

Quantum-Chemical Analysis of C–H...O and C–H...N Interactions in RNA Base Pairs – H–Bond Versus Anti–H–Bond Pattern

Maria Brandl, Michael Meyer and Jürgen Sühnel

E–Mail: mbrandl@imb-jena.de; mmeyer@imb-jena.de; jsuehnel@imb-jena.de

Introduction

Recently, the first examples of base pairs that are linked by only one direct **standard H–bond** and an additional C–H...O [1] or C–H...N [2] contact have been detected in crystal structures of RNA. There has been growing evidence that C–H...acceptor interactions can be considered as weak H–bonds [3]. Thus, observations of C–H bond length contractions upon formation of C–H...O and C–H... π contacts between small organic molecules were rather surprising. Showing characteristics that are contrary to classical H–bonds, such contacts have been classified as anti–H–bonds and claimed to be relevant to biopolymers [4].

In addition to the analysis of global base pair features, quantum chemical studies allow a detailed analysis of the electronic and structural properties of C–H...acceptor contacts. We have determined the equilibrium geometries and interaction energies including electron correlation for the UU (Calcutta) base pair (FIG. 1a) with a C–H...O contact and the AA base pair (FIG. 1c) with a C–H...N contact. To exclude the effects exerted by the neighboring standard hydrogen bonds, model systems including the base C–H donor and formaldehyde (FA) as an acceptor (FIGS. 1b,d) have been studied as well.

FIGURE 1: Chemical formulae of the UU base pair (a), uracil–formaldehyde complex (b), AA base pair (c), adenine–formaldehyde complex (d)

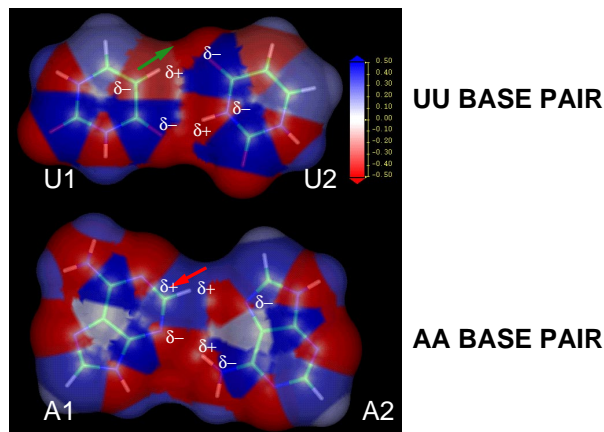
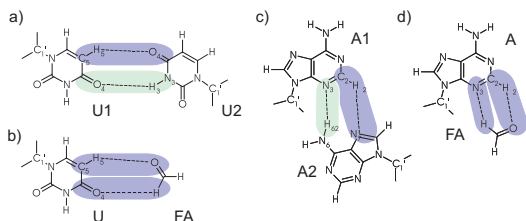


FIGURE 2: Distribution of Mulliken charges in the UU and AA base pairs obtained by MP2/6–31G(d,p)
 → : Stretching of the donor C–H bond
 ← : Compression of the donor C–H bond
 color ramp: Mulliken charges in atomic charge units

TABLE 5 reveals a steady decrease in interaction energy when passing from the UU complex with two standard H–bonds to the AFA system with two C–H...O/N anti–H–bonds. Nonetheless the AFA complex is still stabilized by a weak interaction energy of –1.71 kcal/mol.

ΔE^T [U1U2]	ΔE^T [A1A2]	ΔE^T [UFA]	ΔE^T [AFA]
–8.13	–5.39	–3.68	–1.71

C–H...O/N Interactions in Base Pairs and Model Systems

The optimized geometries (FIG. 2) of both the AA and UU base pairs are close to the ones observed within the respective RNA crystal structure and can thus be regarded as structurally autonomous building blocks of nucleic acids. With 2.19 and 2.41 Å (TAB. 1), both the C–H...O contact in UU and the C–H...N contact in AA, respectively, are shorter than the sums of van der Waals radii for the hydrogen and the acceptor.

Calcutta base pair (U1U2)		AA base pair (A1A2)	
d [H3(U2)...O4(U1)]	1.89	d [H62(A2)...N3(A1)]	2.07
d [N3(U2)...O4(U1)]	2.91	d [N6(A2)...N3(A1)]	3.08
α [N3(U2) H3(U2) O4(U1)]	171	α [N6(A2) H62(A2) N3(A1)]	179
d [H5(U1)...O4(U2)]	2.19	d [H2(A1)...N7(A2)]	2.41
d [C5(U1)...O4(U2)]	3.23	d [C2(A1)...N7(A2)]	3.39
α [C5(U1) H5(U1) O4(U2)]	150	α [C2(A1) H2(A1) N7(A2)]	162

Whereas the C–H bond length in the UU C–H...O contact increases and the wavenumber of its stretching vibration decreases upon formation of the complex (H–bond pattern), the C–H bond of the AA C–H...N interaction is shortened and its stretching vibration shifted to higher wavenumbers (anti–H–bond pattern, TAB. 2). This qualitatively different behaviour correlates with the charge distribution in the C–H donor groups shown in FIG. 2. The carbon involved in the UU C5–H5...O4 motif bears a **negative partial charge**, whereas the carbon in the AA C2–H2...N7 motif is **positively charged**.

U1U2 complex – isolated U		A1A2 complex – isolated A	
d [N3–H3(U2)] – d [N3–H3(U)]	+ 0.014	d [N6–H62(A2)] – d [N6–H6(A)]	+ 0.010
d [C5–H5(U1)] – d [C5–H5(U)]	+ 0.003	d [C2–H2(A1)] – d [C2–H2(A)]	– 0.002
ν [C5–H5(U1)] – ν [C5–H5(U)]	– 34	ν [C2–H2(A1)] – ν [C2–H2(A)]	+ 14

As can be seen in TABS. 3 and 4, the C2–H2 and C5–H5 donor groups of adenine and uracil show quantitatively the same behaviour when complexed with formaldehyde (FA) (FIGS. 1b, 1d). Interestingly, formaldehyde having a positive charge on its carbon, has also been observed to form interactions with an **anti–H–bond pattern**.

UFA complex		AFA complex	
d [H(FA)...O4(U)]	2.33	d [H(FA)...N3(A)]	2.46
d [C(FA)...O4(U)]	3.36	d [C(FA)...N3(A)]	3.40
α [C(FA) H(FA) O4(U)]	155	α [H(FA) C(FA) N3(A)]	143
d [H5(U)...O(FA)]	2.29	d [H2(A)...O(FA)]	2.48
d [C5(U)...O(FA)]	3.34	d [C2(A)...O(FA)]	3.43
α [C5(U) H5(U) O4(FA)]	163	α [H2(A) C2(A) O(FA)]	146

UFA complex – isolated U or FA		AFA complex – isolated A or FA	
d [C–H(FAUFA)] – d [C–H(FA)]	– 0.004	d [C–H(AAFA)] – d [C–H(FA)]	– 0.003
ν [C–H(FAUFA)] – ν [C–H(FA)]	+ 67	ν [C–H(AAFA)] – ν [C–H(FA)]	+ 48
d [C5–H5(UFA)] – d [C5–H5(U)]	+ 0.003	d [C2–H2(AFA)] – d [C2–H2(A)]	– 0.001
ν [C5–H5(UFA)] – ν [C5–H5(U)]	– 32	ν [C2–H2(AFA)] – ν [C2–H2(A)]	+ 15

Conclusions

- Both the Calcutta–UU and the AA base pair are structurally autonomous building blocks of nucleic acids.
- The C–H...O contact in UU exhibits the **H–bond pattern** upon base pair formation (increase of C–H bond length, C–H stretching frequency red shift), whereas the C–H...N contact in AA is of the **anti–H–bond type** (decrease of C–H bond length, C–H stretching frequency blue shift).
- In both the base/base and base/formaldehyde complexes C–H...O contacts with negatively charged donor carbons show the **H–bond pattern** whereas C–H...O/N contacts with positively charged carbons are **anti–H–bonds**.
- Both C–H...O H–bonds and C–H...O/N anti–H–bonds are attractive and need not be enforced by neighbouring interactions.

Methods

Only the base parts of the nucleotides have been used and ribose C1' carbons have been replaced by hydrogens. The geometries of the base pairs have been optimized ("Tight") with inclusion of electron correlation according to second order Møller–Plesset perturbational theory (MP2(FC); FC – frozen core) and the 6–31G(d,p) basis set. Interaction energies ΔE have been corrected for the basis set superposition error by the standard counterpoise method [5]. A further correction was done for the deformation energies (ΔE^{DEF}), which are defined as the energy differences between the structures of the isolated bases and the structures of the bases in the complex. The total interaction energy ΔE^T is thus given by $\Delta E^T = \Delta E + \Delta E^{DEF}$. For comparison, the DFT vibrational frequencies have been calculated at the B3LYP/6–31G(d,p) level. They have been scaled with the factor 0.9613. The interaction energies obtained with the DFT approach are in all cases similar to the corresponding MP2 energies. The charge distribution of the systems under study has been determined by means of a Mulliken population analysis. All quantum–chemical calculations have been performed using Gaussian'94 [6].

References

- Wahl, MC et al 1996. Nat Struct Biol 3:24.
- Brandl, M et al 1999. Theoret Chem Acc 101:103.
- Wahl, MC & Sundaralingam, M 1997. Trends Biochem Sci 22:97.
- Hobza, P et al 1998. J Phys Chem A 102:2501.
- Boys, SB & Bernardi, F 1970. J Mol Phys 17:553.
- Frisch, MJ et al Gaussian 94, revision E1. Gaussian, Inc., Pittsburgh.