Electrostatic investigation of metal cation binding to DNA bases and base pairs

Shridhar R. Gadre,[a] Savita S. Pundlik,[a] Ajay C. Limaye[b] and Alistair P. Rendell[b]

[a] *Department of Chemistry, University of Pune, Pune 411007, India*
[b] *Supercomputer Facility, Australian National University, Canberra, ACT 0200, Australia*

Various binding sites in DNA bases and base pairs as predicted by rigorous analysis of the molecular electrostatic potential (MESP) are explored for coordination with Li^+ and Ca^{2+} cations; the electrostatics is generally seen to provide an explanation for the observed trends in binding upon subjecting the anticipated structures to optimization at a high level *ab initio* theory.

The interactions between metal cation and the bases of DNA are being widely studied for gaining insights into the origin of stabilization or destabilization of DNA due to the presence of such ions.[1] In the polynucleotide DNA sequence, most of the cations predominantly interact with the backbone phosphate groups,[2] the charge neutralization leading to an enhancement of stabilization of the sequence. However, the cation–base interaction is not negligible.[1] In fact, some transition metal ions like Zn^{2+} and Cd^{2+} are found[2] to interact extensively with the bases facilitating renaturation of thermally denatured DNA.

Theoretical Hartree–Fock calculations carried out for the Watson–Crick (WC) base pairs with minimal basis[3] revealed that metal ion binding, in general, leads to an increased stability of the complementary base pairs. It is observed experimentally that the purine–purine–pyrimidine (PuPuPy) type triplexes respond differently to various cations.[4] For example, the GGC triplexes are found to be stabilized by divalent alkaline earth as well as transition metal cations while AAT triads are stabilized exclusively by the latter. The observed differential stabilization also finds support from the recent high level *ab initio* calculations by Sponer *et al*.[5] who have explained the phenomenon on the basis of missing lone-pair interactions with d orbitals in the case of alkaline earth cations.

An important issue concerning DNA–metal cation interactions is the relative energetic preference of the various lone-pair sites in the bases. There seems to be a general consensus that the N^7 site in guanine is the most favored one among all the bases for a given cation,[1,6] although the recent theoretical treatment[1] also emphasizes the influence of the O^6 atom in stabilizing the cation. The larger stabilization in G–cation complexes has been[1] accounted for by the large dipole moment of the base.

The interactions of DNA with the metal cation are mostly driven by electrostatics. However, an analysis of the binding sites in terms of the complete electrostatic description of bases and base pairs is conspicuous by its absence from the earlier literature. The MESP,[7] V, at a point $\boldsymbol{r}$ due to nuclear charges $\{Z_\Lambda\}$ at $\{\boldsymbol{R}_\Lambda\}$ and the electronic charge density $\rho(\boldsymbol{r})$ is defined as

$$V(\boldsymbol{r}) = \sum_{\mathrm{A}} \frac{Z_{\mathrm{A}}}{|\boldsymbol{R}_{\mathrm{A}} - \boldsymbol{r}|} - \int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}' - \boldsymbol{r}|} d^3r'$$

$V(\boldsymbol{r})$ can assume both positive and negative values and can provide useful information regarding electron-rich sites. It is interesting to see[8a,b] whether all the negative valued critical points (CPs) in MESP of the base (and the base pair) turn out to be the binding sites for cations and also whether any of the amino-lone pair sites are accessed by them. Since the reported findings[1,3] are based on calculations with either the effective core potentials or basis sets of limited extent for the metal species, there is a need for employing better quality basis sets including polarization functions. Here, we report the findings of base–cation as well as WC base pair–cation interactions using an electrostatic approach at a high level *ab initio* theory.

Planar optimized structures at 6-31G** basis set are used for the bases A, G, T, C and the WC base[9] pairs AT and GC. A 6-31G* basis set is employed for Li while for Ca, use is made of TZV* basis. The topography of MESP[10] of the organic molecules is used for initial positioning of Li^+ or Ca^{2+}. The cation is docked by keeping the base (or base pair) geometry fixed until the electrostatic interaction energy[11] attains a minimum. The resultant structure obtained by electrostatic docking is consequently subjected to full *ab initio* geometry optimization carried out on Fujitsu VPP300 with the GAUSSIAN94 package tuned for this platform. The single point energy (ΔE_S) values as well as the interaction energies of the fully optimized structures (ΔE_F) are reported in Table 1.

It can be seen that the relative ordering of the interaction energies for Li^+ and Ca^{2+} with various hosts is according to the depth of negative potential at the CP positions. On *ab initio* optimization, it is found that the distance of the cation from the nearest atom in the host is similar to the distance of the

Table 1 Interaction energies [kcal mol^{-1} (1 cal = 4.184 J)] with single point SCF calculations at the electrostatically docked geometry, ΔE_S, and full optimization, ΔE_F for various DNA···cation interactions[a]

Host	Site	ΔE_S with Li^+	ΔE_F with Li^+	ΔE_S with Ca^{2+}	ΔE_F with Ca^{2+}
A	a_1	−42.83	−45.80	−66.52	−70.97
	a_2	−42.85	−46.54	−65.70	−72.35
	a_3	−24.94	−45.80[b]	−41.62	−92.31[c]
	a_4	−39.28	−41.41	−56.15	−61.90
G	g_1	−19.00	−40.85[c]	−38.73	−62.74
	g_2	−27.11	−33.10	−40.31	−47.01
	g_5, g_6, g_7	−68.13	−78.24	−124.31	−133.78
T	t_2	−44.68	−51.90	−70.86	−81.40
	t_4	−46.45	−53.63	−71.25	−86.06
C	c_3, c_4, c_5	−70.20	−76.02	−108.16	−123.33
AT	at_1	−57.49	−62.82	−85.13	−142.63[d]
			(−4.47)		
	at_2	−44.34	−56.98[b]	−67.52	−161.81[b,d]
	at_3	−49.67	−56.98	−72.68	Not
			(−5.81)		converged
	at_4, at_5, at_6	−68.04	−68.97	−99.75	−111.70
			−5.26		−18.35
	at_7	−56.31	−63.77	−84.91	−161.81[d]
			(−10.13)		
GC	gc_1	−58.58	−100.56[d]	−53.62	−138.13[d]
	gc_2	−66.91	−73.12	−62.63	−74.06
			(−13.97)		(−27.05)
	gc_3, gc_4, gc_5	−103.79	−110.57	−142.74	−143.63
			(−6.28)		(−10.45)
	gc_6	−66.89	−79.19	−58.45	−200.05[d]

[a] The relevant CPs in MESP (Fig. 1) of the bases and the WC base pairs are employed as starting positions of the cations. The numbers in parentheses for AT and GC are the base pair stabilization energies on cation binding. [b] These structures are similar to those obtained from different starting guess. [c] Amino group is twisted to facilitate manifold coordination of the cation. [d] Cation is sandwiched between two bases splitting the base pair.

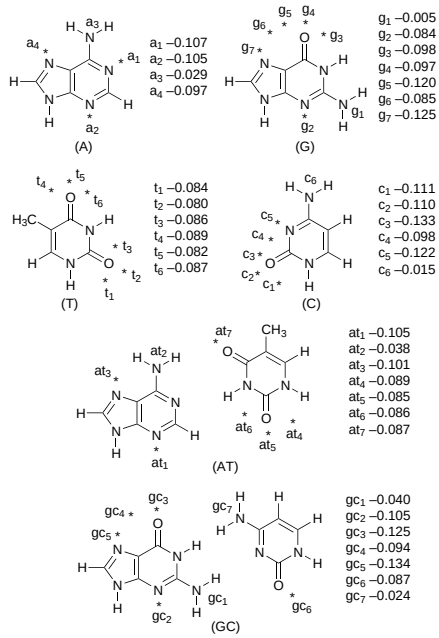

Fig. 1 MESP topographical features of DNA bases A, G, T and C as well as Watson–Crick base pairs AT and GC. The CPs are marked by * and corresponding MESP values given alongside.

corresponding MESP CP. Whereas the binding energies at various sites in A as well as in T are quite close, the region between the O^6 and N^7 atoms in G shows much more binding strength than other sites. From an electrostatic point of view, it is natural to expect that the extensive negative MESP region constituted by a number of CPs stabilizes the cation to a greater extent in G and C. In T, the cation enjoys a surrounding negative MESP region when it is close to t_2 or t_5, the saddles in MESP. Thus, rather than relying upon the partial information such as moments of various orders, the MESP topography leads to a more complete picture of these interactions. This is evident from the fact that all the negative MESP regions which are disjoint give rise to distinct optimum structures.

Electrostatic energy minimization predicts off-plane binding for both the cations with A and G near the amino nitrogen but not for C, probably due to the high negative potential region being concentrated only near O^2 and N^3 in C. Further optimization at the SCF level leads Li^+ to converge to the position a_1 (rather than a_4) in A without affecting the geometry of the NH_2 group. In G, however, full optimization leads Li^+ to the g_2 site with an out of plane twisting of the amino group. The cation is thus able to coordinate with the N^2 as well as N^3 atoms forming a bidentate structure. Using the wavefunction of this twisted amino-G, it is observed that the MESP minimum at N^3 (g_2) deepens to $-0.100\ E_h$. This is reflected in the improvement in the stabilization energy ΔE_F which is much more than the value for the monodentate structure at the N^3 site. The differential behavior of Li^+ towards N^6 in A and N^2 in G could arise due to a relatively deep negative MESP region in the former and a very weak one in the latter (*cf.* Fig. 1). Similar bidentate structure is formed by Ca^{2+} with A (coordinating with N^1 and N^6) while the optimized structure for G–Ca^{2+} has the cation occupying the off-plane site. The discrimination between the two purine bases by the divalent cation could arise due to higher ionic charge as well as polarization effects.

From Table 1, the coordination of cations at various sites in WC base pairs is enhanced as compared to individual bases. This appears to be due to increased charge concentration upon base pairing, which is reflected in the MESP topography (Fig. 1). Thus MESP seems to govern the relative binding affinities of various sites in base pairs. In some cases (site at_7 in AT accessed by Li^+ and gc_4 in GC accessed by Ca^{2+}), the interaction energy at the electrostatically docked geometries is < 2 kcal mol^{-1} ($< 2\%$) away from the value for the final optimized structure. There are other instances when ΔE_S and ΔE_F differ substantially owing to large changes in the geometrical orientation of the bases within the pair upon optimization. These structures, however, may not have much significance in biological systems. Geometry optimization could not be achieved for Ca^{2+} at the N^7 position of A in AT. This observation is consistent with the finding of Sponer *et al.*[5] resulting in the inability of alkaline earth cations to stabilize the AAT triplex.

Base pair stabilization due to cation binding is calculated according to Anwander *et al.*,[3] although the interaction energies are not corrected for basis set superposition error. The earlier calculations[3] are based upon optimization of a cation bound to a single base and the use of inter-base geometry determined by X-ray crystallography, though their trends generally agree with our results for GC. Thus, the largest stabilization occurs at the purine N^3 site in GC for both cations. For AT, the present full optimization at higher level does not lead to binding of Ca^{2+} at N^7 and N^3 positions (the base pair is split) which may be interpreted as a destabilization effect. The AT pair is stabilized due to the cation only if it binds near the O^2 of thymine, reports[3] also predict this site to be the most stabilizing.

The outcome of this work is that the strength of MESP at CPs can be meaningfully employed for predicting the sites of cation coordination in bases and base pairs as well as the respective binding energies. In general, Li^+ and Ca^{2+} do not prefer to occupy amino-nitrogen lone pair sites with the exception of G–Ca^{2+} complex. In all those structures where the base pair geometry does not alter much, the initial site predicted by electrostatics is very close to the final optimized one. In conclusion, electrostatics may be used as a powerful tool for a qualitative and semiquantitative prediction of cation coordination sites in DNA bases and base pairs.

The financial assistance by Council of Scientific and Industrial Research [CSIR 80(0022) 96-EMR-II and 9/137 (271)/95 EMR-I, New Delhi] is gratefully acknowledged by S. R. G. and S. S. P. A. C. L. and A. P. R. thank Fujitsu, Japan for financial assistance.

Notes and References

* E-mail: gadre@parcom.ernet.in

1 J. V. Burda, J. Sponer and P. Hobza, *J. Phys. Chem*., 1996, **100**, 7275.
2 M. Langlais, H. A. Tajmir-Riahi and R. Savoie, *Biopolymers*, 1990, **30**, 743.
3 E. H. S. Anwander, M. M. Probst and B. M. Rode, *Biopolymers*, 1990, **29**, 757.
4 V. N. Potaman and V. N. Soyfer, *J. Biomol. Struct. Dynam*., 1994, **11**, 1035.
5 J. Sponer, J. V. Burda, J. Leszczynski and P. Hobza, *J. Biomol. Struct. Dynam*., 1997, **14**, 613.
6 R. B. Martin, *Acc. Chem. Res*., 1985, **18**, 32.
7 P. Politzer and D. G. Truhlar, *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, Plenum, New York, 1981; J. Tomasi, R. Bonaccorsi and R. Cammi, in *Theoretical Methods of Chemical Bonding*, ed. Z. B. Maksic, Springer, 1990, vol. 3.
8 (*a*) S. R. Gadre, S. A. Kulkarni and I. H. Shrivastava, *J. Chem. Phys*., 1992, **96**, 5253; (*b*) R. N. Shirsat, S. V. Bapat and S. R. Gadre, *Chem. Phys. Lett*., 1992, **200**, 373; (*c*) S. R. Gadre and S. S. Pundlik, *J. Am. Chem. Soc*., 1995, **117**, 9559; (*d*) S. R. Gadre, P. Bhadane, S. S. Pundlik and S. S. Pingale, in *Molecular Electrostatic Potentials: Theory and Applications*, ed. J. S. Murray and K. D. Sen, Elsevier, Amsterdam, 1996, vol. 3.
9 J. Sponer, J. Leszczynski and P. Hobza, *J. Phys. Chem*., 1996, **100**, 1965. The optimized planar structures of WC base pairs are kindly made available by Dr Sponer.
10 S. R. Gadre and S. S. Pundlik, *J. Phys. Chem*., 1997, **101**, 3298.
11 See S. P. Gejji, C. H. Suresh, L. J. Bartolotti and S. R. Gadre, *J. Phys. Chem*., 1997, **101**, 5678 and references therein for a discussion of electrostatic docking.

Received in Cambridge, UK, 20th November 1997; 7/08372F