Sugar Edge/Sugar Edge Base Pairs in RNA: Stabilities and Structures from Quantum Chemical Calculations

Judit E. Šponer,*,† Jerzy Leszczynski,‡ Vladimír Sychrovský,§ and Jiří Šponer*,†,§

Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic, Computational Center for Molecular Structure and Interactions, Department of Chemistry, Jackson State University, Jackson, Mississippi 39217, and Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo náměstí 2, 166 10 Prague 6, Czech Republic

Received: June 22, 2005; In Final Form: August 3, 2005

Cis and trans sugar edge/sugar edge (SE/SE) binding patterns are essential building units of RNAs. For example, SE/SE interactions form the A-minor motifs, the most important tertiary interaction type in functional RNAs. This study provides an in-depth structure and stability analysis for these two base pair families. Gas-phase-optimized geometries are reported for 12 cis and 7 trans SE/SE base pairs and contrasted to their X-ray counterparts. Interaction energies are computed at the RIMP2 level of theory using the density-functional-theory-optimized geometries. There is a good overall agreement between the optimized and X-ray geometries of the cis SE/SE base pairs. In contrast, only three of the seven trans SE/SE binding patterns could be optimized without a significant distortion of the X-ray geometry. Note, however, that many SE/SE base pairs participate in broader networks of interactions; thus it is not surprising to see some of them to deviate from the X-ray geometry in a complete isolation. Computed interaction energies reveal that all 12 known cis SE/SE binding patterns are very stable. Among the trans SE/SE binding patterns, only the rG/rG, rG/rC, and rA/rG base pairs are sufficiently stable in the crystal geometry. Prediction has been made for some structures not yet detected by crystallography, namely, cis rC/rC, rG/rC, rG/rU, and rU/rU and trans rG/rA base pairs. Interestingly, the new cis SE/SE binding patterns are not necessarily isosteric with the remaining 12 members of this family. The trans rG/rA base pair represents a viable option for base pairing in RNA to be identified by future X-ray studies. In a complete lack of structural information, prediction of other unknown members of the trans SE/SE family was not attempted. Analysis of the interaction energies shows a very large electron correlation component of the interaction energy, pointing at the elevated role of dispersion energy as compared to other types of base pairs. This likely is profitable for stabilization of SE/SE binding patterns in polar environments and could be one of the reasons why the A-minor motif is the leading type of tertiary interactions in RNAs.

Introduction

Presence of the 2′-OH hydroxyl group makes RNAs capable of specific interactions involving the ribose part of the polynucleotide structure. The non-Watson−Crick (non-WC) base pairs play a fundamental role in the folding and stability of RNA architectures. They have been characterized by a multitude of experimental techniques, such as X-ray crystallography,[1,2] NMR,[3] and database analysis.[4] A majority of the theoretical studies dealing with non-WC RNA base pairs come from molecular dynamics (MD) simulations,[5] while quantum chemical calculations have been concentrated almost exclusively on standard base pairing. Thus, despite the enormous importance of non-WC interactions in functional RNAs, the relevant quantum mechanical (QM) literature remains quite limited.[6,7]

In a sharp difference with DNA, the structural variability of RNAs is further propagated by the fact that base pairs with similar sizes and directionalities of the glycosidic bonds may often substitute for each other, irrespective of the base sequence. A unifying approach introduced by Leontis and Westhof provided a classification of the RNA base pairing patterns.[4b] They considered that each nucleotide offers three edges (Watson−Crick, Hoogsteen, and sugar edges) for interactions (Figure 1). Then, since the bases can be in either cis or trans positions, they established 12 different basic base pair families according to the binding modes observed in crystal geometries. Each base pair family is further divided into isosteric subgroups made up of those members of the family that exhibit similar steric extension and directionality of the glycosidic bonds.

In our recent QM studies, we have investigated the complete cis and trans Watson−Crick/sugar edge (WC/SE) base pair families (32 base pairs) and found a very good correlation between the crystallographic and the computed geometries.[7] Further, geometries of all WC/SE base pairs not yet seen in the available pool of experimental RNA structures were determined. The computed interaction energies have revealed that the WC/SE binding patterns are intrinsically stable. In summary, the data demonstrated that properties of experimentally observed WC/SE RNA base pairing patterns stem to a large extent from the intrinsic (gas phase) properties of the base pairs. Thus, the isolated WC/SE base pairs represent well-defined basic RNA

* Authors to whom correspondence should be addressed. E-mail: judit@ncbr.chemi.muni.cz; sponer@ncbr.chemi.muni.cz.
† Institute of Biophysics, Academy of Sciences of the Czech Republic.
‡ Jackson State University.
§ Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic.

10.1021/jp053379q CCC: $30.25
Published on Web 09/09/2005

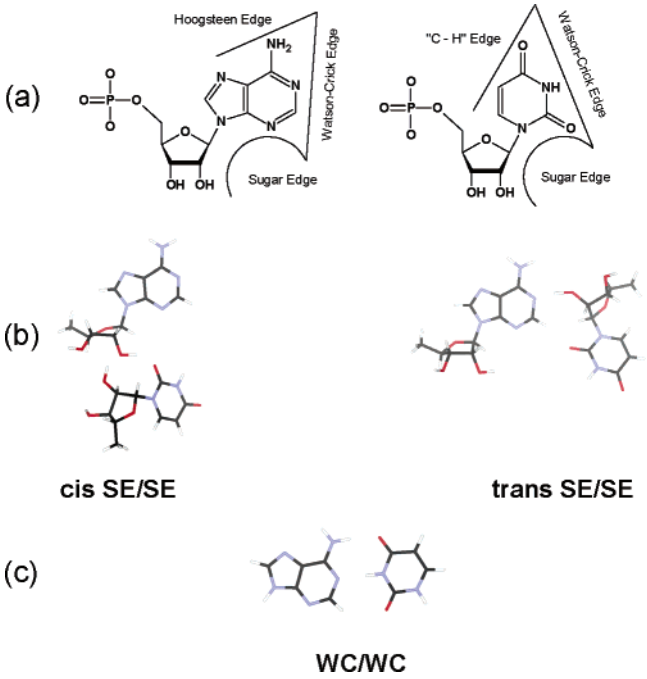

Figure 1. Nomenclature of the RNA base pairs.[4b] (a) Classification of the interaction sites in purine and pyrimidine nucleobases. (b) Cis and trans isomers of the rA/rU sugar edge/sugar edge (SE/SE) base pair (crystal geometries). (c) Standard A−U Watson−Crick/Watson−Crick (WC/WC) base pair.

building blocks. Decomposition of the total interaction energy into base−base and sugar−base terms revealed that a significant part of the intermolecular stabilization comes from the ribose−base contact. Our results have further shown that the WC/SE base pairs have increased electron correlation (dispersion) stabilization compared to standard base pairs. A sound correlation between the computed interaction energies obtained at the RIMP2 level of QM theory and by simple AMBER molecular mechanics calculations validated the force field approaches used for the modeling of RNA.[5]

In the present work, we complete our basic QM studies on RNA base pairing by investigating the perhaps biologically most important type of RNA base pairs, the sugar edge/sugar edge (SE/SE) binding patterns. We will consider the cis and trans isomers at once, although similarly to the previously studied WC/SE base pairing patterns each of them represents a distinct base pair family. QM investigation of the SE/SE families is somewhat more intricate compared with that of the WC/SE families, since an additional sugar needs to be explicitly included. Further, the SE/SE binding patterns in X-ray structures appear to be less strictly structurally defined. The SE/SE base pairs often occur in the context of larger local hydrogen bond networks and are thus less clearly separable from the surrounding parts of RNA. Many SE/SE base pairs play important roles in RNAs. These base pairs form the A-minor motif, the most important tertiary interaction in all functional RNAs. A-minor motifs stabilize interactions between RNA helices, loops and helices, and junctions.[2d] A-minor motifs involve adenosines that usually participate in some other non-Watson−Crick base pairs, leaving their Watson−Crick and sugar edges available for further interactions. These adenosines then interact with the shallow grooves of Watson−Crick base pairs (Figure 2). There is thus an insertion of the smooth, minor groove edges of adenines into the minor groove of neighboring helices, preferentially at C═G base pairs, where they form hydrogen bonds with one or both of the 2′-OH groups of those pairs. The A-minor motif is the most abundant tertiary structure interaction in the large ribosomal subunit and apparently also in other RNAs. At least 186 adenines in 23S and 5S rRNA participate in A-minor interactions, 68 of which are conserved.[2d] Several types of A-minor interactions have been identified. Among the SE/SE interactions, cis rA/rA is utilized in the type 0 A-minor motif, trans rA/rG and cis rA/rC are characteristic for A-minor type I, and cis rC/rA are characteristic for A-minor type II. The SE/SE interaction occurring in the A-minor type III tertiary contact represents a borderline between the cis and the trans rA/rA binding patterns (Figure 2). While the riboses adopt a cis orientation, the hydrogen-bonding scheme reflects the main features of the trans stereoisomer. The variability of A-minor contacts is enhanced by specific hydration. Many of these base pairs can be water-mediated. MD simulations revealed that the cis rA/rC base pair stabilizing the A-minor type I contact can fluctuate between presumably isoenergetic direct and water-mediated conformations on a time scale 1−10 ns. The water-mediated contact involves long-residency water molecules, and the dynamical structures are in full agreement with the range of "static" geometries seen in X-ray studies.[5g]

Having 12 known crystal geometries, the cis SE/SE base pairs belong to one of the best-described RNA base pair families. This base pair family exhibits several genuine features. There is a universal interaction pattern that makes all of its members isosteric. In addition, a direct ribose−ribose binding results in a practically identical C1′−C1′ distance for all binding patterns (5.4−5.9 Å based on X-ray geometries).

The trans SE/SE binding patterns are closely related to their trans WC/SE counterparts, because a rotation by 90° brings the WC base into the SE position. A consequence of this structural similarity is revealed by the crystal geometries of rPur/rB (Pur = A or G, B = A, G, U, or C) base pairs, because the trans SE/SE structures partly utilize their WC edges to establish ribose−base contacts. Only 7 of the 16 possible trans SE/SE structures have been seen experimentally; namely, none of the rPyr/rB (Pyr = C or U) base pairs has been observed so far. The trans SE/SE family is divided into two isosteric subgroups, assigned to the rA/rB and rG/rB type base pairs, respectively.

Present work provides a detailed QM analysis of the SE/SE RNA pairing patterns, exploiting two main advantages of modern ab initio QM methodology. (i) The calculated geometries and energies are reliable and accurate, and (ii) quantum chemistry provides a direct and unambiguous link between molecular structures and energies. This is an important piece of information to properly understand the role of non-Watson−Crick base pairing in the complex architectures of RNAs.

Computational Methods

Geometry Optimizations. When available, initial structures were built up on the basis of crystal geometries, using the structures listed in ref 4b, Figures 12 and 13. A majority of them were thus taken from the large ribosomal subunit of *Haloarcula marismortui* (NDB code RR0033).[2e] Additional crystal data were obtained from PR0022,[2f] RR0015,[2b] UR0004,[2g] and PTR009.[2h] The studied complexes consisted of two nucleosides, with replacement of the 5′-OH groups of the riboses by hydrogen atoms. In most cases, initial positions of the rotatable 2′-OH hydroxyl groups of the riboses were chosen following the hydrogen-bonding patterns suggested by Leontis et al. in ref 4b Figures 12 and 13. In addition, wherever it was feasible (trans rG/rG and rG/rC base pairs), we also considered structural variants with amino-acceptor hydrogen bonds between the 2′-OH and the exocyclic amino group of the purine nucleosides.

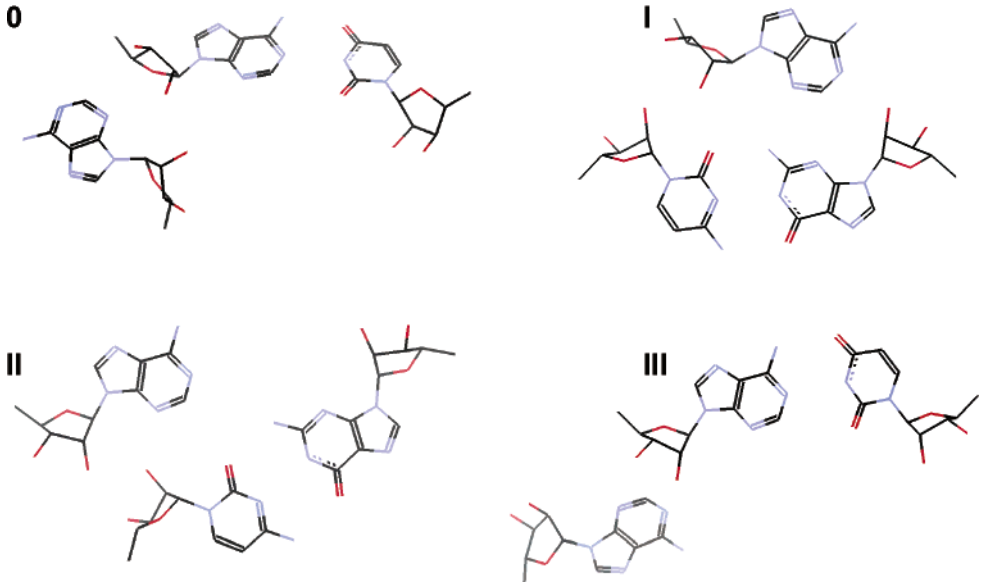

Figure 2. A-minor motifs[2d] from the structure of *H. Marismortui* (NDB code RR0033).[2e]

As in real RNA, the 3′-OH groups are never present, due to esterification with the phosphate group; in the initial geometries, they were oriented in such a way to avoid hydrogen-bond formation with the other nucleoside unit. For those structures not yet seen in crystallographic studies, we have constructed the initial model geometries according to the hydrogen-bonding pattern shown in ref 4b Figures 12 and 13. In all cases, the QM calculations resulted in a substantial refinement of the hydrogen-bonding patterns.

Geometry optimizations were carried out at the density functional theory (DFT) level of theory using the Gaussian 98 program package.[8] The density functional was built up by Becke's three-parameter exchange[9] and Lee, Yang, and Parr's correlation functional (abbreviated as B3LYP).[10] The 6-31G** basis set was used for geometry optimizations. We have shown recently that the B3LYP/6-31G**-optimized structures compare very well with reference RIMP2/cc-pVTZ data.[11]

Interaction Energies. Interaction energies were computed on the B3LYP-optimized structures using the RIMP2 approach combined with a large diffuse aug-cc-pVDZ basis set of atomic orbitals internally stored in the Turbomole code.[12] The same approach has been used in our former study,[7] where we have shown that the aug-cc-pVDZ basis set gives almost as good energies as those computed with the aug-cc-pVTZ basis set. The RIMP2 method for calculating interaction energies has been validated in refs 11 and 13. It is shown that the RIMP2 interaction energies are close to identical to true MP2 interaction energies (within 0.03 kcal/mol) while the RIMP2 method is much faster.[13] For hydrogen-bonded base pairs, interaction energies computed at the B3LYP/6-31G**//B3LYP/6-31G** level agree with those obtained at the RIMP2/aug-cc-pVDZ//RIMP2/cc-pVTZ level within 1.5 kcal/mol and are about 1−2 kcal/mol lower than the data extrapolated to the complete basis set.[11]

The interaction energy of a dimer consisting of two nucleosides Nuc1 and Nuc2 ($\Delta E^{\mathrm{Nuc1Nuc2}}$) is defined as

$$\Delta E^{\mathrm{Nuc1Nuc2}} = E^{\mathrm{Nuc1Nuc2}} - E^{\mathrm{Nuc1}} - E^{\mathrm{Nuc2}} \qquad (1)$$

where E^{Nuc1Nuc2} stands for the electronic energy of the whole system and E^{Nuc1} and E^{Nuc2} are the electronic energies of the isolated subsystems Nuc1 and Nuc2.

The interaction energy (ΔE) has two components: the Hartree−Fock (HF) term (ΔE^{HF}) and the electron correlation term (ΔE^{cor})

$$\Delta E = \Delta E^{\mathrm{HF}} + \Delta E^{\mathrm{cor}} \qquad (2)$$

The ΔE^{HF} term includes mainly the electrostatic interaction energy, short-range exchange repulsion, and polarization/charge-transfer contributions. The ΔE^{cor} term is dominated by the dispersion attraction and further includes the electron correlation correction to the electrostatic energy. The latter term is mostly repulsive since the electron correlation reduces the dipole moments of the monomers.

In contrast to our previous studies,[7] we did not decompose the interaction energy into pairwise sugar−base, base−base, and sugar−sugar contributions. Models of the SE/SE binding patterns include an excess ribose relative to those of the WC/SE base pairs; hence division of the model into two sugar and two base fragments would require cutting of one more chemical bond. This could, however, lead to a considerable error cumulation and may undermine the reliability of the estimated pairwise terms.

All interaction energies are corrected for the basis set superposition error using the standard counterpoise procedure[14] but do not include deformation energies. The main reason for this is the substantial structural alteration of the sugar−base segment upon base pairing in many structures. Thus a direct inclusion of monomer deformations into the interaction energies would rather bias the results. For a detailed discussion regarding the role of the deformation energies in base pairing calculations, see ref 11. Deformation energies can be readily derived considering the optimized geometries available in the Supporting Information.

Results and Discussion

Hydrogen-Bonding Patterns and Base Pair Stability in the Cis SE/SE Family. In addition to the common sugar−base and base−base components, the hydrogen-bonding network of cis SE/SE base pairs includes a mutual contact between the 2′-OH hydroxyls of the riboses (Figure 3 and Scheme 1). This type of ribose−ribose interaction does not occur in any other base pair family, including the trans SE/SE binding pattern, because it

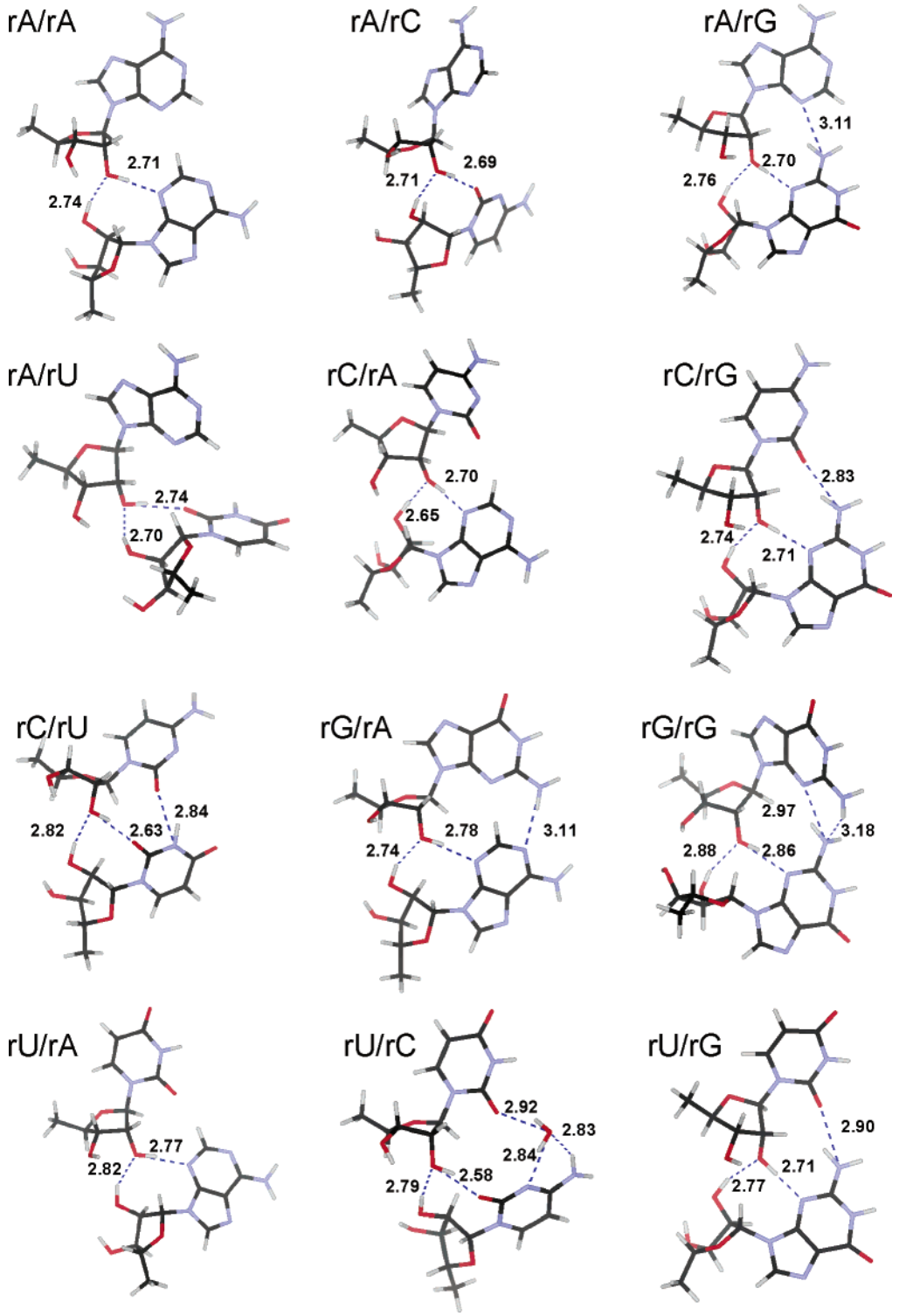

Figure 3. Fully optimized geometries for the 12 known members of the cis SE/SE base pair family. Hydrogen-bonding contacts are indicated with a dashed line. Distances (Å) between the donor and the acceptor atoms are listed explicitly.

requires the sugars to be in a close contact. The acceptor 2′-OH group behaves, in addition, as a hydrogen-bond donor toward the nucleobase of the other nucleotide (Scheme 1). Occasionally, a third interbase hydrogen bond contributes to the stabilization of the binding pattern. If the interacting bases lie far away from each other, then even a water-mediated interbase contact is feasible (Scheme 1).

The two hydrogen bonds brought about by the 2′-OH groups are strong enough to provide a satisfactory intrinsic stability for the cis SE/SE binding pattern (Table 1). The interaction energies for the rA/rU, rU/rU, rC/rU, rA/rA, rA/rC, and rU/rA base pairs range from −15.7 to −19.6 kcal/mol. Interbase hydrogen bonds further enhance the strength of the rA/rG, rC/rG, rG/rA, rG/rG, and rU/rG base pairs, and their computed interaction energies fall into the −23.3 to −26.9 kcal/mol interval. For a comparison, the stabilities of the A−U and G−C Watson−Crick base pairs are, with the same method, −15.3 and −29.4 kcal/mol.

Crystallographic studies suggest water-assisted interbase interaction for two SE/SE binding patterns. In the rU/rC base

SCHEME 1

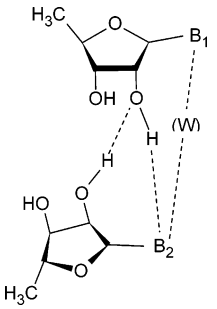

TABLE 1: Interaction Energies (kcal/mol) for cis SE/SE Base Pairs Obtained at the RIMP2/aug-cc-pVDZ Level Using the B3LYP/6-31G-Optimized Geometries**

base pair	ΔE[a]	ΔE^{HF} [a]	ΔE^{cor} [a]
rA/rA	−18.5	−8.8	−9.7
rA/rC	−19.0	−10.9	−8.1
rA/rG	−23.6	−11.1	−12.5
rA/rU	−15.7	−6.6	−9.1
rC/rA	−20.0[b]	−9.2[b]	−10.8[b]
rC/rC[c]	−21.4	−12.9	−8.5
rC/rG	−26.9	−16.5	−10.4
rC/rU	−17.2	−8.8	−8.4
rG/rA	−23.7	−11.4	−12.3
rG/rC[c]	−22.9	−12.9	−10.0
rG/rG	−24.3	−10.3	−14.0
rG/rU[c,d]	−18.9	−9.5	−9.4
rU/rA	−19.6	−10.7	−8.9
rU/rC[e]	−28.2[e]	−17.2[e]	−11.0[e]
rU/rG	−23.3	−12.4	−10.9
rU/rU[c]	−15.8	−7.5	−8.3

[a] ΔE = interaction energy computed at the RIMP2 level of theory; ΔE^{HF} = HF component of ΔE; ΔE^{cor} = correlation component of ΔE. For definition, see Computational Methods section. [b] This value was obtained for the optimized geometry without water. [c] No crystallographic data are available for these structures. [d] Interaction energy was calculated for the O3′-methylated structure. For details, see text. [e] Water-mediated structure; the energies are calculated for a trimer (including the water) and thus cannot be directly compared with those of the other base pairs.

pair, the water bridges the O2 positions of both bases. The strength of this base pair (−28.2 kcal/mol) is very similar to that computed for the trans WC/SE U/rG water-mediated base pair (−26.4 kcal/mol).[7b] Note that interaction energies for water-inserted pairs are formally evaluated for *trimers* consisting of two RNA fragments and the inserted water molecule and thus cannot be directly compared with the other (*dimer*) base pairs. We have also optimized the structure of the rU/rC pair without the water bridge. The resulting optimized structure is very similar to all other cis SE/SE binding patterns; however, the base pair is rather weak (−14.4 kcal/mol). On the contrary, we were not able to optimize the proposed C2(A)−H···W···O2(C) water bridge in rC/rA (Figure 12 in ref 4b). Due to the water insertion in the fully optimized geometry, the base pair opens up in such a way that rC is shifted toward the ribose part of rA and forms a hydrogen bond with 3′-OH. Because the 3′-OH is not present in the real RNA structures, this geometry is to be considered as an artifact of the gas-phase modeling. Nevertheless, we could satisfactorily optimize this base pair without the water bridge (Figure 3). Thus, the resulting binding pattern with a significant interaction energy of −20.0 kcal/mol is used for our analysis. As discussed elsewhere,[5g] SE/SE interactions may commonly exhibit a balance between more compact direct and looser water-mediated substates.

TABLE 2: C1′−C1′ Distances ($d_{C1'-C1'}$, Å) and Angles of the C1′−N Vectors (α, deg) for the cis SE/SE Base Pair Family

	$d_{C1'-C1'}$		α	
base pair	crystal	computed	crystal	computed
rA/rA	5.6	5.7	82.9	71.8
rA/rC	5.7	5.0	87.2	65.0
rA/rG	5.6	5.6	81.5	77.0
rA/rU	5.5	5.1	86.7	93.6
rC/rA	5.5	5.6	90.8	68.1
rC/rC[a]	5.7[b]	5.2		35.9
rC/rG	5.8	5.6	75.3	75.1
rC/rU	5.9	5.4	94.9	27.1
rG/rA	5.5	5.5	77.8	64.9
rG/rC[a]	5.6[b]	5.4		37.6
rG/rG	5.4	5.6	101.5	106.2
rG/rU[a]	5.5[b]	5.8		45.6
rU/rA	5.4	5.9	76.4	83.3
rU/rC	5.4	5.1	63.7	57.7
rU/rG	5.7	5.6	88.5	76.6
rU/rU[a]	5.4[b]	5.3		29.7

[a] No X-ray data are available for these structures. [b] Predicted values from ref 4b.

Isostericity in the cis SE/SE Family. By inspecting the crystal geometries, Leontis et al. assign the entire cis SE/SE family into a single isosteric subgroup.[4b] As we have shown in our previous studies, the C1′−C1′ distances between the interacting nucleosides and the angle of the glycosidic bonds defined by the C1′−N vectors (hereafter signified with α) can be successfully used as a criterion of isostericity. We are aware that this description of the isostericity is not complete, because it requires at least two additional parameters to exactly define the relative positions of the C1′−N vectors. Namely, one can consider the N(Nuc1)−C1′(Nuc1)−C1′(Nuc2) and C1′(Nuc1)−C1′(Nuc2)−N(Nuc2) angles, where N(Nuc1) and N(Nuc2) denote the N-atoms participating in the glycosidic bond inside the nucleoside unit. These two angles have a decisive role in defining the geometry of the coplanar RNA base pairs, because they carry information on the lateral shift of the bases. However, when a ribose is involved in the hydrogen-bonding scheme and, consequently, the base pair becomes non-coplanar, part of the information about the lateral shift of the bases is transformed into the angle of the glycosidic bonds. In such cases, as our experience shows, the angle α and the C1′−C1′ distance is sufficient to discriminate between the isosteric subgroups (see a complete set of isostericity parameters for the trans SE/SE base pairs in Table S1 in the Supporting Information) in those binding patterns involving at least one sugar unit. We evaluated the above parameters for the optimized gas-phase structures (depicted in Figure 3) and contrasted them with those found in RNA crystals (Table 2). Very good correspondence was found between the two sets of structures.

The angle of the C1′−N vectors ranges from 63.7° to 101.5° and 57.7° to 106.2° for the X-ray and gas-phase-optimized geometries, respectively. There is only one base pair, rC/rU, where the computed (27.1°) and X-ray (94.9°) values differ significantly. Here, the geometry optimization resulted in a rotation of rU and an extra hydrogen-bond formation between N3(U) and O2(C). Let us note, however, that this rearrangement of the structure did not influence the C1′−C1′ distance. The computed C1′−C1′ distances (5.0−5.9 Å), similar to their X-ray counterparts (5.4−5.9 Å), fall within a relatively tight range, unequivocally ordering all computed geometries into the same isosteric subfamily. The structural basis of this unprecedented similarity is to be attributed to the well-defined ribose−ribose interaction that is the common structural feature in all of these base pairs (see above).

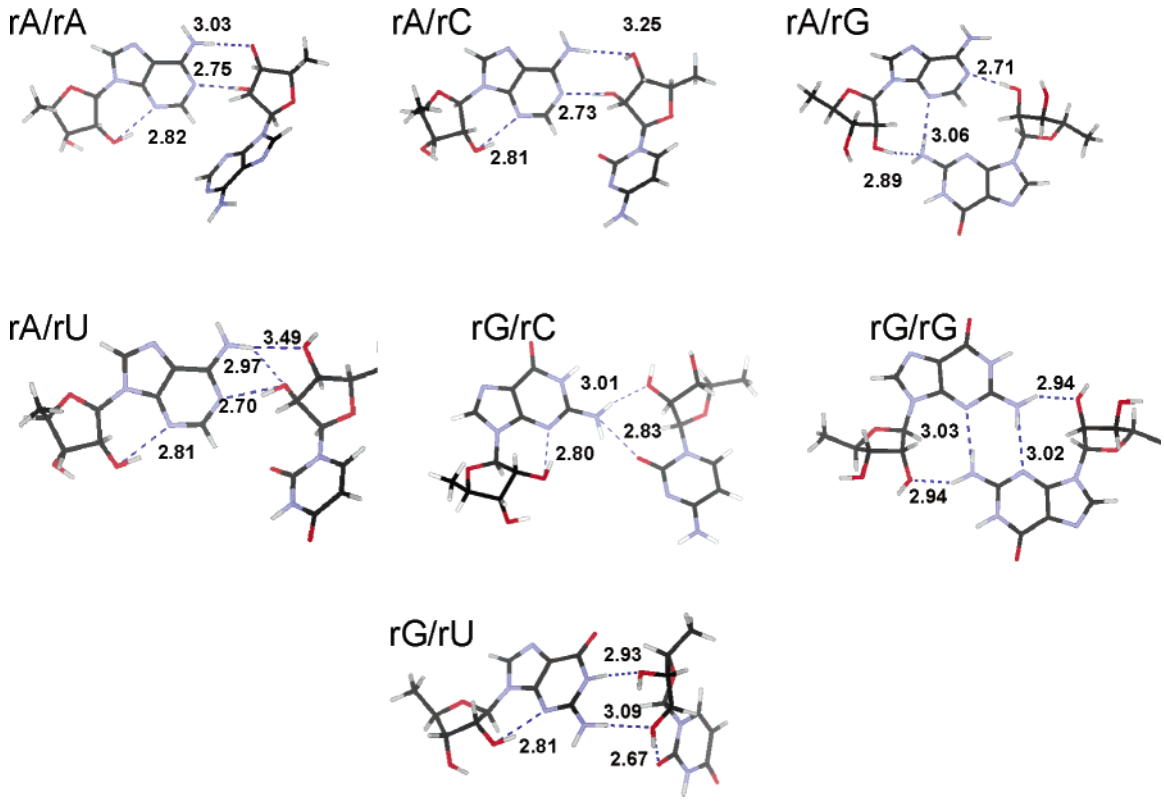

Figure 4. Fully optimized geometries for the seven known members of the trans SE/SE base pair family. Hydrogen-bonding contacts are indicated with a dashed line. Distances (Å) between the donor and the acceptor atoms are listed explicitly.

Hydrogen-Bonding Patterns and Base Pair Stability in the trans SE/SE Family. As mentioned in the Introduction, the hydrogen-bonding network of the trans SE/SE base pairs is always extended toward the WC edge of the participating nucleobases. Particularly, establishing the ribose−base contact requires some involvement of the WC edge in the interactions. The N2(G) amino group participates in the interbase hydrogen bonds in the guanine-containing base pairs (rA/rG, rG/rG, rG/rC, and rG/rU). In the rest of the crystallographically documented binding patterns (rA/rU, rA/rA, and rA/rC, see Figure 13 in ref 4b), C−H···N and C−H···O hydrogen bonds are assumed to stabilize the interbase contacts.

Only three of the seven known crystal geometries (rA/rG, rG/rG, and rG/rC) could be optimized without a significant distortion of the crystal geometry and alteration of the suggested hydrogen-bonding patterns (Figure 4). For the rA/rU, rA/rA, and rA/rC base pairs, however, geometry optimization resulted in a binding pattern with an extra N6(A)−H···O3′ hydrogen bond. Such binding, in fact, cannot be ruled out if it does not distort the base pair geometry too much. As Figure 5 illustrates, the primary interaction sites, i.e., the hydrogen-bond-donor 2′-OH groups, are not much displaced in the optimized geometries relative to their X-ray positions. Therefore, we suggest considering these interaction patterns as a viable option for RNA base pairing. Similarly, the rG/rU base pair did not maintain the suggested hydrogen-bonding pattern either, because full optimization led to a shift of rU toward the WC edge of rG (Figure 5d). This structural rearrangement required a more substantial change of the X-ray geometry than in the previous three cases, making the gas-phase-optimized structures rather unlikely to occur in real RNA architectures.

An intermolecular hydrogen bond between 2′-OH and N2(G) can be formed in two different ways in the rG/rC and rG/rG base pairs: (i) either as a canonical N2(G)−H···O2′ or (ii) as an O2′−H···N2(G) amino-acceptor interaction.[15] Therefore, these two base pairs were optimized considering both alternatives.[7b] The optimized geometries for conventional as well as amino-acceptor binding are depicted in Figures 4 and 6, respectively. For a stereoview of the same structures, refer to Figure S1 in the Supporting Information. The geometrical parameters of the rG/rG base pair were invariant to the directionality of the hydrogen bond (Table 3). In contrast, the optimized geometries obtained for the rG/rC base pair were strikingly different, and the one with canonical N2(G)−H···O2′ hydrogen bond gave a better agreement with the X-ray structure. While in the documented X-ray geometries the ribose−nucleobase interaction is reported between O2′ of cytidine and N2(G) only (see Figure 13 in ref 4b), there is an extra hydrogen bond donated by O2′ of guanosine to N3 of cytidine in the optimized geometry of the amino-acceptor variant (Figure 6). This makes the latter more compact than the reported X-ray geometry. Let us note here that similar duality of the hydrogen bonding has been observed also in the I_1 subfamily of the trans WC/SE base pairs.[7b]

Except for rG/rU, we calculated interaction energies using the fully optimized geometries. In general, the stability of the trans SE/SE structures (−14.3 to −21.4 kcal/mol, see Table 4) is well below (in absolute values) that of their cis counterparts. Among them, rA/rU, rA/rA, and rA/rC are of the lowest stability, despite formation of the extra N6(A)−H···O3′ hydrogen bond. For comparison, we have reoptimized the geometry of rA/rC by fixing the atomic positions of the ribose rings. The intermolecular interaction in the constrained geometry is 5 kcal/mol weaker than that in the fully optimized structure. A similar constrained optimization has also been performed for the rG/rU pair. The subsequent interaction energy calculation, however, revealed that the base pair in the crystal geometry is very weak with an interaction energy of −4.4 kcal/mol. It is thus very likely that the observed geometry is affected by external topological constraints and does not originate in the intrinsic base pairing

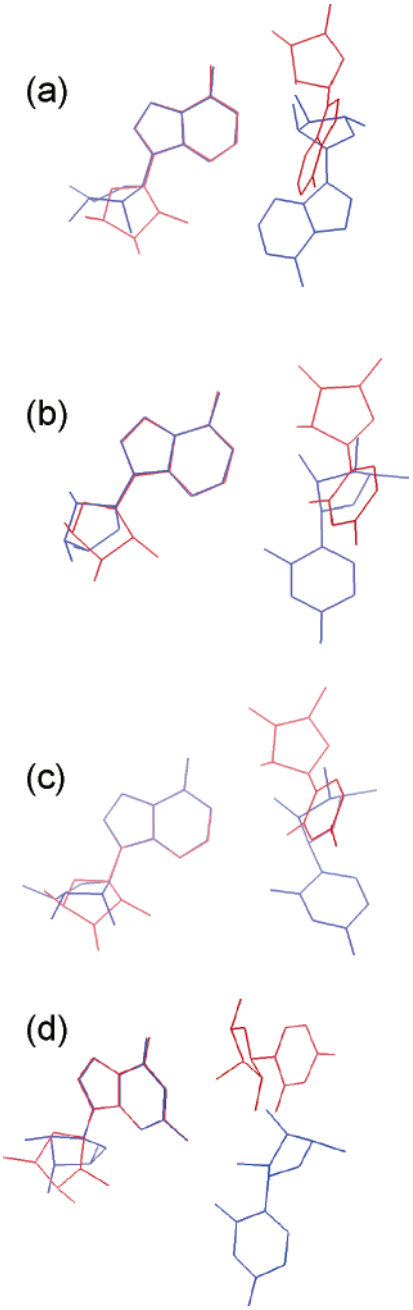

Figure 5. Overlay of the fully optimized (red) and X-ray (blue) geometries for the (a) rA/rA, (b) rA/rU, (c) rA/rC, and (d) rG/rU trans SE/SE base pairs.

stability. (It also cannot be ruled out that the X-ray geometry is affected by resolution of the crystal structure.)

Directionality of the hydrogen bond between O2′ and N2(G) influences the stability of the rG/rC base pair. Because of the extra hydrogen bond between O2′(rG) and N3(rC), in rG/rC the stability of the amino-acceptor structure (−20.9 kcal/mol) significantly exceeds that found for conventional binding (−13.5 kcal/mol). In contrast, there is not much stability difference between the amino-acceptor (−17.9 kcal/mol) and conventional (−21.2 kcal/mol) binding forms in rG/rG, both establishing four hydrogen-bonding contacts between the two nucleoside moieties.

Isostericity in the trans SE/SE Base Pairs. Leontis et al. rank the trans SE/SE base pairs into two isosteric subgroups.[4b] I_1 comprises the rA/rB (B = A, G, C, or U) base pairs, while the rG/rB structures are assigned to the I_2 subfamily. Scrutinizing the crystal geometries, in fact, we found that all base pairs are

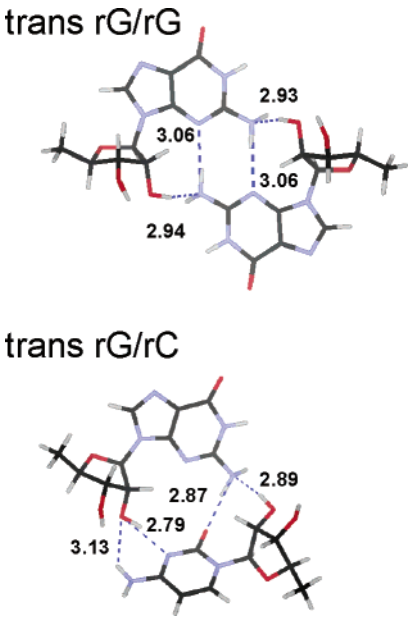

Figure 6. Possible amino-acceptor interactions in the trans SE/SE base pair family. Geometries are optimized at B3LYP/6-31G** level of theory. Hydrogen-bonding contacts are indicated with a dashed line. Distances (Å) between the donor and the acceptor atoms are listed explicitly.

TABLE 3: C1′−C1′ Distances ($d_{C1'-C1'}$, Å) and Angles of the C1′−N Vectors (α, deg) for the trans SE/SE Base Pair Family

	$d_{C1'-C1'}$		α	
base pair	X-ray	computed	X-ray	computed
rA/rA	8.3	9.4	151.1	108.6
rA/rC	8.3	9.3	148.1	121.7
rA/rG	8.2	7.8	146.6	143.2
rA/rU	8.3	9.4	131.8	117.1
rG/rA[a]	8.4	8.0		148.8
rG/rC	9.0	9.1/7.3[b]	156.8	160.0/140.1[b]
rG/rG	8.4	8.4/8.1[b]	159.2	154.5/165.1[b]
rG/rU	8.4	8.0	156.1	100.1

[a] No X-ray geometry is available for this structure. [b] The first value refers to the optimized structure with conventional hydrogen bonds, while the second value was computed for optimized geometries with amino-acceptor interactions.

of very similar dimensions (see columns 1 and 3 in Table 3): The C1′−C1′ distances range from 8.2 to 9.0 Å. Nevertheless, the angle of the glycosidic bonds is slightly different in the two subfamilies, being 131.8−151.1° and 156.1−159.2° for the I_1 and I_2 isosteric subgroups, respectively. Let us note here that the same differences can be observed also in the N(Nuc1)−C1′(Nuc1)−C1′(Nuc2) and C1′(Nuc1)−C1′(Nuc2)−N(Nuc2) angles. (N(Nuc1) and N(Nuc2) indicate the N-atoms involved in the glycosidic bonds; for numerical values refer to Table S1 in the Supporting Information.)

In general, the correlation of the computed and X-ray C1′−C1′ distances is unusually poor for the trans SE/SE structures. While computations provided satisfactory results for rA/rG, rG/rG, and the amino-acceptor form of rG/rC in majority of the computed base pairs, the structural rearrangements (see above) led to a considerable alteration of the C1′−C1′ parameter. The same applies also for the angle of the glycosidic bonds, whose computed values range from 100° to 165° in the fully optimized geometries. The most likely reason is that external forces, such as backbone geometry and other bases, affect the observed geometries of the trans SE/SE base pairs in the real RNA architectures, shifting them thus from the genuine gas-phase

TABLE 4: Interaction Energies (kcal/mol) for the trans SE/SE Base Pairs Obtained at the RIMP2/aug-cc-pVDZ Level Using the B3LYP/6-31G-Optimized Geometries[a]**

base pair	ΔE	ΔE^{HF}	ΔE^{corr}
rA/rA	−16.5	−9.1	−6.4
rA/rC	−15.5/−10.5[c]	−8.6/−4.5[c]	−6.9/−6.0[c]
rA/rG	−21.4	−7.7	−13.7
rA/rU	−14.3	−7.8	−6.5
rG/rA[b]	−21.0	−8.6	−12.4
rG/rC	−13.5/−20.9[d]	−11.1/−10.8[d]	−2.4/−10.1[d]
rG/rG	−21.2/−17.9[d]	−12.0/−4.8[d]	−9.2/−13.1[d]
rG/rU	−4.4[e]	−1.7[e]	−2.7[e]

[a] ΔE = interaction energy; ΔE^{HF} = HF component of the interaction energy; ΔE^{corr} = correlation component of the interaction energy. For definition, see Computational Methods section. [b] No X-ray data are available for this structure. [c] The first value refers to the fully optimized structure, while the second value was obtained for a partially relaxed geometry with fixed ribose rings. [d] The first value refers to the conventional binding; the second value was calculated for the optimized geometry with amino-acceptor hydrogen bonds. For details, see text. [e] The interaction energy was calculated for a geometry obtained from constrained optimization with fixed ribose rings.

minimum. We plan to investigate this issue in the future using larger model complexes and MD simulations.

Predicted Structures. The cis SE/SE base pair family is relatively well-described by X-ray data with 12 known base pairs. A good estimate of the missing four geometries (rC/rC, rG/rC, rG/rU, and rU/rU base pairs) can be easily obtained assuming the hydrogen-bonding patterns presented in Figure 12 of ref 4b. Optimized geometries of all of the predicted base pairs discussed in this paper are depicted in Figure 7. For rC/rC, rG/rC, and rU/rU, the gas-phase-optimized geometries are in line with the structural models suggested by Leontis et al. Thus, all three base pairs incorporate the common structural motif of the cis SE/SE base pairs, i.e., a ribose−ribose and ribose−nucleobase hydrogen bond joined by the same 2′-OH group. While there is no interbase contact in the rC/rC base pair, an additional N3(U)−H···O2(U) hydrogen bond enhances the gas-phase stability of rU/rU. Similarly, N2(G)−H···N3(C) and N4(C)−H···N2(G) hydrogen bonds complement the suggested hydrogen-bonding network of rG/rC in the gas phase. In contrast, geometry optimization of rG/rU substantially altered the proposed hydrogen-bonding scheme. In the optimized structure, the expected ribose−base hydrogen bond is completely absent due to an internal hydrogen-bond formation between 2′-OH and N1 of guanosine and a concomitant switch of the ribose pucker from C3′-endo to C2′-endo. This was the only case when we observed conformational change in the optimized geometry with respect to the proposed structure reported in ref 4b Figure 12. Otherwise, the sugar pucker is C3′-endo in all of the optimized base pairs if the 2′-OH is involved in intermolecular hydrogen bonds. The C2′-endo conformation is thus observed exclusively in those cases when the 2′-OH forms an internal hydrogen bond inside the nucleoside unit. Let us note that we have not computed the interaction energy for the cis rG/rU base pair, because we have noticed a strong hydrogen bond between the 3′-OH groups that is an artifact of the gas-phase modeling and could substantially bias the energy results. Instead, we have reoptimized the geometry by replacing the hydrogen terminations of the 3′-OH groups by CH_3. The methyl substitution has not much changed the hydrogen-bonding pattern of the cis rG/rU base pair, except that it helped to get rid of the interaction between the 3′-OH groups. The overlay of the optimized geometries with and without CH_3 groups is shown in Figure 8. Thus, for the cis rG/rU base pair we have considered the methyl-substituted variant in the interaction energy calculations. The computed interaction energies (from −15.8 to −22.9 kcal/mol) demonstrate that all four unknown cis SE/SE base pairs (rC/rC, rG/rC, rU/rU, and rG/rU) might be, in fact, intrinsically stable. However, an inspection of the main geometrical parameters shows that they are not necessarily isosteric with the crystallographically known members of the family. Albeit the C1′−C1′ distances (from 5.2 to 5.8 Å) fit well into the interval

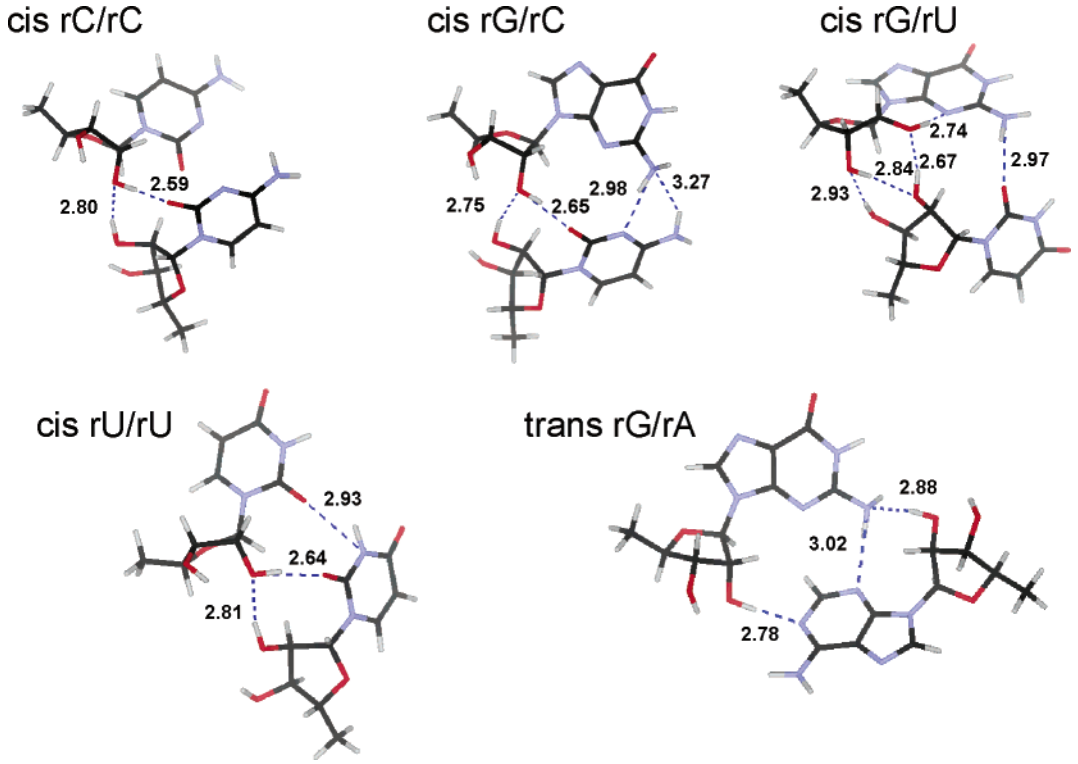

Figure 7. Fully optimized geometries for the predicted cis and trans SE/SE base pairs. Hydrogen-bonding contacts are indicated with a dashed line. Distances (Å) between the donor and the acceptor atoms are listed explicitly.

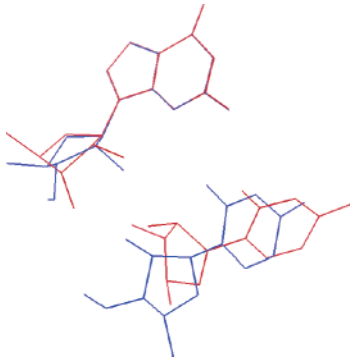

Figure 8. Overlay of the fully optimized geometries of the cis rG/rU base pair with (blue) and without (red) the methyl group

established for the known members of the cis SE/SE base pair family, the angles of the C1′−N vectors (from 29.7° to 45.6°) are significantly smaller than those found for other cis SE/SE base pairs (see above). However, this might be also an artifact of the gas-phase computations, as seen in case of the cis rC/rU base pair, where, due to the formation of an additional interbase hydrogen bond in the gas phase, the angle of the C1′−N vectors was significantly reduced.

Among the many unknown trans SE/SE binding patterns, rG/rA is the only base pair considered in this study. For the trans SE/SE rPyr/rB base pairs, there are no suggested geometries available in the literature, and we also did not find a plausible way to construct these structures. The interaction energy of the predicted rG/rA base pair is good, −21 kcal/mol due to an extensive hydrogen-bonding network involving two sugar−base and two base−base contacts. Note that in addition to the three hydrogen bonds proposed by Leontis et al.[4b] geometry optimization resulted in an extra sugar−base contact between O2′(rG) and N1(A) for this base pair (Figure 7). The computed C1′−C1′ distance (8.0 Å) is only marginally lower than the predicted one (8.4 Å), and the angle of the C1′−N vectors (148.8°) is in sound agreement with the corresponding data of other base pairs from this base pair family. Thus, the rG/rA trans SE/SE base pair is to be considered as a viable option for RNA base pairing.

The SE/SE Base Pairs Are Primarily Stabilized by the Dispersion Energy. In our previous studies,[7] we have shown that ribose−nucleobase hydrogen-bond formation results in an elevated correlation component (ΔE^{cor}) of the total interaction energy in the WC/SE base pairs with respect to the standard WC/WC base pairs. While in the standard WC/WC base pairs ΔE^{cor} does not exceed 30% of the total interaction energy, in some WC/SE base pairs it might reach even 50%.

In the third column of Tables 1 and 4, we list ΔE^{cor} values for the cis and trans SE/SE binding patterns, respectively. In these systems, the role of electron correlation becomes even more apparent. It provides at least 40% of the total interaction energy and, thereby, in majority of the studied cases becomes the leading contribution to the intermolecular stabilization. Note that the aug-cc-pVDZ basis set still modestly underestimates the electron correlation stabilization. In addition, the dispersion attraction usually exceeds the electron correlation stabilization, since the correlation term also contains the electrostatic correlation correction, which is repulsive. The increased role of dispersion energy means that the SE/SE pairing is shifted from predominantly electrostatic interactions toward more hydrophobic (dispersion-controlled) systems. It actually should enhance the stability of such interactions in polar solvent and may be one of the reasons why the SE/SE interactions play such a prominent role in mediating key tertiary contacts in RNA.

Conclusions

We have used advanced QM calculations for theoretical analysis of RNA cis and trans SE/SE binding patterns. This work is intended to fill the gap in the computational literature describing key non-WC base pairing in RNA and, therefore, is the first QM study of such base pairs. The SE/SE interactions belong to the most important biomolecular interactions, because they mediate, for example, leading tertiary interactions in RNA molecules, that is the A-minor motif.

While the cis SE/SE base pair family is well-characterized with 12 known binding patterns, only 7 out of the 16 possible trans SE/SE base pairing patterns have so far been reported by X-ray studies.

Geometry and intrinsic stability of the cis SE/SE base pairs is dictated by a common structural motif, which includes the 2′-OH groups of both riboses and one of the nucleobases. The computed geometries and hydrogen-bonding patterns of the 12 known base pairs are in sound agreement with those proposed by crystallographic studies. In agreement with the database analysis of Leontis et al.,[4b] we also suggest assigning all known cis SE/SE base pairs into the same isosteric subfamily. A prediction has been made for the four missing members of the family using the proposed geometries from ref 4b. Computations reveal large intrinsic stability for all cis SE/SE binding patterns with an interaction energy varying from −15.7 to −26.9 kcal/mol. However, one cannot unambiguously declare the predicted cis rG/rC, rC/rC, rU/rU, and rG/rU base pairs to be isosteric with the crystallographically documented members of this base pair family.

No significant alteration of the X-ray geometry was observed in the fully optimized geometries of trans rA/rG, rG/rG, and rG/rC. These three base pairs have been found to be intrinsically stable with interaction energies of −18 to −21 kcal/mol. Optimization of the less stable trans SE/SE base pairs (rA/rA, rA/rC, rA/rU, and rG/rU) led to a considerable distortion of the X-ray geometries, affecting mainly the angle of the C1′−N vectors. Assuming the crystal geometry, intrinsic stability of these four base pairs has been found to be very poor, about −4 to −10 kcal/mol. The observed geometries are thus likely stabilized by additional RNA parts surrounding the base pair and are not stable intrinsically. In contrast to cis and trans WC/SE and cis SE/SE structures, the trans SE/SE base pairs likely do not form genuine (self-structured) RNA building blocks and need to be considered as a part of larger building blocks and motifs. Calculations verified the proposed geometry of the rG/rA base pair. Due to its distinct intrinsic stability, this base pair is proposed to be identified by future X-ray works on RNA.

Analysis of the interaction energies has revealed a very significant role of the correlation component to the interaction energy in stabilizing the SE/SE pairing patterns. In many SE/SE base pairs, the correlation stabilization exceeds the HF term. In other words, the SE/SE interactions are assumed to be considerably more hydrophobic compared to standard base pairs, with the dispersion energy being the leading stabilization force in many SE/SE base pairs. This makes the SE/SE base pairs, together with their structural versatility, natural candidates for efficient tertiary interactions in RNAs.

Acknowledgment. This study was supported by a Wellcome Trust International Senior Research Fellowship in Biomedical Science in Central Europe (Grant No. GR067507; J.Š.) and the Grant Agency of the Czech Republic (Grant No. 203/05/0388; V.S., J.Š., J.E.Š.). J.L., J.Š., and J.E.Š. further acknowledge the financial support from the National Institutes of Health (Grant

No. S06 GM008047), the National Science Foundation Centers of Research Excellence in Science and Technology (Grant No. HRD-0318519), and the Office of Naval Research (Grant No. N00034-03-1-0116). The Institute of Biophysics and Institute of Organic Chemistry and Biochemistry are supported by grants AVO Z5 004 0507 and Z4 055 0506 from the Ministry of Education of the Czech Republic.

Supporting Information Available: Cartesian coordinates of all structures depicted in Figures 3, 4, 6, and 7, the isostericity parameters, stereofigures of the cis rA/rA, trans rG/rC, and trans rG/rG SE/SE base pairs. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) (a) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K. H.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678−1685. (b) Ferre-D'Amare, A. R.; Zhou, K. H.; Doudna, J. A. *Nature* **1998**, *395*, 567−574. (c) Scott, W. G.; Murray, J. B.; Arnold, J. R. P.; Stoddard, B. L.; Klug, A. *Science* **1996**, *274*, 2065−2069.

(2) (a) Moore, P. B.; Steitz, T. A. *Annu. Rev. Biochem.* **2003**, *72*, 813−850. (b) Wimberly, B. T.; Brodersen, D. E.; Clemons, W. M.; Morgan-Warren, R. J.; Carter, A. P.; Vonrhein, C.; Hartsch, T.; Ramakrishnan, V. *Nature* **2000**, *407*, 327−339. (c) Klein, D. J.; Schmeing, T. M.; Moore, P. B.; Steitz, T. A. *EMBO J.* **2001**, *20*, 4214−4221. (d) Nissen, P.; Ippolito, J. A.; Ban, N.; Moore, P. B.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4899−4903. (e) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 905−920. (f) Lewis, H. A.; Musunuru, K.; Jensen, K. B.; Edo, C.; Chen, H.; Darnell, R. B.; Burley, S. K. *Cell* **2000**, *100*, 323−332. (g) Su, L.; Chen, L.; Egli, M.; Berger, J. M.; Rich, A. *Nat. Struct. Biol.* **1999**, *6*, 285−292. (h) Arnez, J. G.; Steitz, T. A. *Biochemistry* **1996**, *35*, 14725−14733.

(3) (a) Vallurupalli, P.; Moore, P. B. *J. Mol. Biol.* **2003**, *325*, 843−856. (b) Finger, L. D.; Trantírek, L.; Johansson, C.; Feigon, J. *Nucleic Acids Res.* **2003**, *31*, 6461−6472. (c) Nixon, P. L.; Rangan, A.; Kim, Y. G.; Rich, A.; Hoffman, D. W.; Hennig, M.; Giedroc, D. P. *J. Mol. Biol.* **2002**, *322*, 621−633. (d) Kolk, M. H.; van der Graaf, M.; Wijmenga, S. S.; Pleij, C. W. A.; Heus, H. A.; Hilbers, C. W. *Science* **1998**, *280*, 434−438.

(4) (a) Leontis, N. B.; Westhof, E. *Q. Rev. Biophys.* **1998**, *31*, 399−455. (b) Leontis, N. B.; Stombaugh, J.; Westhof, E. *Nucleic Acids Res.* **2002**, *30*, 3497−3531.

(5) (a) Reblová, K.; Špačková, N.; Stefl, R.; Csaszar, K.; Koča, J.; Leontis, N. B.; Šponer, J. *Biophys. J.* **2003**, *84*, 3564−3582. (b) Auffinger, P.; Bielecki, L.; Westhof, E. *J. Mol. Biol.* **2004**, *335*, 555−571. (c) Csaszar, K.; Špačková, N.; Stefl, R.; Šponer, J.; Leontis, N. B. *J. Mol. Biol.* **2001**, *313*, 1073−1091. (d) Schneider, C.; Brandl, M.; Suhnel, J. *J. Mol. Biol.* **2001**, *305*, 659−667. (e) Zacharias, M. *Curr. Opin. Struct. Biol.* **2000**, *10*, 311−317. (f) Rážga, F.; Špačková, N.; Reblová, K.; Koča, J.; Leontis, N. B.; Šponer, J. *J. Biomol. Struct. Dyn.* **2004**, *22*, 183−194. (g) Rážga, F.; Koča, J.; Šponer, J.; Leontis, N. B. *Biophys. J.* **2005**, *88*, 3466−3485.

(6) (a) Šponer, J.; Mokdad, A.; Šponer, J. E.; Špačková, N.; Leszczynski, J.; Leontis, N. B. *J. Mol. Biol.* **2003**, *330*, 967−978. (b) Brandl, M.; Meyer, M.; Suhnel, J. *J. Biomol. Struct. Dyn.* **2001**, *18*, 545−555. (c) Brandl, M.; Meyer, M.; Suhnel, J. *J. Phys. Chem. A* **2000**, *104*, 11177−11187. (d) Hobza, P.; Šponer, J.; Cubero, E.; Orozco, M.; Luque, F. J. *J. Phys. Chem. B* **2000**, *104*, 6286−6292. (e) Pan, Y.; Priyakumar, D.; MacKerell, A. D., Jr. *Biochemistry* **2005**, *44*, 1433−1443.

(7) (a) Šponer, J. E.; Špačková, N.; Kulhánek, P.; Leszczynski, J.; Šponer, J. *J. Phys. Chem. A* **2005**, *109*, 2292−2301. (b) Šponer, J. E.; Špačková, N.; Leszczynski, J.; Šponer, J. *J. Phys. Chem. B* **2005**, *109*, 11399−11410.

(8) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. J.; Cui Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 2001.

(9) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(10) (a) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789. (b) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200−206.

(11) Šponer, J.; Jurečka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142−10151.

(12) (a) Eichkorn, K.; Treutler, O.; Oehm, H.; Haeser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *242*, 652−660. (b) Weigend, F.; Haeser, M. *Theor. Chem. Acc.* **1997**, *97*, 331−340. (c) Weigend, F.; Haeser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143−152.

(13) Jurečka, P.; Nachtigall, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4578−4582.

(14) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.

(15) (a) Luisi, B.; Orozco, M.; Šponer, J.; Luque, F. J.; Shakked, Z. *J. Mol. Biol.* **1998**, *279*, 1123−1136. (b) Šponer, J.; Hobza, P. *J. Am. Chem. Soc.* **1994**, *116*, 709−714.