ARTICLE

OBC
www.rsc.org/obc

Modelling protein–RNA interactions: an electron density study of the guanidinium and formate complexes with RNA bases†

Isabel Rozas,*[a] Ibon Alkorta[b] and José Elguero[b]

[a] *Department of Chemistry, Trinity College Dublin, Dublin 2, Ireland. E-mail: rozasi@tcd.ie; Fax: 353 1671 2826; Tel: 353 1608 3731*

[b] *Instituto de Química Médica, CSIC, Juan de la Cierva 3, E-28006, Madrid, Spain*

Received 13th October 2004, Accepted 15th November 2004
First published as an Advance Article on the web 15th December 2004

The complexes formed by the double interaction established between RNA bases and guanidinium and formate ions, as a model for the interacting groups of arginine and glutamic or aspartic amino acid side chains, have been theoretically studied. A density functional theory method (B3LYP/6-31 + G**) has been used for this study. The range of interaction energies obtained allowed for a distinction between bidentate and bifurcate hydrogen bond interactions. The analysis of the electron density and the natural bond orbital analysis shows that these complexes are bound by double hydrogen bonds established between the donor and acceptor groups of guanidinium and formate respectively and those of the RNA bases. Comparisons are made with the results obtained in some previous theoretical and experimental studies.

Introduction

This article presents the continuation of our previously reported[1] research on the characteristics of the double hydrogen bonds (HBs) formed between small molecules mimicking amino acid side chains and RNA bases. The multiple functions performed by RNA (*e.g.* gene replication and expression) involve the interaction of this macromolecule with proteins. Hence, the study of RNA recognition by peptides has become an extremely important subject. Different reviews, analyses and computational studies on protein–RNA interactions have been reported in the literature.[2–6]

Hydrogen bonding[7] is the most important interaction established between RNA bases and amino acids and, particularly, it has been suggested that glutamic acid (Glu) and arginine (Arg) form a substantial amount of HB contacts with RNA bases.[8] In the work of Hermann,[4] it is proposed that RNA bases interact with the side chains of the amino acid residues of proteins, although the main interaction between both macromolecules occurs through the phosphate linkages of RNA. These phosphate–amino acid interactions have been thoroughly studied at a theoretical level by Leszczynski *et al*.[9] Our group has also performed a theoretical study of the complexes formed by HB interactions between a protein backbone model and nucleic acid bases.[10]

In the more recent works of Cheng *et al.*[6] the actual HBs between the RNA bases and some groups representing proteins (functional groups of the side chains of amino acids and the peptide bond) were analyzed. In these studies, the possible arrangements between amino acids bonding by HB to unpaired bases have been calculated. In their last work they found 21 possible complexes that involve bidentate HBs to the four unpaired bases.[6a]

The inherent donor–acceptor arrangements of nucleic bases give rise to many possible interactions. If one looks at the ability of a small molecule simulating the amino acid side chain to form bidentate HBs three types of ligands can be identified. Those with donor–donor HB groups such as guanidinium (as in Arg), those with acceptor–acceptor HB groups such as carboxylate (as in Glu and Asp) and those with donor–acceptor HB groups such as formamide (as in Gln and Asn). In a previous study[1] we examined the last case and analysed formic acid and formamide, which present a donor–acceptor profile. In the present study, complexes with the first two types of small ligands (donor–donor (guanidinium cation) and acceptor–acceptor (formate anion)) are explored.

As starting points we chose all those possible donor–acceptor interactions between both the donor–donor and acceptor–acceptor ligands (guanidinium and formate ions) and the four RNA bases (adenine, cytosine, guanine and uracil). In the present study, we have explored all the possible positions for the formation of bidentate HB in the four bases (as donor or as acceptor). Uracil has no adjacent donor–donor positions; therefore, no possible bidentate complexes with formate ions can be formed. However, carboxylic O atoms could be considered, in principle, as acceptor–acceptor groups by themselves, owing to their two lone pairs. This type of interaction has been described previously for complexes of urea.[11] Thus, cytosine and uracil can establish interactions with the guanidinium cation by means of their =O groups in positions 2 and 4. Adenine only has a donor–donor position at the NH_2 group in position 4 of the ring and no adjacent acceptor–acceptor positions; therefore, only complexes with formate are possible. Looking at the small molecule mimicking the amino acid, the guanidinium cation has two possible donor–donor approaches. Thus, this cation could form bidentate HBs by interacting with the two H atoms of one of the NH_2 groups or with two H atoms belonging to two different NH_2 groups. Both approaches have been considered.

Experimental

The geometries of all the monomers and complexes have been fully optimised with the program Gaussian-98[12] using the hybrid method Becke3LYP[13] with the 6-31 + G**[14] basis set. In all cases the nature of the complexes as a potential energy minimum has been established by verifying that all the corresponding frequencies were positive.

Interaction energies, E_Is, have been calculated as the difference in the total energy of the complex and the sum of the isolated monomers and have been corrected for basis set superposition error (BSSE) using the counterpoise method.[15]

The topological properties of the electron charge density have been characterized using the atoms in molecules (AIM)[16] methodology. An electron density of 0.001 e a.u.$^{-3}$ has been used to define the atomic volume. Using the AIM formalism, we have located the bond critical points (*i.e.*, points where the

† This article is part 2 of the same study. For part 1 see ref. 1.

DOI: 10.1039/b415768k

electron density function, ρ(bcp), is minimum along the bond path and maximum in the other directions) because the values of the charge density and its Laplacian at these critical points give useful information regarding the strength of the linkages. The Laplacian of the density, $\nabla^2\rho$(bcp), identifies regions of space wherein the electronic charge is locally depleted [$\nabla^2\rho$(bcp) > 0] or built up [$\nabla^2\rho$(bcp) < 0]. The former situation is typically associated with interactions between closed shell systems (*e.g.* ionic bonds, hydrogen bonds, and van der Waals molecules (VDW)), whereas the latter characterizes covalent bonds where the electron density is concentrated in the internuclear region.

The natural bond orbital (NBO) analysis, implemented in Gaussian-98,[17] was used to determine the nature of the interactions in the formation of the complexes.

The nomenclature of the complexes is expressed by the letter corresponding to the RNA base, then the guanidinium cation (CN_3H_6) or formate anion (HCO_2) and finally the number of the RNA bases atoms that interact with the guanidinium or formate ions. Thus, for example, the purine base complex C:HCO_2(4/5) will be formed between one of the H atoms of the NH_2 in position 4 of cytosine and one of the O atoms of formate and between the C–H of cytosine in position 5 and the other O atom of formate (for numeration of purine and pyrimidine rings, see Fig. 1).

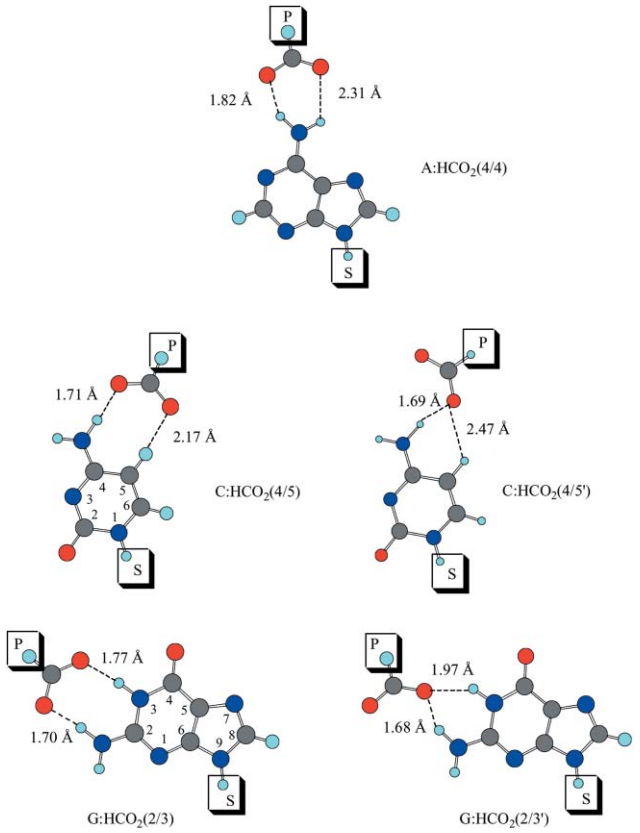

Fig. 1 B3LYP/6-31 + G** optimised structures of the RNA bases–formate complexes. The position of the sugar (**S**) in the RNA nucleoside and the rest of the amino acid (**P**) in the formate moiety are indicated.

Results and discussion

All of the possible HB complexes between the purine and pyrimidine RNA bases and guanidinium and formate were optimised and their structures of minimum energy are presented in Figs. 1 and 2 respectively. These complexes are formed by a double interaction between both monomers since guanidinium and formate exhibit two HB donor (N–H) and two HB acceptor (=O) groups, respectively, and RNA bases can also act as HB donors or acceptors depending on the orientation of the molecule.

In the case of adenine, there is only a possible donor–donor position at the NH_2 group in position 4, therefore only complex A:HCO_2(4/4) has been explored. The distances and angles between the carboxylate O atoms and the amino protons are in agreement with the possible formation of double HBs that are not equal (see Fig. 1).

Cytosine shows two possible donor–donor positions, the two protons of the NH_2 group in 4 and one of these protons and the C–H group in position 5. Protons belonging to C–H groups have shown to be able to establish weak HB interactions[18] and, for that reason, should be taken into account. When considering the interaction between formate and the NH_2 group at 4, we found that the formate molecule rotates and only one of the carboxylate O atoms remains near one of the amino protons and to the C–H group (complex C:HCO_2(4/5′) in Fig. 1), possibly forming a type of bifurcated HB. In a previous article[19] we studied other type of bifurcated HB, named "three-centred interactions", in which one H atom was binding, simultaneously, two HB acceptors. In the present case one HB acceptor (a carboxylate O atom) is interacting at the same time and asymmetrically with two H atoms. Thus, the O atom of the carboxylate is nearer one of the H of the amino group in 4 than of the CH very likely because amino groups are better HB donors than CH. The distance and angle between the amino H atom and the carboxylate C–O group is in agreement with a HB, while between the =O and the C–H is almost in the limit for a HB.

Finally, the formate complexes with guanine yielded complexes G:HCO_2(2/3) and G:HCO_2(2/3′) (see Fig. 1). Guanine can form complexes with an acceptor–acceptor ligand only through the amino group in 2 and the NH in 3. Even though there is an NH and a CH in positions 9 and 8 of the purine ring, they could not be involved in an interaction with formate since position 9 would be occupied by the sugar in the real RNA. We explored the possible interactions between the carboxylate group and the two H atoms of the amino group in 2 but, as happened with cytosine, the formate molecule rotated forming a bifurcated HB between one of the carboxylate O and one H from the amino group in 2 and the NH in 3 (complex G:HCO_2(2/3′)). In this case, the O atom is nearer to the H of the amino group in 2 than to the NH in 3, probably due to secondary interactions[20] between the other O atom of the formate and the primary amine. The H · · · O distances and N–H · · · O angles obtained for both complexes are in agreement with the formation of HB interactions.

In general, we found the same complexes with the carboxylate group that Cheng *et al.* reported in their recent article,[6a] however they did not identify the bifurcated complexes nor the bidentate complex with cytosine C:HCO_2(4/5) represented in Fig. 1, which are all energy minima.

Regarding the complexes with the donor–donor ligand, the guanidinium cation, only complexes with cytosine, guanine and uracil could be formed (Fig. 2). As previously mentioned adenine residues do not show two contiguous acceptor–acceptor positions; therefore, no possible complex with guanidinium can be formed. In this kind of interaction another point should be taken into account. It has been proven in different theoretical studies that the structure of the guanidinium cation is better described not as a planar structure but as a propeller that is interchanging between two mirror image structures passing through a planar transition state with a very low cost of energy.[21] In the present study, all the energy minimum complexes obtained show the guanidinium cation as a non planar structure.

Cytosine exhibits two acceptor–acceptor positions, the O atom in position 2 and the N atom in position 3 of the ring, and the guanidinium cation shows two possible donor–donor approaches as mentioned before. Thus, complex C:CN_3H_6(2/3) is formed by the interaction of the H atoms of two different amino groups of guanidinium and the distances and angles obtained for this bidentate interaction are in accordance to the

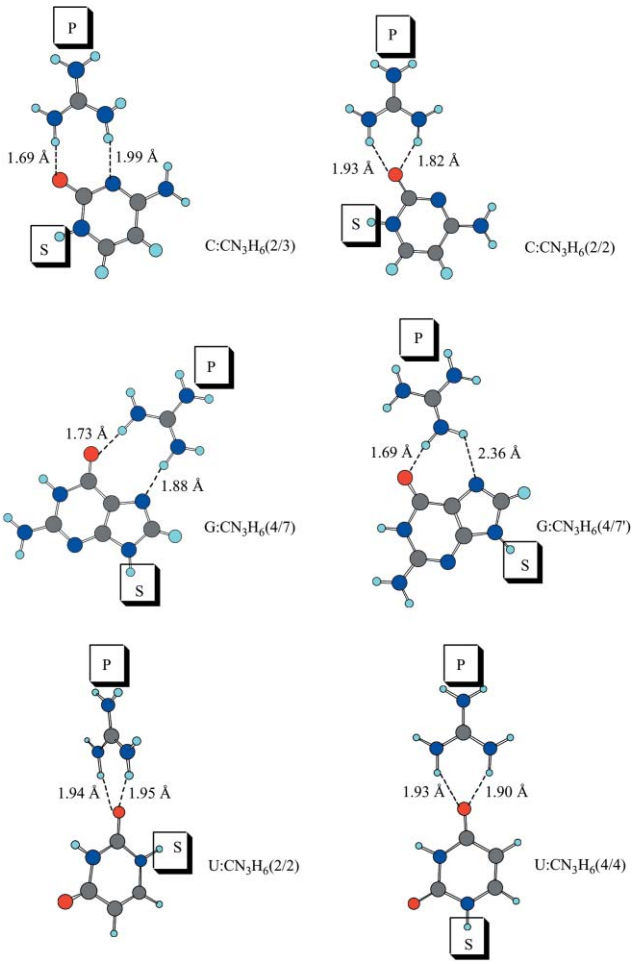

Fig. 2 B3LYP/6-31 + G** optimised structures of the RNA bases–Guanidinium complexes. The position of the sugar (**S**) in the RNA nucleoside and the rest of the amino acid (**P**) in the guanidinium moiety are indicated.

formation of HBs. When trying to obtain the corresponding cytosine complex with the two H atoms of one of the amino groups of guanidinium, the cation turned forming a bifurcated interaction between two H atoms of two different amino groups and the O atom in position 2 (see complex $C:CN_3H_6(2/2)$ in Fig. 2). This was one of the complexes we were expecting by taking the carbonyl group as a acceptor–acceptor group and, as in the previous bifurcated complex, these HBs are not symmetrical. Again, a possible explanation could be the existence of secondary interactions attractive in the case of the N in position 3 of the ring and repulsive in the case of the NH in position 1.[20] The distances and angles obtained for this bifurcated interaction are in agreement with a HB.

Guanine presents also two acceptor–acceptor positions, one at the O atom in 4 and the other between the O in 4 and the N atom in position 7 of the ring. Again, considering the two possible approaches of the guanidinium cation we obtained the bidentate complexes $G:CN_3H_6(4/7)$ and $G:CN_3H_6(4/7')$, the interaction distances and angles of which would correspond to a HB (Fig. 2). When trying to obtain the corresponding bifurcated complex between guanidinium and the carbonyl O in position 4 of guanine, the cation rotated yielding complex $G:CN_3H_6(4/7)$.

Uracil shows two carbonyl groups in positions 2 and 4 of the ring and then both O atoms can act as acceptor–acceptor groups independently. Approaching the guanidinium cation to both positions yielded the corresponding $U:CN_3H_6(2/2)$ and $U:CN_3H_6(4/4)$ complexes (Fig. 2). In complex $U:CN_3H_6(2/2)$ the guanidinium cation has rotated around itself and is interacting with the =O in a perpendicular way. In both complexes the

Table 1 Total energies (E_T, hartrees) and interaction energies including BSSE corrections (BSSE, E_{I+BSSE}, kcal mol^{-1}) for the complexes formed between RNA bases, guanidinium and formate calculated at B3LYP/6-31 + G** computational level

	E_T	BSSE	E_{I+BSSE}
$A:HCO_2(4/4)$	−656.6051997	0.52	−19.10
$C:HCO_2(4/5)$	−584.2325653	0.67	−29.85
$C:HCO_2(4/5')$	−584.2295157	0.59	−28.02
$G:HCO_2(2/3)$	−731.8717064	0.82	−36.95
$G:HCO_2(2/3')$	−731.8676198	0.71	−34.49
$C:CN_3H_6(2/3)$	−600.8066625	0.77	−35.04
$C:CN_3H_6(2/2)$	−600.8054817	0.69	−34.38
$G:CN_3H_6(4/7)$	−748.4384728	0.71	−37.75
$G:CN_3H_6(4/7')$	−748.4313835	0.57	−33.44
$U:CN_3H_6(2/2)$	−620.6658724	0.65	−19.49
$U:CN_3H_6(4/4)$	−620.6714240	0.55	−23.07

distances and angles obtained between the H atoms and the HB acceptors are those expected for a HB interaction.

All the bidentate complexes computationally found with the guanidinium cation are in agreement with those reported by Cheng and Frankel.[6a] However, we have found complexes with bifurcated interactions (some of which are energetically very similar to the corresponding bidentated complexes) that have not been reported before.

The BSSE corrected interaction energies obtained for all these complexes are gathered in Table 1. All of them exhibit large interaction energies in agreement with a dual type of interaction, ionic (due to the charged small ligands) and bidentated or bifurcated HBs.

In most cases, the interactions with the formate anion seem to be slightly weaker than those with the guanidinium cation. Thus, when considering cytosine we found E_{I+BSSE}s between −28 and −30 kcal mol^{-1} for formate, while with guanidinium these E_{I+BSSE}s are around −35 kcal mol^{-1}. In the case of the guanine complexes with formate, the interaction energies range from 35 to 37 kcal mol^{-1}, whereas with guanidinium these energies are between 34 and 38 kcal mol^{-1}.

In general, the interaction energies of the bidentate donor–donor–acceptor–acceptor complexes are larger than those corresponding to the complexes with bifurcated HBs such as $C:HCO_2(4/5')$, $C:HCO_2(2/3')$, $C:CN_3H_6(2/2)$, $U:CN_3H_6(2/2)$ and $U:CN_3H_6(4/4)$. Thus, when comparing bifurcated and bidentate HBs in the same systems we observed that the E_{I+BSSE} of the bifurcated are smaller than those of the bidentate. Thus, the E_{I+BSSE} of the bifurcated complex $C:HCO_2(4/5')$ is around 2 kcal mol^{-1} weaker that the corresponding bidentate system $C:HCO_2(4/5)$; the interaction energy of $G:HCO_2(2/3')$ is 2.5 kcal mol^{-1} weaker than that of $G:HCO_2(2/3)$, and the interaction in $C:CN_3H_6(2/2)$ is 0.6 kcal mol^{-1} weaker than that in $C:CN_3H_6(2/3)$.

The most stable complex is $G:CN_3H_6(4/7)$ as predicted by Seeman *et al.*,[3] in agreement with the findings of Allers and Shamoo[5] and Luscombe *et al.*[22] who concluded that this interaction is one of those most frequently found in their studies, accounting for 43% of all double interactions. Following on in stability are $G:HCO_2(2/3)$, $C:CN_3H_6(2/3)$, $G:HCO_2(2/3')$, $C:CN_3H_6(2/2)$, $G:CN_3H_6(4/7')$, both complexes between cytosine and formate in positions 4 and 5 and finally the complex with adenine is the weakest. It is particularly interesting that the order in interaction strength corresponds to the same order of the dipole moment of the isolated bases (A: 2.45, C: 6.89 and G: 7.07 D, respectively). Considering that the interactions studied here are established between charged and neutral species, the corresponding charge–dipole contribution should be rather important in justifying the rather weak binding to adenine.

The energy order that we obtain does not correspond totally with that computed by Cheng and Frankel,[6a] maybe because neither the structures nor the computational methods used are

Table 2 Electron density (a.u.) and Laplacian at the bcp of the HB found in the complexes optimised at B3LYP/6-31 + G** level. The HB distances (Å) are also shown

		ρ(bcp)	$\nabla^2\rho$(bcp)	d(X···H)
A:HCO_2(4/4)	O···HN	0.0360	0.0956	1.82
	O···HN	0.0144	0.0470	2.31
C:HCO_2(4/5)	O···HN	0.0451	0.1230	1.71
	O···HC	0.0181	0.0469	2.17
C:HCO_2(4/5′)	O···HN	0.0486	0.1279	1.69
	O···HC	0.0087	0.0334	2.47
G:HCO_2(2/3)	O···HN	0.0467	0.1229	1.70
	O···HN	0.0384	0.1047	1.77
G:HCO_2(2/3′)	O···HN	0.0489	0.1359	1.68
	O···HN	0.0228	0.0763	1.97
C:CN_3H_6(2/3)	NH···O	0.0447	0.1342	1.69
	NH···N	0.0275	0.0652	1.99
C:CN_3H_6(2/2)	NH···O	0.0357	0.1021	1.82
	NH···O	0.0234	0.0844	1.93
G:CN_3H_6(4/7)	NH···O	0.0460	0.1311	1.70
	NH···N	0.0158	0.0488	2.28
G:CN_3H_6(4/7′)	NH···O	0.0399	0.1230	1.73
	NH···N	0.0346	0.0829	1.88
U:CN_3H_6(2/2)	NH···O	0.0239	0.0788	1.94
	NH···O	0.0234	0.0776	1.95
U:CN_3H_6(4/4)	NH···O	0.0274	0.0855	1.90
	NH···O	0.0251	0.0811	1.93

Table 3 Charge transfer (e), and orbital interaction energy (kcal mol^{-1}) calculated at the B3LYP/6-31 + G** level with the NBO method

	Charge transfer	E(2)
A:HCO_2(4/4)	−0.075	N–H···O 17.32
		N–H···O 2.45
C:HCO_2(4/5)	−0.106	N–H···O 26.85
		C–H···O 5.69
C:HCO_2(4/5′)	−0.097	N–H···O 29.37
		C–H···O 1.35
G:HCO_2(2/3)	−0.151	N–H···O 21.23
		N–H···O 27.74
G:HCO_2(2/3′)	−0.110	N–H···O 27.07
		N–H···O 10.19
C:CN_3H_6(2/3)	0.114	N–H···O 21.63
		N–H···N 16.21
C:CN_3H_6(2/2)	0.074	N–H···O 14.70
		N–H···O 9.67
G:CN_3H_6(4/7)	0.113	N–H···O 17.88
		N–H···N 23.61
G:CN_3H_6(4/7′)	0.082	N–H···O 24.42
		N–H···N 4.95
U:CN_3H_6(2/2)	0.046	N–H···O 7.15
		N–H···O 7.08
U:CN_3H_6(4/4)	0.059	N–H···O 8.25
		N–H···O 7.76

the same. However, the most stable complexes they obtained with aspartic acid and arginine with guanine correspond to our most stable complexes, even though in a different order.

Positions 2 and 3 of guanine have been described to be simultaneously interacted with by glutamic or aspartic acid, but these contacts are less frequent than those between positions 6 and 7 of guanine and arginine[5] which is in agreement with the stability order we found. We have found a complex between cytosine and formate involving a C–H group as a HB donor [C:HCO_2(4/5)]. Even though this RNA base–amino acid interaction has not been previously reported and, in general, C–H···X are weak interactions, the total interaction energy is larger than that obtained for the adenine complex. The complexes of adenine with formate and uracil with guanidinium are the weakest of those studied here and correspond to a bidentate complex with an amino group and also to two bifurcated complexes with the lone pairs of a carbonyl O atom. Contacts between the NH_2 group in position 4 of adenine and C=O are frequently found; however, the specific contacts with glutamic or aspartic acid are very rare. In the case of uracil, contacts between both =O atoms and arginine or NH groups are relatively frequent.[5]

Regarding the AIM analysis of these complexes (summarized in Table 2) a bond critical point (bcp) was found in all the cases between the atoms involved in a possible HB. The topological characteristics of these bcps (electron density and its Laplacian) correspond to those of "closed-shell" interactions and more particularly to HBs (ρ(bcp) around 10^{-2} a.u. and positive Laplacians). Overall, these topological characteristics are in agreement with strong HBs except in the case of the C–H···O interactions, found in the C:HCO_2(4/5) and C:HCO_2(4/5′) complexes, with ρ(bcp) around 18 and 9×10^{-3} a.u., more in accordance with a weak HBs. This is also reflected in the smaller interaction energies obtained for those complexes (see Table 1).

The usual logarithmic correlation was found between the electron density at the bcp and the HB distance: $d(\mathrm{H}\cdots\mathrm{X}) = -0.475\ \ln[\rho(\mathrm{bcp})] + 0.226$, $R^2 = 0.96$, $n = 22$.[23] If we consider only those complexes with =O as a HB acceptor the resulting correlation improves: $d(\mathrm{H}\cdots\mathrm{X}) = -0.465\ \ln[\rho(\mathrm{bcp})] + 0.253$, $R^2 = 0.98$, $n = 19$. When considering only the bidentate HBs, on one hand $\{d(\mathrm{H}\cdots\mathrm{X}) = -0.537\ \ln[\rho(\mathrm{bcp})] + 0.039$, $R^2 = 0.99$, $n = 12\}$, and the bifurcate HBs on the other $\{d(\mathrm{H}\cdots\mathrm{X}) = -0.441\ \ln[\rho(\mathrm{bcp})] + 0.321$, $R^2 = 0.98$, $n = 10\}$ the correlations also improve. This fact indicates differences in the nature of these interactions. As we have previously shown, bifurcated HBs are energetically weaker interactions than two regular HBs, and consequently, the HB distances become longer.[18]

The analysis of the orbital interactions by means of the NBO approach provides two important indicators of the strength and nature of the interactions established, the charge transferred and the orbital interaction energy, E(2). Both parameters have been computed and the results are shown in Table 3.

In terms of the charge, in the case of the formate anion complexes there is a transfer from such an ion to the neutral RNA base. This transfer is larger than that between neutral molecules and it amounts from −0.07 to −0.15 e. In the case of the guanidinium ion, the transfer occurs from the neutral base to the cation and is larger than that between neutral species, ranging from 0.05 to 0.11 e. These large charge transfers are in agreement with the ionic nature of one of the species involved and will confirm a dual nature in the intermolecular interaction in these complexes with a HB component and an ionic one. In all cases the charge transferred within the complex is larger in the case of bidentate complexes than in the case of bifurcated ones, which is in agreement with the interaction energies obtained. The larger the interaction energy the larger the charge transferred.

Upon analysing the actual interactions it was found that all of them occur between the lone pair of the HB donor (=O or =N) and one unoccupied (N/C)–H molecular orbital. The largest E(2)s obtained correspond to the shortest HBs indicating a strong interaction between the two species of the complex. This is the case with the bidentate complexes which displays the largest interaction energies, G:HCO_2(2/3) and G:CN_3H_6(4/7). In the case of complexes with bifurcated HBs (C:HCO_2(4/5′), G:HCO_2(2/3′), C:CN_3H_6(2/2), U:CN_3H_6(2/2) and U:CN_3H_6(4/4)), the orbital interaction energies are larger for one of the interactions than for the other, in agreement with the asymmetry of these bifurcated HBs.

We have computed the atomic charges of the atoms involved in the HB interactions for the bidentate complexes, by both AIM and NBO methods, to compare the performance of these methods. The results are shown in Table 4. In general, the atomic charges calculated by the AIM approach are larger than those computed by using the NBO method except for the atomic charge of the C atom involved in a C–H···O bond in complex C:HCO_2(4/5) that is much smaller. This difference was extremely noticeable for the N and O atoms and it was minimum in the case of H atoms. For these H atoms it was possible to find a good correlation between both sets of charges (Q[NBO] = 0.415

Table 4 Atomic charges (*e*) calculated at the B3LYP/6-31 + G** level with the AIM and NBO methods for the bidentated complexes

	X···H	*Q*[AIM]	*Q*[NBO]
A:HCO_2(4/4)	**N**–(H)	−1.319	−0.795
	(N)–**H**	0.548	0.473
	(H)···**O**	−1.281	−0.797
	(N)–**H**	0.484	0.458
	(H)···**O**	−1.290	−0.793
C:HCO_2 (4/5)	**N**–(H)	−1.311	−0.823
	(N)–**H**	0.559	0.471
	(H)···**O**	−1.283	−0.804
	C–(H)	−0.010	−0.379
	(C)–**H**	0.147	0.305
	(H)···**O**	−1.286	−0.773
G:HCO_2 (2/3)	**N**–(H)	−1.240	−0.674
	(N)–**H**	0.550	0.487
	(H)···**O**	−1.269	−0.768
	N–(H)	−1.294	−0.862
	(N)–**H**	0.558	0.465
	(H)···**O**	−1.275	−0.785
C:CN_3H_6(2/3)	**N**–(H)	−1.295	−0.829
	(N)–**H**	0.569	0.483
	(H)···**O**	−1.234	−0.691
	N–(H)	−1.274	−0.850
	(N)–**H**	0.529	0.467
	(H)···**N**	−1.205	−0.646
G:CN_3H_6(4/7)	**N**–(H)	−1.289	−0.833
	(N)–**H**	0.562	0.482
	(H)···**O**	−1.221	−0.675
	N–(H)	−1.286	−0.841
	(N)–**H**	0.544	0.474
	(H)···**N**	−1.207	−0.522
G:CN_3H_6(4/7′)	**N**–(H)	−1.326	−0.821
	(N)–**H**	0.556	0.473
	(H)···**O**	−1.221	−0.682
	(N)–**H**	0.498	0.465
	(H)···**N**	−1.189	−0.512

Q[AIM] + 0.247, $R^2 = 0.98$). This trend has been previously observed by Wiberg and Rablen[24] and in our previous article,[1] concluding that the AIM charges of electronegative atoms are much larger than those calculated with the NBO approximation. The AIM charges can be quite unrealistic as was previously shown;[24] however, we include them for the sake of comparison, since they provide a generic view of the tendency of certain atoms to accept or donate HBs.

As in our previous article,[1] we have found a good correlation between the electron density at the bcp of the HB interactions and the corresponding orbital interaction energy. This correlation seems coherent since the ρ(bcp) gives an idea of the strength of the interaction established whereas the $E(2)$ indicates the nature of the bond studied. Thus, when comparing both sets of data in the complexes studied the following equation was found: $E(2) = 708.71\ \rho(\text{bcp}) - 7.27$, $R^2 = 0.92$, $n = 22$. Again, the nature of the HB acceptor seems to be more important that the nature of the HB donor and, thus, the correlation improved when considering only those interactions where the HB acceptor is an O atom: $E(2) = 716.99\ \rho(\text{bcp}) - 8.14$, $R^2 = 0.95$, $n = 19$. Correlations between these two descriptors have not previously been reported elsewhere.

Conclusions

All of the possible complexes formed between three of the RNA bases and the HB donor–donor and acceptor–acceptor ions, guanidinium and formate, by means of double interactions have been optimised at B3LYP/6-31 + G** level. The topological characteristics of the electron density at the bcp have been evaluated, as well as the nature of the interactions by using the NBO approach. All these studies have confirmed that the interactions found in these complexes are HBs of medium strength, established by the interaction between a lone pair of the HB acceptor and an unoccupied bond orbital of the HB donor.

The charge transfer observed in these RNA bases–guanidinium/formate complexes is large in agreement with the high interaction energies computed for these complexes. This could be explained by the ionic nature of one of the molecules involved in the complexes which would reinforce the intermolecular binding forming charge-assisted hydrogen bonds.[25]

Geometry, energy, electron density and natural bond orbital results show that the G:CN_2H_6(4/7) and G:HCO_2(2/3) complexes are the most stable and this is in agreement with the experimental observation showing that the HB contacts found with highest frequency are those between guanine and arginine and glutamic acid. Cytosine shows a medium frequency distribution of HB contacts with glutamic acid, but its complexes with formate are low in stability. Our computations have yielded RNA bases–formate and guanidinium complexes involving a type of bifurcated HB in which =O atoms act as acceptor–acceptor groups. Some of these complexes are energetically equivalent to the corresponding bidentate ones and have not been reported previously.

In addition, we have found cytosine–guanidinium complexes involving a C–H group as a HB donor. This kind of RNA base–amino acid interaction has not been studied previously but, even though weaker than N–H···X interactions, should not been ignored since it provides rather stable complexes. Some bifurcated complexes have been found which show lower stability than the corresponding bidentated ones.

Good correlations have been found between the logarithm of the electron density at the bcp and the HB distance, as well as between the ρ(bcp) and the orbital interaction energy $E(2)$. The correlation between electron density and orbital interaction energy that was first reported by our group in a previous article is a very relevant finding since both parameters provide information on the nature and strength of a bond but from very different sources.

Acknowledgements

The authors thank the University of Dublin, Trinity College (Start-up Project No. 504/R01056) and the Spanish DGICYT (Project No. BQU-2000-0906) for financial support.

References

1 I. Rozas, I. Alkorta and J. Elguero, *J. Phys. Chem. B*, 2004, **108**, 3335.
2 D. E. Draper, *J. Mol. Biol.*, 1999, **293**, 255.
3 N. C. Seeman, J. M. Rosenberg and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1976, **73**, 804.
4 T. Hermann, *Angew. Chem., Int. Ed.*, 2000, **39**, 1890.
5 J. Allers and Y. Shamoo, *J. Mol. Biol.*, 2001, **311**, 75.
6 (*a*) A. C. Cheng and A. D. Frankel, *J. Am. Chem. Soc.*, 2004, **126**, 434; (*b*) A. C. Cheng, W. W. Chen, C. N. Fuhrmann and A. D. Frankel, *J. Mol. Biol.*, 2003, **327**, 781.
7 M. E. Vazquez, A. M. Caamaño and J. L. Mascareñas, *Chem. Soc. Rev.*, 2003, **32**, 338.
8 S. Jones, D. T. A. Daley, N. M. Luscombe, H. M. Berman and J. M. Thornton, *Nucleic Acids Res.*, 2001, **29**, 943.
9 (*a*) V. V. Murashov and J. Leszczynski, *J. Phys. Chem. B*, 1999, **103**, 8391; (*b*) A. Pelmenschikov, X. Yin and J. Leszczynski, *J. Phys. Chem. B*, 2000, **104**, 2148.
10 I. Alkorta and J. Elguero, *J. Phys. Chem. B*, 2003, **107**, 5306.
11 (*a*) A. Masunov and J. J. Dannenberg, *J Phys. Chem. A*, 1999, **103**, 178; (*b*) A. Masunov and J. J. Dannenberg, *J. Phys. Chem. A*, 2000, **104**, 806.
12 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul,

B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle and J. A. Pople, *GAUSSIAN 98 (Revision A.11)*, Gaussian, Inc., Pittsburgh, PA, 2001.
13 (*a*) A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648; (*b*) C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.
14 P. A. Hariharan and J. A. Pople, *Theor. Chim. Acta*, 1973, **28**, 213.
15 S. B. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553.
16 (*a*) R. F. W. Bader, *Atoms in Molecules. A Quantum Theory*, ed. J. S. Rowlinson, Oxford University, Oxford, 1994; (*b*) F. W. Bieger-Konig, R. F. W. Bader and T. H. Tang, *J. Comput. Chem.*, 1982, **3**, 317–328.
17 A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899.
18 I. Alkorta, N. Campillo, I. Rozas and J. Elguero, *J. Org. Chem.*, 1998, **63**, 7759.
19 I. Rozas, I. Alkorta and J. Elguero, *J. Phys. Chem. A*, 1998, **102**, 9925.
20 C. Alvarez-Rua, S. Garcia-Granda, S. Goswami, R. Mukherjee, S. Dey, R. M. Claramunt, M. D. Santa Maria, I. Rozas, N. Jagerovic, I. Alkorta and J. Elguero, *New J. Chem.*, 2004, **28**, 700.
21 (*a*) A. Gobbi and G. Frenking, *J. Am. Chem. Soc.*, 1993, **115**, 2362; (*b*) I. Rozas, P. E. Kruger, unpublished results, 2004.
22 N. M. Luscombe, R. A. Laskowski and J. M. Thorton, *Nucleic Acids Res.*, 2001, **29**, 2860.
23 (*a*) I. Alkorta, I. Rozas and J. Elguero, *Struct. Chem.*, 1998, **9**, 243; (*b*) I. Alkorta, L. Barrios, I. Rozas and J. Elguero, *THEOCHEM*, 2000, **496**, 131; (*c*) E. Espinosa, M. Souhassou, H. Lachekar and C. Lecomte, *Acta Crystallogr., Sect. B: Struct. Sci.*, 1999, **55**, 563; (*d*) O. Knop, K. N. Rankin and R. J. Boyd, *J. Phys. Chem. A*, 2001, **105**, 6552.
24 K. B. Wiberg and P. R. Rablen, *J. Comp. Chem.*, 1993, **14**, 1504.
25 (*a*) D. Braga, K. Rubini and L. Maini, *CrystEngComm*, 2004, **6**, 236; (*b*) V. Ferretti, V. Bertolasi and L. Pretto, *New J. Chem.*, 2004, **26**, 646.