THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Structural and Energetic Impact of Non-Natural 7-Deaza-8-Azaadenine and Its 7-Substituted Derivatives on H-Bonding Potential with Uracil in RNA Molecules

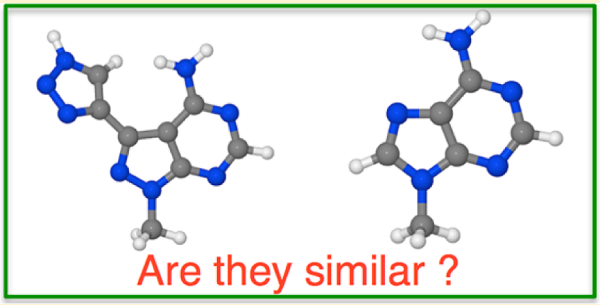
Mohit Chawla,[†] Raffaele Credendino,[†] Romina Oliva,[*,‡] and Luigi Cavallo[*,†]

[†]Physical Sciences and Engineering Division, Kaust Catalysis Center, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

[‡]Department of Sciences and Technologies, University Parthenope of Naples, Centro Direzionale Isola C4, I-80143, Naples, Italy

Supporting Information

ABSTRACT: Non-natural (synthetic) nucleobases, including 7-ethynyl- and 7-triazolyl-8-aza-7-deazaadenine, have been introduced in RNA molecules for targeted applications, and have been characterized experimentally. However, no theoretical characterization of the impact of these modifications on the structure and energetics of the corresponding H-bonded base pair is available. To fill this gap, we performed quantum mechanics calculations, starting with the analysis of the impact of the 8-aza-7-deaza modification of the adenine skeleton, and we moved then to analyze the impact of the specific substituents on the modified 8-aza-7-deazaadenine. Our analysis indicates that, despite of these severe structural modifications, the H-bonding properties of the modified base pair gratifyingly replicate those of the unmodified base pair. Similar behavior is predicted when the same skeleton modifications are applied to guanine when paired to cytosine. To stress further the H-bonding pairing in the modified adenine−uracil base pair, we explored the impact of strong electron donor and electron withdrawing substituents on the C7 position. Also in this case we found minimal impact on the base pair geometry and energy, confirming the validity of this modification strategy to functionalize RNAs without perturbing its stability and biological functionality.



INTRODUCTION

During the past two decades, the interest of biochemists and chemical biologists in RNA has progressively increased, as RNA molecules have proved to fulfill a variety of biological functions, including storage of information, regulation of gene expression, and catalysis.[1−6] This large functional repertoire is also due to the fact that, in addition to the four canonical nucleotides, adenine (A), uracil (U), guanine (G), and cytosine (C), naturally occurring RNAs may feature a variety of over 100 post-transcriptional modifications,[7] greatly enhancing their chemical information and functional capability. Modification of the nucleobases[8−11] and of the ribose−phosphate[12−15] backbone can have an impact on the stability, kinetics, and resistance to enzymatic degradation of nucleic acid molecules.[16,17] Recently, non-natural (synthetic) modifications have been introduced to complement the natural ones. Synthetically modified nucleosides are being used in numerous applications, such as expanding the genetic code, probing biological interactions and functions,[18] imparting favorable properties on small interfering RNAs (siRNAs),[19,20] and even creating a semisynthetic organism with increased potential for information storage and retrieval.[21,22] Thus, ribonucleoside non-natural modifications further extend the structural repertoire and, in turn, the functional properties of RNA molecules.[23]

In this scenario, a few azapurine and deazapurine analogues have been pointed out recently for several applications.[19,24−29] In particular, 7-deazapurines, as well as 7-deaza-8-azapurine analogues, have been reported to be suitable for modification by copper-catalyzed azide/alkyne cycloaddition (CuAAC or click) chemistry, and used to modify siRNA molecules and to probe RNA editing by adenosine deaminases.[18,19,30,31] This explains why several efforts have been aimed at the synthesis of different purine analogues including 7-deaza-8-azapurine and its 7-substituted derivatives. In particular, Beal's group focused on exploring modifications of nucleic acid residues that project different substituents either in the minor groove (the sugar edge) or in the major groove (the Hoogsteen edge) of a siRNA duplex.[19,30,31] Indeed, the synthesis of nucleobase analogues that retains the "Watson−Crick" like pairing and that places the substituent in the major-groove is of particular interest, since they preserve the duplex stability and do not alter recognition by the nuclease of the RNA interference (RNAi) pathway.[31]

In this context, a very recent study[31] reported the impact on RNA duplexes of two modifications located on the major

Received: July 16, 2015
Revised: September 20, 2015

DOI: 10.1021/acs.jpcb.5b06861
J. Phys. Chem. B XXXX, XXX, XXX−XXX

groove (the "Hoogsteen edge") of the adenine moiety, namely, 7-ethynyl-7-deaza-8-aza adenine (from here on 7-E-7-DAA) and 7-triazole-7-deaza-8-azaadenine (7-T-7-DAA) (Figure 1).

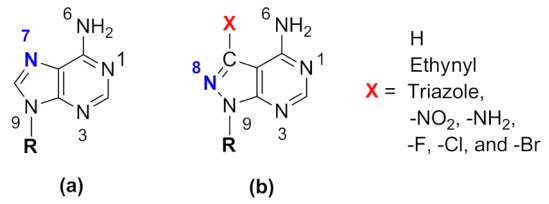


Figure 1. Structure of (a) adenine and (b) 7-substituted-7-deaza-8-aza-adenine with conventional notation of purines.

Looking at these modifications, it is clear that the H-bonding network on the Watson−Crick edge is unperturbed, thus allowing the modified adenines to pair with uracil in a canonical A:U *cis* Watson−Crick geometry. Indeed, it has been shown that the introduction of Br, Cl, or methyl as 7-substituents into 7-deazapurines results in a stabilization of oligonucleotide duplexes with antiparallel or parallel chain orientation.[29,32,33,24,27,33,34] To gain structural insight on the impact of the above-defined modifications in the context of a real RNA structure, Beal and colleagues reported crystallographic structures for two 16-base pair RNA duplexes incorporating the above modifications and concluded that, as expected, 7-E-7-DAA and 7-T-7-DAA can canonically pair to uracil and are well-accommodated within a A-form helix.[31] Notably, they also showed that the two modified nucleosides are read as adenosine in a template strand by avian myoblastosis virus reverse transcriptase (AMV-RT).[31]

To complement these structural studies, and to rationalize the apparently negligible structural effect on the RNA structure of the above quite complex modifications, we decided to perform quantum mechanics calculations. In particular, we aimed to shed light on the impact of these modifications on the H-bond pairing of the A−U. Following an approach we used to evaluate the impact of natural modifications on H-bonding base pairing,[35] we initially focused on the effect of the modifications on the geometry and energetics of the base pairs they are involved in, and compared them with the standard A−U Watson−Crick pairs, by analyzing the impact of the structural modification of the purine skeleton (consequent to the translocation of the N7 and C8 atoms). We anticipate here that the severe 7-aza-8-deaza modification has a limited impact on the geometry and stability of the modified base pair. Then, we analyzed the impact of the ethynyl and triazole substituents on the C7 atom of 8-aza-7-deazaadenine. Furthermore, to have a better understanding of the response of the H-bonding capability of the modified base to the nature of the substituent on the C7 position, we also analyzed modified bases bearing either the powerful electron withdrawing NO_2 group, or the powerful electron donating NH_2 group. These two groups are at the extremes of the Hammett scale,[36] which measures the electron donor and withdrawing capability of substituents, and are known to modify remarkably the properties of the aromatic ring to which they are bonded. Finally, the analysis is completed by considering the impact of a halide substituent, F, Cl, and Br, on the C7 atom, again to examine the influence of groups that can differently impact the electron density on the modified base skeleton. For all these systems we considered the classic Watson−Crick (cWW) and the reverse Watson−Crick (tWW) geometries, both in gas phase and in water. The reverse

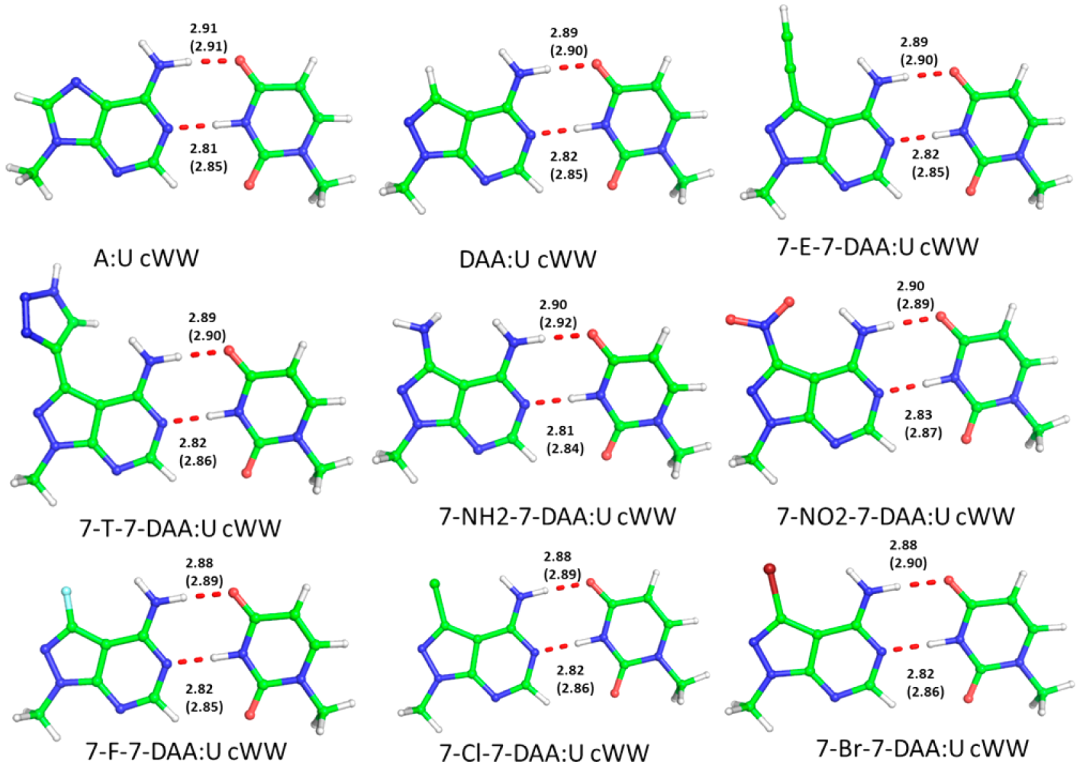


Figure 2. Stick representation of the base pairs including a modified adenine H-bonded to uracil in the cWW geometry. The values in parentheses correspond to the optimized distances in water, and values without parentheses correspond to optimized distances in the gas phase. All distances are in Å.

DOI: 10.1021/acs.jpcb.5b06861
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Watson−Crick (tWW) geometry, characteristic of parallel stranded helices, has also been considered, as 7-substituents in 7-deazapurine moieties have been introduced, and relative stability has been experimentally investigated in both antiparallel and parallel stranded double helices.

MODELS AND COMPUTATIONAL DETAILS

To investigate the stability of the unnatural modified base pairs under study, we have modeled the H-bonded modified base pairs and calculated their interaction energy using density functional theory and post-HF method.

Modeling the Interaction System. We first focused on studying the impact on H-bonding of a heterocyclic ring with a simple translocation of the N7−C8 atoms of adenine. For this, we just took into account 7-deaza-8-azaadenine (DAA) H-bonded to uracil (U) (i.e., DAA:U cWW). The base pair system was modeled by preserving the H-bonding pattern as in A:U cWW. Next, we studied the impact of different C7 substituents on the "DAA" purine moiety that is H-bonded to uracil. The studied substituents include ethynyl, 1,2,3-triazole, $-NO_2$, $-NH_2$, −F, −Cl, and −Br present at the C-7 position (see Figure 1) giving rise to 7-E-7-DAA:U, 7-T-7-DAA:U, 7-NO2-7-DAA:U, 7-NH2-7-DAA:U, 7-F-7-DAA:U, 7-Cl-7-DAA:U, and 7-Br-7-DAA:U base pairs (see Figure 2). The coordinates of the modeled geometries of the modified base pairs corresponding to 7-E-7-DAA:U cWW and 7-T-7-DAA:U cWW base pairing geometries were built from the X-ray structures with PDB IDs 4NFP and 4NFQ, respectively.[31] For other base pairs, respective functionalities were introduced at C7 position of DAA. Finally, to model A:U cWW, the modified "adenine" was replaced by a canonical "adenine" preserving the H-bonding pattern. Thus, in total, we modeled 8 different geometries of the non-natural modified base pair systems. The glycosidic bonds of the described base pairs are oriented in "*cis*" orientation, which corresponds to the geometries pertinent to antiparallel stranded (aps) RNA structure.

Due to the ambiguous H-bonding properties of the A:U base pair, it can be well-accommodated also in parallel stranded nucleic acid structures, in addition to the above-described modeled geometries we also considered the similar modifications now considering the "*trans*" orientation of glycosidic bonds. This geometry will correspond to a parallel stranded helix, which is common to RNA structures. Also, in this case a total of 8 modified base pairing geometries were modeled, taking the initial model from A437:U205 tWW, PDB ID 1JJ2. For all the model systems described above, the base pairing geometries are truncated at the C1′ atom of the ribose. This is the standard approach used in literature.

QM Calculations. A density functional theory approach, based on the hybrid B3LYP functional,[37−39] as implemented in the Gaussian09 package, in connection with the cc-pVTZ basis set[40] was used for all the geometry optimizations in gas phase as well as in water, modeled with the C-PCM continuum solvation model.[41] Since dispersion interactions might contribute differently to the stability of the non-natural base pairs, we also added the Grimme D3 term to the electronic energy.[42] Interaction energies were calculated on the B3LYP-D3/cc-pVTZ optimized geometries at the second order Moeller−Plesset (MP2)[43] level of theory using the more extended aug-cc-pVTZ basis set. For these calculations, we took advantage of the faster RIMP2[44] method as implemented in the Turbomole 6.1 package, with water modeled with the continuum solvation model COSMO.[41] All the interaction energies were corrected for the basis set super position error (BSSE),[45] using the counterpoise procedure.[45] Thus, the interaction energy E_{int} is calculated as in eq 1

$$E_{int} = E_{complex} - (E_{M1} + E_{M2}) + BSSE \quad (1)$$

where $E_{complex}$ is the electronic energy of the optimized M1:M2 base pair, and E_{M1} and E_{M2} are the electronic energies of the isolated geometries of the M1 and M2 bases in the geometry they have in the base pair, and BSSE is the basis set superposition error. The total interaction energy, which includes the deformation energy of the isolated M1 and M2 bases to the geometry they have in the base pair, is also calculated; this information can be found in Table S1. This is a rather standard approach used in this kind of calculation.[46−52] In the present study we also derived the interaction energy in water, which was calculated using the same recipe as suggested by Sponer and co-workers.[46,53]

To have an immediate and intuitive understanding of the impact of a specific modification, we introduce the modification energy, E_{mod}, defined as the energy difference between the interaction energy of the modified and of the corresponding natural base pair, as shown in eq 2.

$$E_{mod} = E_{int}(\text{modified base pair}) - E_{int}(\text{natural base pair}) \quad (2)$$

Within this definition, positive and negative E_{mod} values indicate modifications that decrease or increase the stability of a specific base pair, respectively.

Electron Density Analysis. Comparative analysis of the electron density of a single adenine base with that of the modified adenine [DAA], as well as with that of the other modified bases, was performed as follows. First, the geometry of the DAA base was optimized at the B3LYP/cc-pVTZ level of theory. For the sake of easier analysis,[54] C_s symmetry, with the symmetry plane coincident with the purine plane, was imposed to the systems, and the electron density analysis could be performed in the symmetry plane. After optimization, we compared the RI-MP2/aug-cc-pVTZ electron densities of the modified DAA base, ρ^{DAA}, and that of the nonmodified A base with the geometry it has in the DAA base, $\rho^{A/DAA}$. In other words, we took the optimized geometry of the DAA base, we replaced C7 with a N, and we removed the H7 atom; plus, we replaced N8 with a C atom, and we added a H atom at 1.10 Å from C8. Further, the heavy atoms of the 6-membered ring and of the N6 amino group were frozen to the geometry they have in adenine, while the other atoms were relaxed. With this approach, the heavy atom skeleton of the 6-membered ring of DAA and A bases, involved in the H-bond interaction, is identical and can be perfectly superimposed. This is fundamental to avoid noise in the analysis of the electron density difference, $\rho^{DAA-A/DAA} = \rho^{DAA} - \rho^{A/DAA}$, in the region of the base involved in the H-bond interaction. A similar procedure was followed for the other modifications where different functionalities (X), see Modeling the Interaction System section, were introduced at the C7 position of DAA. In this case, we compared the RI-MP2/aug-cc-pVTZ electron densities of DAA-X, of ρ^{DAA-X}, and of the A base obtained by removing the "X" group from the optimized DAA-X base, and replacing the C7 atom of DAA-X with a N atom, and replacing N8 atom with C, $\rho^{A/7DAA-X}$. The plotted electron density difference is $\rho^{DAA-X-A/DAA-X} = \rho^{DAA-X} - \rho^{A/DAA-X}$. We adopted this procedure to have perfect overlap between the skeletons of the A and DAA-X bases and, consequently, meaningful electron

DOI: 10.1021/acs.jpcb.5b06861
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Table 1. Geometry and Interaction Energy Values for the cWW and tWW A:U and Modified A:U Base Pairs[a]

system	geometry	E_{HF}	E_{corr}	ΔEint-gas	E_{mod} (gas)	ΔE-water	E_{mod} (water)
A:U	cWW	−10.09	−6.36	−16.45	0.00	−7.90	0.00
DAA:U	cWW	−10.79	−6.17	−16.96	−0.51	−8.00	−0.10
7-E-7DAA:U	cWW	−10.49	−6.20	−16.69	−0.24	−7.91	−0.01
7-T-7-DAA:U	cWW	−10.71	−6.18	−16.89	−0.44	−7.93	−0.03
7-F-7-DAA:U	cWW	−10.68	−6.11	−16.79	−0.34	−7.88	0.02
7-Cl-7-DAA:U	cWW	−10.54	−6.16	−16.70	−0.25	−7.90	0.00
7-Br-7-DAA:U	cWW	−10.47	−6.20	−16.67	−0.22	−7.91	−0.01
7-NO2-7-DAA:U	cWW	−10.23	−5.92	−16.15	0.30	−7.71	0.19
7-NH2-7-DAA:U	cWW	−10.50	−6.18	−16.68	−0.23	−8.06	−0.16
A:U	tWW	−9.44	−6.30	−15.74	0.00	−7.67	0.00
DAA:U	tWW	−9.93	−6.16	−16.09	−0.35	−7.75	−0.08
7-E-7-DAA:U	tWW	−9.70	−6.19	−15.89	−0.15	−7.66	0.01
7-T-7-DAA:U	tWW	−9.77	−6.19	−15.96	−0.22	−7.70	−0.03
7-F-7-DAA:U	tWW	−9.89	−6.09	−15.98	−0.24	−7.63	0.04
7-Cl-7-DAA:U	tWW	−9.76	−6.15	−15.91	−0.17	−7.64	0.03
7-Br-7-DAA:U	tWW	−9.72	−6.16	−15.88	−0.14	−7.65	0.02
7-NO2-7-DAA:U	tWW	−9.32	−6.19	−15.51	0.23	−7.45	0.22
7-NH2-7-DAA:U	tWW	−9.68	−6.16	−15.84	−0.10	−7.81	−0.14

[a]All interaction energy values are reported in kcal/mol. E_{mod} is the difference between the interaction energy of the modified base pair and of the reference pair, i.e., A:U cWW for cWW geometries, and A:U tWW for tWW geometries. Negative and positive values of E_{mod} indicate that the modified base pair is more stable or less stable than the A:U base pair.

density differences. A similar procedure was adopted for the base pairs where the electron density difference maps were evaluated.

RESULTS AND DISCUSSION

Eight different modifications of adenine residue on the "Hoogsteen edge" have been investigated, all of them being derivatives of 8-aza-7-deazaadenine. Optimal geometries and energetics of the base pairs they form with uracil, both in a classical *cis* Watson−Crick (cWW) and in the reversed Watson−Crick geometry (tWW), have been calculated in the gas phase and in water. Table 1 summarizes calculated energies for the investigated base pairs. Optimal geometry for the *cis* Watson−Crick base pairs are shown in Figure 2, where also H-bonding distances in gas and in water are shown.

Impact of the Change in the Heterocyclic Ring: Translocation of N7 and C8 Atoms. The geometry of the optimized DAA:U cWW pair is almost identical to the corresponding unmodified A:U cWW pair, with differences in H-bond lengths within 0.02 Å. The interaction of the modified DAA:U cWW pair is slightly stronger as compared to the A:U cWW pair, with E_{mod} of −0.51 kcal/mol and −0.10 kcal/mol in the gas phase and in water, respectively, clearly indicating the better stability of the modified pair. To shed light on the increased stability of the modified system, a consequence of the translocation of the N7 and C8 atoms, we compared the electron densities of the modified and nonmodified systems. Figure 3 reports the difference in the electron density in the gas phase between DAA and a simple A base, as well as between the DAA:U and the A:U cWW base pair (see Models and Computational Details section). Inspection of Figure 3A,B clearly indicates that the presence of the 7-deaza-8-aza modification substantially increases electron density around the N1 atoms of DAA which increases its H-bond accepting ability, while there is a depletion of electron density at the N6 atom, which reduces the H-bond donating ability of the N6−H group. These differences, consistent with an increased natural population analysis (NPA) charge on the N1 atom by −0.02e,

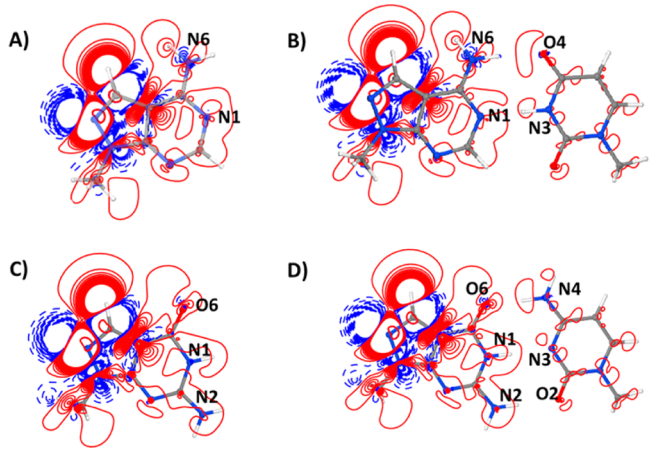


Figure 3. Electron density difference, in the base plane, (A) between the DAA and A bases, (B) between the DAA:U and A:U base pairs, (C) between the DAG and G base, and (D) between the DAG:C and G:C base pair. Density difference curves are plotted between −0.02 and 0.02 au, with a spacing of 0.001 au. Blue (red) lines that refer to negative (positive) density difference curves, i.e., to areas where the system including DAA presents reduced (increased) electron density as compared to the system including A.

and a reduced NPA charge on the H6 atom, by +0.01, explains the increased stability of the DAA:U cWW base pair system.

As for the tWW optimized geometry of the DAA:U pair, it is also very similar to the A:U tWW, both in gas and in water, with differences in H-bonds within 0.02 Å. Again, the stability of the base pair including DAA is higher than that of the base pair including the nonmodified base both in the gas phase and in water, with E_{mod} of −0.35 and −0.08 kcal/mol, respectively. The increased interaction strength of the base pair including the modified base can also be explained on the basis of the increased electron densities around atoms involved in H-bonding.

A similar analysis is also performed on 7-deaza-8-aza modification of the guanine base (DAG) of the canonical

DOI: 10.1021/acs.jpcb.5b06861
J. Phys. Chem. B XXXX, XXX, XXX−XXX

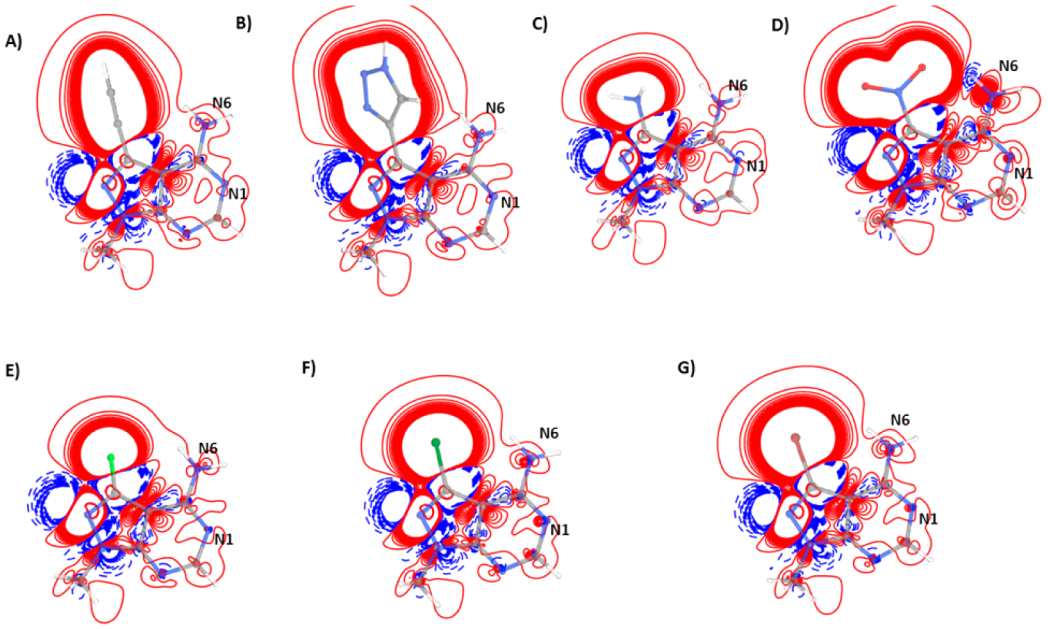


Figure 4. Electron density difference, in the base plane, between DAA bases presenting different substituents on the C7 atom, and the A base: (A) 7-E-7-DAA, (B) 7-T-7-DAA, (C) 7-NH2-7-DAA, (D) 7-NO2-7-DAA, (E) 7-F-7-DAA, (F) 7-Cl-7-DAA, (G) 7-Br-7-DAA. Density difference curves are plotted between −0.02 and 0.02 au, with a spacing of 0.001 au. Blue (red) lines refer to negative (positive) density difference curves, i.e., to areas where the substituted DAA presents reduced (increased) electron density as compared to the natural A base.

G:C cWW base pair. The optimized geometry of DAG:C pair is similar to G:C cWW, both in gas phase and in water, with differences in H-bonds within 0.02 Å. However, the G:C cWW base pair is more stable with respect to DAG:C cWW with E_{mod} of 0.41 and 0.04 kcal/mol. A slight decrease in interaction strength of DAG:C system, Figure 3C,D, could be explained in terms of the decrease in electron density around O6, which decreases its H-bonding accepting ability.

Impact of Different Substituents on the C7 of 7-Deaza-8-azaadenine (DAA). The first two introduced functional groups correspond to those whose structure in the context of a RNA duplex has been recently characterized experimentally: ethynyl, resulting in 7-ethynyl-7-deaza-8-azaadenine (7-E-7-DAA), and triazole, resulting in 1,2,3 trizole-7-deaza-8-azaadenine (7-T-7-DAA). In agreement with the experimental study,[31] the optimized geometries of 7-E-7-DAA:U cWW and 7-T-7-DAA:U cWW pairs are virtually unaffected by the modification (compared to the nonmodified A:U cWW pair). As for the interaction energies, the 7-E-7-DAA:U and 7-T-7-DAA:U cWW pairs, with E_{mod} of −0.24 and −0.44 kcal/mol, respectively, are slightly more stable than the nonmodified pair, in the gas phase. This slight increase in the interaction strength can be explained on the same grounds used to explain the higher stability of the base pair containing the parent DAA modified base, with increased electron density around the H-bond acceptor N6 atom. Different from the base pairs containing the parent DAA base, which was appreciably more stable than the nonmodified base pair in water, the stability of the 7-E-7-DAA:U cWW base pair in water is substantially identical to that of the nonmodified A:U base pair, with an E_{mod} of −0.01 kcal/mol, while the 7-T-7-DAA:U base pair has E_{mod} of −0.03 kcal/mol. As a remark, we also tested an alternative geometry of the 7-T-7-DAA base, with the triazole ring rotated by roughly 180° from the orientation it has in Figure 2, thus presenting the N-side of the ring oriented toward the N6 amino group, see Figure S2. This alternative geometry is more stable by 10.0 kcal/mol over the conformation depicted in Figure 2, essentially due to a H-bond between the N6 amino group and one of the N atoms of the triazole. However, this alternative conformation of the triazole ring would orient differently the R-substituent of the triazole, see again Figure S2. This different orientation would push the R-substituent of the triazole toward the phosphate backbone, resulting in steric clashes. Since this alternative conformation is not consistent with the X-ray structure, we do not discuss it further.

The tWW geometry is also unaffected by the two modifications, and the interaction between the two bases in the gas phase, with E_{mod} of −0.15 and −0.22 kcal/mol for 7-E-7-DAA and 7-T-7-DAA, respectively, is virtually unchanged by the modification. However, consistent with the gas phase calculations, the water interaction energy, with E_{mod} of −0.08 and +0.01 kcal/mol for 7-E-7-DAA:U and 7-T-7-DAA:U base pair, respectively, does not reflect an increased stabilization of the modified base pair. Comparison of the electron density difference maps involving the DAA:U base pair on one side, Figure 3B, and the 7-E-7-DAA:U and 7-T-7-DAA:U pairs on the other side, Figure 4A,B, indicates that the electron density from the substituent on the C7 atom involves also the amino N6 group, which can have an impact on proper solvation of the 7-E-7-DAA:U and 7-T-DAA:U base pairs around this area.

Intrigued by the geometric and energetic robustness of the A:U interaction to the severe modification investigated above, we decided to stress more the basic DAA:U skeleton by considering the effect of a strong electron withdrawing group, $-NO_2$, and of a strong electron donating group, $-NH_2$, on the C7 atom of DAA, Figure 4C,D. To have a more comprehensive picture, we also investigated the effect of a halogen substituent, −F, −Cl, and −Br, on the C7 atom, Figure 4E−G. Focusing on the optimized geometry of the base pairs, we again found very small deviations, H-bonds within 0.02 Å as compared to the A:U base pair, with all the introduced functional groups ($-NO_2$, $-NH_2$, −F, −Cl, −Br). With consideration of the

DOI: 10.1021/acs.jpcb.5b06861
J. Phys. Chem. B XXXX, XXX, XXX−XXX

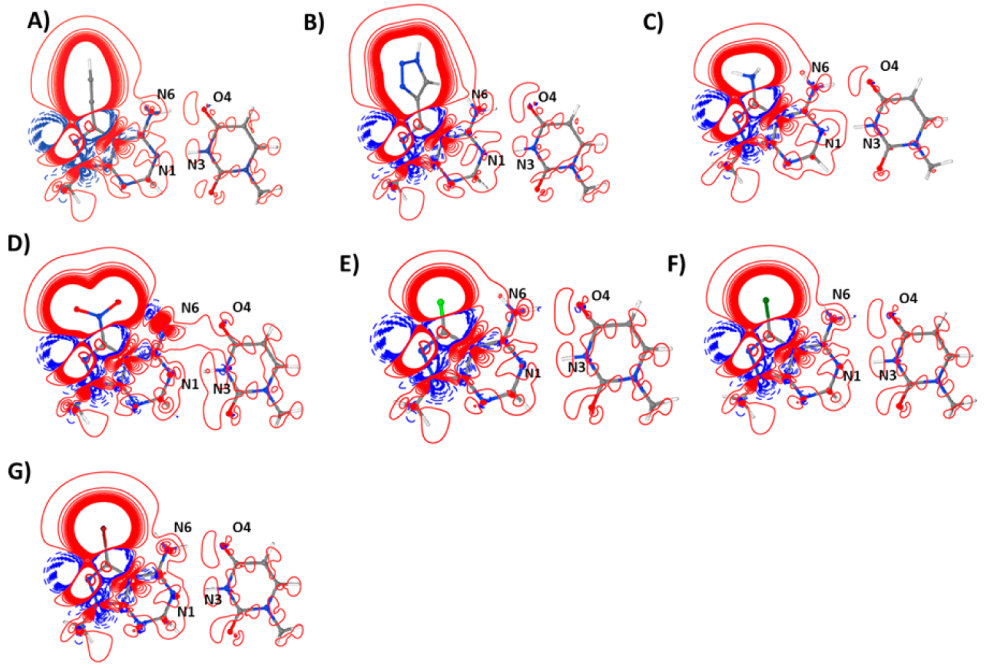


Figure 5. Electron density difference, in the base plane, between DAA:U base pairs presenting different substituents on the C7 atom, and the A:U base pair: (A) 7-E-7-DAA:U, (B) 7-T-7-DAA:U, (C) 7-NH2-7-DAA:U, (D) 7-NO2-7-DAA:U, (E) 7-F-7-DAA:U, (F) 7-Cl-7-DAA:U, (G) 7-Br-7-DAA:U. Density difference curves are plotted between −0.02 and 0.02 au, with a spacing of 0.001 au. Blue (red) lines refer to negative (positive) density difference curves, i.e., to areas where the base pair including the substituted DAA presents reduced (increased) electron density as compared to the A:U base pair.

remarkably different electronic properties of the considered groups, this clearly indicates that the A:U skeleton is remarkably robust to both the skeleton 7-deaza-8-aza modification as well as to the electronic effect due to the substituent on the C7 atom of the DAA parent skeleton, Figure 5A−G. Moving to gas phase energies, a small stabilization effect with E_{mod} within −0.34 kcal/mol is observed for the base pairs presenting a halogen at the C7 position of DAA. In contrast, both $-NH_2$ and $-NO_2$ at C7 position also results in very similar stabilization, E_{mod} of −0.23 and 0.30 kcal/mol, when compared to the A:U cWW base pair.

This indicates that the impact of the substituent on the C7 position is totally unrelated to its electronic property, as transmitted through the σ-bond skeleton. Rather, interaction can occur directly between the N6 amino group and the C7-substituent, as indicated by the modification in the electron density around the N6 amino group in Figures 4 C,D and 5 C,D consistently with our previous analysis on the G:C pair involving the archaeosine modification.[54] However, E_{mod} values in water are consistently around −0.16/+0.19 kcal/mol for all these modifications, Figure 5A−G. Interestingly, a similar trend in the E_{mod} is also observed for these base pairs in the "*trans*" tWW orientation.

As a final test to highlight similarity or differences between the original A:U pair and the modified base pairs considered in this work, we split the overall E_{int} into ΔE_{HF} and ΔE_{MP2} contributions. The former has been usually taken as an indicator of the H-bond contribution to the base pair stability, while the second has been usually taken as an indicator of the contribution of dispersion interactions to the base pair stability.[46−48,50,51] Inspection of Figure 6 clearly shows that basically all the base pairs including a modified adenine base behave quite similarly to the unmodified A base, confirming

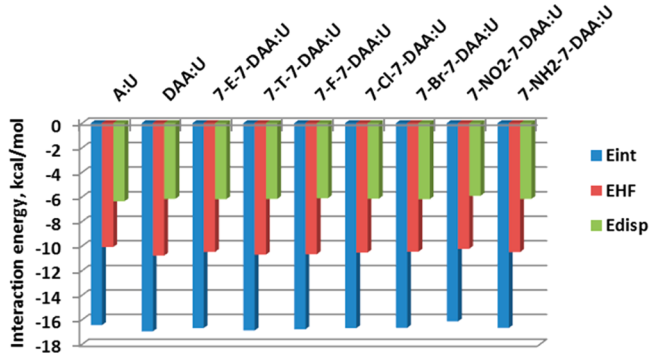


Figure 6. Total interaction energy (blue), HF contribution (red), and dispersion energies (green) between modified adenine (annotated on *X*-axis) with H-bonded uracil with glycosidic bonds oriented in cWW orientation. All values are given in kcal/mol. Dispersion contribution was calculated as the difference between the MP2 and HF energies.

again they can mimic very well the base pairing properties of the unmodified A base.

CONCLUSIONS

In this work we have examined the impact that a series of non-natural modifications, starting from the 8-aza-7-deaza modification of the adenine skeleton, followed by inclusion of a series of different substituents on the C7 position of the modified skeleton, can have on the geometry and stability of the A:U base pair. According to our calculations, the 7-deaza-8-aza modification and its 7-substituted derivatives have a negligible impact both on the geometry and the energetics of the base pairs, with H-bond distances basically matching those of the unmodified base pair, and with interaction energies usually differing by less than 1 kcal/mol. This is observed even

DOI: 10.1021/acs.jpcb.5b06861
J. Phys. Chem. B XXXX, XXX, XXX−XXX

when strong electron withdrawing and donating substituents as $-NO_2$ and $-NH_2$ have been considered. Similar conclusions are reached both for *trans* and *cis* geometries of the base pair, which implies that the modified bases can be easily accommodated in both parallel and antiparallel RNA strands. Further, the 8-aza-7-deaza modification of the guanine skeleton results in a similar impact on the G:C pair. Considering the severe modification of the base skeleton, and of the strong capability of the C7 groups to alter the electronic properties of the aromatic ring to which they are bonded, this should be considered a rather minimal impact. In conclusion, our work confirms that replacement of natural adenines in nucleic acids by a modified base belonging to this family should be transparent in terms of structure and stability, confirming the validity of this modification strategy to functionalize RNAs without perturbing its stability and biological functionality.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b06861.

Cartesian coordinates of all the structures discussed in this work (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: oliva@uniparthenope.it.

*E-mail: luigi.cavallo@kaust.edu.sa. Phone: +966-12-8080709.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to the reviewers for very useful comments that helped to improve the paper. The research presented in this work is supported by the King Abdullah University of Science and Technology.

■ REFERENCES

(1) Woese, C. R.; Dugre, D. H.; Dugre, S. A.; Kondo, M.; Saxinger, W. C. On the Fundamental Nature and Evolution of the Genetic Code. *Cold Spring Harbor Symp. Quant. Biol.* **1966**, *31*, 723−36.

(2) Woese, C. R.; Dugre, D. H.; Saxinger, W. C.; Dugre, S. A. The Molecular Basis for the Genetic Code. *Proc. Natl. Acad. Sci. U. S. A.* **1966**, *55*, 966−74.

(3) Crick, F. H. The Origin of the Genetic Code. *J. Mol. Biol.* **1968**, *38*, 367−79.

(4) Nahvi, A.; Sudarsan, N.; Ebert, M. S.; Zou, X.; Brown, K. L.; Breaker, R. R. Genetic Control by a Metabolite Binding Mrna. *Chem. Biol.* **2002**, *9*, 1043−49.

(5) Orgel, L. E. Evolution of the Genetic Apparatus. *J. Mol. Biol.* **1968**, *38*, 381−93.

(6) Winkler, W. C.; Cohen-Chalamish, S.; Breaker, R. R. An Mrna Structure That Controls Gene Expression by Binding Fmn. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15908−13.

(7) Machnicka, M. A.; Milanowska, K.; Osman Oglou, O.; Purta, E.; Kurkowska, M.; Olchowik, A.; Januszewski, W.; Kalinowski, S.; Dunin-Horkawicz, S.; Rother, K. M.; Helm, M.; Bujnicki, J. M.; Grosjean, H. Modomics: A Database of Rna Modification Pathways−2013 Update. *Nucleic Acids Res.* **2013**, *41*, D262−7.

(8) Avino, A.; Fabrega, C.; Tintore, M.; Eritja, R. Thrombin Binding Aptamer, More Than a Simple Aptamer: Chemically Modified Derivatives and Biomedical Applications. *Curr. Pharm. Des.* **2012**, *18*, 2036−47.

(9) Gros, J.; Rosu, F.; Amrane, S.; De Cian, A.; Gabelica, V.; Lacroix, L.; Mergny, J. L. Guanines Are a Quartet's Best Friend: Impact of Base Substitutions on the Kinetics and Stability of Tetramolecular Quadruplexes. *Nucleic Acids Res.* **2007**, *35*, 3064−75.

(10) Lech, C. J.; Cheow Lim, J. K.; Wen Lim, J. M.; Amrane, S.; Heddi, B.; Phan, A. T. Effects of Site-Specific Guanine C8-Modifications on an Intramolecular DNA G-Quadruplex. *Biophys. J.* **2011**, *101*, 1987−98.

(11) Sagi, J. G-Quadruplexes Incorporating Modified Constituents: A Review. *J. Biomol. Struct. Dyn.* **2014**, *32*, 477−511.

(12) Haaima, G.; Hansen, H. F.; Christensen, L.; Dahl, O.; Nielsen, P. E. Increased DNA Binding and Sequence Discrimination of Pna Oligomers Containing 2,6-Diaminopurine. *Nucleic Acids Res.* **1997**, *25*, 4639−4643.

(13) Li, Z.; Lech, C. J.; Phan, A. T. Sugar-Modified G-Quadruplexes: Effects of Lna-, 2′f-Rna- and 2′f-Ana-Guanosine Chemistries on G-Quadruplex Structure and Stability. *Nucleic Acids Res.* **2014**, *42*, 4068−79.

(14) Paul, A.; Sengupta, P.; Krishnan, Y.; Ladame, S. Combining G-Quadruplex Targeting Motifs on a Single Peptide Nucleic Acid Scaffold: A Hybrid (3 + 1) Pna-DNA Bimolecular Quadruplex. *Chem. - Eur. J.* **2008**, *14*, 8682−9.

(15) Peng, C. G.; Damha, M. J. G-Quadruplex Induced Stabilization by 2′-Deoxy-2′-Fluoro-D-Arabinonucleic Acids (2′f-Ana). *Nucleic Acids Res.* **2007**, *35*, 4977−88.

(16) Hernandez, A. R.; Peterson, L. W.; Kool, E. T. Steric Restrictions of Risc in Rna Interference Identified with Size-Expanded Rna Nucleobases. *ACS Chem. Biol.* **2012**, *7*, 1454−61.

(17) Layzer, J. M.; McCaffrey, A. P.; Tanner, A. K.; Huang, Z.; Kay, M. A.; Sullenger, B. A. In Vivo Activity of Nuclease-Resistant Sirnas. *RNA* **2004**, *10*, 766−771.

(18) Grammel, M.; Hang, H. C. Chemical Reporters for Biological Discovery. *Nat. Chem. Biol.* **2013**, *9*, 475−84.

(19) Ibarra-Soza, J. M.; Morris, A. A.; Jayalath, P.; Peacock, H.; Conrad, W. E.; Donald, M. B.; Kurth, M. J.; Beal, P. A. 7-Substituted 8-Aza-7-Deazaadenosines for Modification of the Sirna Major Groove. *Org. Biomol. Chem.* **2012**, *10*, 6491−7.

(20) Terrazas, M.; Kool, E. T. Rna Major Groove Modifications Improve Sirna Stability and Biological Activity. *Nucleic Acids Res.* **2009**, *37*, 346−353.

(21) Dhami, K.; Malyshev, D. A.; Ordoukhanian, P.; Kubelka, T.; Hocek, M.; Romesberg, F. E. Systematic Exploration of a Class of Hydrophobic Unnatural Base Pairs Yields Multiple New Candidates for the Expansion of the Genetic Alphabet. *Nucleic Acids Res.* **2014**, *42*, 10235−44.

(22) Li, L.; Degardin, M.; Lavergne, T.; Malyshev, D. A.; Dhami, K.; Ordoukhanian, P.; Romesberg, F. E. Natural-Like Replication of an Unnatural Base Pair for the Expansion of the Genetic Alphabet and Biotechnology Applications. *J. Am. Chem. Soc.* **2014**, *136*, 826−9.

(23) Peacock, H.; Fucini, R. V.; Jayalath, P.; Ibarra-Soza, J. M.; Haringsma, H. J.; Flanagan, W. M.; Willingham, A.; Beal, P. A. Nucleobase and Ribose Modifications Control Immunostimulation by a Microrna-122-Mimetic Rna. *J. Am. Chem. Soc.* **2011**, *133*, 9200−3.

(24) Seela, F.; Becher, G.; Zulauf, M. 8-Aza-7-Deazapurine DNA: Synthesis and Duplex Stability of Oligonucleotides Containing 7-Substituted Bases. *Nucleosides Nucleotides* **1999**, *18*, 1399−1400.

(25) Seela, F.; Kroschel, R. The Base Pairing Properties of 8-Aza-7-Deaza-2′-Deoxyisoguanosine and 7-Halogenated Derivatives in Oligonucleotide Duplexes with Parallel and Antiparallel Chain Orientation. *Nucleic Acids Res.* **2003**, *31*, 7150−7158.

(26) Seela, F.; Peng, X.; Li, H. Base-Pairing, Tautomerism, and Mismatch Discrimination of 7-Halogenated 7-Deaza-2′-Deoxyisoguanosine: Oligonucleotide Duplexes with Parallel and Antiparallel Chain Orientation. *J. Am. Chem. Soc.* **2005**, *127*, 7739−51.

(27) Seela, F.; Ramzaeva, N.; Zulauf, M. Duplex Stability of Oligonucleotides Containing 7-Substituted 7-Deaza- and 8-Aza-7-Deazapurine Nucleosides. *Nucleosides Nucleotides* **1997**, *16*, 963−966.

DOI: 10.1021/acs.jpcb.5b06861
J. Phys. Chem. B XXXX, XXX, XXX−XXX

(28) Seela, F.; Wei, C. F.; Kazimierczuk, Z. Substituent Reactivity and Tautomerism of Isoguanosine and Related Nucleosides. *Helv. Chim. Acta* **1995**, *78*, 1843−1854.
(29) Seela, F.; Zulauf, M. 7-Deazaadenine-DNA: Bulky 7-Iodo Substituents or Hydrophobic 7-Hexynyl Chains Are Well Accommodated in the Major Groove of Oligonucleotide Duplexes. *Chem. - Eur. J.* **1998**, *4*, 1781−1790.
(30) Phelps, K.; Morris, A.; Beal, P. A. Novel Modifications in Rna. *ACS Chem. Biol.* **2012**, *7*, 100−9.
(31) Phelps, K. J.; Ibarra-Soza, J. M.; Tran, K.; Fisher, A. J.; Beal, P. A. Click Modification of Rna at Adenosine: Structure and Reactivity of 7-Ethynyl- and 7-Triazolyl-8-Aza-7-Deazaadenosine in Rna. *ACS Chem. Biol.* **2014**, *9*, 1780−7.
(32) Aubert, Y.; Perrouault, L.; Helene, C.; Giovannangeli, C.; Asseline, U. Synthesis and Properties of Triple Helix-Forming Oligodeoxyribonucleotides Containing 7-Chloro-7-Deaza-2 ′-Deoxyguanosine. *Bioorg. Med. Chem.* **2001**, *9*, 1617−1624.
(33) Ramzaeva, N.; Seela, F. Duplex Stability of 7-Deazapurine DNA: Oligonucleotides Containing 7-Bromo- or 7-Iodo-7-Deazaguanine. *Helv. Chim. Acta* **1996**, *79*, 1549−1558.
(34) Lin, W. Q.; Xu, K. Y.; Seela, F. 7-Substituted 8-Aza-7-Deazapurines and 2,8-Diaza-7-Deaza-Purines: Synthesis of Nucleosides and Oligonucleotides. *Nucleosides, Nucleotides Nucleic Acids* **2005**, *24*, 869−873.
(35) Chawla, M.; Oliva, R.; Bujnicki, J. M.; Cavallo, L. An Atlas of Rna Base Pairs Involving Modified Nucleobases with Optimal Geometries and Accurate Energies. *Nucleic Acids Res.* **2015**, *43*, 6714−29.
(36) Hansch, C.; Leo, A.; Taft, R. W. A Survey of Hammett Substituent Constants and Resonance and Field Parameters. *Chem. Rev.* **1991**, *91*, 165−195.
(37) Becke, A. D. Density-Functional Thermochemistry 0.3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(38) Becke, A. D. Density-Functional Thermochemistry. *Abst. Pap. Am. Chem. Soc.* **1996**, *212*, 112 COMP.
(39) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(40) Dunning, T. H. Gaussian-Basis Sets for Use in Correlated Molecular Calculations 0.1. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(41) Klamt, A.; Schuurmann, G. Cosmo - a New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799−805.
(42) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (Dft-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(43) Moller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 0618−0622.
(44) Weigend, F.; Haser, M. Ri-Mp2: First Derivatives and Global Consistency. *Theor. Chem. Acc.* **1997**, *97*, 331−340.
(45) Boys, S. F.; Bernardi, F. Calculation of Small Molecular Interactions by Differences of Separate Total Energies - Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553−66.
(46) Sponer, J. E.; Reblova, K.; Mokdad, A.; Sychrovsky, V.; Leszczynski, J.; Sponer, J. Leading Rna Tertiary Interactions: Structures, Energies, and Water Insertion of a-Minor and P-Interactions. A Quantum Chemical View. *J. Phys. Chem. B* **2007**, *111*, 9153−64.
(47) Sponer, J. E.; Spackova, N.; Kulhanek, P.; Leszczynski, J.; Sponer, J. Non-Watson-Crick Base Pairing in Rna. Quantum Chemical Analysis of the Cis Watson-Crick/Sugar Edge Base Pair Family. *J. Phys. Chem. A* **2005**, *109*, 2292−301.
(48) Sponer, J. E.; Spackova, N.; Leszczynski, J.; Sponer, J. Principles of Rna Base Pairing: Structures and Energies of the Trans Watson-Crick/Sugar Edge Base Pairs. *J. Phys. Chem. B* **2005**, *109*, 11399−410.
(49) Sharma, P.; Sharma, S.; Chawla, M.; Mitra, A. Modeling the Noncovalent Interactions at the Metabolite Binding Site in Purine Riboswitches. *J. Mol. Model.* **2009**, *15*, 633−49.
(50) Chawla, M.; Sharma, P.; Halder, S.; Bhattacharyya, D.; Mitra, A. Protonation of Base Pairs in Rna: Context Analysis and Quantum Chemical Investigations of Their Geometries and Stabilities. *J. Phys. Chem. B* **2011**, *115*, 1469−84.
(51) Chawla, M.; Abdel-Azeim, S.; Oliva, R.; Cavallo, L. Higher Order Structural Effects Stabilizing the Reverse Watson-Crick Guanine-Cytosine Base Pair in Functional Rnas. *Nucleic Acids Res.* **2014**, *42*, 714−26.
(52) Oliva, R.; Cavallo, L.; Tramontano, A. Accurate Energies of Hydrogen Bonded Nucleic Acid Base Pairs and Triplets in Trna Tertiary Interactions. *Nucleic Acids Res.* **2006**, *34*, 865−879.
(53) Zirbel, C. L.; Sponer, J. E.; Sponer, J.; Stombaugh, J.; Leontis, N. B. Classification and Energetics of the Base-Phosphate Interactions in Rna. *Nucleic Acids Res.* **2009**, *37*, 4898−918.
(54) Oliva, R.; Tramontano, A.; Cavallo, L. Mg2+ Binding and Archaeosine Modification Stabilize the G15 C48 Levitt Base Pair in Trnas. *RNA* **2007**, *13*, 1427−36.

DOI: 10.1021/acs.jpcb.5b06861
J. Phys. Chem. B XXXX, XXX, XXX−XXX