman, B.: *Theoret. Chim. Acta (Berl.)* **43**, 207 (1977)

Received September 30, 1977