J Mol Model (2009) 15:633–649
DOI 10.1007/s00894-008-0384-y

ORIGINAL PAPER

Modeling the noncovalent interactions at the metabolite binding site in purine riboswitches

Purshotam Sharma · Sitansh Sharma · Mohit Chawla · Abhijit Mitra

Received: 1 September 2008 / Accepted: 8 October 2008 / Published online: 10 January 2009
© Springer-Verlag 2009

Abstract We present gas phase quantum chemical studies on the metabolite binding interactions in two important purine riboswitches, the adenine and guanine riboswitches, at the B3LYP/6-31G(d,p) level of theory. In order to gain insights into the strucutral basis of their discriminative abilities of regulating gene expression, the structural properties and binding energies for the gas phase optimized geometries of the metabolite bound binding pocket are analyzed and compared with their respective crystal geometries. Kitaura-Morokuma analysis has been carried out to calculate and decompose the interaction energy into various components. NBO and AIM analysis has been carried out to understand the strength and nature of binding of the individual aptamer bases with their respective purine metabolites. The Y74 base, U in case of adenine riboswitch and C in case of guanine riboswitch constitutes the only differentiating element between the two binding pockets. As expected, with W:W cis G:C74 interaction contributing more than 50% of the total binding energy, the interaction energy for metabolite binding as calculated for guanine (-46.43 Kcal/mol) is nearly double compared to the corresponding value for that of adenine (-24.73 Kcal/mol) in the crystal context. Variations in the optimized geometries for different models and comparison of relative contribution to metabolite binding involving four conserved bases reveal the possible role of U47:U51 W:H trans pair in the conformational transition of the riboswitch from the metabolite free to metabolite bound state. Our results are also indicative of significant contributions from stacking and magnesium ion interactions toward cooperativity effects in metabolite recognition.

Keywords Adenine · Guanine · Morokuma · Noncovalent · Purine metabolites · Riboswitches

P. Sharma · S. Sharma · M. Chawla · A. Mitra (✉)
Center for Computational Natural Sciences and Bioinformatics,
International Institute of Information Technology,
Hyderabad 500032, India
e-mail: abi_chem@iiit.ac.in

Introduction

Riboswitches are highly structured regions within the 5′-untranslated region (5′-UTR) of several mRNA molecules that modulate their expression by allosteric interconversion between two alternative secondary structures in their evolutionarily conserved sensing region, the aptamer domain, in response to the binding of a specific metabolite [1–3]. This metabolite induced structural transition in the aptamer region in turn induces switching of the expression platform domain present immediately downstream to it [4], thereby up- or down-regulating the gene expression. Recently many riboswitches, which bind small metabolites, have been identified. The range of metabolites include purine bases [5, 6], amino acids [7–9], coenzyme B12 [10], flavin mononucleotide [11, 12], thiamine pyrophosphate (TPP) [11, 13], S-adenosylmethionine (SAM) [14–16]. It has been estimated that in Bacillus subtilis and its related species, over 2.5% of the genes are regulated by interaction of riboswitches with small metabolites [17]. This highlights the importance of these RNA genetic control elements in normal cellular metabolism [17]. Moreover, complex riboswitch architecture with tandem aptamer domains have been shown to detect two different metabolites (S-adenosylmethionine and thiamine pyrophosphate) at two different positions, suggesting that more complex regulatory elements can also be assembled from individual riboswitches [18]. This and

other related observations also suggest multifaceted biotechnological possibilities which can evolve out of studies on the mechanism of riboswitch functioning. As the metabolite concentration is directly related to gene expression, new structural variants of riboswitch binding metabolites can be used to regulate gene expression, and can thus act as powerful antibacterial drugs [19]. Due to their remarkable ligand binding ability and specificity, these riboswitches can also be used as 'natural biosensors'[20].

Purine riboswitches, constitute a family of natural genetic control RNA elements that bind purines such as adenine [6], guanine [5], hypoxanthine [20] etc. to regulate the genes involved in their metabolism. These riboswitches are known to regulate a number of purine biosynthetic and efflux genes in Gram positive bacteria [21]. There are three distinct types of mechanisms by which purine riboswitches are known to regulate the gene expression in response to metabolite binding. The xpt-pbuX Bacillus subtilis Operon (xpt), for example is regulated by purines such as guanine, hypoxanthine and xanthine [5]. Metabolite binding in this case leads to the stabilization of the terminator fold during transcription, which in turn gets inhibited [5]. Conversely, in case of adenine binding riboswitch (which regulates the ydhL gene in Bacillus subtilis), the metabolite binding leads to stabilization of the antiterminator fold. This initiates the transcription process. A-riboswitch regulating the add-gene encoding for adenine deaminase enzyme in Clostridium perfringens, on the other hand, does not affect transcription. Adenine binding to the aptamer domain of the riboswitch results in unpairing of the otherwise paired nucleobases of the Shine Dalgarno sequence, preparing the mRNA for translation [22]. Although differing in their mechanism of gene regulation, there are common secondary structure elements in the aptamer region of the purine riboswitches. The ligand bound aptamer fold consists

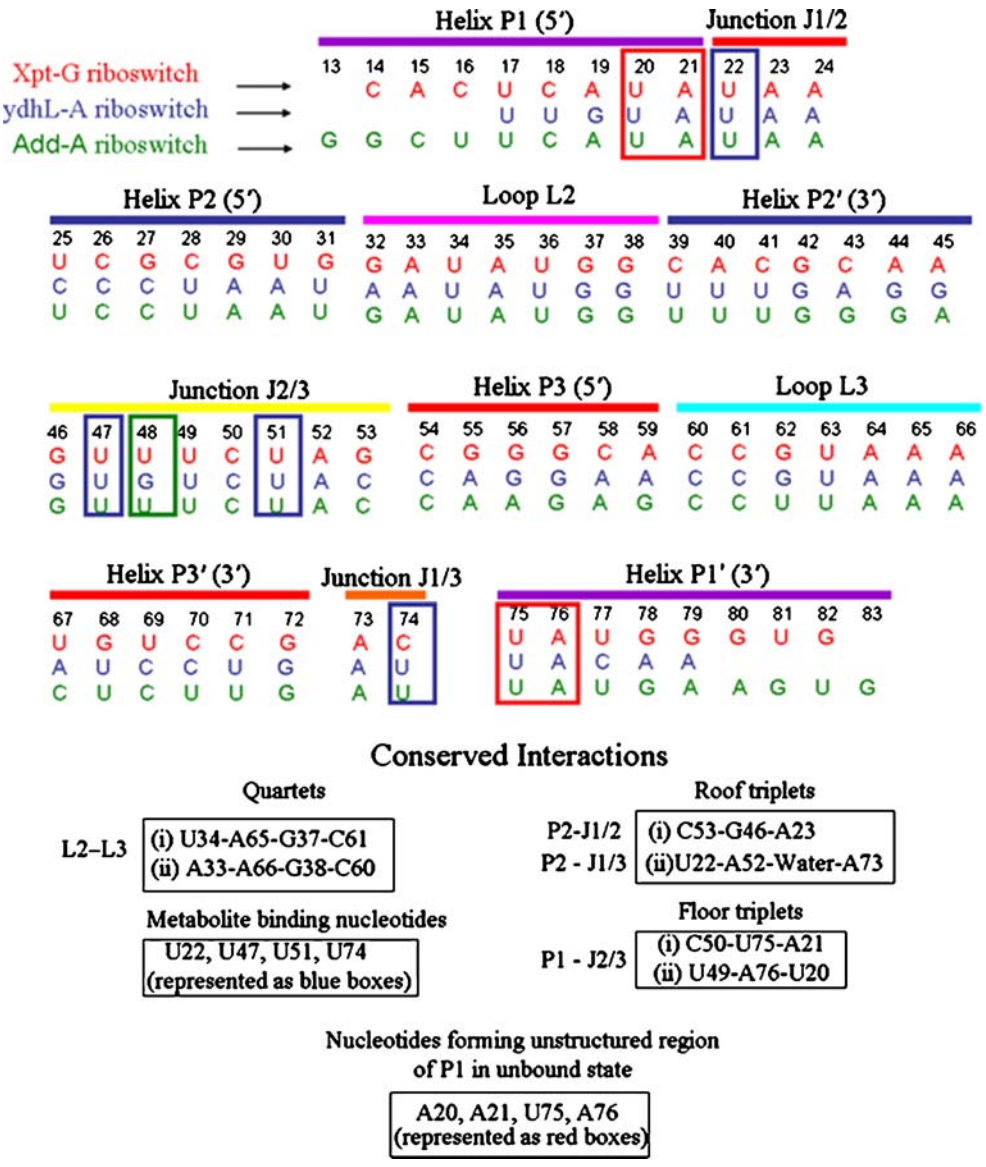

Fig. 1 Comparison of nucleotide sequences in different secondary structural elements of the aptamer domain of 1. xpt-G riboswitch involved in transciption termination (red) [5], 2. ydhL-A riboswitch involved in transcription antitermination (blue) [6] and 3. add A-riboswitch involved in translaional activation (green) [22]. The 48 position of J2/3 loop (represented in a green box) is never adenine in A-riboswitch and is never guanine in G-riboswitch [25]. U22 (J1/2), U47 and U51 (J2/3) and Y74 residues (in blue boxes) interact with the metabolite

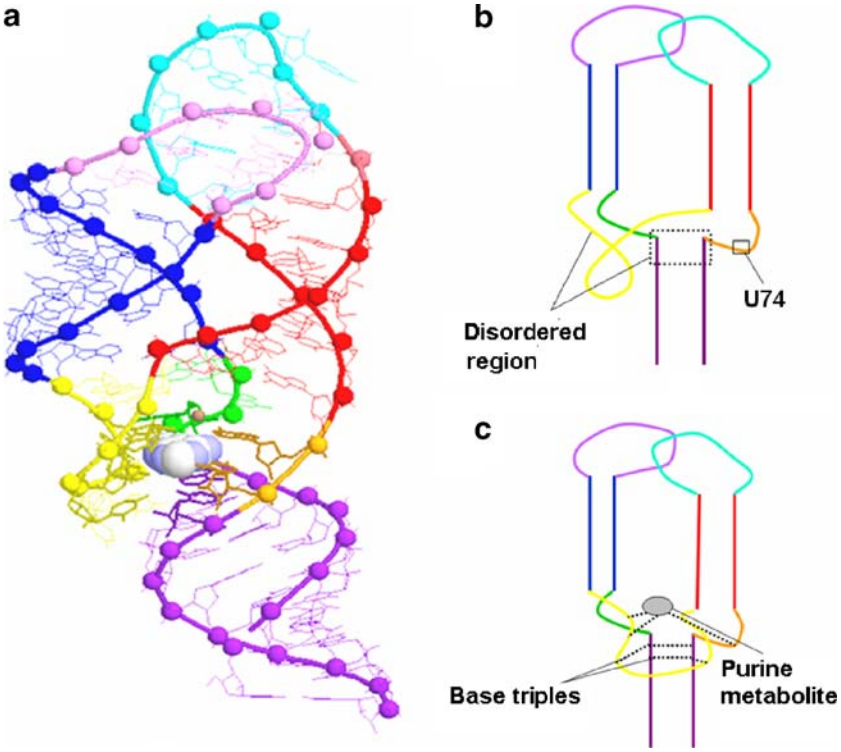

Fig. 2 (**a**) Representation of secondary structure elements of the aptamer domain of purine riboswitch. P1, P2 and P3 are stem regions, L2 and L3 are the terminal loops of P2 and P3 stems respectively. J1/2, J1/3 and J2/3 are the junctions connecting P1 and P2, P1 and P3 and P2 and P3 stems respectively. The phosphorus atoms of each nucleotide (represented by spheres) are joined to trace the backbone topology. The regions are colored in the manner described in Fig. 1. The purine metabolite is shown in spacefill model. (b and c) Schematic representation of the riboswitch in ligand unbound (**b**) and bound (**c**) states. The metabolite is represented as an ellipse. Note the disordered region around the binding pocket and formation of base triples on interaction with the metabolite

of three stems P1, P2, and P3 connected by three loops J1/2, J1/3 and J2/3 forming a three way junction, where the ligand binds (Fig. 1 and Fig. 2a). The stems P2 and P3 are parallel to each other and end in hairpin loops L2 and L3 respectively. The P1 stem consists of the initial stretch beginning from the 5′ end of the riboswitch pairing up with the 3′ end of the aptamer domain. Immediately downstream of the aptamer domain is the expression platform domain of the riboswitch (not shown in figure) which, depending on the specific nature of switching action, consists of characteristic terminator or antiterminator fold (for transcriptional regulation) or alternatively, masked or unmasked stretch of the Shine Dalgarno sequence (for translational regulation) [22]. Evidences gathered so far, indicate that regions ranging from the downstream strand of P3 stem, the J1/3 loop and the P1 stem, present in the metabolite bound structure, interact with stretches in the expression platform in the metabolite unbound state, thus stabilizing a default conformational fold [23, 24]. Metabolite binding leads to stabilization of the aptamer domain in a conformation where its sequence stretch can no longer pair with the expression platform domain. This forces the default fold in the expression platform to switch to an alternative fold. Depending on the specific riboswitch, this may imply switching either from on to off state or vice versa. It is also widely held that in the presence of Mg^{2+} ions, a prerequisite to proper ligand recognition, the binding pocket is mostly preformed [25]. The P2 and P3 stems are involved in several interactions involving base triples, and the loops L2 and L3 participate in a complex loop-loop interaction motif consisting of at least two base quadruples [22]. The binding pocket in the metabolite bound state, has the metabolite surrounded by O2′ of rU22 of J1/2, U47 and U51 of J2/3 and Y(C or U)74 of J1/3. In addition, two base triplets, water mediated triplet U22-A52-A73 and C53-G46-A23 triplet, directly above the bound metabolite, form the roof of the binding pocket. Two more triplets C50-U75-A21 and U49-A76-U20 formed by interaction of J2/3 loop bases with two adjoining base pairs of the P2 stem, form the floor of the binding pocket. The details of the aptamer interactions are shown in Fig. 1 and Fig. 2, while those related to the bound metabolite are shown in Fig. 3. As explained in Fig. 2b, it is held that prior to metabolite binding, the junction J2/3 and the upper two base pairs of stem P1 are in a disordered state [25], and provide the structural link between conformational changes in aptamer domain and that in the expression platform. Metabolite binding involves the movement of U47 and U51 nucleotides of J2/3 junction loop which can assist the latter to form base triples involving U49 and C50 of J2/3 and the upper portion

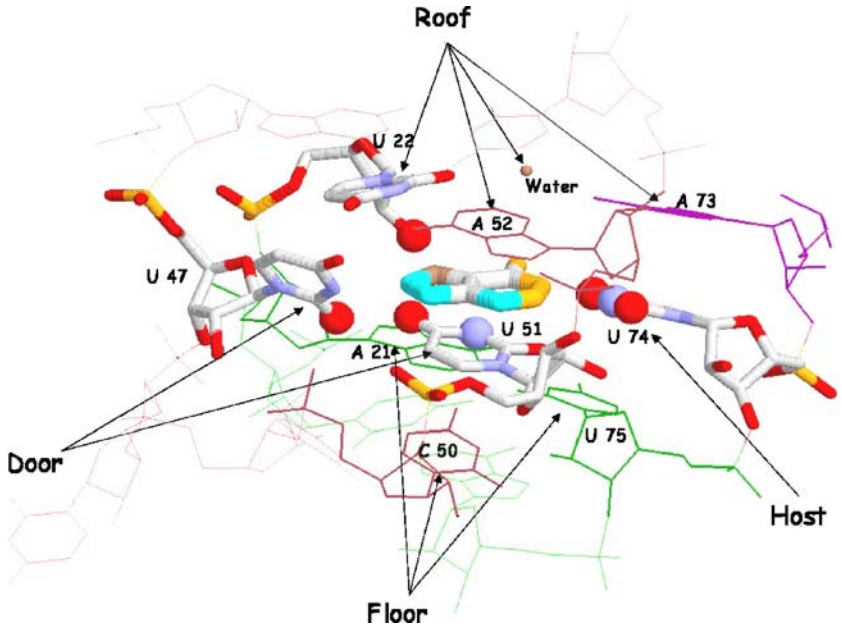

Fig. 3 Representation of interactions in the periphery of the bound metabolite forming roof, door, floor and host interactions around the adenine metabolite. The atoms of the surrounding nucleotides forming hydrogen bonds with the metabolite are shown as spheres. In the guanine metabolite, the uracil at position 74 is replaced by cytosine

of the P1 stem (Fig. 1 and Fig. 2c). This possibly stabilizes the complete P1 stem and allows conformational switching in the expression platform. Analysis of molecular dynamics simulation trajectories also shows that the root mean square fluctuation (RMSF) of the upper part of the P1 and junction J2/3 come down significantly in the metabolite bound state as compared to the unbound state [26].

Adenine and guanine aptamers are highly specific to their respective metabolite binding, despite the fact that the secondary structure of the aptamer domain is similar (RMSD of 1.7 Å before and after superposition of phosphates [27]). The A-riboswitch discriminates against guanine by more than four orders of magnitude [5], whereas the G-riboswitch discriminates against adenine by six orders of magnitude [6]. The question concerning the structural basis of different functionalities and high discrimination among purine riboswitches has been addressed by various groups by carrying out structural, biochemical and genetic (mutational) experiments introducing variations in the aptamer sequences as well as in the structure of the metabolite [22, 27, 28]. Comparison of A- and G-riboswitch sequences revealed that pyrimidine at 74 position (U in case of adenine riboswitch and C in case of guanine riboswitch) of the J1/3 loop is the sole determinant of specificity to ligand binding [22–24]. This base forms a W:W cis canonical interaction with the metabolite. It may be mentioned that in natural RNA structures, we often find G:U W:W cis [29, 30] base pair (commonly known as G:U Wobble base pair) in place of G:C canonical base pair. The G:U interaction energy in W:W cis arrangement is higher than interaction energy of the A:U W:W cis base pair. Thus, in this context, it is counter intuitive that G-riboswitch can be changed to A-riboswitch and vice versa by interconverting C74 and U74 [5].

Apart from the metabolite specificity issue, other larger unresolved questions concerning metabolite binding in purine riboswitches are:

1. What are the essential structural differences between the metabolite free and metabolite bound state?
2. How large and how fast are the changes in the metabolite free to metabolite bound conformations of the aptamer domain and how exactly the aptamer region interacts with the expression platform? As the time gap between initiation of transcription activity of RNA polymerase and binding of metabolite to aptamer domain is very small, it has been suggested that most of the aptamer region is preformed before ligand binding, and the metabolite binding takes place by an induced fit mechanism [25]. Whereas a general consensus about these issues has been arrived at, detailed answers to the above questions are dependant on two other issues which are in turn interrelated with each other. These are:
    1. What are the most significant stabilizing interactions at the metabolite binding site?
    2. What is the importance of different forces, such as Mg^{2+} ions, ordered water molecules, hydrogen bonding and stacking, in the stabilization of the binding pocket and that of ligand within the binding pocket?

Several structural, biochemical, kinetic and thermodynamic studies have been carried out recently in order to

address both the above issues. The reported crystal structures of different purine-aptamer complexes [22] also offer detailed interaction patterns of the aptamer region and its metabolite binding site at the atomic level. In fact, purely from geometric consideration of the crystal structure, it has been suggested that intermolecular hydrogen bonding interactions play a greater role in ligand stabilization as compared to stacking [1, 22]. The authors thereby imply that the stacking interactions provided by the "roof" and "floor" triples of the binding pocket are negligible compared to the hydrogen bonding energies involved in these and those involving the metabolite with U47, U51, rU22 and Y74.

Be that as it may, clearly, the above approaches cannot go beyond a certain point without knowing the exact structure of the disordered or metabolite unbound state. Several experiments have been carried out to differentiate between paired and unpaired regions formed in the solution phase, in the presence of Mg^{2+} ions and metabolite at different respective concentrations in order to address this issue. The major among these are:

1. In line probing assays of patterns of spontaneous RNA cleavage: Given that single stranded regions fragment more easily than paired regions, the differences in fragmenatation pattern shows that the aptamer region undergoes major structural reorganization on ligand binding [5, 6].
2. Single molecule FRET studies demonstrated a mixture of folded and unfolded states of adenine riboswitch lacking loop-loop interactions [31, 32].
3. Imino proton NMR spectra showed that except the U20 and U75 positions of upper two P1 stem base pairs, all stem imino proton resonances are observed both for ligand bound and unbound state [25]. The junction nucleotides in the free state, especially those belonging to J2/3 on the other hand do not give detectable imino proton resonance due to proton exchange with water, suggesting a disordered arrangement in the absence of ligand.
4. Magnetic field induced residual dipolar coupling (mRDCs) based analysis of the imino N-H and C-H mRDCs of the stem bases also suggest that the global fold and stem orientations remain similar in metabolite free and metabolite bound states [25].

From these evidences, it appears that even in the metabolite unbound state, the loops L2 and L3 interact with each other holding the recognition box more or less in position. The junction loops J1/2 and J1/3 comprising of only 3 and 2 bases respectively, are not large enough to allow major geometric changes on ligand binding. The eight base long J2/3 loop however has the potential to act as a raised trapdoor, which on the entry of the metabolite can close the box initiating a cascade of conformational changes which affect the expression platform. According to this mechanism, the U47 and U51 nucleotides of J2/3 loop play a crucial role, forming hydrogen bonding interactions with the metabolite. The C50 and U49 bases of J2/3 also interact with the upper two base pairs of P1 stem forming base triplets, and the U48 base flips out in this process. Particularly notable in this regard is the unique role of U47 which, apart from participating in hydrogen bonding interactions with the U51 and the metabolite, is held in its geometry through a large number of stabilizing interactions within its crystal environment (Fig. 6). The implications of this are discussed in the Results and discussion section.

Detailed sequence analysis and mutational studies, on the activities of adenine and guanine ribowitches, show that the nucleobase at 48 position is not conserved. However, there is exclusion of G48 in guanine riboswitch and A48 in adenine riboswitch [24]. On introducing G and A respectively, the activity has been reported to be lost, possibly because in that case, the J2/3 loop would have the potential to completely sterically block the access of cognate metabolite into the binding pocket because of interaction of the base in position 48 with Y74.

Experimental investigations described above, have outlined the possible geometry of interactions of metabolite and the binding pocket. However, they do not clearly reveal the extent to which each aptamer nucleotide, interacting with the metabolite, contributes to the overall binding energy. In general, X-ray crystallography gives us structural models of macromolecules, but does not provide information regarding energies and stabilities of various structural possibilities. The interpretation of interactions, based purely on X-ray data, can therefore be misleading [33]. A detailed interaction energy characterization of the binding pocket is thus indispensable for a better understanding of the chemistry of aptamer-metabolite interaction. Even the characterization of the metabolite-aptamer interactions in terms of strength and nature of hydrogen bonds can add to such understanding. Such details can be correctly described only by nonempirical theoretical calculations.

Despite the crucial importance of natural RNA aptamers in gene regulation, and the fact that computational techniques can significantly contribute to the understanding of structure and dynamics of biological molecules, very few theoretical studies on metabolite binding to aptamers have been carried out till date. In an effort toward systematically understanding the physicochemical basis of metabolite-aptamer interactions at a desired level of theory, we carried out gas phase quantum chemical studies on the models consisting of purine metabolite (adenine and guanine) and conserved nucleotides of mRNA aptamer that stabilize the metabolite through hydrogen bonding interactions. Though in such models, some important features (such as role of solvent molecules, metal ions etc.) of the real complex biological system are neglected, they allow the accurate determination of major individual contributing effects. In

any case, our approach, which involves comparison of interaction energies and bonding parameters of the optimized geometry with those of the crystal geometry, provides us with definitive clues to the nature and extent of environmental effects which are neglected in our models.

Nucleic acids base pairing has been studied earlier at a variety of theoretical levels including HF/6-31G(d,p) [34–36], B3LYP/6-31G(d,p) [37–41] as well as RIMP2/aug-cc-pVdz [42]. While the computationally least expensive HF/6-31G(d,p) level does not explicitly include electron correlation, the RIMP2/6-31G(d,p) level is computationally too expensive for carrying out optimizations on higher order hydrogen bonded structures, as for example, the geometry of the metabolite and hydrogen bonded bases taken together. Use of density functional theory (Hybrid B3LYP functional along with 6–31G(d,p) basis set of atomic orbitals) for geometry optimization has been shown to give reliable geometries with much reduced computational time. These geometries are reliable enough for understanding details of interbase hydrogen bonding in RNA base pairs [37–42], and correlate very well with the reference RIMP2/aug-cc-pVTZ geometries [42]. A cautionary note is added here. Hybrid density functional (B3LYP) fails to correctly describe the dispersion component of interaction energy, as the nonlocal dispersion energy is not included in local DFT energy [43–45]. However, it has been suggested that the DFT optimizations are biased by the missing dispersion component (resulting in underestimation of total interaction energy), and missing BSSE correction (which overestimates the total interaction energy), both of which by and large compensate each other [43]. This error compensation produces sufficient additional attractive contribution to account for missing dispersion in the case of hydrogen bonded complexes, but is not sufficient in the case of complexes which are stabilized only through dispersion energy, such as stacked nucleobases [43]. These problems with current DFT exchage-correlation functionals have recently been addressed by various groups by designing new density functionals [46–49] reparametrizing the existing functionals [50–52] or by adding empirical dispersion term separately to the DFT energy [53, 54], although most of these codes are still not available in standard quantum chemical packages. However, the accurate interaction energies can only be evaluated with the proper inclusion of the dispersion term.

Methodology

Modeling of the binding pocket

The initial geometries of the adenine and guanine metabolites, along with the corresponding interacting nucleotides of their respective binding pockets, were extracted from crystal structures (PDB code: 1Y26 and 1Y27 respectively). The PDB files contain only the heavy atom coordinates. The coordinates of hydrogen atoms were initially added, on the basis of default bond lengths, to the extracted heavy atom geometries.

The hydrogen added structures were further geometry optimized in two different ways, one with applied constraints freezing the heavy atoms (hydrogen optimization) and the other without any applied constraint (full optimization). Three models (I-III), differing in terms of representation of sugar moeities and level of constraints used for geometry optimization, were prepared for each of the two types of riboswitches: a. Adenine riboswitch and b. Guanine riboswitch (Fig. 4). The constitution and rationale for the models used for our interaction energy studies are described below:

Hydrogen optimized models of metabolite+rU22+U47+U51+Y74: models Ia and Ib

In the crystal geometry, apart from Y74, U51 and U47, the O2′ from U22 also interacts with the metabolite. Thus, though the metabolite does not interact with its nucleobase portion, the U22 nucleoside was included in the model alongwith its ribose sugar. The 5′ and 3′ phosphate groups of rU22 were replaced by hydrogen atoms in order to reduce the complexity of the model as well as computational time required for optimization. For the same reasons, for the other three interacting bases, where the sugar moeities were not interacting with the metabolite, the C1′ atoms of the sugar moeities were also replaced by hydrogen atoms. Such models have been earlier used in literature for studying RNA base pairing and are justified in the present context [38, 39]. The complete hydrogen added initial models were hydrogen optimized as described in Geometry optimizations and are labeled as Ia(adenine binding pocket) and Ib(guanine binding pocket).

Full optimized models of metabolite+U47+U51+Y74+CH_3OH (mimicking 2'-OH of U22: models IIa and IIb)

Apart from hydrogen optimization, full optimization of the models described in Modeling of the binding pocket was also carried out. In the resulting geometries, an additional unreal hydrogen bond is formed between O2 of U22 and N1-H of U47. In order to remove this artifact, we replaced the rU22 with a CH_3OH molecule with its oxygen atom fixed at the crystallographic coordiantes of the O2′ of U22. Addition of hydrogen atoms to the remaining heavy atoms were carried out as usual before subjecting them to a full optimization protocol and the final models were labeled as IIa(adenine binding pocket) and IIb(guanine binding pocket)

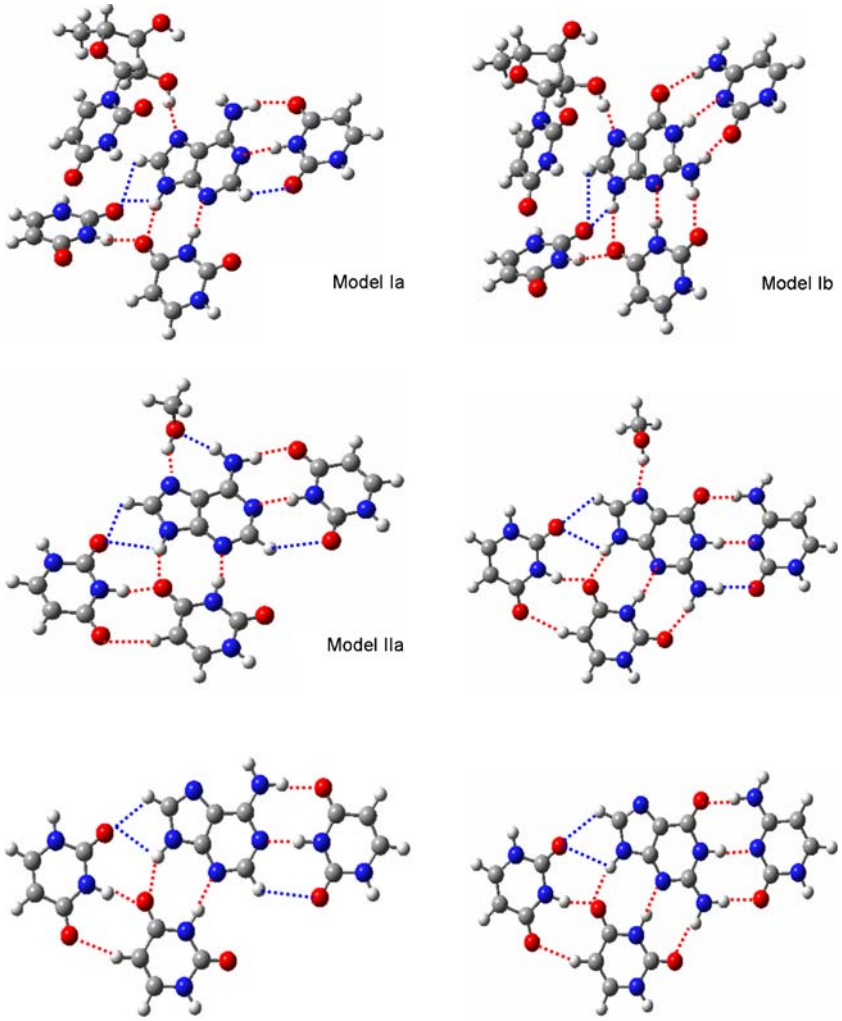

Fig. 4 Optimized geometries of the interaction models at the B3LYP/6-31G(d,p) level. Dashed lines represent the hydrogen bonds. Blue dashed lines represent interactions not assigned previously in crystal geometries

Full optimized models of metabolite+U47+U51+Y74: models IIIa and IIIb

Full optimization of the binding pocket along with the metabolite was also carried out removing all the constraints. U22 was however not taken into account in these models. The final models were labeled as IIIa (adenine binding pocket) and IIIb(guanine binding pocket)

Computational methods

Geometry optimizations

All the models were optimized at the B3LYP/6-31G(d,p) level of theory using Gaussian [55] and Gamess [56] suites of quantum chemical programs. In the first category of models(Ia and Ib), the coordinates of heavy atoms were frozen and hydrogens were allowed to relax during optimization, using IFREEZ option of Gamess, using the same procedure as described in our earlier papers [34–36].

Interaction energy evaluations

Interaction energies of the studied models were decomposed to analyze the contributions of electrostatic, exchange repulsion, polarization, charge transfer and other higher order coupling terms at the HF/6-31G(d,p) level according to Kitaura-Morokuma scheme [57]. Although HF single point energy calculations are less accurate, but they do not introduce fundamental errors into the relative orders of magnitude of interaction energy, which is the primary focus of the present paper, i.e., how much each residue contributes to the overall binding energy. Kitaura-Morokuma energy decomposition scheme is commonly used to understand the intermolecular interactions in biomolecules [36, 58, 59]. Gamess program was used to carry out energy decomposition.

Analysis of noncovalent interactions

Hydrogen bonding interactions between the metabolite and surrounding mRNA bases in the optimized models were

analyzed and characterized using natural bond orbital (NBO) analysis scheme [60] at the B3LYP/6-31G(d,p) level of theory. Atoms in molecules (AIM) analysis [61] using Bader's AIM theory was also carried out to characterize the noncovalent interactions in the binding pocket. NBO and AIM analysis was carried out using the Gaussian package.

Results and discussion

Table 1 describes the geometric parameters corresponding to hydrogen bonding interactions in the studied models consisting of purine metabolite and its surrounding bases. The BSSE corrected interaction energies of these models were calculated and decomposed into components using the Kitaura-Morokuma energy decomposition scheme, and the results are described in Table 2. Percentage contribution of pairwise interacions of individual nucleotides with the metabolite, toward the total binding energy of the metabolite-aptamer complex in each of the models is presented in Table 3. Cooperative effects arising out of interaction of metabolite and surrounding bases are described in Table 4. The interactions between filled donor and unfilled acceptor natural bonding orbitals corresponding to each of the hydrogen bonds in the studied models have been descibed using NBO analysis in Table 5. The electron densities as well as the Laplacian of electron densities at the hydrogen bond critical points (HBCPs) derived from AIM analysis are also described in Table 5.

Interaction energies from hydrogen bonding of metabolites in the binding pocket

Guanine binds more strongly to aptamer binding pocket than adenine

Both adenine and guanine metabolites interact, with the surrounding bases of their respective riboswitch aptamer domains, with reasonably high total interaction energies (E_{int}). However, the interaction energy for guanine binding is much larger than that of adenine (Table 2). The most significant reason for this difference, in interaction energies in the respective purine-aptamer complexes, is the relatively higher contribution of G:C74 binding in the aptamer-guanine interaction, to the overall binding energy, compared to A:U74 interaction in adenine-aptamer complex. In

Table 1 Relevant geometrical parameters for hydrogen bonding interactions in the studied models

	Hydrogen optimized geometry			Fully optimized geometry		
bond	r(D-A) (Å)	r(H...A) (Å)	∠D-H-A (°)	r(D-A) (Å)	r(H...A) (Å)	∠D-H-A (°)
	Model Ia			Model IIa (Model IIIa in parentheses)		
O2′-H(U22)...N7(A)	2.781	1.837	168.86	2.794	1.860	156.90
N6-H(A)...O2′ (U22)	-	-	-	2.910	1.914	163.76
C8-H(A)...O2(U47)	3.309	2.669	118.00	3.154(3.190)	2.455(2.513)	121.32(119.80)
N9-H(A)...O2(U47)	3.260	2.681	116.92	(3.183)	(2.650)	(112.58)
N3-H(U47)...04(U51)	2.709	1.721	167.68	2.800(2.796)	1.774(1.770)	171.20(171.05)
C5-H(U51)...04(U47)	-	-	-	3.401(3.404)	2.360(2.363)	160.44(160.34)
N3-H(U51)...N3(A)	2.769	1.752	176.04	2.863(2.860)	1.815(1.813)	175.87(175.89)
N9-H(A)...04(U51)	2.927	1.978	156.73	2.889(2.916)	1.930(1.971)	155.38(153.10)
N3-H(U74)...N1(A)	2.648	1.627	175.50	2.856(2.852)	1.808(1.803)	179.71(179.93)
N6-H(A)...04(U74)	2.873	1.909	160.91	2.969(2.932)	1.949(1.912)	176.41(175.06)
C2-H(A)...O2(U74)	3.261	2.429	133.34	3.617(3.645)	2.790(2.837)	132.57(131.11)
	Model Ib			Model IIb (Model IIIb in parentheses)		
O2′-H(U22)...N7(G)	2.620	1.660	174.82	2.816	1.845	168.86
C8-H(G)...O2(U47)	3.273	2.709	112.03	3.225(3.315)	2.535(2.617)	120.91(121.72)
N9-H(G)...O2(U47)	3.128	2.501	120.37	3.248(3.322)	2.722(2.780)	112.36(113.45)
N3-H(U47)...04(U51)	2.810	1.832	164.18	2.797(2.802)	1.773(1.776)	171.43(172.31)
C5-H(U51)...04(U47)	-	-	-	3.425(3.393)	2.390(2.356)	158.98(159.57)
N2-H(G)...O2(U51)	2.927	1.951	165.73	3.066(3.053)	2.054(2.042)	173.73(173.76)
N3-H(U51)...N3(G)	2.777	1.757	176.56	2.819(2.813)	1.762(1.754)	178.97(179.24)
N9-H(G)...04(U51)	3.028	2.111	151.44	2.938(2.953)	1.981(1.997)	155.48(155.60)
N4-H(C74)...06(G)	2.957	1.950	176.32	2.793(2.779)	1.751(1.740)	179.51(179.63)
N1-H(G)...N3(C74)	2.919	1.928	166.93	2.927(2.938)	1.892(1.905)	177.09(177.38)
N2-H(G)...O2(C74)	2.745	1.746	175.65	(2.779)	(1.740)	(179.6)

[a]

[a] D stands for hydrogen bond donor atom, A stands for hydrogen bond acceptor. r(D-A) is the distance between donor and acceptor atom forming the hydrogen bond. r(H...A) is the hydrogen bond distance between acceptor and hydrogen atom. D-H-A is the hydrogen bonding angle.

Table 2 Results of interaction energy decomposition from Kitaura-Morokuma Scheme for the studied models

Interaction energy decomposition								
Model	E_{int}	BSSE	IE	E_{elec}	E_{ex}	E_{pol}	E_{CT}	E_{HOC}
Ia	-24.73	8.48	-33.21	-78.17	83.59	-16.86	-26.17	4.41
Ib	-46.43	8.88	-55.31	-98.19	89.64	-22.80	-25.92	1.95
IIa	-36.72	7.59	-44.31	-78.05	69.05	-14.42	-21.10	0.21
IIb	-54.92	6.97	-61.89	-95.23	76.29	-19.91	-22.73	-0.31
IIIa	-28.76	5.16	-33.92	-56.54	48.09	-10.94	-14.91	0.38
IIIb	-46.97	5.82	-52.61	-80.15	63.24	-16.08	-19.19	-0.44

[a] E_{int} is the total interaction energy given by E_{int}=IE+BSSE. BSSE stands for basis set superposition error. IE is the interaction energy excluding BSSE. E_{elec} , E_{ex}, E_{pol}, E_{CT} and E_{HOC} represent the electrostatic, exchange repulsion, polarization, charge transfer and higher order coupling component of interaction energy respectively. All values are expressed in Kcal/mol. IE is decomposed into components accordng to Kitaura-Morokuma Scheme [57].

fact, in the crystal geometry (Model Ia and Ib), the G:C74 interaction energy in case of guanine aptamer is about four times higher than A:U74 interaction energy of adenine aptamer. Moreover, G:U51 interaction with three interbase hydrogen bonds also possess greater interaction energy than A:U51 interaction with two hydrogen bonds.

Relative roles of different interactions in ligand stabilization are dissimilar

In case of guanine aptamer, the canonical interaction between guanine and C74 is the major interaction which helps in stabilization of metabolite in the binding pocket of aptamer, and contributes about 50% of the total binding energy of the aptamer-guanine complex in all three models (Table 3). U51, on the other hand, interacting through its Watson-Crick edge with the sugar edge of guanine, contributes around 30% of the overall binding energy. Contrastingly, in the case of adenine aptamer, noncanonical A:U51 interaction contributes around 50% of the overall binding energy. Canonical A:U74 interaction, being relatively weaker, stands as the second major interaction stabilizing the metabolite in the binding pocket of the aptamer domain. In fact, in the crystal geometry, A:U74 interaction contributes only around 26% to the overall binding energy. This relatively smaller role of U74 is compensated by greater contribution of rU22, which provides about 20% of aptamer-adenine binding energy.

U47 has very low interaction energy with the metabolite

The metabolite:U47 interaction, as calculated in all the models is very small and thus apparently has a very small contribution to the overall binding energy of the aptamer-metabolite complexes. This is not surprising since in our models, we have only considered the interaction energy contribution arising out of hydrogen bonding of U47 with the metabolite. Arguably, U47 is essential to metabolite binding, more in terms of stabilization of the overall complex rather than in terms of its direct contribution through hydrogen bonding with the metabolite.

2'-OH of U22 plays a greater role in the adenine riboswitch

As evident from the crystal structure of adenine and guanine bound aptamers, the interaction of U22 takes place

Table 3 Percentage contribution of pairwise interaction energy of individual nucleotides sur- rounding the metabolite to the total interaction energy of the aptamer-metabolite complex

Base pair	model Ia	model Ib	model IIa	model IIb	model IIIa	model IIIb
Y74-met	26.2(-6.05)	56.3(-24.97)	32.66(-11.58)	51.54(-27.38)	41.2(-11.47)	58.8(-27.38)
U51-met	46.1(-10.65)	30.9(-13.69)	39.26(-13.92)	31.89(-16.94)	49.7(-13.83)	36.6(-17.06)
U47-met	8.1(-1.88)	4.8(-2.11)	6.60(-2.34)	4.1(-2.15)	9.2(-2.55)	4.6(-2.14)
U22-met	19.6(-4.54)	8.1(-3.59)	21.49(-7.62)	12.51(-6.65)	-	-

The actual interaction energy values are given in parentheses.

[a] Y stands for pyrimidine base, which is uracil in adenine binding riboswitch and cytosine in guanine binding riboswitch. met stands for metabolite (Adenine for model Ia, IIa and IIIa and guanine for model Ib, IIb and IIIb). The values in parentheses are the actual contribution of pairwise interaction energies to the total interaction energy in each case.

Table 4 Cooperativity effects revealed through interaction energy calculations for studied models

Model	E_{int} (Kcal/mol)	Sum of pairwise E_{int} (Kcal/mol)	difference (Kcal/mol)
Ia	-24.73	-23.12	-1.61
Ib	-46.43	-44.36	-2.07
IIa	-36.72	-35.46	-1.26
IIb	-54.92	-53.12	-1.80
IIIa	-28.76	-27.85	-0.91
IIIb	-46.79	-46.58	-0.39

[a]

[a] The values in the last column are evaluated as: $(E_{(U22+U47+U51+Y74)\text{-met}})-(E_{(U22\text{-met})} + E_{(U47\text{-met})} + E_{(U51\text{-met})} + E_{(Y74\text{-met})})$ where met stands for adenine and gua-nine metabolites respectively.

through 2′-OH group of its ribose sugar. The base itself forms a canonical hydrogen bonding interaction with neighboring A52, and cannot interact with the metabolite. Therefore, models IIa and IIb were prepared in order to account for effect of 2^1-OH of U22. It is interesting to note that, although the role of U22 is thus limited to providing a single -OH group for hydrogen bonding with the metabolite, the absolute interaction energy is reasonably high. In fact, it is much higher than that for U47 which incidentally also interacts with only one acceptor atom with the metabolite. Another interesting observation is that the interaction of this -OH group of U22 with the metabolite is about 1 Kcal/mol higher with adenine as compared to with guanine. This may be explained for models IIa and IIb in terms of the fact that in case of adenine, the -OH acts as a donor in hydrogen bonding with N7, and also as an acceptor in its interaction with N6. This explanation, however, does not hold for the interaction energy difference in crystal geometry models Ia and Ib, since the additional bond with N6 is missing in Ia (Table 1).

Strength and specificity of metabolite-aptamer interaction depends on nature of interacting donor/acceptor groups

We had earlier proposed that base pairs with two or more N-H...O hydrogen bonds are more rigid than other types of base pairs [36]. The relatively higher values of electron density at hydrogen bond critical points for these bonds in AIM analysis also substantiate this proposition (Table 5), which helps us to understand the strength of aptamer-metabolite interaction. In this context, it is interesting to observe the large number of uracils occuring around the binding site. It is possible that uracils with two carbonyl groups on their Watson-Crick edge can interact with metabolite through formation of strong N-H...O hydrogen bonds, which clearly occur in case of interaction of U47 and U51 with the metabolite, as also in case of interaction of U74 with adenine metabolite. It may be argued that the extra stabilization of guanine binding to the guanine aptamer is also far greater than corresponding values for adenine aptamer because of extra N-H...O bonds present in the former case, both for interaction with Y74 as well as with U51.

Sum of pairwise interaction energies between individual binding pocket bases and the metabolite is not very different from the total interaction energy

The difference between overall interaction energy and sum of individual pairwise interaction energies between individual binding pocket bases and the metabolite for each model indicates very little cooperative effects between aptamer bases surrounding the binding pocket (Table 4). This difference, reflecting the extent of cooperativity, decreases as we move from model I to model II to model III, and is indicative of a significant role of O2′ of U22 in overall cooperativity of ligand binding. However, the magnitude of indicated cooperativity does not seem to justify the massive conformational changes that are triggered off as a result of metabolite binding. This anomaly has been discussed in Greater compactness of the aptamer domain in the crystal structure is indicative of large cooperativity in noncovalent interactions.

Variation in geometry on full optimization is conspicuous

The orientation of ligands in plane with the purine appears to be undergoing a significant change in geometry on free optimization. There are three major observations regarding change in geometry on optimization:

1. The orientation of all interaction bases change on optimization, resulting in decrease in values of base pair parameters such as buckle, propellar etc. (unreported results) indicating increase in planarity of the bases with respect to the metabolite and with respect to each other.
2. Figure 5 clearly shows that the maximum change in geometry is for U47. Quite significantly, it was

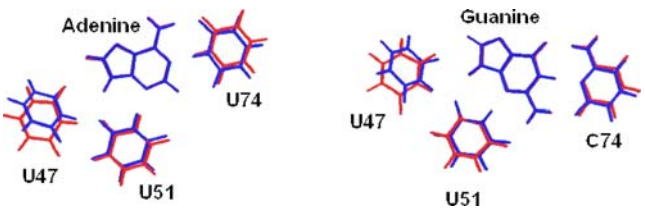

Fig. 5 Superposition of (**a**) model Ia and IIa of the adenine-aptamer complex and (**b**) model Ib and IIb of the guanine-aptamer complex. In each case, the metabolite molecules are superposed and the dispositions of the respective hydrogen bases are compared. The structures in blue correspond to the crystal geometry, while the red ones represent the optimized geometry. The position of rU22 is not shown in the figure. The maximum variaion is observed for U47 (see text)

nonplanar with respect to metabolite in the crystal geometry. On optimization, it undergoes a major shift in geometry, gaining greater planarity.

3. If we make an average of all the metabolite-aptamer hydrogen bond distances in the optimized geometry, we consistently find that the average bond distance increases as we move from crystal geometry models of category I to optimized geometry models II and III. These variations clearly arise due to isolation of our model system from its native crystal environment, and have been discussed in Difference between crystal and optimized geometry provides insights into the significance of environment factors.

Morokuma decomposition of interaction energies

Significance of the components

The relative balance of interaction energy components plays a significant role in stabilization of hydrogen bonded RNA systems [36]. To obtain a clearer view of aptamer metabolite interactions, Kitaura-Morokuma energy decomposition was carried out to analyze the contribution of electrostatic, exchange repulsion, charge transfer, polarization and higher order coupling terms. In case of noncovalent interactions, the total interaction energy has been shown to be largely determined by the attractive electrostatic energy term (E_{elec}) and the repulsive exchange repulsion term (E_{ex}) [52, 62], both of which generally have larger values compared to the contribution from other components. In case of hydrogen bonding, in general both E_{elec} as well as E_{ex} increase as the donor and acceptor comes close to each other. The optimum geometry is defined by the net balance of the two terms- one positive and the other negative. The polarization (E_{pol}) and charge transfer terms (E_{CT}) describe the interaction between filled and unfilled orbitals of the complex. Whereas polarization arises from mixing of occupied and empty orbitals of the same fragments, the charge transfer term represents the mixing of orbitals of one fragment with the other in a complex.

Morokuma decomposition shows higher exchange repulsion component in crystal geometry

As mentioned above, the optimum geometry and the corresponding interaction energy of hydrogen bonded systems are determined by the variation in E_{elec} and E_{ex} with donor-accpetor distances. Morokuma decomposition analysis of interaction energies in our models inevitably showed lower total energies and higher E_{ex}, for the crystal geometry models Ia and Ib, when compared to those for the optimized models of categories II and III respectively (Table 3). Thus, inspection of energy components for all the studied models suggests that electrostatic components are the largest attractive component in all the models. The exchange repulsion, the only repulsive component of interaction energy is relatively smaller in magnitude compared to E_{elec} in the optimized models of category II and III. This observation ties up neatly with the fact that, in our studies on geometry variation on optimization, we inevitably found that average hydrogen bonding distances were lower in the crystal geometry a compared to those in the corresponding optimized geometry [36].

E_{pol} dominates in guanine-aptamer complex

Charge transfer and polarization components also play a significant role in aptamer- metabolite interaction. In general, the magnitude of charge transfer interaction is greater than the polarization term. However, the polarization forces seem to be more dominating in the guanine-aptamer interaction. The main reason for this seems to be the high dipole moment of guanine and C74. In case of adenine-aptamer interaction, the relatively smaller dipole moments of adenine as well as that of uracil result in a complex with smaller dipole moments where individual monomers are less polarized. On the other hand, the G:C interaction (where the dipole moments of both G and C are mutually aligned) results in a complex with relatively higher dipole moment, which results in increased contribution of polarization term to the total interaction energy. In our earlier studies also, we observed higher contribution of polarization term in case of base pairs involving cytosine as one of the monomer [36]. The high contribution of polarization term to the pairwise C74-G interaction energy also substantiates the fact that high dipole moment of cytosine is responsible for greater contribution of E_{pol} to the total interaction energy.

NBO and AIM analysis of hydrogen bonding interactions of surrounding aptamer nucleotides with the metabolite

Significance of the calculated parameters

Analysis of strength and nature of noncovalent interactions is an important aspect in the study of hydrogen bonded complexes. In order to characterize the hydrogen bonds in the aptamer-metabolite complexes, natural bond orbital (NBO) analysis and atoms in molecules (AIM) analysis have been carried out. Using the natural Lewis type picture of chemical bonding, natural bond orbital (NBO) analysis looks at the hydrogen bonding in terms of charge transfer

Table 5 Results of natural bond orbital (NBO) analysis and atoms in molecules (AIM) analysis of hydrogen bonding interactions of purine metabolites with aptamers in the studied models

Model	hydrogen bond	NBO analysis			AIM analysis	
		occupancy (e)			ρ (e/a^3_0)	∇^2 (e/a^5_0)
		n(A)	σ* (D-H)	E(2) (Kcal/mol)		
Ia	O2′ -H(U22)...N7(A)	1.892(1.927)	0.054(0.014)	22.39	0.038	0.096
	C8-H(A)...O2(U47)	1.973(1.977)	0.020(0.019)	0.30	0.006	0.024
	N9-H(A)...O2(U47)	1.973(1.977)	0.040(0.017)	0.57	0.006	0.025
	N3-H(U47)...04(U51)	1.939(1.977)	0.056(0.015)	14.88	0.040	0.139
	N3-H(U51)...N3(A)	1.886(1.913)	0.079(0.015)	32.11	0.047	0.112
	N9-H(A)...04(U51)	1.939(1.977)	0.041(0.017)	6.49	0.024	0.068
	N3-H(U74)...N1(A)	1.845(1.908)	0.099(0.015)	45.57	0.064	0.135
	N6-H(A)...04(U74)	1.870(1.977)	0.041(0.010)	8.48	0.028	0.082
	C2-H(A)...02(U74)	1.848(1.977)	0.027(0.026)	1.16	0.011	0.035
Ib	O2′ -H(U22)...N7(G)	1.869(1.927)	0.079(0.014)	38.62	0.035	0.280
	C8-H(G)...O2(U47)	1.974(1.977)	0.020(0.017)	0.13	0.005	0.024
	N9-H(G)...O2(U47)	1.974(1.976)	0.039(0.019)	0.88	0.007	0.034
	N3-H(U47)...04(U51)	1.950(1.977)	0.046(0.015)	10.19	0.024	0.135
	N2-H(G)...O2(U51)	1.845(1.976)	0.032(0.007)	6.60	0.021	0.096
	N3-H(U51)...N3(G)	1.853(1.899)	0.082(0.015)	32.93	0.032	0.209
	N9-H(G)...04(U51)	1.950(1.977)	0.034(0.019)	4.34	0.015	0.062
	N4-H(C74)...06(G)	1.863(1.978)	0.044(0.011)	8.49	0.019	0.094
	N1-H(G)...N3(C74)	1.869(1.895)	0.055(0.019)	18.71	0.025	0.122
	N2-H(G)...02(C74)	1.853(1.976)	0.051(0.006)	15.14	0.030	0.184
IIa	O2′ -H(U22)...N7(A)	1.894(1.927)	0.053(0.014)	21.34	0.026	0.148
	N6-H(A)...O2′ (U22)	1.930(1.975)	0.042(0.010)	13.74	0.023	0.113
	C8-H(A)...O2(U47)	1.973(1.977)	0.019(0.019)	1.00	0.009	0.036
	N3-H(U47)...04(U51)	1.866(1.977)	0.063(0.015)	15.03	0.027	0.164
	N3-H(U51)...N3(A)	1.866(1.913)	0.080(0.015)	29.03	0.028	0.170
	N9-H(A)...04(U51)	1.940(1.977)	0.044(0.017)	9.63	0.021	0.098
	N3-H(U74)...N1(A)	1.860(1.908)	0.081(0.015)	29.55	0.029	0.176
	N6-H(A)...04(U74)	1.862(1.977)	0.041(0.010)	9.11	0.020	0.094
	C2-H(A)...O2(U74)	1.842(1.977)	0.024(0.026)	0.44	0.004	0.019
IIb	O2′ -H(U22)...N7(G)	1.892(1.927)	0.062(0.014)	22.70	0.026	0.153
	C8-H(G)...O2(U47)	1.974(1.977)	0.018(0.017)	0.75	0.007	0.031
	N9-H(G)...O2(U47)	1.847(1.976)	0.040(0.019)	0.54	0.005	0.024
	N3-H(U47)...04(U51)	1.865(1.977)	0.061(0.015)	14.56	0.027	0.164
	N2-H(G)...O2(U51)	1.843(1.976)	0.029(0.007)	6.05	0.017	0.071
	N3-H(U51)...N3(G)	1.845(1.899)	0.092(0.015)	34.89	0.031	0.201
	N9-H(G)...04(U51)	1.943(1.977)	0.040(0.019)	7.92	0.019	0.084
	N4-H(C74)...06(G)	1.848(1.978)	0.065(0.011)	18.08	0.027	0.172
	N1-H(G)...N3(C74)	1.861(1.895)	0.065(0.019)	22.36	0.026	0.134
	N2-H(G)...O2(C74)	1.849(1.976)	0.042(0.006)	10.78	0.022	0.107
IIIa	C8-H(A)...O2(U47)	1.974(1.977)	0.019(0.019)	0.78	0.008	0.032
	N9-H(A)...O2(U47)	1.974(1.977)	0.041(0.017)	0.69	0.006	0.027
	N3-H(U47)...04(U51)	1.865(1.977)	0.063(0.015)	15.76	0.038	0.112
	C5-H(U51)...04(U47)	1.864(1.977)	0.019(0.009)	2.23	0.012	0.034
	N3-H(U51)...N3(A)	1.866(1.913)	0.080(0.015)	29.19	0.041	0.092
	N9-H(A)...04(U51)	1.943(1.977)	0.041(0.017)	8.51	0.024	0.070
	N3-H(U74)...N1(A)	1.859(1.908)	0.082(0.015)	30.25	0.043	0.094
	N6-H(A)...04(U74)	1.862(1.977)	0.044(0.010)	10.23	0.028	0.077
	C2-H(A)...O2(U74)	1.975(1.977)	0.024(0.026)	0.36	0.005	0.017
IIIb	C8-H(G)...O2(U47)	1.973(1.976)	0.019(0.017)	0.58	0.006	0.025
	N9-H(G)...O2(U47)	1.973(1.976)	0.034(0.019)	0.44	0.005	0.020
	N3-H(U47)...04(U51)	1.864(1.977)	0.061(0.015)	14.71	0.037	0.111
	C5-H(U51)...04(U47)	1.865(1.977)	0.020(0.009)	2.25	0.012	0.035
	N2-H(G)...O2(U51)	1.844(1.976)	0.030(0.007)	6.22	0.021	0.057

Table 5 (continued)

Model	hydrogen bond	NBO analysis occupancy (e)			AIM analysis	
		n(A)	σ* (D-H)	E(2) (Kcal/mol)	ρ (e/a$^3_0$)	∇^2 (e/a$^5_0$)
	N3-H(U51)...N3(G)	1.843(1.899)	0.095(0.015)	36.05	0.048	0.100
	N9-H(G)...04(U51)	1.944(1.977)	0.039(0.019)	7.52	0.023	0.065
	N4-H(C74)...06(G)	1.847(1.978)	0.069(0.011)	19.93	0.041	0.117
	N1-H(G)...N3(C74)	1.862(1.895)	0.063(0.019)	21.44	0.034	0.080
	N2-H(G)...O2(C74)	1.849(1.976)	0.041(0.006)	10.21	0.027	0.074

[a]

[a] n(A) is the donor NBO (lone pair of electrons localized on acceptor atom in this case) and σ* (D-H) is the acceptor NBO (antibonding orbital of D-H bond associated with hydrogen bond formation in each case. The occupancies of n(A) and σ* (D-H) are measured in terms of number of electrons. The corresponding values in parentheses are for free monomeric form, and are given for comparison. The occupancy of the donor NBO (n(A)) decreases and the occupancy of the acceptor NBO (σ* (D-H)) increases as compared to the values for the monomer written in parentheses. E(2) is the second order stabilization energy gained after interaction between the donor NBO and acceptor NBO (measured in Kcal/mol). ρ stands for electron density at hydrogen bond critical point (measured in atomic units) and ∇^2 ρ is the Laplacian of electron density at hydrogen bond critical point (measured in atomic units).

interactions from lone pair orbital of the acceptor atom (n(A)) to the antibonding orbital of the corresponding donor-hydrogen bond (σ*(D-H)). The nonbonding orbital of the acceptor possesses smaller occupancy on complex formation compared to the corresponding value in the free monomeric form. Similarly, the σ*(D-H) orbital gains electron density and its occupancy increases on complex formation. The magnitude of stabilization energy (E(2)) associated with the delocalization of electron density as a result of this charge transfer indicates the extent of interaction between the orbitals involved in hydrogen bond formation. These characteristics for all the hydrogen bonding interactions in our studied models are described in Table 5.

Atoms in molecules calculations with evaluation of electron densities at hydrogen bond critical points (HBCPs) have also been carried out to characterize the hydrogen bonding interactions in these complexes. It is generally accepted that the formation of a hydrogen bond is associated with the appearance of HBCP between the hydrogen atom and the acceptor atom [61]. Being a closed shell interaction, hydrogen bonding is characterized by relatively low value of charge density (ρ) at HBCP and positive value of Laplacian of charge density ($\nabla^2\rho$). In general, for hydrogen bonded complexes, the values of ρ and $\nabla^2\rho$ lie between 0.002–0.340 a.u. and 0.016–0.130 a.u. respectively [63]. The values of ρ for the aptamer-metabolite hydrogen bonds are found to lie between 0.006–0.041 a.u. for N-H...O hydrogen bonds, 0.025–0.064 a.u for N-H...N hydrogen bonds and 0.004–0.011 a.u for C-H...N/O hydrogen bonds, The values of $\nabla^2\rho$ lie between 0.017–0.280 a.u. for the studied hydrogen bonds, and are within the proposed limits for hydrogen bonding interactions.

A:U74 interaction possess extra C-H...O bond

A noteworthy point is the presence of C2-H(A)...O2(U74) interaction in crystal geometry of adenine aptamer which becomes weaker in the optimized models IIa and IIIa (Table 1 and Table 5). It has been earlier proposed that in the crystal geometries, A:U W:W cis interaction possesses three hydrogen bonds including one C2-H(A)...O2(U) hydrogen bond [30]. This interaction is not earlier assigned in the crystal geometry of the binding pocket. The presence of this interaction is supported by the ρ and $\nabla^2\rho$ values at HBCPs in these models. It may be argued that the presence of stronger C-H...O interaction in crystal also increases the open angle of the A:U47 interaction than in the optimized geometry (results not shown). The higher deformation of base pair in its crystal geometry allows the strengthening of C-H...O interaction.

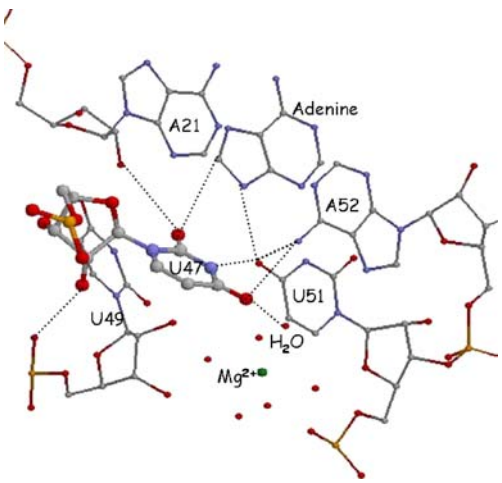

Fig. 6 Different types of interactions around U47 in the crystal environment

U47-metabolite interaction shows bifurcated hydrogen bonding pattern

There is a controversy on whether U47 forms a hydrogen bond with metabolite or not. Sergnov et al. assigned a N9-H (metabolite)...O2(U47) hydrogen bond on the basis of crystal structures [22]. However, this interaction was ruled out in later studies arguing that N9-H acts as a donor for hydrogen bond with O4 of U51 [1]. AIM analysis of hydrogen bonds involving U47, U51 and metabolite shows the presence of HBCPs for both N9-H(met)...O2(U47) as well as N9-H(met)...O4(U51) hydrogen bonds (Table 5). Additionally critical point corresponding to C8-H(met)...O4 (U47) interaction of U47 with metabolite and C5-H(U51)... O4(U51) is also detected in the optimized geometries (models of categories II and III) in the absence of crystal environment. The 2′O-H(U22)...N7(metabolite) interaction also seems to be strong enough in case of crystal geometries, with ρ value of 0.038 a.u. for interaction with adenine and 0.035 a.u. for interaction with guanine.

NBO analysis of hydrogen bonds

Analysis of E(2) values suggest that the charge transfer interaction for C2-H(A)...O2(U74) is greater in crystal geometries than in optimized geometries. On the other hand, E(2) value for N6-H(A)...O4(U74) interaction is smaller in crystal geometry than in the optimized geometry. This is consistent with our earlier proposition that in the crystal geometry, the base pair get deformed in order to form this C-H...O interaction. In general, it is seen that E(2) values for N-H...N hydrogen bonds are greater than N-H... O hydrogen bonds as seen earlier in the case of RNA base pairs [36]. Since N is less electronegative than O, and can participate in charge transfer with greater ease.

Difference between crystal and optimized geometry provides insights into the significance of environmental factors

In Variation in geometry on full optimization is conspicuous, we observed a relative change in geometries of the bases on optimization of the crystal geometry. We had observed this phenomenon earlier in our studies on noncanonical RNA base pairs [36] where we suggested a possible role of environmental effects, including that of dispersion energies. We also explained the high buckle and propellar, in base pairs in crystals, in terms of minimization of exchange repulsion at smaller donor-acceptor distances. The fact that optimization of geometry of models, isolated from the crystal context, led to both greater planarity and larger donor-acceptor distances and was accompanied by an increase in interaction energies, clearly points toward a significant role of other stabilizing influences present in the crystal.

The orientation of U47 base undergoes a major shift on free optimization, as compared to the rest of the bases. Clearly, this is an effect of isolating our model system from its native crystal environment. In this context, the interactions with U47 present in the crystal (Fig. 6) and our observed behavior of U47 on free optimization (Fig. 5) is noteworthy. Analysis of the crystal environment shows a number of interactions of U47, which are neglected in our models. The major interactions present in the crystal around U47 are: (a) Presence of Mg^{2+} ion (id: 101) in the vicinity of U47. One of the water molecules present in the hydration sphere of this Mg^{2+} interacts with the O4 atom of U47. The oxygen of this water molecule is present at a distance of 2.6 Å from O4 atom of U47 in the crystal geometry. (b) Interaction of the amino group of A52 with O4 atom of U47 with a D-A distance of 2.9 A. (c) Interaction of O2 of U47 with 2′-OH group of A21. (d) Van der Waals interaction between nonbonded phosphate group of U49 with 2′-OH group of U47 [1] (Fig. 6). In the crystal context, these interactions hold the uracil away from its normally observed W:H trans geometry [29, 30]. On optimization of the model in the absence of these effects, the standard U:U W:H trans geometry is restored. Free optimization of the model not only helps U47 to attain planar geometry, it also enhances its interaction with U51 by forming an additional C5-H(U51)...O4(U47) hydrogen bond (Table 1 and 5). To summarize, we find a substantial movement in position of U47. The possible implications of this movement to the ligand unbound conformation of junction loop J2/3 and consequent facilitation of the entrance of metabolite into the possibly preformed A/G box can be rationalized as follows:

U47 forms a bad base pairing geometry with the metabolite. In the absence of crystal environment, it reverts to a W:H trans geometry with U51. It is suggested that U51 bonds strongly to the metabolite and U47 strongly paired as above, accompanies it. This helps in conformational changes in J2/3. Among other advantages, the positioning of U47 possibly facilitates the flipping out of the base in position U48 by stabilizing the sharp turn in the phosphate backbone through large number of interactions with its environment [1]. It is reasonable to speculate that the flipping out of base U48, and the consequent conformational change in the backbone might be essential for allowing the nucleotides 49 and 50 to bind with the upper two base pairs of the P1 stem. It has been discussed earlier how these two base triples thus formed, may hold the key to the switching observed in the expression platform. It may be mentioned here that though some of the weak interactions are not properly reflected in the molecular mechanics force fields, the essential observation described above

seems to be validated by our preliminary molecular dynamics studies on add-riboswitches [26].

Specificity of metabolite interaction may be attributed to geometry of binding pocket

In earlier reports, explanation for high discrimination ability and specificity of adenine and guanine riboswitches toward their respective metabolites of A-box against guanine has been attributed only to the strength of canonical metabolite: Y74 base pair recognition [22]. We feel that this explanation is not sufficient. It is well established that the interaction energy of G:U W:W cis base-pair is nearly half of the corrosponding value for G:C W:W cis interaction. If the guanine metabolite binds to the adenine aptamer, it will thus be stabilized by a much weaker G:U74 W:W cis interaction [29, 30], which though much less than that of G:C W:W cis, has higher interaction energy than A:U74 W:W cis interaction energy. Thus, a mere decrease of interaction energy with Y74 cannot explain why guanine metabolite does not bind to the A-box of the adenine riboswitch. We feel that the reason for selectivity clearly lies with the Wobble geometry of W:W cis interaction of G:U74, which can disturb the geometry of U47-U51 intiated loop closure. Here the geometry of base pairing is more significant than interaction energy and is reminiscent of the importance of isostericity of base pairs for covariaion in RNA structures [30]. Along the same lines, binding of adenine to guanine aptamer may result in unfavorable A:C W:W cis geometry which is unstable due to the absence of suitable donor-acceptor interactions necessary for hydrogen bond formation.

Greater compactness of the aptamer domain in the crystal structure is indicative of large cooperativity in noncovalent interactions

We have shown from our studies (Variation in geometry on full optimization is conspicuous) that the models Ia and Ib are more compact than their geometry optimized counterparts among models II and III. It is fairly well established that positively cooperative binding involving noncovalent interaction between a ligand and a macromolecule, such as an enzyme, leads to compaction of the macromolecule [64]. This is associated with the phenomenon of enthalpy-entropy compensaion in molecular recognition processes. Extension of this principle can be used to interpret the compactness of the geometry of crystal structures. However, we have shown in Sum of pairwise interaction energies between individual binding pocket bases and the metabolite is not very different from the total interaction energy that the cooperativity in binding, as observed in our models, does not appear to be very high. The magnitude of resultant conformational changes required for the switching should require greater cooperative effects. From the discussions in Difference between crystal and optimized geometry provides insights into the significance of environmental factors and Specificity of metabolite interaction may be attributed to geometry of binding pocket above, we can infer that larger cooperativity arises through formation of extra hydrogen bonding networks which get stabilized as a result of binding.

As mentioned earlier, apart from its interaction with U51 and metabolite, U47 also interacts with A21, A52, U49 and water molecule in the hydration shell of the surrounding magnesium ion while participating in the closing J2/3 trap by sealing the binding pocket from side. Also, there are two major interactions present above and below the metabolite in the crystal. They are:

1. A water mediated triplet at the top comprising of U22, A52 and A73 nucleotides along with another triplet immediately above it, forms the roof of the binding pocket (Fig. 3).
2. A triplet comprising of A21, U75 and C50 forms the floor of binding pocket along with another triplet below it (Fig. 1). It may be mentioned here that the importance of these stacking interactions has earlier been undermined in favor of hydrogen bonding with the conserved nucleotides of aptamer domain [1]. However, the argument was based on parallely displaced geometry of metabolite with aptamer as observed in the X-ray crystal structures. It is important to note that other bases in the periphery of metabolite are ordered and stacking interactions from above and below play an important role in their ordered arrangement in the pocket.

Conclusion

Adenine and guanine riboswitches, belonging to smallest and phylogenetically highly conserved family of purine riboswitches, have been studied with the help of interaction energy evaluations, Morokuma energy decomposition, NBO analysis and AIM analysis. Several questions relating to the strucutral basis of riboswitch functioning, including their discriminative ability and their exact mechanism of switching have been addressed and insights have been gained. In addition to analysis of geometry, we have addressed the question of specificity of metabolite binding, and have approached the problem through ab-initio studies of the metabolite binding pocket of adenine and guanine aptamers. Some hitherto unreported hydrogen bonds have been detected and their viability has been validated with the help of NBO and AIM analysis. On the basis of variation in geometry on free optimizaion of the crystal geometry, we

have indirectly assessed the possible role of environmental factors in the crystal context.

On the basis of Morokuma decomposition, we have proposed the possibility of extensive cooperativity leading to closer approach of conserved bases to the metabolite in the crystal, as compared to when fully optimized. In conjunction with basic principles of positive cooperativity in ligand binding processes, our studies have evaluated the importance of weakly binding U47 in molecular recognition and switching of the riboswitches. Our studies, and analysis of results obtained, clearly highlight the crucial importance of detailed quantitative understanding of different types of noncovalent interactions in the context of molecular recognition and switching processes.

In the future, we wish to carry out further quantitative estimates of dispersion energy due to stacking and electrostatics of Mg^{2+} binding.

Acknowledgements AM wishes to thank Department of Biotechnology (DBT), Government of India for research grant (NO. BT/PR5451/BID/07/111/2004. PS and SS thank the CSIR, New Delhi for research fellowships. Travel grants from DBT (for PS) and DST (for SS) for attending MDMM-2008, Piechwice, Poland are gratefully acknowledged.

References

1. Gilbert SD, Stoddard CD, Wise SJ, Batey RT (2006) J Mol Biol 359:754–768
2. Nudler E, Mironov AS (2004) Trends Biochem Sci 29:11–17
3. Mandal M, Breaker RR (2004) Nat Rev Mol Cell Biol 5:451–463
4. Winkler W, Nahvi A, Breaker RR (2002) Nature 419:952–956
5. Mandal M, Boese B, Barrick JE, Winkler WC, Breaker RR (2003) Cell 113:577–586
6. Mandal M, Breaker RR (2004) Nat Struct Mol Biol 11:29–35
7. Grundy FJ, Lehman SC, Henkin TM (2003) Proc Natl Acad Sci USA 100:12057–12062
8. Sudarsan N, Wickiser JK, Nakamura S, Ebert MS, Breaker RR (2003) Genes Dev 17:2688–2697
9. Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, Ruzzo WL (1992) Ann Rev Phys Chem 43:257–282
10. Nahvi A, Sudarsan N, Ebert MS, Zou X, Brown KL, Breaker RR (2002) Chem Biol 9:1043–1049
11. Mironov A, Gusarov I, Rafikov R, Lopez LE, Shatalin K, Kreneva RA, Perumov DA, Nudler E (2002) Cell 111:747–756
12. Winkler W, Cohen-Chalamish S, Breaker RR (2002) Proc Natl Acad Sci USA 99:15908–15913
13. Winkler W, Nahvi A, Breaker RR (2002) Nature 419:952–956
14. McDaniel BAM, Grundy FJ, Artsimovitch I, Henkin TM (2003) Nat Struct Biol 10:701–707
15. Epshtein V, Mironov AS, Nudler E (2003) Proc Natl Acad Sci USA 100:3083–3088
16. Winkler W, Nahvi A, Sudarsan N, Barrick JE, Breaker RR (2003) Nat Struct Biol 10:701–707
17. Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, Collins J, Lee M, Roth A, Sudarsan N, Jona I (2004) Proc Natl Acad Sci USA 101:6421–6426
18. Sudarsan N, Hammond MC, Block KL, Welz R, Barrick JE, Roth A, Breaker RR (2006) Science 314:300–304
19. Lea CR, Piccirilli JA (2007) Nat Chem Biol 3:16–17
20. Batey RT, Gilbert SD, Montage RK (2004) Nature 432:411–415
21. Ebbole DJ, Zalkin H (1987) J Biol Chem 262:8274–8287
22. Serganov A, Yuan Y, Pikovskaya O, Polonskaia A, Malinina I, Phan AT, Hobartner C, Micura R, Breaker RR, Patel DJ (2004) Chem Biol 11:1729–1741
23. Mulhbacher J, Lafontaine DA (2007) Nucleic Acids Res 35:5568–5580
24. Lemay J-F, Lafontaine DA (2007) RNA 13:39–350
25. Ottink OM, Rampersad SM, Tssari M, Zaman GJR, Heus HA, Wijmenga SS (2007) RNA 13:2202–2212
26. manuscript under preparation
27. Gilbert SD, Mediatore SJ, Batey RT (2006) J Am Chem Soc 128:14214–14215
28. Gilbert SD, Love CE, Edwards AL, Batey RT (2007) Biochemistry 46:13297–13309
29. The Leontis and Westhof nomenclature for nucleic acids base pairing is employed using the following format: Base1:Base2 (A/U/G/C) Edge1:Edge2 (W/H/S) Glycosidic bond orienta- tion (Cis/Trans) where A, U, G, C, W, H and S stand for adenine, uracil, guanine, cytosine, Watson-Crick edge, Hoogsteen edge and sugar edge respectively. See Ref. [30] for details
30. Leontis NB, Westhof E (2001) RNA 7:499–512
31. Batey RT, Gilbert SD Chem Biol 13:805–807
32. Lemay J-F, Penedo JC, Tremblay R, Lilley DMJ, Lafontaine DA (2006) Chem Biol 13:857–868
33. Šponer J, Hobza P (2003) Collect Czech Chem Comm 68:2231–2282
34. Bhattacharyya D, Koripella SC, Mitra A, Rajendran VB, Sinha B (2007) J Biosci 32:809–825
35. Sharma P, Mitra A, Sharma S, Singh H (2007) J Chem Sci 119:525–531
36. Sharma P, Mitra A, Sharma S, Singh H, Bhattacharyya D (2008) J Biomol Struct Dyn 25:709–732
37. Sharma P, Singh H, Mitra A (2008) Lec Notes in Comp Sci 5102:379–386
38. Šponer JE, Spackova N, Leszczynski J, Šponer J (2005) J Phys Chem B 109:11399–11410
39. Šponer JE, Spackova N, Kulhanek P, Leszczynski J, Šponer J (2005) J Phys Chem A 109:2292–2301
40. Šponer JE, Spackova N, Sychrovsky', Leszczynski J, Šponer J (2005) J Phys Chem B 109:18680–18689
41. Roy A, Panigrahi S, Bhattacharyya M, Bhattacharyya D (2008) J Phys Chem B 112:3786–3796
42. Šponer J, Jurečka, Hobza P (2004) J Am Chem Soc 126:10142–10151
43. Černý J, Hobza P (2007) Phys Chem Chem Phys 9:5291–5302
44. Kristyan S, Puly P (1994) Chem Phys Lett 229:175–180
45. Hobza P, Šponer J, Reschel T (1995) J Comput Chem 16:1315–1325
46. Hesselmann A, Jansen G, Schutz M (2005) J Chem Phys 122:014103
47. Misquitta AJ, Jeziorski B, Szalewicz K (2003) Phys Rev Lett 91:033201
48. Kohn W, Meir Y, Makarov DE (1998) Phys Rev Lett 80:4153–4156
49. Dion M, Rydberg H, Schroder E, Langreth DC, Lundqvist BI (2004) Phys Rev Lett 92:246401
50. Grimme S (2006) J Chem Phys 124:034108
51. Antony J, Grimme S (2006) Phys Chem Chem Phys 8:5287–5293
52. Grimme S (2006) J Comput Chem 27:1787–1799
53. Grimme S (2004) J Comput Chem 25:1463–1473
54. Morgado C, Vincent MA, Hillier IH, Shan X (2007) Phys Chem Chem Phys 9:448–451
55. Frisch MJ et al (2003) Gaussian03 revison B.05 Gaussian Inc Pittsburgh PA

56. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen J, Koseky S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupius M, Montgomery JA (1993) J Comput Chem 14:1347–1363
57. Kitaura K, Morokuma K (1976) Int J Quantum Chem 10:325–331
58. Wang J, Gu J, Leszczynski (2005) J Phys Chem B 109:13761–13769
59. Wang J, Gu J, Leszczynski (2006) J Biomol Struct Dyn 24:139–148
60. Reed AE, Schleyer PvR (1987) J Am Chem Soc 109:7362–7373
61. Bader RFW (1990) Atoms in molecules. A quantum theory. The Clarendon Press, Oxford
62. Grimme S, Antony J, Schwabe T, Lichtenfeld CM (2007) 0rg Biomol Chem 5:741–758
63. Kolandaivel P, Nirmala V (2004) J Mol Struct 694:33–38
64. William S, Dudley H, Stephens E, O'Brien DP, Min Z (2004) Angew Chem 43:6596–6616