Supporting information for Evidence for hidden involvement of N3 protonated guanine in RNA structure and function

Antarip Halder [a],† Saurabh Vemuri,† Rohit Roy,† Jayanth Katuri,† Dhananjay Bhattacharyya,‡ and Abhijit Mitra*,†

†*Center for Computational Natural Sciences and Bioinformatics (CCNSB), International Institute of Information Technology, Hyderabad (IIIT-H), Gachibowli, Hyderabad 500032, Telangana, India.*

‡*Computational Science Division, Saha Institute of Nuclear Physics (SINP), 1/AF, Bidhannagar, Kolkata 700064, India.*

E-mail: abi_chem@iiit.ac.in

[a]Current address: Solid State and Structural Chemistry Unit (SSCU), Indian Institute of Science (IISc), Bangalore 560012, Karnataka, India.

Contents

S1 Details of the computational methods **S3**

RNA crystal structure dataset S3

BPFIND S3

Explicit solvent model S4

S2 Possible sites of class II protonation in different RNA base pairs **S5**

S3 Neighboring residues of G:C W:W Trans and G:G W:H Cis pairs **S6**

S4 HF and correlation components of MP2 interaction energy **S7**

S5 Frontier orbitals **S8**

S6 Analysis of rotational and translational parameters of optimized base pairs **S9**

S7 PDB Ids of the RNA crystal structures studied in this work **S11**

S7.1 HD-RNAS (167 files) S11

S7.2 NDB (838 files) S11

References **S13**

S1 Details of the computational methods

RNA crystal structure dataset

We have analyzed two different non-redundant sets of RNA crystal structures, one is provided by the Nucleic Acid Database (NDB) (*1*) maintained by BGSU server and the other is provided by the HD-RNAS (*2*) database. We have applied a resolution cut-off of 3.5 Å and shortlisted 838 structures available in the version 1.89 of the NDB database. This set of 838 structures will be referred as 'NDB dataset' in the manuscript. The representative structures present in HD-RNAS are decided upon after taking into account length, R-factor, resolution and sequence similarity. We have further shortlisted 167 RNA crystal structures after applying a resolution cut-off of 3.5 Å and length cut-off of 30 nucleotides in order to exclude the small synthetic RNA constructs. We have referred to this set of 167 structures as 'HD-RNAS dataset' in the manuscript. All the PDB Ids of the crystal structures of the corresponding data sets are listed at the end of the file. However, it should be noted that these two non-redundant datasets are not mutually exclusive to each other.

BPFIND

Occurrences of G:C W:W Trans and G:G W:H Cis base pairs within these sets of RNA crystal structures have been identified using the BPFIND software. BPFIND is a well accepted (*3–5*) precursor atom based algorithm, which identifies two nucleobases as a base pair if there are at least two conventional hydrogen bonds (N-H$\cdots$N, N-H$\cdots$O, O-H$\cdots$N, O-H$\cdots$O, C-H$\cdots$N and C-H$\cdots$O type) present between them. We have applied the following cut-offs to detect only 'good' base pairing interactions – (i) cutoff distance of 3.8 Å between the acceptor and donor atoms, (ii) cutoff angle of 120.0° for checking planarity of precursor atoms and linearity of the hydrogen bonds and (iii) cutoff 'E-value' of 1.8 to signify the overall distortion and maintain a good base pairing geometry. On the basis of E-value, the least distorted examples of G:C W:W Trans (515G:548C; 1N78 chain C) and G:G W:H Cis

(251G:254G; 3KNH chain A) have been selected for QM calculations.

Explicit solvent model

We have calculated the impact of a polar solvent environment ($\varepsilon = 78.4$) over the gas phase interaction energies by implementing a computationally efficient and robust conductor-like polarizable continuum model (CPCM) (*6*, *7*), which uses united atom topological model to define the atomic radii. This was found to be more appropriate for polar liquids, where the electrostatic potential goes to zero on the surface and is considered to be less sensitive to outlying charge error (*8*).

S2 Possible sites of class II protonation in different RNA base pairs

Table S1: List of base pairs belonging to six different geometric families and their possible site of *Class II* protonation. Some base pairs show multi modality which is specified within the parenthesis. The details of the multiple modes have been discussed in the "Terms and Definitions" page of RNAP COGEST database (*9*).

Protonation Site	Geometric Family	Base Pairs
Adenine-N1	WW	–
	WH	A:A W:H Cis, A:A W:H Trans, U:A W:H Cis
	WS	A:rA W:S Cis, A:rA W:S Trans, A:rC W:S Trans, A:rU W:S Trans
	HH	A:A H:H Cis, A:A H:H Trans, A:G H:H Cis, A:U H:H Cis, A:U(I) H:H Trans, A:U (II) H:H Trans
	HS	A:rA H:S Cis (Adjacent), A:rA H:S Cis (Near Adjacent), A:rA H:S Trans, A:rG H:S Cis (Adjacent), A:rG H:S Cis (Distant), A:rG H:S Trans, A:rC H:S Cis (Adjacent), A:rC H:S Cis (Distant), A:rU H:S Cis (Adjacent), A:rU H:S Cis (Distant), A:rU H:S Trans
	SS	rA:rA S:S Cis, rA:rA S:S Trans, rA:rG(I) S:S Cis, rA:rC(I) S:S Cis, rA:rC(II) S:S Cis, rA:rU(I) S:S Cis, rA:rU(II) S:S Cis
Adenine-N3	WW	A:A W:W Cis, A:A W:W Trans, A:G W:W Cis, A:U W:W Cis, A:U W:W Trans, A:C W:W Trans
	WH	A:A W:H Cis, A:A W:H Trans, A:U W:H Cis, U:A W:H Cis
	WS	A:rA W:S Cis, A:rA W:S Trans, A:rC W:S Cis, A:rC W:S Trans, A:rU W:S Cis, A:rU W:S Trans,
	HH	A:A H:H Cis, A:A H:H Trans, A:G H:H Cis, A:U H:H Cis, A:U(I) H:H Trans, A:U (II) H:H Trans
	HS	A:rA H:S Cis (Adjacent), A:rA H:S Cis (Near Adjacent), A:rA H:S Trans, A:rG H:S Cis (Adjacent), A:rG H:S Cis (Distant), A:rG H:S Trans, A:rC H:S Cis (Adjacent), A:rC H:S Cis (Distant), A:rC H:S Trans, A:rU H:S Cis (Adjacent), A:rU H:S Cis (Distant), A:rU H:S Trans
	SS	rA:rA S:S Trans, rA:rC(I) S:S Cis, rA:rC S:S Trans, rA:rU S:S Trans
Adenine-N7	WW	A:A W:W Cis, A:A W:W Trans, A:G W:W Cis, A:U W:W Cis, A:U W:W Trans, A:C W:W Trans
	WH	A:A W:H Cis, A:A W:H Trans, A:U W:H Cis
	WS	A:rA W:S Cis, A:rA W:S Trans, A:rC W:S Cis, A:rU W:S Cis
	HH	–
	HS	A:rA H:S Cis (Adjacent), A:rA H:S Cis (Near Adjacent), A:rA H:S Trans, A:rC H:S Trans
	SS	rA:rA S:S Cis, rA:rA S:S Trans, rA:rG(I) S:S Cis, rA:rG(II) S:S Cis, rA:rG S:S Trans, rA:rC(I) S:S Cis, rA:rC(II) S:S Cis, rA:rC S:S Trans, rA:rU(I) S:S Cis, rA:rU(II) S:S Cis, rA:rU S:S Trans
Guanine-N3	WW	A:G W:W Cis, G:G W:W Trans, G:C W:W Cis, G:C W:W Trans, G:U W:W Cis, G:U W:W Trans
	WH	G:G W:H Cis, G:G W:H Trans, G:U W:H Trans
	WS	G:rC W:S Cis, G:rU W:S Trans
	HH	A:G H:H Cis, G:C H:H Cis, G:C(I) H:H Trans, G:C(II) H:H Trans
	HS	A:rG H:S Cis (Distant), G:rG H:S Cis
	SS	rG:rC S:S Trans, rG:rU(I) S:S Cis
Guanine-N7	WW	A:G W:W Cis, G:G W:W Trans, G:C W:W Cis, G:C W:W Trans, G:U W:W Cis, G:U W:W Trans
	WH	G:G W:H Cis, G:G W:H Trans, G:U W:H Trans
	WS	G:rC W:S Cis, G:rU W:S Trans
	HH	–
	HS	A:rG H:S Cis (Adjacent), A:rG H:S Cis (Distant), A:rG H:S Trans, G:rG H:S Cis
	SS	rA:rG(I) S:S Cis, rA:rG(II) S:S Cis, rA:rG S:S Trans, rG:rG S:S Cis, rG:rG S:S Trans, rG:rC(I) S:S Cis, rG:rC(II) S:S Cis, rG:rC S:S Trans, rG:rU(I) S:S Cis, rG:rU(II) S:S Cis
Cytosine-N3	WW	–
	WH	C:C W:H Cis, C:C W:H Trans
	WS	A:rC W:S Cis, C:rC W:S Cis
	HH	G:C H:H Cis, G:C(I) H:H Trans, G:C(II) H:H Trans
	HS	A:rC H:S Cis (Adjacent), A:rC H:S Cis (Distant), A:rC H:S Trans, C:rC H:S Cis (Near Adjacent), C:rC H:S Cis (Adjacent), C:rC H:S Cis (Distant) C:rC H:S Trans, C:rU H:S Cis (Near Adjacent), C:rU H:S Trans
	SS	rA:rC(I) S:S Cis, rA:rC(II) S:S Cis, rA:rC S:S Trans, rG:rC(II) S:S Cis, rG:rC S:S Trans

S3 Neighboring residues of G:C W:W Trans and G:G W:H Cis pairs

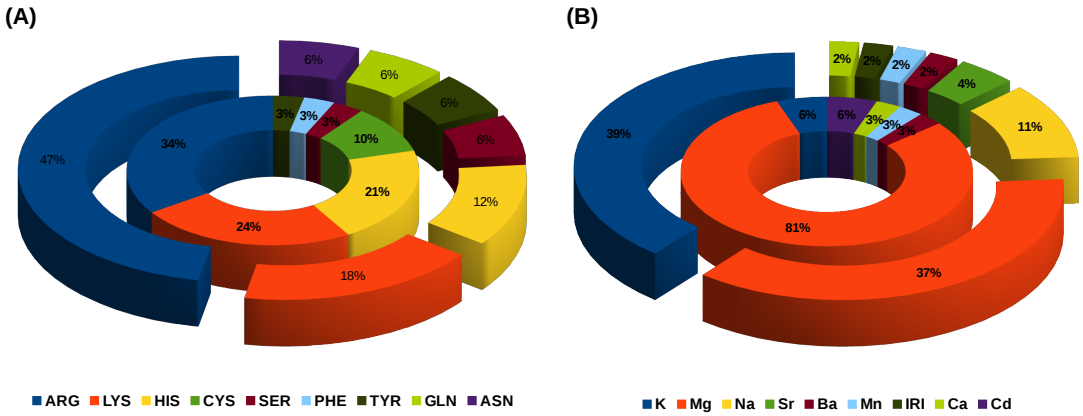


Figure S1: Frequency distribution of different (A) amino acid residues and (B) metal cations found within 6Å distance of the N3 atom of guanine residues which form G:C W:W Trans and G:G W:H Cis base pairs through their respective WC edges. The outer and inner rings represent the distribution for NDB and HDRNAS datasets, respectively.

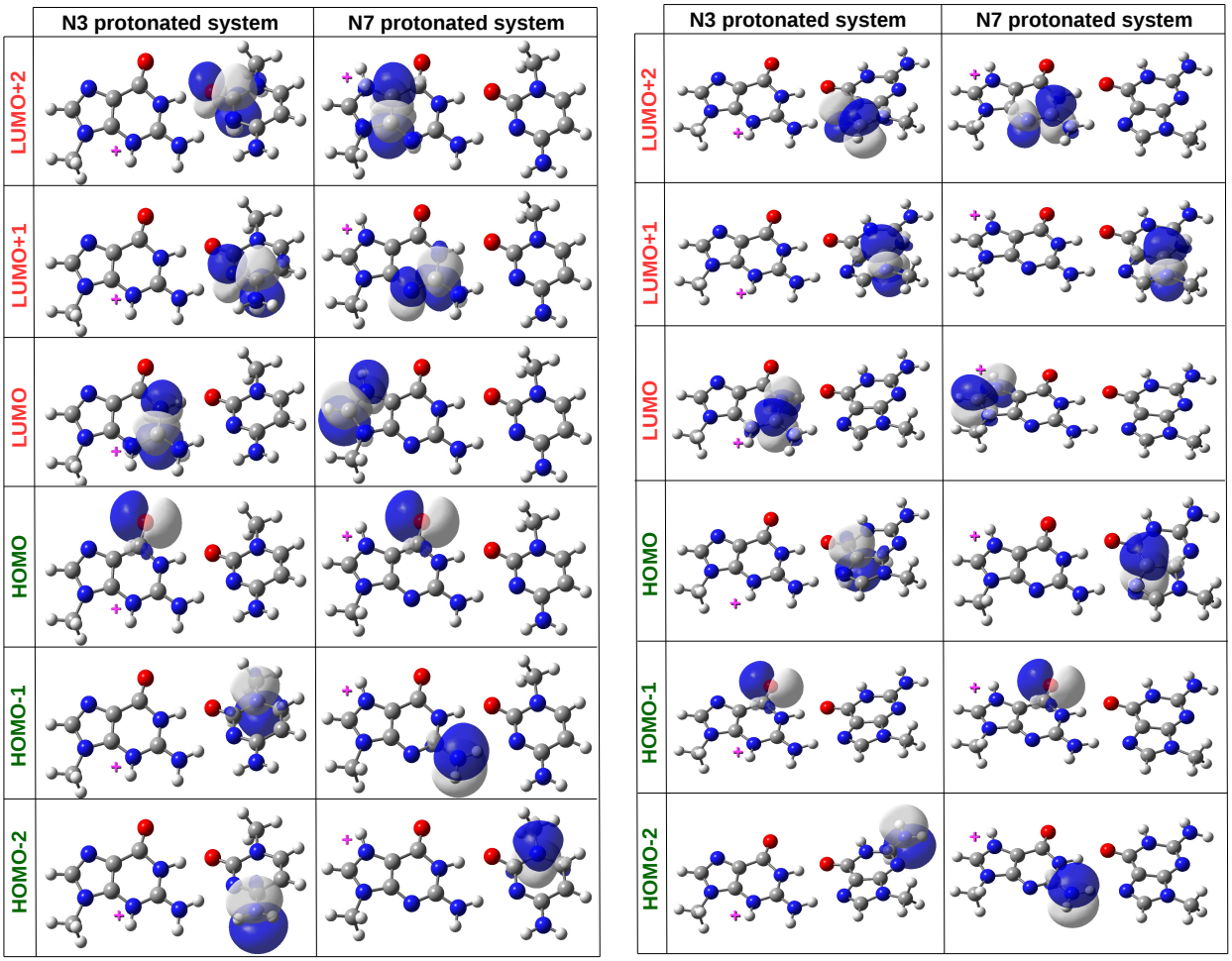
S4 HF and correlation components of MP2 interaction energy

Table S2: Interaction energies of the H_{OPT} and F_{OPT} geometries of the G:C W:W Trans and G:G W:H Cis base pairs with different protonation states of guanine. Geometry optimizations have been performed at B3LYP/6-31+G(d,p) level and interaction energies have been calculated at MP2/aug-cc-pVDZ level (E_{int}^{MP2}). Each E_{int}^{MP2} values have two components, the Hrtree-Fock component (E_{int}^{HF}) and the correlation component (E_{int}^{corr}). All the energy values are reported in kcal mol^{-1}.

Base Pair	Protonation state of guanine	Optimization	E_{int}^{MP2}	E_{int}^{HF}	E_{int}^{corr}	%HF	% corr
G:C W:W Trans	Neutral	H_{OPT}	-16.9	-11.0	-5.9	65.2	34.8
	N3 protonated		-37.2	-31.4	-5.8	84.4	15.6
	N7 protonated		-31.8	-25.4	-6.5	79.7	20.3
	Neutral	F_{OPT}	-20.1	-14.5	-5.6	72.3	27.7
	N3 protonated		-39.7	-32.6	-7.1	82.2	17.8
	N7 protonated		-34.1	-26.6	-7.5	78.1	21.9
G:G W:H Cis	Neutral	H_{OPT}	-15.2	-9.1	-6.1	59.7	40.3
	N3 protonated		-40.7	-34.6	-6.1	85.0	15.0
	N7 protonated		-34.9	-27.9	-6.9	80.1	19.9
	Neutral	F_{OPT}	-17.4	-11.2	-6.1	64.8	35.2
	N3 protonated		-41.7	-33.7	-8.0	80.9	19.1
	N7 protonated		-34.9	-26.6	-8.2	76.4	23.6

It is important to note that, solvent phase calculations essentially reduces only the HF component (*i.e.*, the electrostatic part) of the interaction energy.

S5 Frontier orbitals



(i) G:C W:W Trans (ii) G:G W:H Cis

Figure S2: Frontier orbitals of N3 and N7 protonated (i) G:G W:W Trans and (ii) G:G W:H Cis base pairs.

S6 Analysis of rotational and translational parameters of optimized base pairs

Figure S3: Comparison of the three rotational parameters (Buckle, Open & Propeller twist) of the full optimized geometries of (A-C) G:C W:W Trans and (D-F) G:G W:H Cis base pairs. Optimization has been performed at five different levels of theory, HF, B3LYP, B3LYP-D3, M05-2X and MP2. Buckle parameters except for the base pairs optimized at MP2 level are also shown in inset.

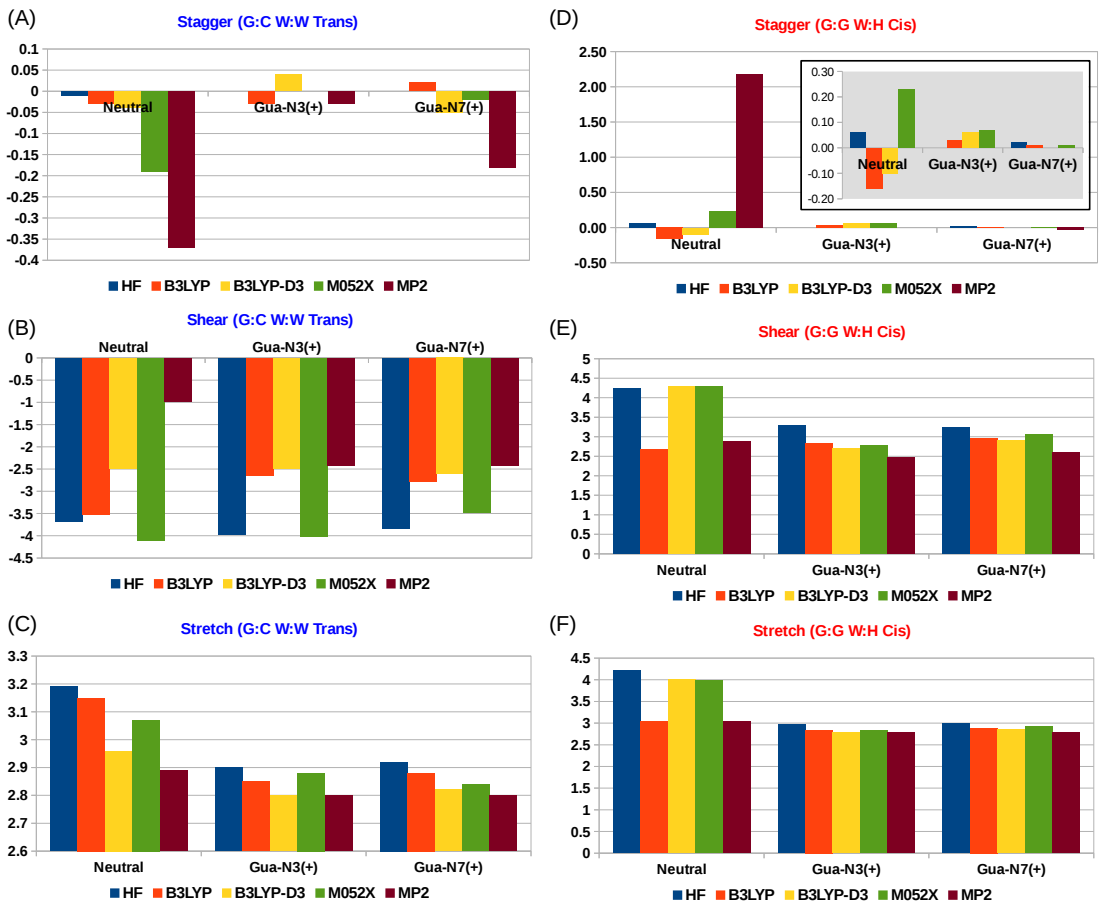


Figure S4: Comparison of the three translational parameters (Stagger, Shear, & Stretch) of the full optimized geometries of (A-C) G:C W:W Trans and (D-F) G:G W:H Cis base pairs. Optimization has been performed at five different levels of theory, HF, B3LYP, B3LYP-D3, M05-2X and MP2. For the G:G W:H Cis pair, buckle parameters except for the base pair optimized at MP2 level are also shown in inset.

S7 PDB Ids of the RNA crystal structures studied in this work

S7.1 HD-RNAS (167 files)

URL: http://www.saha.ac.in/biop/www/HD-RNAS.html

2du6, 2zue, 1f7u, 3kfu, 1b23, 2fmt, 1yfg, 4arc, 3vjr, 2bte, 2azx, 3eph, 1o0c, 3knl, 3akz, 1euy, 1zjw, 3amt, 1ffy, 1qu2, 2zni, 3foz, 1evv, 2zm5, 3tup, 1ehz, 3a3a, 3hl2, 3add, 2fk6, 1qf6, 2dr2, 2y10, 2xqd, 3am1, 1c0a, 1il2, 2der, 1gtr, 1n78, 1qtq, 1wz2, 3tvf, 3uye, 4as1, 2ct8, 3knh, 4dh9, 2csx, 2du5, 1h4s, 3rg5, 1ser, 4gaq, 3uz6, 1j1u, 3uz8, 1j2b, 1gax, 1i6u, 2vqe, 3r8n, 3u5b, 2zjr, 3r8s, 1mms, 1vqo, 3v2d, 3u5d, 1un6, 2xg0, 1u8d, 1xok, 2vpl, 1zho, 2nz4, 1hr2, 1gid, 1c9s, 1sj3, 3nkb, 3g78, 2z75, 2r8s, 2gcv, 1m5o, 3cul, 3d2v, 3f2q, 4enc, 2xnz, 3npq, 3irw, 3v7e, 4fe5, 1y26, 3q3z, 3gx5, 3suh, 3dio, 2pxv, 1l9a, 1lng, 3ktw, 1hq1, 2pxb, 2a64, 1u9s, 1p6v, 3r9w, 3r9x, 2b57, 3npb, 1gtn, 1gtf, 3snp, 3pu0, 3pu1, 2xd0, 2zzn, 2zzm, 2ozb, 2gje, 3ptx, 3pu4, 1fir, 3moj, 3hhz, 357d, 3knn, 1s03, 3siv, 2xdb, 2xdd, 2il9, 3nmu, 1kxk, 1vfg, 3pyu, 3kiq, 3kir, 3kit, 3pyn, 3u4m, 3ciy, 1jbr, 3nvk, 3hax, 3ovb, 3ovs, 3p22, 1kh6, 3pla, 1ddy, 3r4f, 3rw6, 2zh1, 3icq, 2gtt, 2nue, 2hvy, 1u6b, 2hw8, 2jea, 3lwv, 1xjr, 3ds7

S7.2 NDB (838 files)

URL: http://rna.bgsu.edu/rna3dhub/nrlist/

1urn, 3iab, 1ooa, 3r4f, 3zp8, 3snp, 4a3e, 4pqv, 1kxk, 4bxx, 3siu, 3la5, 4tzx, 1cvj, 2bh2, 4a3d, 4a3k, 3b5f, 4lx6, 4q5s, 1p6v, 3egz, 364d, 1t0d, 1yfg, 3nvk, 1hys, 4boc, 3hk2, 1i9v, 1mji, 2tra, 1y26, 4n0t, 2du3, 3moj, 2du6, 1jbr, 2d2l, 3cw5, 1h4s, 1qc0, 4ngf, 3nkb, 1f7v, 1fir, 4c4w, 1vfg, 4jf2, 3vjr, 4ol8, 2ake, 4a93, 2dlc, 1vby, 3tup, 1u0b, total 3ouy, 2ozb, 3e2e, 1qu2, 2g3s, 4kr2, 2p7d, 2hoj, 1j1u, 1i6u, 3wqy, 3po2, 1qf6, 1q96, 1b23, 2rfk, 1VY9, 3amu, 2zuf, 4k4w, 4p5j, 3umy, 2hvy, 1eiy, 1qrs, 4mcf, 1VY8, 4qi2, 2zzm, 3cz3, 4lvz, 4db2, 4kr6, 2gjw, 3hjw, 3gtl, 4kr7, 4k4u, 1qbp, 3am1, 3hax, 3m3y, 1h3e, 3qsy, 4aob, 2i82, 4b5r, 1feu, 3dh3, 3al0, 3bnq, 2nok, 1s03, 2pn4, 3gx5, 4kzd, 4ts2, 3fu2, 1d4r, 4l81, 1tfy, 4wkj, 1tfw, 4o26, 4q0b, 3sux, 1nuv, 3p59, 3zla, 1lng, 3htx, 3ciy, 1yyw, 2vpl, 4w92, 1u63, 3hou, 4pqu, 1ykv, 3w3s, 3ucz, 4wfl, 1duq, 4qlm, 4H6F, 1egk, 3a6p, 4jrc, 4qyz, 3f4g, 4qk9, 4kqy, 1qa6, 4g7o, 1zx7, 4csf, 4mgm, 1fjg, 4qk8, 3j7y, 2AW7, 4un5, 4un3, 1un6, 4v19, 4ilm, 3rw6, 1mfq, 3ivn, 3icq, 2o5i, 2oiu, 3slq, 3j7a, 3siv, 1hmh, 4uyk, 4u7u, 3eph, 2quw, 2zzn, 2ho7, 4W23, 1h38, 3U5F, 2der, 2zni, 2zm5, 3pla, 2azx, 1efw, 1il2, 2qbz, 2wj8, 3wfs, 2qux, 1u9s, 1j2b, 2y9h, 2csx, 3d2v, 3q3z, 2qwy, 2d6f, 4k4y, 2cv1, 1gax, 3d0u, 3ol8, 2bte, 4aq7, 2gdi, 3ol7, 4gxy, 4frg, 4rge, 3ol9, 3rg5, 3dir, 3owz, 1et4, 1wz2, 1g1x, 3add, 3adc, 4p9r, 3cun, 3iwn, 3ktw, 4k50, 3p22, 4frn, 2gtt, 4FY3, 2xd0, 4plx, 4uyj, 1m5o, 1u6b, 1VW3, 4eya, 4oo8, 2czj, 4ioa, 4m6d, 3v7e, 1s72, 4QCN, 2il9, 3ivk, 2QBG, 3rkf, 1nbs, 3r1c, 4qjd, 2a64, 3U5H, 3kfu, 1kog, 1VX6, 1gid, 4mgn, 3hga, 3ok4, 1rmv, 4ejt, 3j06, 4lck, 4W21, 1rxa,

4ang, 3ahu, 3pdm, 1vtm, 2x1a, 1mdg, 1laj, 1osu, 2c4z, 2c50, 1p79, 2fz2, 4afy, 4alp, 2xnr, 4lq3, 4jiy, 4n2s, 333d, 3u2e, 2x1f, 1h2c, 4ohy, 4n2q, 3qjp, 2r7u, 2r7v, 4qu6, 4lmz, 4faw, 2r7s, 3o8c, 3mj0, 2tmv, 4r3i, 3hsb, 4ms9, 4ohz, 2r7t, 283d, 3pf4, 3qsu, 3gpq, 3rtj, 2vop, 1g2j, 3b0u, 3nj6, 4dwa, 3q0q, 3q0r, 3q0s, 2a1r, 4jzv, 4jvh, 3k5q, 3k5y, 3k5z, 3k61, 3k62, 3k64, 4tu0, 3pf5, 3rc8, 3t3o, 1b2m, 2voo, 4i67, 2xs7, 4rkv, 3qgc, 1l3z, 1kq2, 4hot, 4hos, 3avt, 4nku, 4lj0, 2asb, 4j7m, 4qu7, 3nna, 2g91, 1n1h, 413d, 3o8r, 3v71, 3v6y, 2xs2, 3bsb, 4j7l, 4oau, 1nb7, 4hor, 1fxl, 2bbv, 1si3, 4nl3, 3nma, 1i5l, 1pgl, 3mij, 3nmr, 4jab, 255d, 2q66, 2r7w, 4oav, 3t5n, 3v74, 3l26, 1uvk, 4rcj, 1g2e, 4kre, 3nnc, 4d25, 1uvi, 3ice, 3gib, 2c0b, 2po1, 4ola, 3pey, 2q1r, 3g9y, 2vnu, 4am3, 3mdg, 3mdi, 4e59, 4u35, 2v7r, 4u34, 259d, 402d, 4cs1, 3iev, 1wpu, 3sqw, 1uvj, 2jea, 3czw, 4fts, 4kxt, 1cwp, 4e78, 2xs5, 4c8y, 4g9z, 3r1e, 4al7, 3nd3, 3q0o, 3i5x, 3i5y, 1wmq, 1av6, 2a8v, 4qm6, 4ii9, 4mdx, 2xgj, 4krf, 4c8z, 2vuq, 2a0p, 3p4c, 3p4d, 4c40, 4f1n, 2xzo, 2b2d, 3jxr, 3oin, 2xli, 1kd5, 438d, 1kfo, 3ie1, 1uvm, 3q0p, 2bx2, 1ddl, 3o7v, 2jlu, 2jlw, 2vrt, 4g0a, 3bx3, 2g4b, 3aev, 3er9, 2von, 1n38, 4gv9, 3kmq, 2f8k, 3knc, 4o8j, 3rer, 1uvl, 1pvo, 4ba2, 4jvy, 4fsj, 2vod, 1zse, 1zdj, 4e58, 2gxb, 3gvn, 421d, 479d, 4msr, 3bnt, 1bmv, 4ht9, 4n48, 3bx2, 4dzs, 4b3g, 4ftb, 3kms, 3nl0, 422d, 472d, 1dqh, 4jrd, 377d, 3ibk, 1m8y, 4ngd, 3qrr, 3bsx, 4nha, 4f3t, 3uld, 3vnu, 1wne, 165d, 2y8y, 439d, 4l8h, 3eqt, 2r22, 2jlz, 1xpo, 1xpr, 1xpu, 1knz, 3fht, 2izn, 2y8w, 3avu, 2v6w, 4fvu, 3l25, 4j39, 3pkm, 4jk0, 4w5n, 3qjj, 3qjl, 1zh5, 3qrp, 1i9x, 1j6s, 4j5v, 2jlx, 3d2s, 2atw, 1n35, 1uvn, 1ytu, 3vnv, 4nlf, 3fs0, 1lnt, 157d, 3cjz, 4jrt, 4knq, 2ix1, 354d, 3avv, 4e6b, 2pjp, 2r20, 433d, 4pei, 4f02, 3bso, 1fuf, 3boy, 2py9, 3m85, 3q2t, 3bsn, 3cgp, 3cgr, 3cgs, 3ssf, 1rna, 4kyy, 2gje, 3dd2, 3m7n, 3nnh, 2ann, 2e9r, 4c9d, 4w5r, 3adi, 1l2x, 4h5p, 1b7f, 2e9z, 387d, 4tv0, 3hxm, 3r2d, 4ay2, 353d, 1csl, 2a43, 466d, 435d, 405d, 3q50, 2nz4, 4pmw, 1f8v, 2izm, 1k8w, 3avw, 1mhk, 397d, 4gha, 2val, 1m8w, 4peh, 3r2c, 4bw0, 1zbi, 1zbl, 4oq8, 3avx, 2bgg, 3ncu, 3lrr, 3k49, 3avy, 2j0s, 7msf, 3ho1, 3tmi, 2xsl, 4pcj, 2w89, 420d, 3og8, 1jid, 1jzv, 4w5t, 2iz9, 4ato, 4kji, 3hm9, 3sn2, 4h8k, 406d, 398d, 1f1t, 1j9h, 3p6y, 4bwm, 2zy6, 1zev, 2jlt, 4m59, 4w5q, 2bu1, 3glp, 2oe5, 4m7d, 4w5o, 1u1y, 2e9t, 3hjf, 1f27, 3lrn, 3tzr, 3sj2, 1o9m, 2ao5, 3r9w, 2db3, 4oo1, 4m7a, 4ill, 1t0k, 4qil, 2ec0, 1e7k, 1sa9, 2hw8, 1saq, 4pdb, 4e5c, 4ijs, 1UTF, 3mqk, 3szx, 1t0e, 4fxd, 4ghl, 2bq5, 5msf, 4jah, 2dr8, 361d, 3syw, 1mwl, 3bnl, 4qqb, 6msf, 3ova, 2fqn, 2az2, 3td0, 4k31, 2grb, 4jng, 3bt7, 1zbh, 1xok, 2o3x, 3td1, 4ktg, 3mei, 2qek, 4bhh, 4u8t, 1br3, 1ec6, 1ntb, 4hkq, 4oog, 2o3v, 4kq0, 409d, 1duh, 1sds, 3rzo, 1yzd, 1q2r, 4gg4, 4nfo, 4qik, 2zi0, 1nlc, 2awe, 2gic, 3ptx, 3pu0, 4iqs, 3s1m, 3s1n, 2f8s, 3bnp, 2zko, 1xpe, 280d, 4j1g, 1gtn, 1gtf, 4o41, 4mce, 1rpu, 2qk9, 4erd, 3erc, 4lgt, 3npn, 3loa, 1yz9, 3ks8, 1xjr, 2nue, 3p4b, 3pu1, 3zc0, 1qcu, 3agv, 4wsb, 3r9x, 4wsa, 4wrt, 3b31, 4py5, 3ftm, 1kh6, 1utd, 3s1r, 3ts0, 1dfu, 429d, 4ig8, 1r9f, 2q1o, 4k27, 3s2d, 1rlg, 4enb, 3wbm, 4oji, 2f8t, 3ts2, 2i91, 3s14, 2ply, 1sdr, 2xb2, 1VW4, 3s49, 3s15, 3ftf, 2nug, 2pxv, 3s16, 1i6h, 3r1d, 3s17, 3e5c, 2b3j, 1qln, 488d, 3nvi, 3nj7, 3po3, 4k4v, 1mzp, 4m30, 3t5q, 4a3g, 4gcw, 1jbs, 4m4o, 1m8v, 2w2h, 1s77, 1yvp, 2ez6, 4kz2, 3fo4, 1msw, 4e48, 3f73, 1r3e, 1f7y, 4c7o, 1e8o, 3bsu, 1ser, 4ifd, 1a9n, 4b3o

References

1. Leontis, N. B., and Zirbel, C. L. *Nonredundant 3D structure datasets for RNA knowledge extraction and benchmarking*; Springer, 2012; Vol. 27; pp 281–298.

2. Ray, S. S., Halder, S., Kaypee, S., and Bhattacharyya, D. (2012) HD-RNAS: An automated hierarchical database of RNA structures. *Front. Genet. 3*, 1–10.

3. Olson, W. K., Esguerra, M., Xin, Y., and Lu, X.-J. (2009) New information content in RNA base pairing deduced from quantitative analysis of high-resolution structures. *Methods 47*, 177–186.

4. Otero-Navas, I., and Seminario, J. M. (2012) Molecular electrostatic potentials of DNA base–base pairing and mispairing. *J. Mol. Model 18*, 91–101.

5. Srinivasadesikan, V., Sahu, P. K., and Lee, S.-L. (2012) Quantum Mechanical Calculations for the Misincorporation of Nucleotides Opposite Mutagenic 3, N 4-Ethenocytosine. *J. Phys. Chem. B 116*, 11173–11179.

6. Barone, V., and Cossi, M. (1998) Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A 102*, 1995–2001.

7. Cossi, M., Rega, N., Scalmani, G., and Barone, V. (2003) Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem. 24*, 669–681.

8. Klamt, A., and Schuurmann, G. (1993) COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2 0*, 799–805.

9. Bhattacharya, S., Mittal, S., Panigrahi, S., Sharma, P., Preethi, S., Paul, R., Halder, S., Halder, A., Bhattacharyya, D., and Mitra, A. (2015) RNABP COGEST: a resource for investigating functional RNAs. *Database 2015*, bav011.