Theoret. Chim. Acta (Berl.) 43, 207–214 (1977)

THEORETICA CHIMICA ACTA

Cation-Binding to Biomolecules

IV. An *ab initio* Study on the Interaction of Na^+ with the Purine and Pyrimidine Bases of the Nucleic Acids

David Perahia, Alberte Pullman and Bernard Pullman

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

Ab initio SCF MO computations are performed on the binding of the Na^+ ion to the purine and pyrimidine bases of the nucleic acids. The results are compared in particular with those of previous studies on proton affinities. The computations indicate two principal differences between these two types of interactions: 1) the increased significance in cation binding of the individual oxygen binding sites in bases containing both oxygen and nitrogen; 2) the appearance in cytosine and guanine of "bridged" positions between a nitrogen and a carbonyl oxygen as the preferred binding sites. The theoretical results are compared with the available experimental data. Their general significance for cation binding is discussed.

Key words: Purines – Pyrimidines – Cation-binding to biomolecules

1. Introduction

As a continuation of our study on the interaction of cations with biomolecules and in particular with the constituents of the nucleic acids [1–4] we present here an *ab initio* investigation on the interaction of Na^+ with the purine and pyrimidine bases of the nucleic acids. We would like to underline from the start that these computations, which involve in each case the interaction between a single base and a single cation, must be considered essentially as model investigations for this particular situation which may rarely be encountered. Real cation-ligand interactions in this field (and *a fortiori* with nucleosides, nucleotides or nucleic acids) imply generally a more complex scheme, involving a larger number of competing binding sites, anionic forms of bases, hydrated or chelated cations, cross linkings etc. Nevertheless, it appears that this model study is able to provide some fundamental information about the intrinsic cation binding tendencies of the bases which may be related to some aspects of the interactions occurring in the more complex real situations. From that point of view the investigation appears as a useful if not indispensable first step on a road leading to the understanding of the nature of metal-ligand interactions involving nucleic acids.

2. Procedure

The computations have been performed within the SCF LCAO *ab initio* procedure using Gaussian basis functions. The Gaussian orbitals adopted for the nucleic acid bases consist of a (7*s*, 3*p*/3*s*) basis contracted to a minimal set with the exponents and contraction coefficients utilized by Clementi *et al.* [5] in their computations on the DNA bases. The basis set used for Na^+ was derived [6] from the STO 3G basis given for the sodium atom [7] by reoptimization of the 1*s*, 2*sp* and 3*sp* exponents with respect to the energy of the cation and suppression of the 3*p* empty orbitals. It has been shown elsewhere [6] in a study of the binding of the alkali and alkaline-earth ions to water that the use of this type of basis for the cation together with the (7*s*, 3*p*/3*s*) set on the ligand yields reasonable values of both the binding energies and the distances of approach, although these latter are slightly too short. In each case the components of the interaction energy (Coulombic, exchange and delocalization, this last one representing polarization + charge transfer) have been computed following the procedure of Dreyfus and Pullman [8] in order to analyse better the origin of the binding preferences observed.

The geometries adopted for the bases as input data, correspond to the mean values given by Arnott [9].

3. Results and Discussion

3.1. Interaction of Na^+ with Uracil

We have already studied previously [2] in some detail the interaction of uracil with the Mg^{++} cation. The choice of this particular system was dictated by the existence of a similar study performed by the CNDO/2 method, the results of which (preferential binding to the C_5=C_6 double bond of uracil) seemed *a priori* surprising and were in fact shown to represent an artifact of this method. Because of the high computational time required in this type of calculation we have limited the present study to what appeared from the previous one as corresponding most probably to the preferential sites of binding and have explored essentially two directions of approach of the cation: one making an angle of 10° with the C_4=O_4 bond on the side of C_5, the other located along the C_2=O_2 bond. These directions are oriented towards the minima of the molecular electrostatic potential map constructed for uracil [2] or thymine [10].

Fig. 1a indicates the location of the preferred binding sites and further information on the results is contained in the corresponding part of Table 1. The optimum distances between the cation and the carbonyl oxygens are about 2 Å in these computations, and the binding energies of the order of 28–33 kcal/mole, O_4 being the preferred binding site. The decomposition of the interaction energy into its constituent terms, indicated in Table 1, shows the predominance of the Coulombic term. In fact at the equilibrium distances with respect to O_4 and O_2, the exchange and delocalization energies, which are of opposite sign, practically equilibrate each other, so that the energy of binding is practically equal to the Coulombic term. As a result of this situation, the molecular

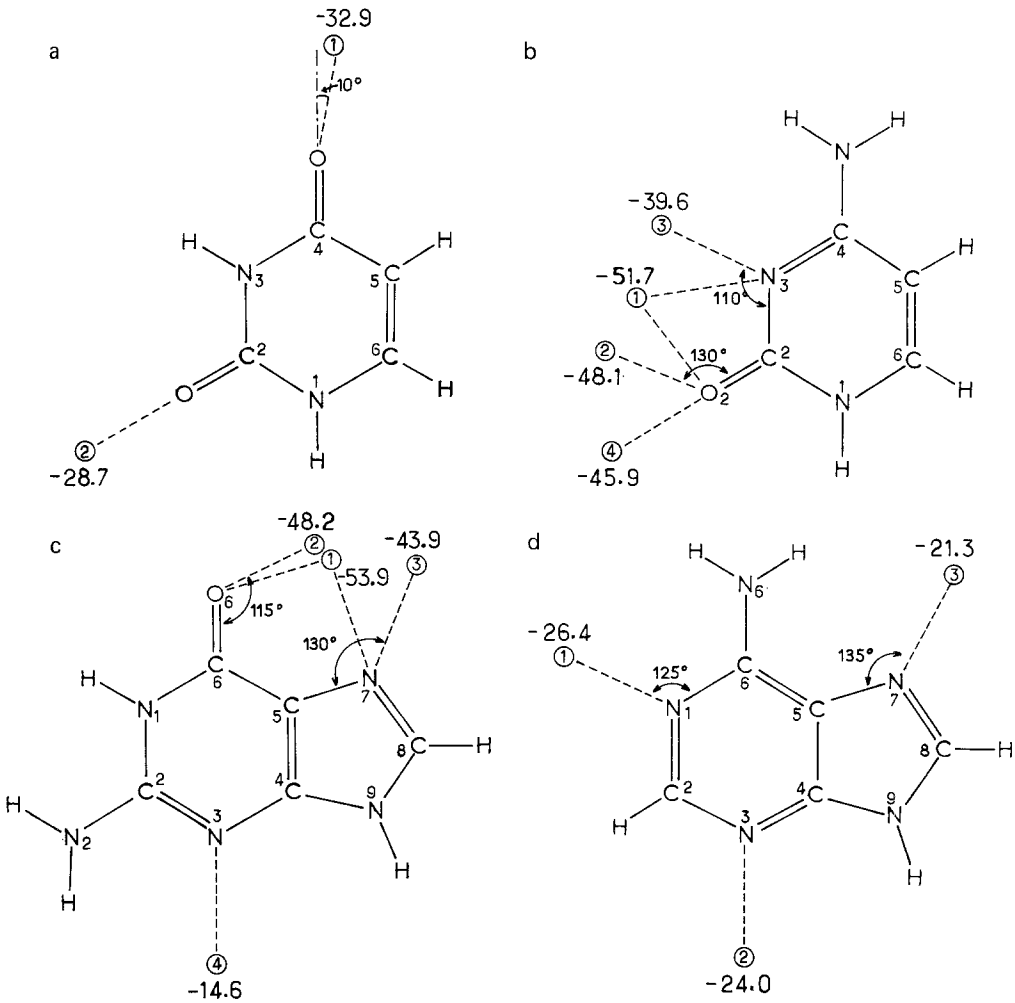

Fig. 1. *Ab initio* interaction energies (kcal/mole) of Na^+ with a) uracil, b) cytosine, c) guanine and d) adenine

electrostatic potential map of uracil is sufficient to predict with a fairly good approximation the axis of preferred location of the binding sites and, at the equilibrium distance, the interaction energy.

It may be useful at this point to compare these results with those found previously for the interaction of uracil with Mg^{++}. The preferred sites of binding were located along the same axis of approach. The optimum distance of approach between the divalent cation and O_2 or O_4 was, however, shorter: 1.75 Å. On the contrary, the energies of binding were much larger: −103.8 kcal/mole for binding to O_4 and −92.4 kcal/mole for binding to O_2. O_4 is thus the preferred binding site in both cases. On the other hand, in binding with Mg^{++} the total binding energy exceeds the Coulombic term by 15–20 kcal/mole because in this case the delocalization effect outweighs the exchange term by about this amount. The Coulombic component remains, however, the pre-

Table 1. Binding energies of Na^+ to nucleic acid bases

Sites of Cation Binding (See Fig. 1)	Ligand...Cation Distance (A)	Energy of Cation Binding (kcal/mole)			
		Total	Coulombic	Exchange	Delocalization
a) *Binding to uracil*					
1	$O_4...Na^+ = 2.15$	−32.70	−28.25	4.75	−9.2
	$O_4...Na^+ = 2.00$	−32.92	−32.83	11.13	−11.22
	$O_4...Na^+ = 1.85$	−28.69	−39.59	24.77	−13.87
2	$O_2...Na^+ = 2.00$	−28.72	−29.28	10.90	−10.34
b) *Binding to cytosine*					
1	$N_3...Na^+ = 2.15$, $O_2...Na^+ = 2.00$	−48.25	−58.52	26.03	−15.76
	$N_3...Na^+ = 2.30$, $O_2...Na^+ = 2.15$	−51.73	−51.08	11.66	−12.31
2	$O_2...Na^+ = 2.00$	−48.09	−50.12	12.93	−10.90
	$O_2...Na^+ = 2.15$	−47.90	−44.35	5.58	−9.13
3	$N_3...Na^+ = 2.00$	−35.95	−51.62	26.15	−10.48
	$N_3...Na^+ = 2.15$	−39.59	−43.60	12.75	−8.74
	$N_3...Na^+ = 2.30$	−39.46	−37.96	5.93	−7.43
4	$O_2...Na^+ = 2.00$	−45.87	−44.83	10.92	−11.96
c) *Binding to guanine*					
1	$N_7...Na^+ = 2.90$, $O_6...Na^+ = 2.15$	−53.89	−52.08	12.05	−13.86
2	$O_6...Na^+ = 2.00$	−48.20	−50.03	13.81	−11.98
3	$N_7...Na^+ = 2.15$	−43.88	−48.43	12.77	−8.22
4	$N_3...Na^+ = 2.15$	−14.62	−17.96	12.86	−9.52
d) *Binding to adenine*					
1	$N_1...Na^+ = 2.15$	−26.35	−29.84	12.64	−9.15
2	$N_3...Na^+ = 2.15$	−24.04	−28.04	12.89	−8.89
3	$N_7...Na^+ = 2.15$	−21.34	−24.75	13.89	−10.48

dominant one in this interaction so that here again the molecular electrostatic potential map, which *per se* gives direct information on proton affinities, indicates also the essential lines of cation affinity.

3.2. Interactions of Na^+ with Cytosine

Fig. 1b indicates the *ab initio* binding energies of Na^+ to cytosine in a series of significant positions numbered 1 to 4. In Table 1 may be found the details on the components of the binding energies. The binding of Na^+ separately to N_3 and O_2 has been explored along directions making respectively an angle of 110° with C_2-N_3 (position 3) and 130° with C_2-O_2 (position 2). These directions point towards the two minima of the molecular electrostatic potential map of cytosine [10]. The maximum binding energies obtained along these directions are −48.1 kcal/mole for site 2 and −39.6 kcal/mole for site 3, corresponding, however, to different optimum ligand-cation distances which are respectively 2.15 Å and 2 Å. The carbonyl oxygen exerts a stronger attraction (by about 8.5 kcal/mole) than the nitrogen N_3 towards the cation. An examination of

Table 1 shows that in binding to a single site and for equal distances of Na^+ to O_2 or N_3, the respective Coulombic components of the interaction energies have almost similar magnitudes. The same is also true for the delocalization energies. On the contrary, the exchange term corresponding to the binding of Na^+ to N_3 is almost twice as large as that corresponding to the binding to O_2. This situation explains why the optimum distance found between Na^+ and N_3 (2.15 Å) is larger than the optimum distance between Na^+ and O_2 (2 Å). It may thus be said that it is mainly the exchange component of the interaction energies which is responsible for a less favourable binding of Na^+ to N_3 than to O_2. This situation creates a different relation between the cation-binding sites and the molecular electrostatic potential map than in the case of uracil. In the former, in which the preferred sites were uniformly carbonyl oxygens, the electrostatic potential map constituted a good guide towards the location and ordering of the cation binding sites. In the case of cytosine the electrostatic potential energy map (Fig. 2 of Ref. [10]), which shows a *minimum minimorum* in the region of N_3 and a secondary minimum in the region of O_2, is not sufficient for predicting the preferred binding site for the cation, the Coulombic component of the interaction energies being no longer the determining factor for the binding preferences when different kinds of binding sites (here nitrogen and oxygen) are present on the ligand, in spite of the fact that it is the dominant part of the energy quantity wise.

Moreover, we have explored the possibility in which Na^+ would be simultaneously bound to N_3 and O_2. Such a bridge position of the cation with respect to the two atoms has been computed for different ligand-cation distances. When Na^+ is placed simultaneously at the optimum (individual) distances with respect to N_3 and O_2 ($N_3 \ldots Na^+$ = 2.15 Å, $O_2 \ldots Na^+$ = 2 Å) the binding energy of −48.25 kcal/mole obtained is quite similar to that of site 2 (see Table 1). When the ligand-cation distances are increased in the bridge position, to $N_3 \ldots Na^+$ = 2.30 Å and $O_2 \ldots Na^+$ = 2.15 Å, the binding energy becomes −51.73 kcal/mole. Thus the bridge position is now the energetically preferred one. The result is due principally to the rapid decrease of the repulsive exchange component, greater than the concomitant decrease of the attractive Coulombic and delocalization components.

3.3 Interactions of Na^+ with Guanine

The *ab initio* results obtained on the interactions of Na^+ with guanine are indicated in Fig. 1c and the corresponding part of Table 1. Site 1, representing a simultaneous binding of the cation to N_7 and O_6, appears as the preferred binding position, the corresponding interaction energy being −53.9 kcal/mole. This situation is similar to that obtained for cytosine. The other binding sites explored are situated on the axis directed towards the minima of the electrostatic potential map of guanine [11]. These minima were located around atoms O_6, N_7 and N_3, which thus represent the "individual" poles of attraction for a bare positive charge. In the results of Fig. 1c, O_6 appears to be more attractive than N_7 towards Na^+: this represents an inversion with respect to the affinities of these sites for a proton. The binding of the cation to N_3 (site 4) is much less favourable (as it was also for proton affinity), due essentially to the much smaller value of the Coulombic term for the interaction with this site.

3.4. Interactions of Na^+ with Adenine

The interaction energies of Na^+ with adenine have been computed for three positions of the cation situated in the regions of N_1, N_3 and N_7 along the axis pointing towards the minima of the *ab initio* electrostatic potential map of this base [10, 11]. The results are indicated in Fig. 1d and the last section of Table 1. The preferred binding site for Na^+ is N_1, followed by N_3 and N_7. The binding energies are altogether low when compared to those of the three other bases. This situation is due mainly to the low values of the Coulombic component of the interaction energies (see Table 1).

3.5. General Discussion and Conclusions

The above results lead to some general observations among which three deserve to be underlined, in particular in relation with previous results on the molecular electrostatic potentials of the bases which indicated their preferential proton affinities.

1) The first concerns the increased significance of the individual oxygen binding sites in compounds containing both oxygen and nitrogen. For example, in cytosine, interaction energies at sites 2 or 4 representing individual binding of the cation to the carbon carbonyl oxygens are greater than at site 3, representing individual binding to a nitrogen atom. The reverse was true for proton interaction energies. A similar situation concerns binding at sites 2 and 3 in guanine. This inversion in the intrinsic binding ability for a proton and for Na^+ is due to the exchange repulsive component of the binding energy which, for the same distance, is larger for nitrogen than for oxygen. This is understandable on the ground that the valence shell orbitals of the oxygen atom extend less in space than those of nitrogen, thus giving rise to a smaller overlap with those of the cation at the same distance, and hence to a smaller repulsion.

2) The appearance in cytosine and guanine of "bridged" sites 1, involving simultaneously a nitrogen and a carbonyl oxygen, as the preferred binding position. In the electrostatic potential energy maps of these compounds these N and O atoms were generally associated with separate minima and an intermediate bridged position for a proton was less favourable. It must however be observed that this is due to the fact that the equilibrium distance for a proton is much smaller than that for a sodium ion so that the proton can move into the potential minima located close to the heteroatoms.

3) The intrinsic complexing ability of the free bases towards the sodium ion as measured by the values of the interaction energies is greater for guanine and cytosine than for uracil and adenine.

Concerning the first observation, although it is generally considered (see e.g. [12]) that alkali metal ions bind in solution exclusively to the phosphate moieties of ribonucleotides and DNA (a situation understandable in view of the much more favourable interaction energies of Na^+ with the phosphate group, see Refs. [1] and [4]), a few recent crystallographic results indicate Na^+ co-ordination to the carbonyl oxygens of uracil, cytosine and guanine. Such is the case for example in the crystal structure of cytidine-5′-diphosphocholine [13], sodium β-cytidine-2′,3′-cyclic phosphate [14], the sodium

salt of adenosyl-3′,5′-uridine phosphate [15], tymidylyl-(5′-3′)-thymidylate-5′ [16] and disodium deoxyguanosine-5′-phosphate tetrahydrate [17].

Concerning the second observation, a bridged positioning of the Na^+ cation between O_2 and N_3 is observed in the two above quoted Refs. [13] and [14] concerned with X-ray studies on derivatives of cytidine. Also, in the crystal of disodium deoxyguanosine-5′-phosphate tetrahydrate [17] one of the two cations is bound directly to O_6 of the guanine moiety and through a water bridge to its N_7 (the second cation is linked to O_6 and to oxygens of the phosphate).

This last example displays a degree of binding complexity which goes beyond the simple model study considered here by involving also a water molecule. By doing so it indicates the obvious limits of the direct applicability of our model. Most of the experimental situations studied in relation to the interaction of cations with the nucleic acid base involve supplementary elements such as water, counterion, complex derivatives of the bases, chelated cations etc. A large number of studies have been made with cations belonging to the transition metals. This situation results in various binding schemes which *a priori* need not be those described here. Nevertheless it appears that the principal results obtained by us have a general significance for such bindings. For example, a recent critical review of crystallographic data involving transition metal cations by Kistenmacher [18] stresses the importance of the "bridged" binding, involving the N and O atoms in cytosine and guanine derivatives and this is also indicated by a number of solution studies with different cations (see e.g. [12]). The bridge may be partly an indirect one, through the medium of water (or other ligands) attached to the cation, but altogether seems to represent one of the essential tendencies of cation binding to these bases. The involvement of water may be ascribed to the high energy of water binding to the cation, frequently competitive with the binding energy to the base. (The energy of $Na^+ \ldots H_2O$ binding computed with the same approximations as its binding with the purines and pyrimidines in the present study amounts to 29 kcal/mole [6].) On the other hand, it is obvious that greater complexities of the ligand skeleton may introduce binding schemes and preferences different from those obtained in a model study of the type carried out here. Thus, for example, the presence of an oligophosphate chain on a base, as in ATP, may produce specific binding preferences by making possible the simultaneous binding of the cation to the phosphate group and selected atom of the adenine ring. In fact the X-ray crystal structure of ATP [19], consisting of two independent molecules with four sodium ions in the asymmetric unit cell, shows that two of the sodium ions co-ordinate the two molecules through the phosphate oxygens and N_7 of the adenine ring. (The two remaining cations in the unit cell are bound only to phosphates and hydroxyl groups of the sugar rings.) The choice of N_7 instead of N_1 or N_3 in the binding may be considered as brought about by the possibility of co-ordination with the phosphate group.

Although each situation and each cation have certainly to be considered individually, it nevertheless appears that the model study of the interaction between Na^+ and the purine and pyrimidine bases produces information on the intrinsic cation binding abilities of these ligands, at least for alkali and possibly alkaline earth cations, which may form a useful basis for the understanding of the factors involved in such interactions. We are proceeding presently to the exploration of bindings involving more complex ions.

References

1. Pullman, B., Gresh, N., Berthod, H.: Theoret. Chim. Acta (Berl.) **40**, 71 (1975)
2. Perahia, D. Pullman, A., Pullman, B.: Theoret. Chim. Acta (Berl.) **42**, 23 (1976)
3. Pullman, A., Berthod, H.: Chem. Phys. Letters (in press)
4. Pullman, A., Pullman, B.: Metal-ligand interactions in organic and biochemistry. In: Proceedings of the 9th Jerusalem Symposium in Quantum Chemistry and Biochemistry, Pullman, B., Goldblum, N. Eds. Dordrecht, Holland: Reidel Publishing Comp. (in press)
5. Clementi, E., André, J. M., André, M. Cl., Klint, D., Hahn, D.: Acta Phys. Acad. Sci. Hungaricae **27**, 493 (1969)
6. Berthod, H., Gresh, N., Pullman, A.: Intern. J. Quantum Chem. **S10** (1976) (in press)
7. Hehre, W. J., Ditchfield, R., Stewart, R. F., Pople, J. A.: J. Chem, Phys. **52**, 2769 (1970)
8. Dreyfus, M., Pullman, A.: Theoret. Chim. Acta (Berl.) **19**, 20 (1970)
9. Arnott, S.: Prog. Biophys. Mol. Biol. **21**, 267 (1970)
10. Bonaccorsi, R., Pullman, A., Scrocco, E., Tomasi, J.: Theoret. Chim. Acta (Berl.) **24**, 51 (1972)
11. Bonaccorsi, R., Pullman, A., Scrocco, E., Tomasi, J.: Theoret. Chim. Acta (Berl.) **36**, 339 (1975)
12. Izatt, R. M., Christensen, J. J., Rytting, J. H.: Chem. Rev. **71**, 439 (1971)
13. Viswamitra, M. A., Seshadri, T. P., Post, M. L., Kennard, O.: Nature **258**, 497 (1975)
14. Coulter, C. L.: J. Am. Chem. Soc. **95**, 570 (1973)
15. Rosenberg, J. M., Seeman, N. C., Kim, J. J. P., Suddath, F. L., Nicolas, H. B., Rich, A.: Nature **243**, 150 (1973)
16. Camerman, N., Fawcett, J. K., Camerman, A.: Science **182**, 1142 (1973)
17. Young, D. W., Tollin, P., Wilson, H. R.: Acta Cryst. **B30**, 2012 (1974)
18. Kistenmacher, T. J., Mayzilli, L. G. in: Metal-ligand interaction in organic and biochemistry. Proceedings of the 9th Jerusalem Symposium in Quantum Chemistry and Biochemistry, Pullman, B., Goldblum, N. Eds. Dordrecht, Holland: Reidel Publishing Comp. (in press)
19. Kennard, O., Isaacs, N. W., Motherwell, W. D. S., Coppola, J. C., Wampler, D. L., Larson, A. C., Watson, D. G.: Proc. Roy. Soc. **A325**, 401 (1971)

Received May 31, 1976