Supplementary materials and methods

Determination of the structure of N7-guanosine using NMR methods

Structures of N7-guanosine (7G) were confirmed by 1D ^{1}H and ^{13}C NMR experiments and chemical shifts for 7G were in good agreement with literature data (Figure S8 and S9) (Wawrzyniak et al. 2020). Additionally, assignment of ^{1}H, ^{13}C and ^{15}N resonances was confirmed by the analysis of 2D spectra: ^{1}H-^{1}H correlation spectroscopy (COSY) as well as ^{1}H-^{13}C gradient heteronuclear single quantum coherence (gHSQC) and gradient heteronuclear multiple bond coherence (gHMBC) spectroscopies and compared with the data obtained for guanosine (Tables S4 and S5, Figure S10-S14).

As previously reported, 9→7 isomerization of guanosine caused significant differences in ^{13}C and ^{1}H resonances, especially for carbons of base moiety, *i.e.* C4, C5, and C8, as well as for protons H8, H1' and NH_2 (Table S4 and S5) (Baranowski et al. 2022). Conspicuous upfield shift by 10 ppm for C5 and downfield shift by 8 and 6 ppm for C4 and C8, respectively, were observed, which was a characteristic pattern of changes of ^{13}C resonances for N7 isomer of guanosine compared with N9 isomer. Smaller upfield shifts *i.e.*2-3 ppms were also observed for C2 and C6. For ^{1}H resonances, signals of H8 and H1' underwent a downfield shift by 0.2-0.3 ppm. On the other hand, the signal of NH_2 showed upfield shifts by 0.3 ppm.

The characteristic heteronuclear long range correlations *i.e.* C5-H1' and C8-H1' observed on ^{1}H-^{13}C gHMBC spectrum were crucial for unambiguous determination of D-ribofuranosyl substitution at N7 position of guanosine (Figure S12).

Calculation of free energies of the N7-guanosine and guanosine tautomers

Density functional theory (DFT) calculations at B3LYP/6-311G(d) level of theory were performed on N7-methylated guanine (m^7Gua) representing the 7G-tautomers obtained from the QM optimized structures (at HF/6-311G(d) level of theory) of the 6 different keto-enol tautomers (replacing the ribose sugar with a methyl group) to obtain the free energies of these tautomers (m^7Gua tautomers)

(Figure S4b). Similar calculations were performed on N9-methylated guanine (m^9Gua) tautomers representing guanosine (G)-tautomeric forms (Figure S4a). The m^9Gua tautomers were prepared using MOLDEN (Schaftenaar and Noordik 2000) and subjected to optimization at HF/6-311G(d) theory and the energy calculations (at B3LYP/6-311G(d) level of theory) were performed using the optimized structures. The free energy values were estimated both in the gas phase and the solvent phase (Figure S5). For the calculations in the solvent phase water environment by the polarizable continuum model (PCM) was considered.

Reoptimization of glycosidic torsion parameters for 7G-N1H

Four initial geometries (conformational schemes SC1-4) of the 7G-N1H residue were prepared following Yildirim et al. (Yildirim et al. 2010) and subjected to geometry optimization followed by a gas phase potential energy surface (PES) scan executed around the glycosidic torsion angle (χ+180° i.e. O4′-C1′-N7-C8) with an increase in the value of (χ+180°) by 5° resulting in 72 conformations (for each scheme) following the protocol described in Yildirim et al. (Yildirim et al. 2010). The QM energies of each of the 72 conformations (for each scheme) were calculated using the MP2/6-31G* level of theory (Figure S15). The quantum mechanically (at HF/6-31G* level of theory) optimized lowest energy conformer was used for the calculation of partial atomic charges using RESP fitting using the R.E.D. version III.52 perl program Vanquelef et al, (Vanquelef et al. 2011). Next, the molecular mechanical (MM) energy minimization was performed following the protocol described earlier (Dutta et al. 2023) for the development of glycosidic torsion parameters within the AMBER FF99 force field (Wang et al. 2000). We also followed a similar protocol for fitting of glycosidic torsion potential as described earlier (Dutta et al. 2023). In this approach, the potential energy due to the glycosidic torsion (χ) is obtained from the differences between the QM energies (E_{QM}) and the MM energies (E_{MM}) i.e. from the equation: $E_{\chi} = E_{QM} - E_{MM}$. 4 x 72=288 E_{χ} values calculated for the 288 conformations using the aforementioned equation, were fitted to the Fourier series shown in the

following equation: $E(\phi_1,\phi_2)=\sum_{n=1}^{4} V_{n1}[1+\cos(n\phi_1)] + V_{n2}[1+\cos(n\phi_2)]$ (here, V_{n1}, V_{n2}, ϕ_1 and ϕ_2 are varied). ϕ_1 and ϕ_2 represent the (O4′-C1′-N7-C8) and (C2′-C1′-N7-C8) torsion angles respectively and V_{n1} and V_{n2} represent the potential energy barriers around these torsion angles. A separate fitting was performed for calculating V_{n1} and V_{n2}. The QM and MM energy profiles (with χ set to zero and with the newly derived χ_{KOL0} parameters) for 7G-N1H are shown in Figure S16.

For 7G-N3H a similar protocol as described for 7G-N1H was followed for geometry optimization, PES scan, QM energy calculations was followed and the lowest energy conformer was subjected to further optimization, followed by RESP fitting. The lowest energy conformer of 7G-N1H (based on the QM energy calculations) was modified in MOLDEN to prepare the 7G-N1H,N2-imino tautomer, followed by geometry optimization at HF/6-31G* level of theory and subsequent calculation of partial atomic charges using RESP fitting. As mentioned in the main text, for the 7G-N3H and 7G-N1H,N2-imino tautomers, the glycosidic torsion parameters developed for the 7G-N1H tautomer were applied during MD simulations of duplexes.

Calculation of hybridization energies of duplexes using MM/3D-RISM

As mentioned in the main text, the duplex hybridization energies (binding free energies) for the duplexes in this study were estimated using the molecular mechanics/three-dimensional reference interaction site model (MM/3D-RISM) approach and only the internal seven base pairs were considered during the calculations (i.e. prior to the calculations, both the terminal base pairs were removed). In this approach, 3D-RISM (three-dimensional reference interaction site model) is utilized for the description of the three-dimensional map of solvent density distribution around the solute molecule (Kovalenko and Hirata, 2003; Luchko et al. 2010). Here, in the 3D space (r), the solvent structure is represented using the probability density $\rho_\gamma g_\gamma(r)$ (which approaches the solvent bulk density ρ_γ with $g_\gamma(r) \rightarrow 1$, where $g_\gamma(r)$ is the 3D solvent site distribution function) of finding the solvent sites (γ) around the solute (Luchko et al. 2010). Additionally, we have also utilized the 3D version of

Kovalenko-Hirata closure (3D-KH closure) (Genheden et al. 2010) molecular solvation theory during the MM/3D-RISM calculations. The 3D-RISM-KH approach has been previously utilized for the calculation of binding free energies for short DNA oligonucleotides by Yesudas et al, (Yesudas et al. 2015) and was observed to produce good agreement with experimental results. This approach has also been used for the dsRNA oligonucleotides by Yildirim et al, (Yildirim et al. 2023). In our calculations we used a similar script as provided in Yildirim et al, (Yildirim et al. 2023) and we used the parameters as shown below: /&rism closure="kh", rism_verbose=0, polardecomp=0, thermo="gf"/.

Supporting Tables

Table S1. (A) Calculated properties of the duplexes containing 7G-A base pair from MD simulations

	5'G7GC/3'CAG duplex		**5'C7GG/3'GAC duplex**	
	G7GC (7G(N1H)-A) (GM1)	**G7GC (7G(N1H,N3H,N2-imino)-A) (GM3)**	**C7GG (7G(N1H)-A) (GM2)**	**C7GG (7G(N1H,N3H,N2-imino)-A) (GM3)**
RMSD (Å)	1.5 (0.4)	1.3 (0.3)	1.5 (0.4)	1.4 (0.3)
C1'-C1' distance (Å)	10.9 (0.3)	10.8 (0.2)	10.6 (0.3)	10.7 (0.2)
7G5 (C1'-C1'-N7) (°)	82.7 (4.4)	85.6 (3.3)	30.7 (3.4)	85.2 (11.9)
A14 (C1'-C1'-N9) (°)	65.1 (7.8)	64.1 (3.9)	90.3 (5.0)	65.9 (7.8)
NOE distances (Å)				
A14-H2...7G-H1 Population with distance <3 Å **(strong NOE)**	2.5 (0.7) 94%	2.8 (0.3) 83%	5.4 (0.2) 0.0%	2.8 (0.2) 83%
A14-H2 ...C6-H1′ Population with distance <4.5 Å **(medium NOE)**	4.3 (0.5) 66%	4.2 (0.5) 71%	NA	NA
A14-H2 ...G6-H1′ Population with distance <4.5 Å **(medium NOE)**	NA	NA	6.9 (0.6) 0.01%	4.2 (0.4) 72%
A14-H2 ...G6-H1 Population with distance <6 Å **(Weak NOE)**	NA	NA	4.3 (0.6) 99%	4.6 (0.4) 99%

A14-H2 ...C15-H1′ Population with distance <4.5 Å **(medium NOE)**	4.3 (0.6) 64%	4.2 (0.6) 70%	NA	NA
A14-H2 ...G15-H1′ Population with distance <4.5 Å **(medium NOE)**	NA	NA	6.2 (0.5) 0.08%	4.3 (0.6) 65%
A14-H2 ...G15-H1 Population with distance <6 Å **(Weak NOE)**	NA	NA	5.0 (0.6) 94%	5.3 (0.4) 95%
A14-H2 ... G13-H1 Population with distance <4.5 Å **(medium NOE)**	3.7 (0.5) 96%	3.8 (0.4) 95%	NA	NA
A14-H2 ... C13-H41 Population with distance <6 Å **(weak NOE)**	NA	NA	4.7 (0.5) 99%	6.7 (0.5) 6%
The values within parentheses correspond to standard deviations.				

(B) Calculated properties of the duplexes containing 7G-G base pair from MD simulations

Duplexes	**5'G7GC/3'CGG duplex**			**5'C7GG/3'GGC duplex**		
	G7GC (7G(N1H)-G) (GM1)	**G7GC (7G(N1H)-G) (GM2)**	**G7GC (7G(N3H)-G) (GM3)**	**C7GG (7G(N1H)-G) (GM1)**	**C7GG (7G(N1H)-G) (GM2)**	**C7GG (7G(N3H)-G) (GM3)**
RMSD (Å)	1.7 (0.3)	1.6 (0.4)	1.9 (0.4)	1.9 (0.5)	1.4 (0.3)	1.5 (0.3)
C1'-C1' distance (Å)	11.1 (0.3)	10.6 (0.3)	11.1 (0.2)	11.2 (0.3)	10.4 (0.3)	11.1 (0.2)
7G5 (C1'-C1'-N7) (°)	90.3 (3.8)	29.8 (4.3)	84.2 (3.3)	90.4 (3.8)	30.7 (3.7)	84.4 (3.4)
G14 (C1'-C1'-N9) (°)	46.7 (4.3)	98.9 (4.6)	60.5 (3.3)	43.8 (4.1)	101.3 (4.1)	61.0 (3.4)
NOE distances (Å)						

G14_H1…H1_G6 Population with distance <6 Å **(weak NOE)**	NA	NA	NA	4.9 (0.4) 99%	5.6 (0.5) 81%	4.5 (0.4) 100%
G14_H1…H8_G6 Population with distance <6 Å **(weak NOE)**	NA	NA	NA	4.6 (0.5) 99%	8.7 (0.6) 0.01%	5.5 (0.5) 88%
G14_H1…H8_G15 Population with distance <6 Å **(weak NOE)**	NA	NA	NA	4.9 (0.4) 99%	6.6 (0.4) 8%	5.0 (0.4) 99%
The values within parentheses correspond to standard deviations.						

Table S2. (A) Energetics of the duplexes with 7G-A pair

Duplexes	**Duplex - GGC (G-A) Anti-Syn Syn-Anti Anti-Anti (Geometry-1 95%) Anti-Anti (Geometry-2 3%)**	**Duplex - G7GC (7G(N1H)-A) (GM1)**	**Duplex - G7GC (7G(N1H,N3H,N2-imino)-A) (GM3)**	**Duplex - CGG (G-A) Anti-Syn Syn-Anti (200-360ns) Anti-Anti (Geometry-1 81%) Anti-Anti (Geometry-2 13%)**	**Duplex - C7GG (7G(N1H)-A) (GM2)**	**Duplex - C7GG (7G(N1H,N3H,N2-imino)-A) (GM3)**
Hybridization energies (ΔG) and changes in hybridization energies (ΔΔG) upon modification						
ΔG (MM/3D-RISM) (kcal/mol)	-33.2 (4.6)[a] -27.0 (4.3) -27.1 (4.2) -30.1 (4.7)	-30.5 (4.5)	-36.8 (4.4)	-32.6 (4.4)[a] -26.7 (4.0) -28.5 (4.4) -31.4 (4.3)	-34.1 (4.09)	-37.7 (4.3)
$\Delta\Delta G_{calc}$ (MM/3D-RISM) (kcal/mol)	0 0 0 0	2.7[b] -3.5 -3.4 -0.4	-3.7[b] -9.8 -9.7 -6.7	0 0 0 0	-1.5[b] -7.4 -5.6 -2.7	-5.1[b] -11.0 -9.2 -6.3
$\Delta\Delta G^{\circ}_{37}$ (kcal/mol) (Experiment)	0	-0.86		0	-1.76	
Energies corresponding to the molecular mechanics base pairing and base stacking interactions						
Base pairing energy (kcal/mol)	-13.2 (1.8)[a] -5.8 (1.4) -1.8 (0.7) -12.5 (1.7)	-7.9 (1.7)	-9.5 (2.4)	-8.3 (4.7)[a] -5.9 (1.4) -1.5 (1.0) -12.8 (1.6)	-5.2 (1.5)	-9.2 (2.9)
Stacking energy Step 1 (kcal/mol)	-9.9 (1.7)[a] -9.7 (2.2) -6.5 (1.8) -12.6 (3.6)	-5.4 (1.9)	-9.4 (1.69)	-11.3 (4.6)[a] -5.4 (2.2) -5.0 (2.2) -10.6 (2.6)	-4.8 (1.7)	-6.9 (2.1)
Stacking energy Step 2 (kcal/mol)	-14.3 (2.3)[a] -8.8 (2.6) -12.6 (2.1) -13.6 (2.1)	-14.8 (2.4)	-13.1 (1.8)	-9.9 (3.4)[a] -7.9 (2.1) -8.3 (1.8) -8.4 (2.2)	-10.1 (1.7)	-6.7 (1.9)

The values within parentheses correspond to standard deviations. $\Delta\Delta G_{calc}$ (7G-A) = ΔG (7G-A) – ΔG (G-A). a - these numbers refer to the different geometries of G-A base pairs in the unmodified duplex. b - $\Delta\Delta G_{calc}$ calculated relative to the reference unmodified duplexes with different geometries of G-A base pair

(B) Energetics of the duplexes with 7G-G pair (energies in kcal/mol)

Duplexes	**GGC (G-G) Anti-Syn Syn-Anti**	**G7GC (7G(N1H)-G) (GM1)**	**G7GC (7G(N1H)-G) (GM2)**	**G7GC (7G(N3H)-G) (GM3)**	**CGG (G-G) Anti-Syn Syn-Anti**	**C7GG (7G(N1H)-G) (GM1)**	**C7GG (7G-G) (GM2))**	**C7GG (7G(N3H)-G) (GM3)**
Hybridization energies (ΔG) and changes in hybridization energies (ΔΔG) upon modification								
ΔG (MM/3D-RISM) (kcal/mol)	-30.7 (4.3)[a] -30.5 (4.4)	-35.8 (4.5)	-34.6 (4.5)	-40.9 (4.2)	-31.6 (4.3)[a] -30.2 (4.3)	-35.4 (4.2)	-35.3 (4.2)	-40.2 (4.1)
$\Delta\Delta G_{calc}$ (MM/3D-RISM) (kcal/mol)	0 0	-5.1[b] -5.3	-3.9[b] -4.1	-10.2[b] -10.4	0 0	-3.9[b] -5.2	-3.7[b] -5.1	-8.6[b] -9.9
$\Delta\Delta G^{\circ}_{37}$ (kcal/mol) (Experiment)	0	-0.95			0	-1.76		
Energies corresponding to the molecular mechanics base pairing and base stacking interactions								
Base pairing energy (kcal/mol)	-10.2 (2.6)[a] -10.2 (2.8)	-21 (2.9)	-12.0 (2.3)	-24.2 (2.8)	-13.8 (2.4)[a] -13.5 (2.6)	-19.6 (3.2)	-12.4 (2.2)	-23.9 (2.9)
Stacking energy Step 1 (kcal/mol)	-4.0 (2.1)[a] -11.6 (2.6)	-11.6 (1.9)	-14.2 (2.8)	-17.1 (1.8)	-13.6 (3.1)[a] -5.6 (2.5)	-12.8 (1.9)	-9.6 (1.9)	-9.9 (1.8)
Stacking energy Step 2 (kcal/mol)	-11.5 (2.5)[a] -3.9 (2.2)	-11.2 (1.8)	-13.2 (2.1)	-8.1 (1.6)	-5.6 (2.4)[a] -13.8 (3.1)	-10.2 (2.2)	-16 (1.8)	-12.2 (2.1)

The values within parentheses correspond to standard deviations. $\Delta\Delta G_{calc}$ (7G-G) = ΔG (7G-G) – ΔG (G-G). a - these numbers refer to the different geometries of G-G base pairs in the unmodified duplex. b - $\Delta\Delta G_{calc}$ calculated relative to the reference unmodified duplexes with different geometries of G-G base pair.

Table S3. (A) Occurrences (in %) of hydrogen bonds between the 7G-A base pair observed in MD simulations

Duplex	Hydrogen bonds (Donor–Hydrogen---Acceptor)	Occurrence (%)
Duplex - G7GC (7G(N1H)-A) (GM1)	(7G5)N1-HN1---N1(A14)	61
Duplex - G7GC (7G(N1H,N3H,N2-imino)-A) (GM3)	(7G5)N1-HN1---N1(A14)	62
	(A14)N6-H62---N2(7G5)	49
Duplex - C7GG (7G(N1H)-A) (GM2)	(7G5)N2-H21---N3(A14)	44
Duplex - C7GG (7G(N1H,N3H,N2-imino)-A) (GM3)	(7G5)N1-HN1---N1(A14)	68
	(A14)N6-H62---N2(7G5)	46

(B) Occurrences (in %) of hydrogen bonds between the 7G-G base pair observed in MD simulations

Duplex	Hydrogen bonds (Donor–Hydrogen---Acceptor)	Occurrence (%)
Duplex - G7GC (7G(N1H)-G) (GM1)	(7G5)N1-HN1---O6(G14)	88

	(G14)N1-H1---O6(7G5)	59
Duplex - G7GC (7G(N1H)-G) (GM2)	(G14)N2-H21---N3(7G5)	73
	(7G5)N2-H22---N3(G14)	12
Duplex - G7GC (7G(N3H)-G) (GM3)	(G14)N2-H21---O6(7G5)	88
	(G14)N1-H1---N1(7G5)	76
	(7G5)N2-H21---O6(G14)	65
Duplex - C7GG (7G(N1H)-G) (GM1)	(7G5)N1-HN1---O6(G14)	90
	(G14)N1-H1---O6(7G5)	46
Duplex - C7GG (7G(N1H)-G) (GM2)	(G14)N2-H21---N3(7G5)	72
	(7G5)N2-H22---N3(G14)	21
Duplex - C7GG (7G(N3H)-G) (GM3)	(G14)N2-H22---O6(7G5)	88
	(G14)N1-H1---N1(7G5)	73
	(7G5)N2-H21---O6(G14)	63

Table S4. ^{13}C NMR chemical shifts for 7G and G (in ppm).

	C6	C2	C4	C8	C5	C1’	C4’	C2’	C3’	C5’
7G	154.36	152.91	160.71	142.44	107.73	89.18	85.24	74.48	69.75	61.20
G[a]	157.82	154.66	152.33	136.67	117.72	87.46	86.26	74.75	71.43	62.45

Table S5. ^{1}H NMR chemical shifts for 7G and G (in ppm).

	H8	NH_2	N^1H	H1’	H2’	H3’	H4’	H5’/H5”	2’OH/3’OH	5’OH
7G	8.28 (s)[c]	6.19 (s)	10.90 (brs)	5.96 (d)	4.35 (q)	4.06 (q)	3.88 (q)	3.65 (dt) 3.53 (ddd)	5.37 (d) 5.12 (d)	5.04 (t)
G[b]	7.97	6.52	10.75	5.72	4.43	4.11	3.90	3.64 3.55	5.45 5.20	5.10

a: SDBSWeb; https://sdbs.db.aist.go.jp/sdbs/cgibin/direct_frame_disp.cgi?sdbsno=831&spectrum_type=CNMR&fname=CDS03115

b: SDBSWeb https://sdbs.db.aist.go.jp/sdbs/cgi-bin/direct_frame_disp.cgi?sdbsno=831&spectrum_type=HNMR&fname=HSP49445

c: peak multiplicity labels: s - singlet, brs - broad singlet, d - doublet, t – triplet, q – quartet, dt – doublet of triplets, ddd – doublet of doublets of doublets

Table S6. Conformational preferences of 7G-N1H nucleoside

Conformational properties	Sugar Pucker		Base orientation		γ torsion
	SOUTH population (%)	NORTH population (%)	SYN population (%)	ANTI population (%)	g^+ population (%)
FF99_χ_{KOL0}_γ_{bsc0}	55.6	44.4	44.54	55.25	93.96
NMR (Baranowski et al. 2022)	62-67[a] 59[b]	33-38[a] 41[b]	ANTI		g^+ 62[c]

[a] Reported in (Baranowski et al. 2022).
[b] Calculated from coupling constants ($J_{1'2'}/J_{3'4'}$) by the formulas %N = 100[$J_{3'4'}$ /($J_{1'2'}$ + $J_{3'4'}$)], %S = 100[$J_{1'2'}$ /($J_{1'2'}$ + $J_{3'4'}$)] from experimental (NMR) data reported in (Baranowski et al. 2022).
[c] Calculated from coupling constants $J_{4'5'}$=3.80 Hz $J_{4'5''}$=3.70 Hz by the formulas gg+gt+tg=1, $J_{4'5'}$=2.4gg+2.6gt+10.6tg, $J_{4'5''}$=1.3gg+10.5gt+3.8tg

Supporting Figures

Figure S1. The aromatic/amino region of ^{1}H NMR spectrum of the **7G**-A containing duplexes was recorded at different temperatures with the resonance belonging to the A14 amino group marked.

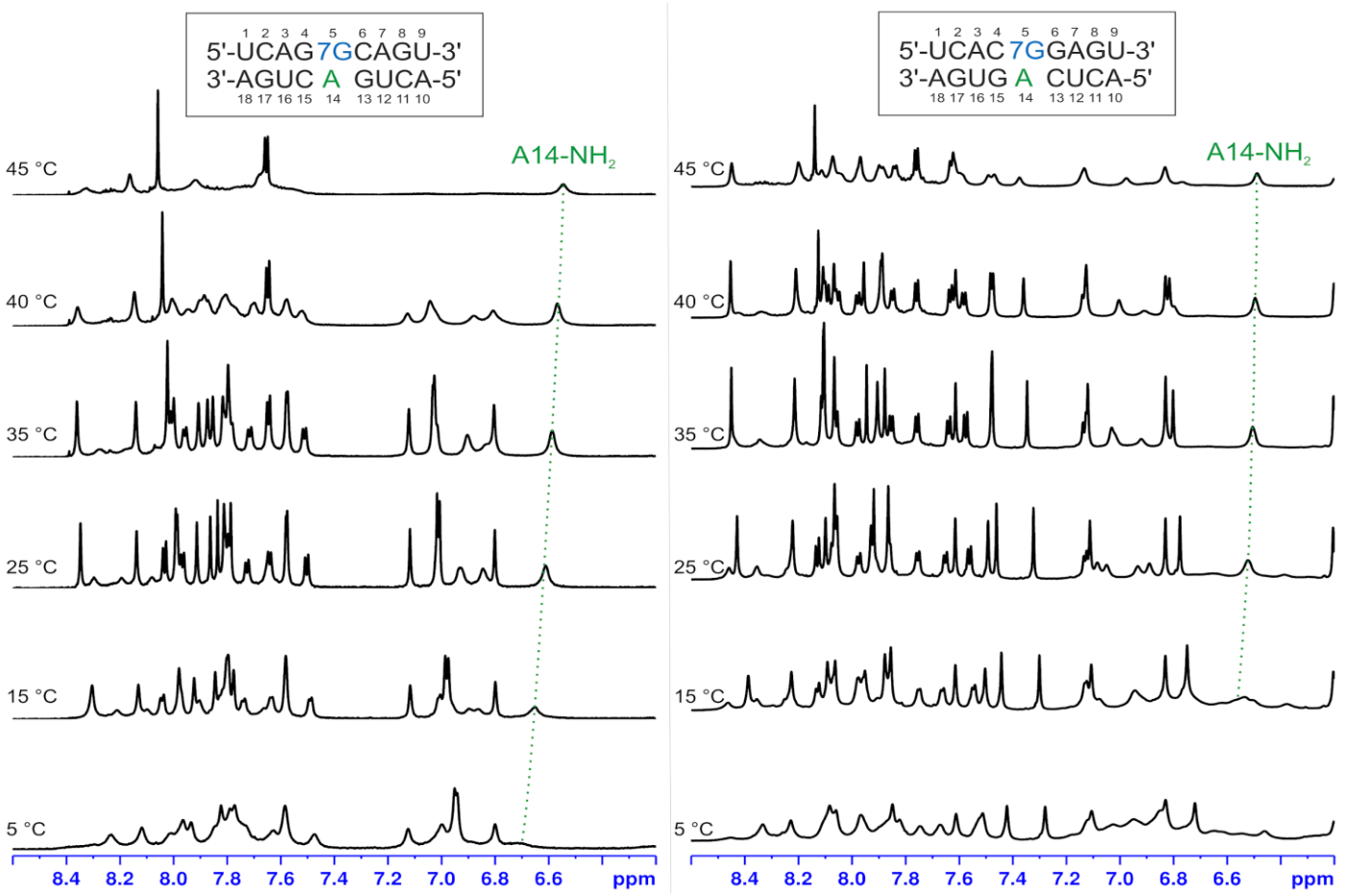


Figure S2. The amino portion of the ^{1}H-^{15}N SOFAST-HMQC spectrums of the **7G**-A containing duplexes. Positive contour levels in black, negative in red. Some small negative peaks visible are due to a limited signal-to-noise ratio.

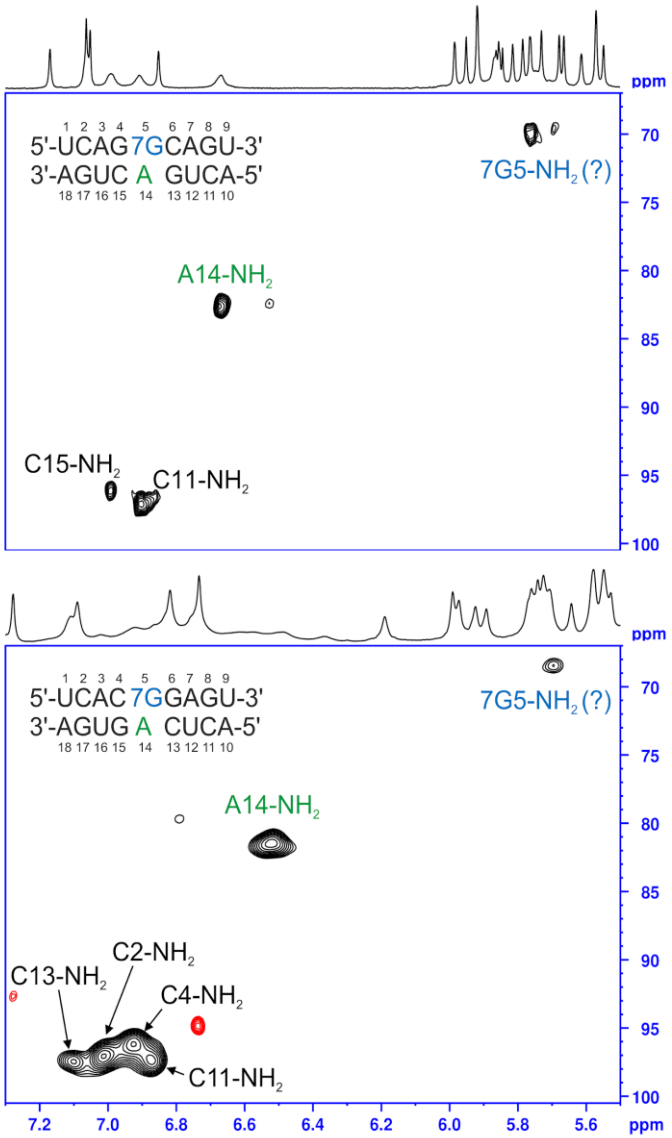


Figure S3. The imino regions of the standard (top) and ^{15}N-selective ^{1}H NMR spectra of one of the duplexes containing the **7G**-G mismatch with the residue G14 site-specifically enriched in ^{15}N. Spectra recorded at 25 °C. Resonances belonging to a second, minor spectral form marked in red stars.

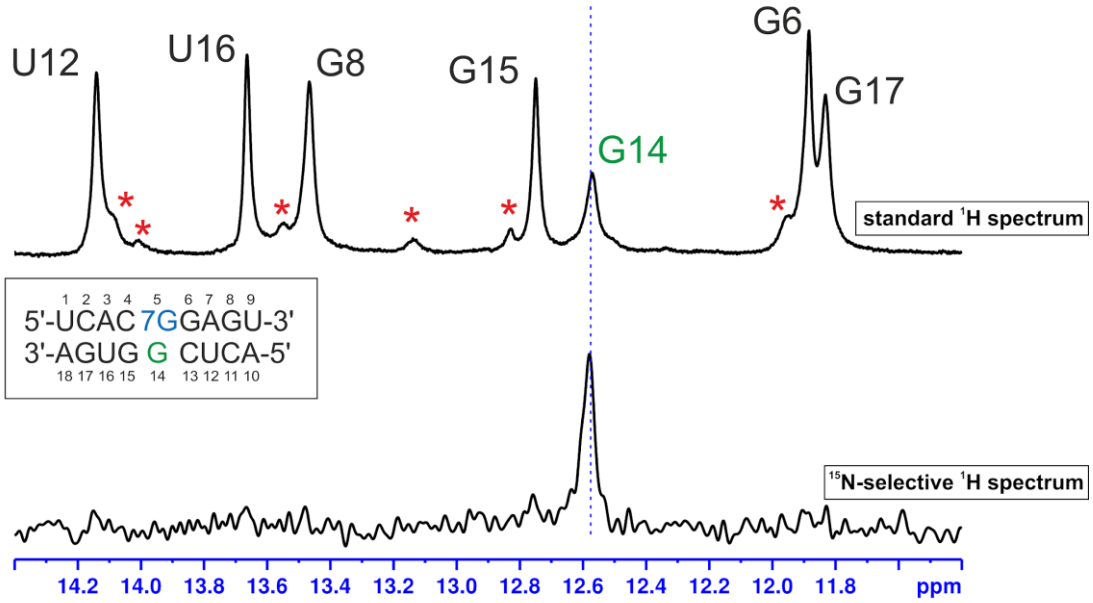


Figure S4. Structures of the tautomers of (A) m^9Gua, (B) m^7Gua studied in this work: Keto-N1H (T1), Keto-N3H (T2), Cis-enol (T3), Trans-enol (T4), N1H,N3H,N2-imino-1 (T5) and N1H,N3H,N2-imino 2 (T6).

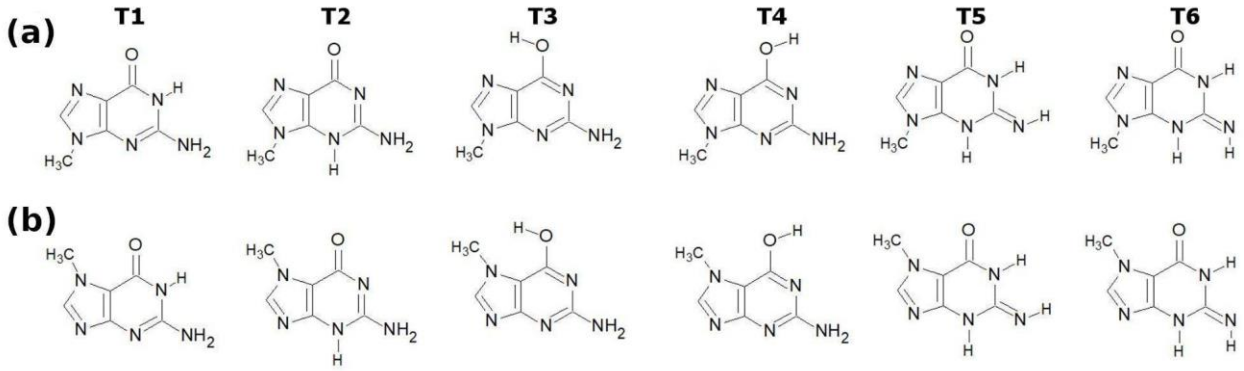


Figure S5. Comparison of the Zero-point-energy (ZPE)-corrected relative free energies of m^7Gua, m^9Gua tautomers with respect to keto-N1H(T1) (the most stable tautomer) in the (a) gas phase and (b) in the solvent phase.

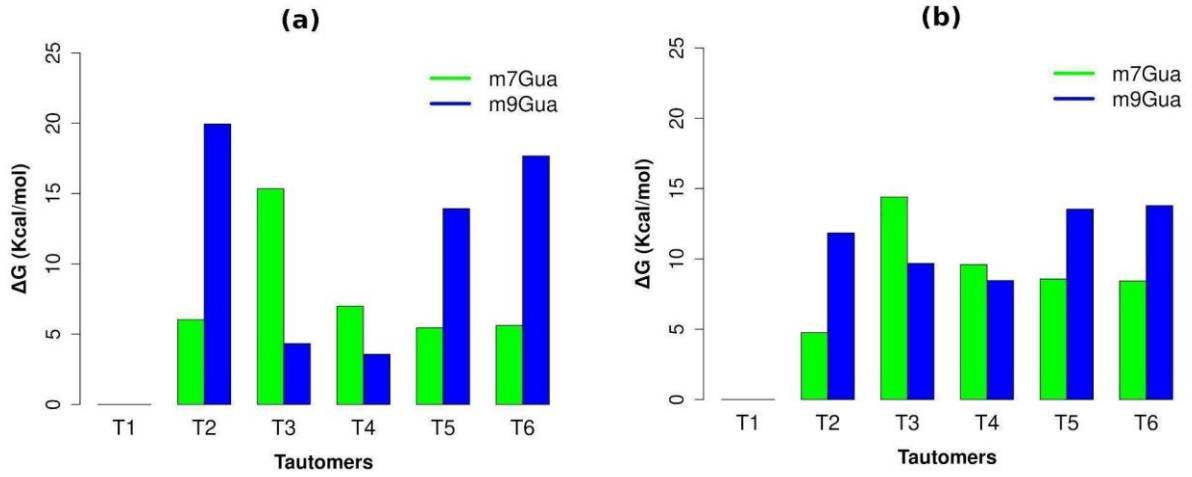
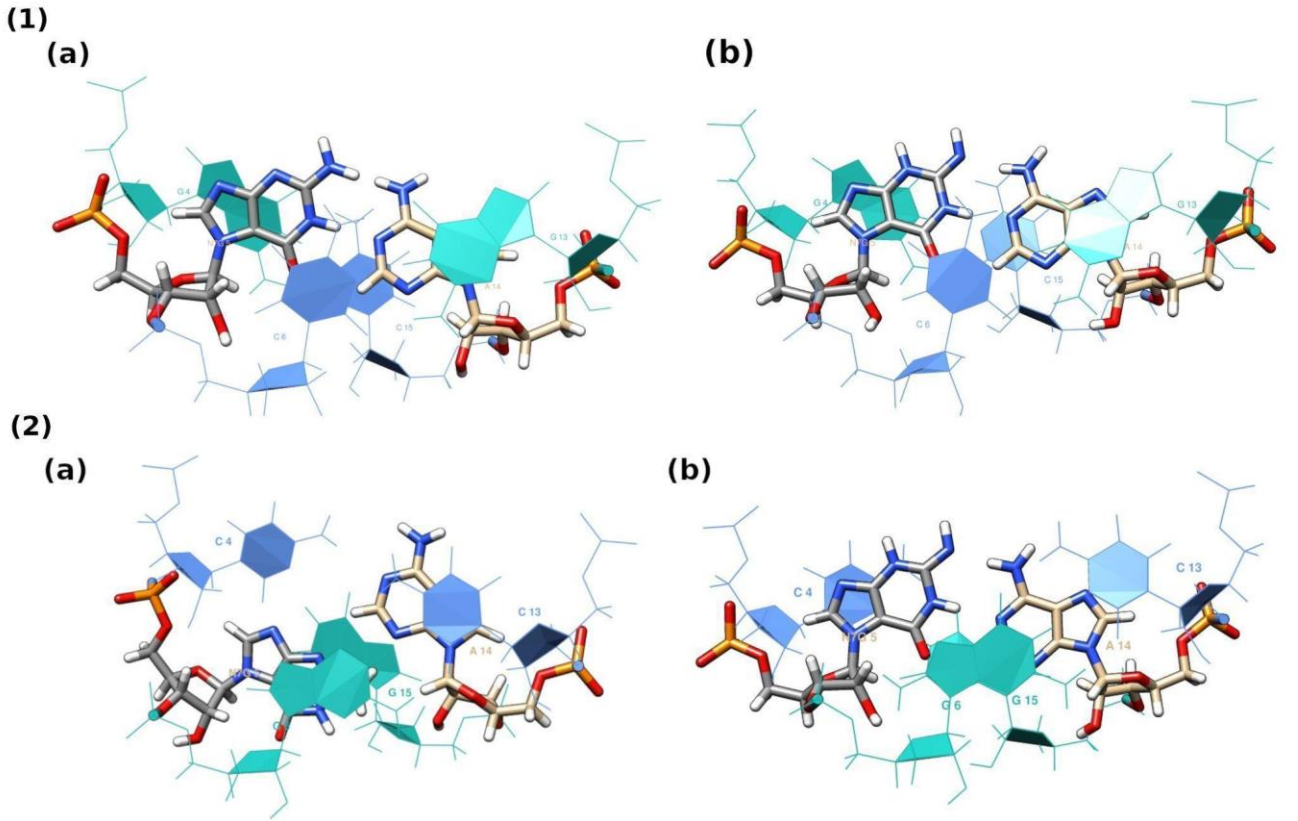


Figure S6. (A) Observed stacked geometries of (1) (a) 5’G7G(N1H)C-3’CAG (GM1) and (b) 5’G7G(N1H,N3H,N2-imino)C-3’CAG (GM3) and (2) (a) 5’C7G(N1H)G-3’GAC (GM2) and (b) 5’C7G(N1H,N3H,N2-imino)G-3’GAC (GM3). 7G5 and A14 residues are shown in sticks (carbons are shown in grey for 7G15 and in tan for A14, nitrogens in blue, oxygens in red and hydrogens in white), for residues 4,6,13 and 15, G-residues are shown in light seagreen, and C-residues are shown in cornflower blue color with base pair 4-15 in the back, 5-14 in the middle and 6-13 in the front.



(B) Observed stacked geometries of (a) 7G(N1H)-G (GM1); (b) 7G(N1H)-G (GM2); (c) 7G(N3H)-G (GM3) for (1) 5'G7GC-3'CGG and (2) 5'C7GG-3'GGC contexts. 7G5 and G14 residues are shown in sticks (carbons are shown in grey for 7G15 and in tan for A14, nitrogens in blue, oxygens in red and hydrogens in white), for residues 4,6,13 and 15, the G-residues are shown in light seagreen, and C-residues are shown in cornflower blue color with base pair 4-15 in the back, 5-14 in the middle and 6-13 in the front.

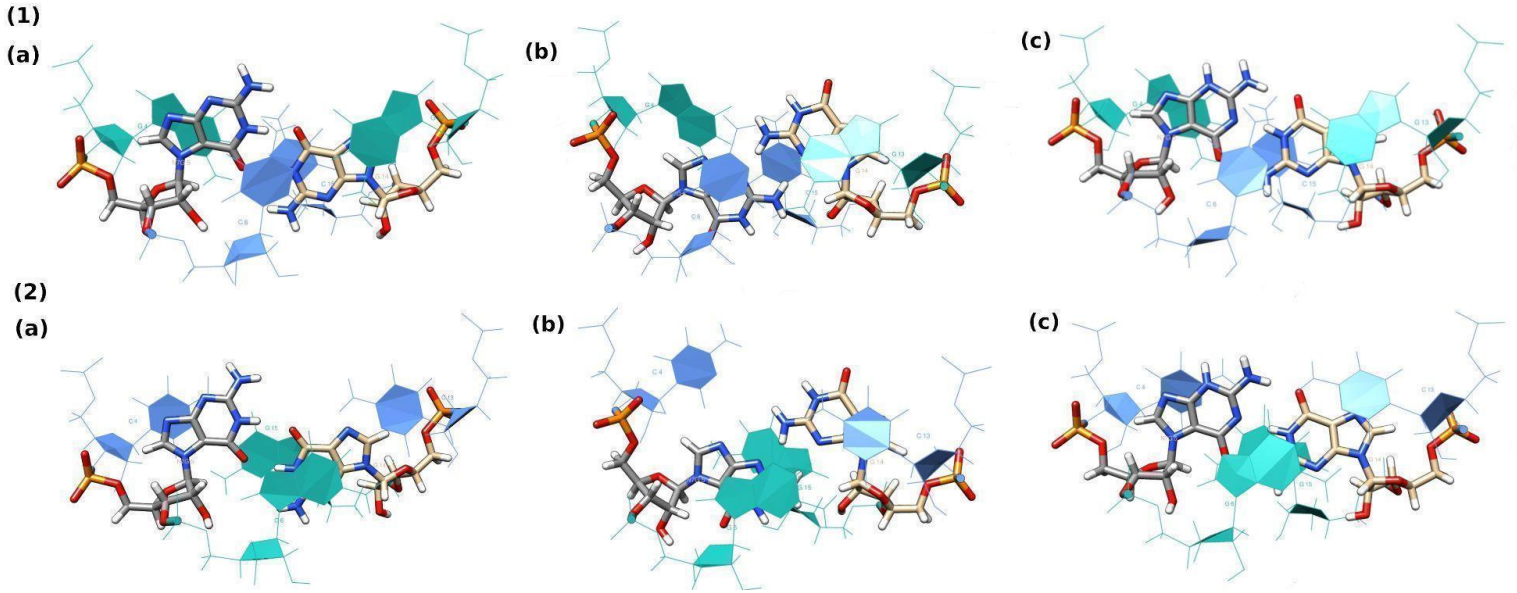
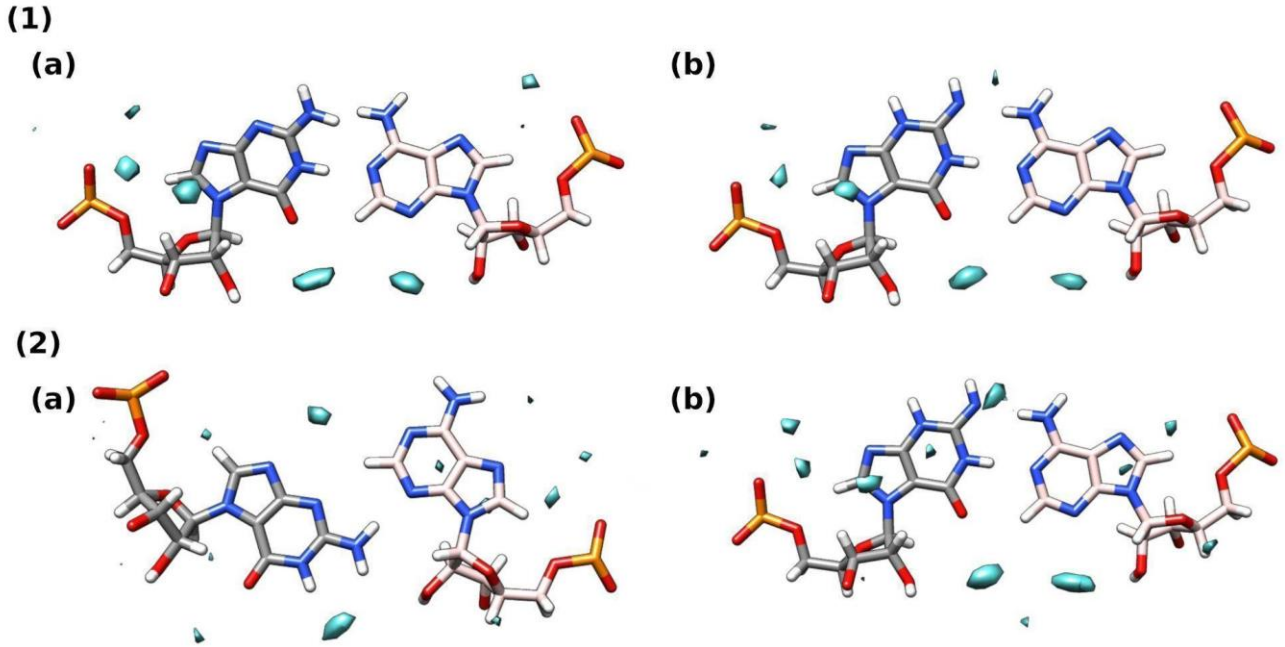


Figure S7. (A) Observed water occupancy maps around (1) (a) 5’G7G(N1H)C-3’CAG (GM1) and (b) 5’G7G(N1H,N3H,N2-imino)C-3’CAG (GM3) and (2) (a) 5’C7G(N1H)G-3’GAC (GM2) and (b) 5’C7G(N1H,N3H,N2-imino)G-3’GAC (GM3). 7G is shown in grey.



(B) Observed water occupancy maps around (a) 7G(N1H)-G (GM1); (b) 7G(N1H)-G (GM2); (c) 7G(N3H)-G (GM3) for (1) 5'G7GC-3'CGG and (2) 5'C7GG-3'GGC contexts. 7G is shown in grey.

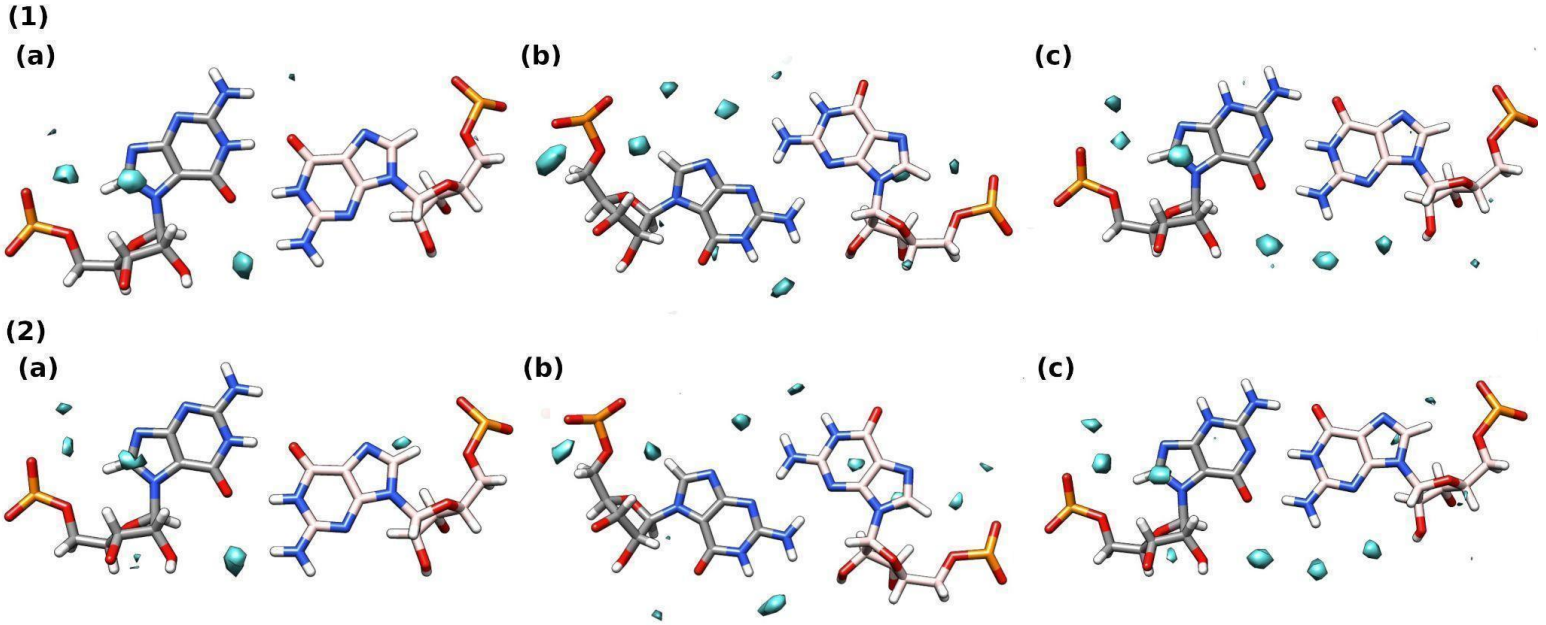


Figure S8. ^{1}H spectrum of 7G in DMSO-d6 at 25 °C

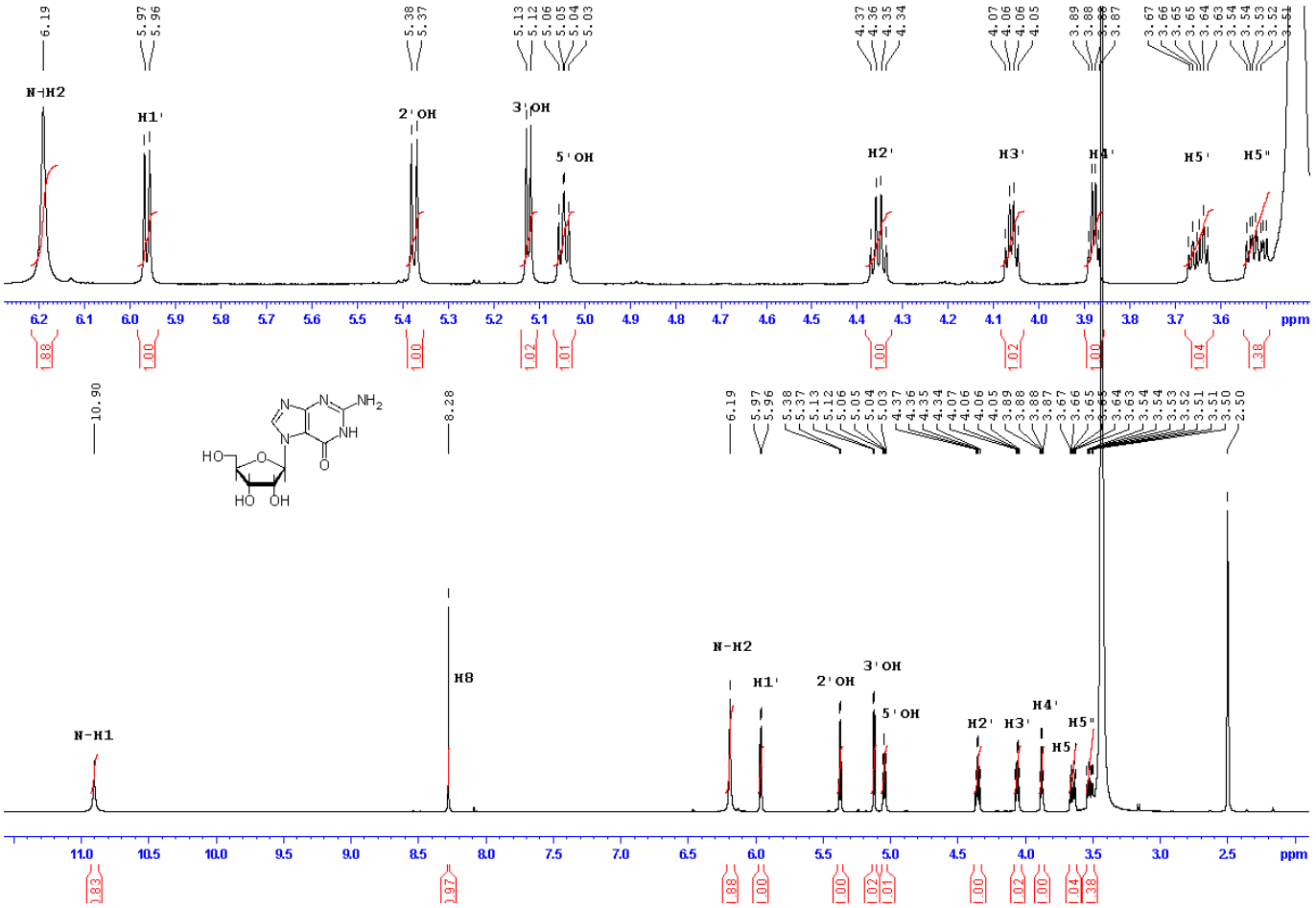


Figure S9. ^{13}C spectrum of 7G in DMSO-d6 at 25 °C

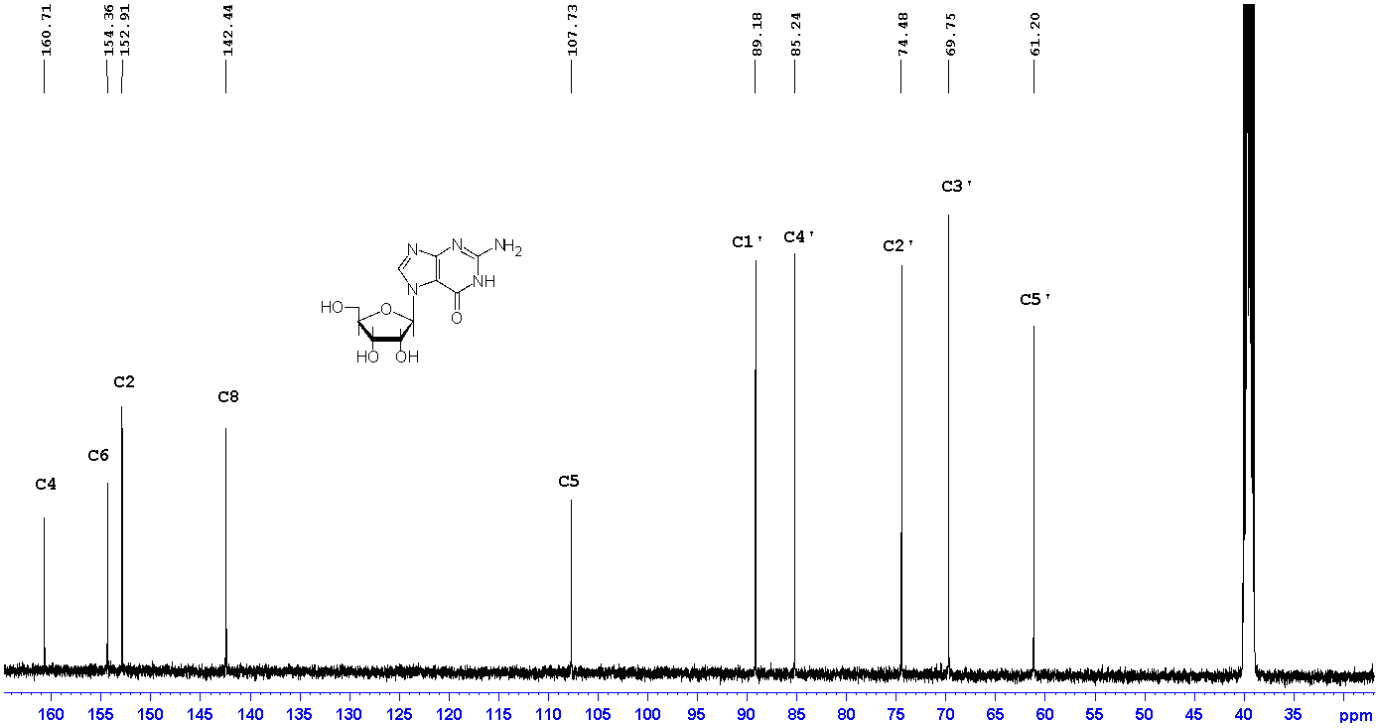


Figure S10. ^{1}H-^{1}H COSY spectrum of 7G in DMSO-d6 at 25 °C

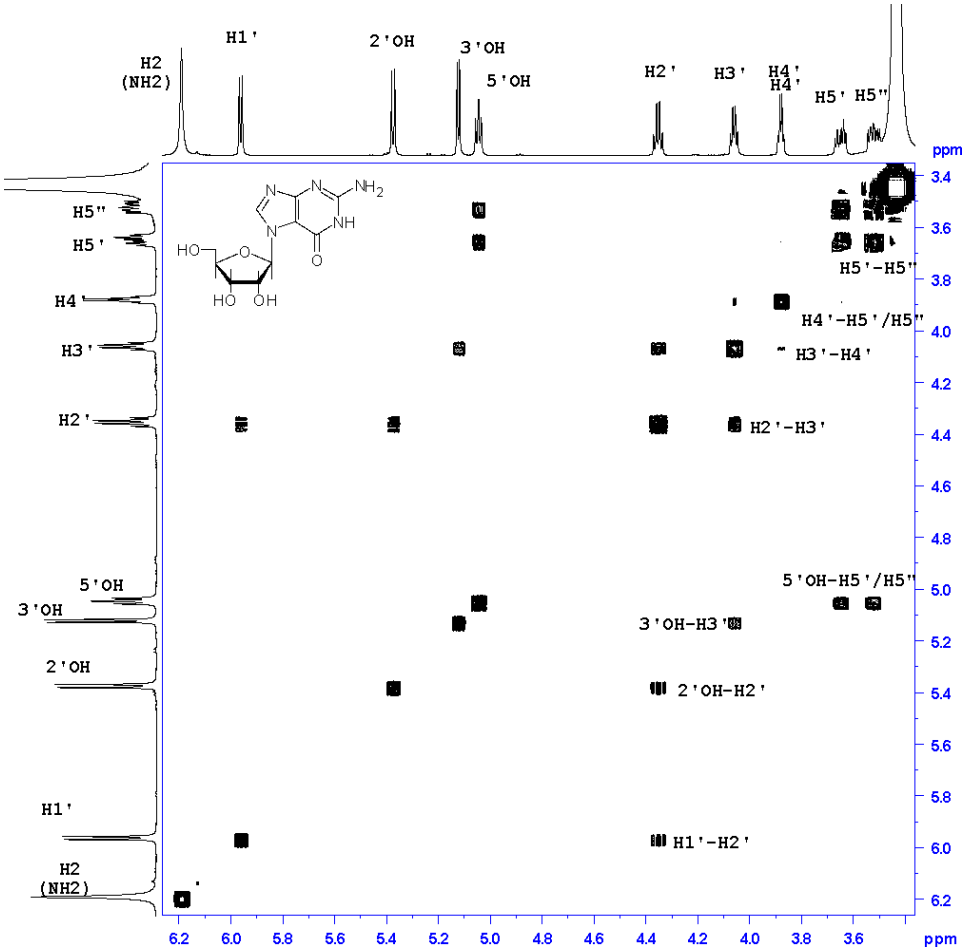


Figure S11. ^{1}H-^{13}C gHSQC spectrum of 7G in DMSO-d6 at 25 °C

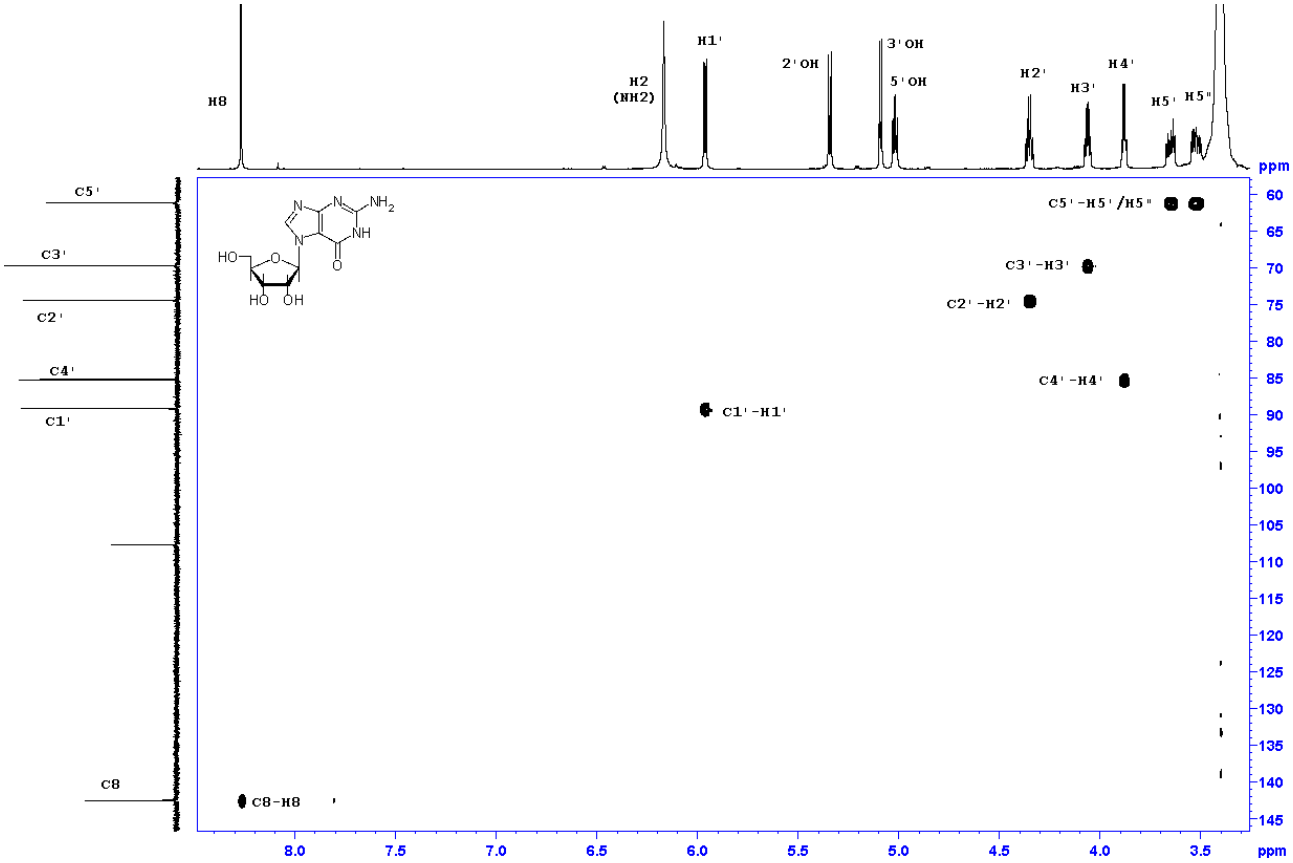


Figure S12. ^{1}H-^{13}C gHMBC spectrum of 7G in DMSO-d6 at 25 °C

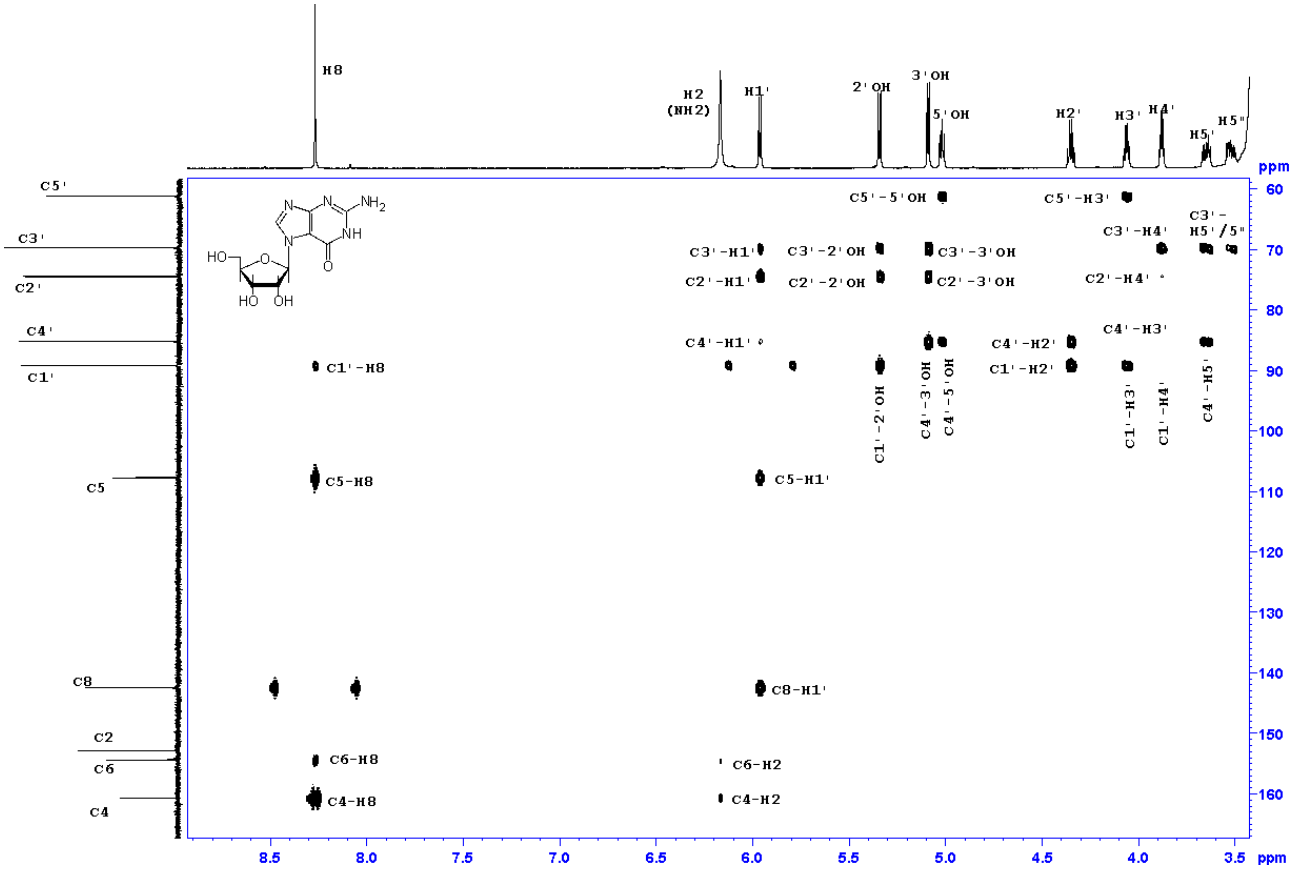


Figure S13. ^{1}H-^{15}N gHSQC spectrum of 7G in DMSO-d6 at 25 °C

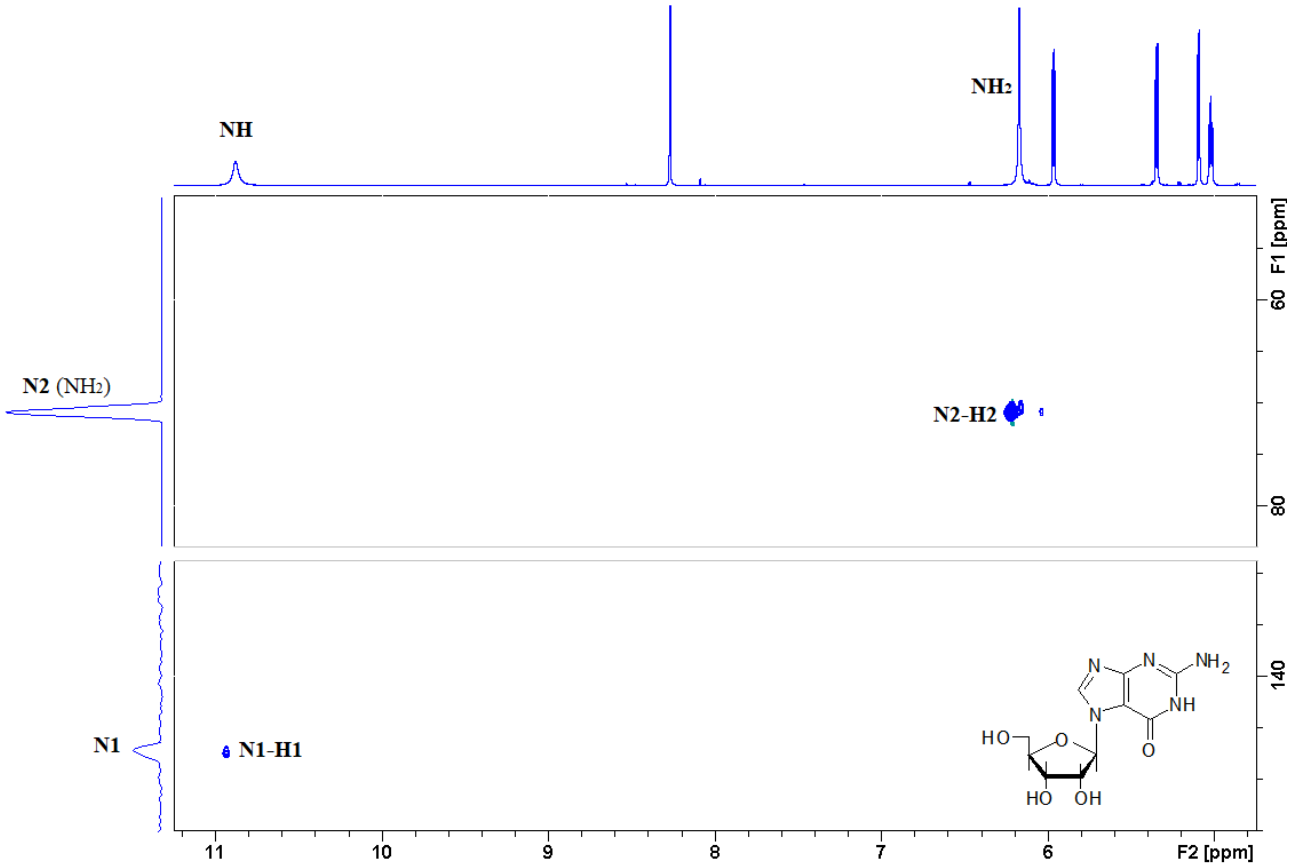


Figure S14. ^{1}H-^{15}N gHMBC spectrum of 7G in DMSO-d6 at 25 °C

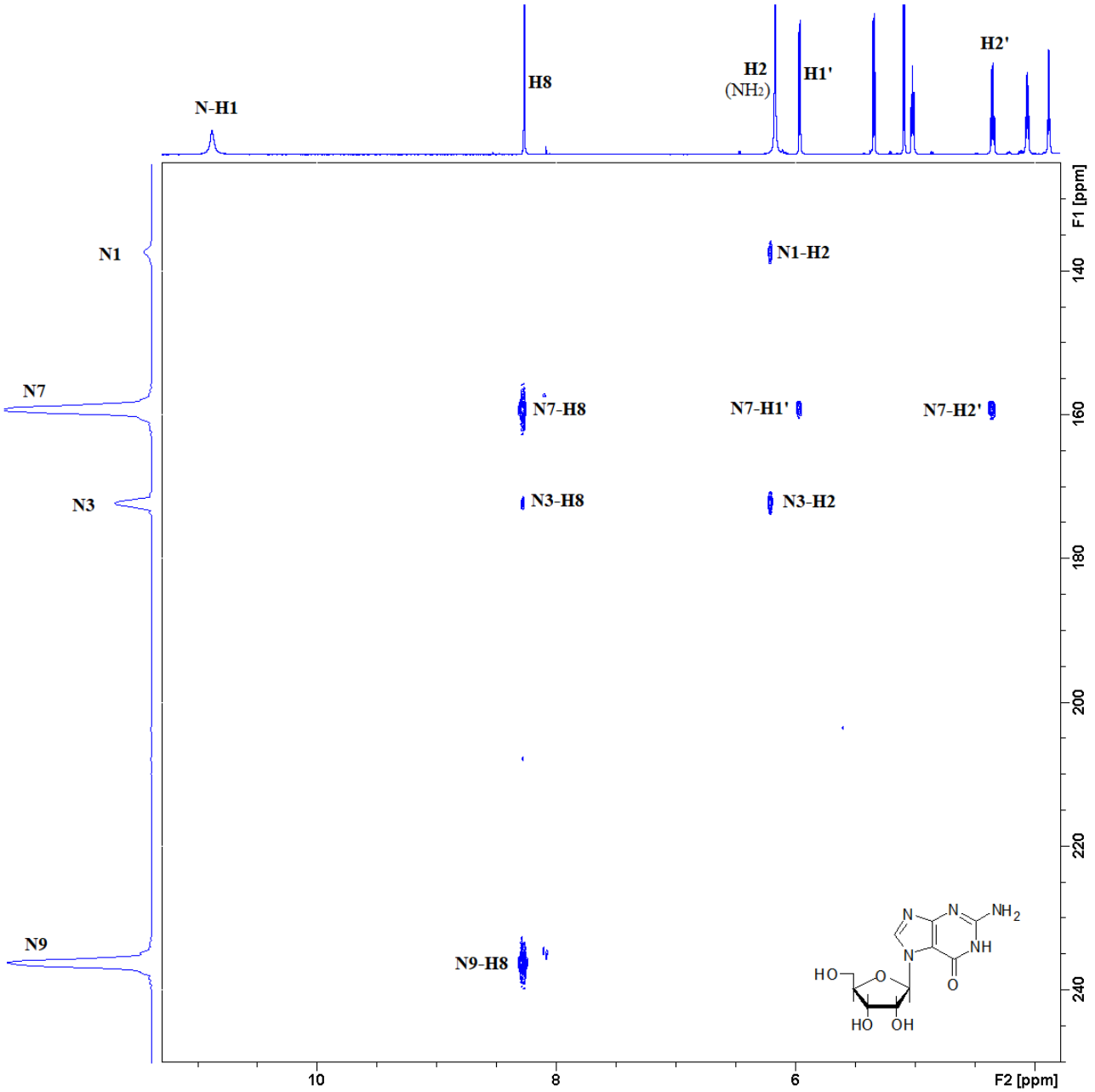


Figure S15. Quantum mechanical energy profiles around **χ** torsional angles (O4'-C1'-N7-C5) for 7G-N1H corresponding to the four conformational schemes, respectively.

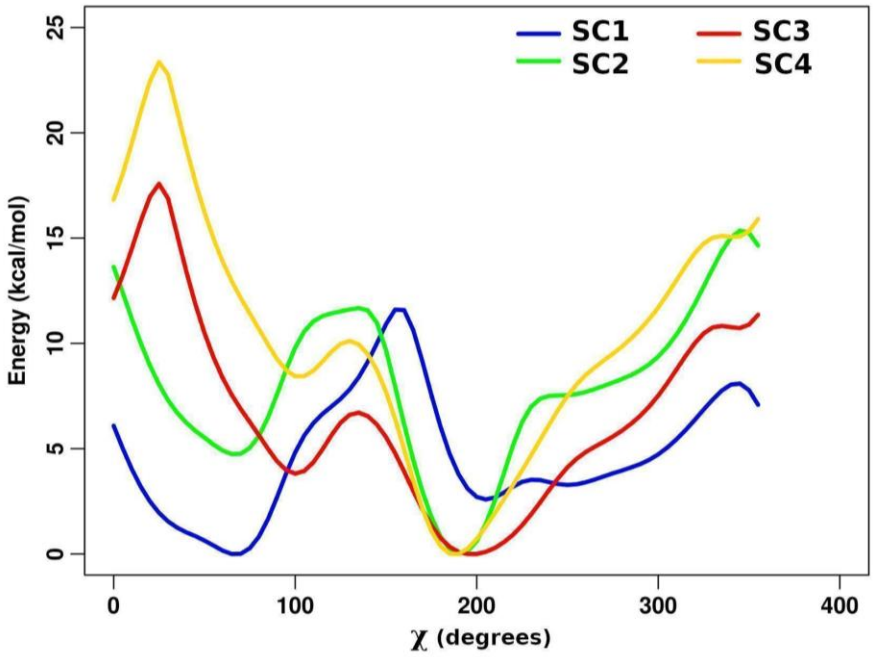


Figure S16. Energy profiles of the **χ**+180° (O4'-C1'-N7-C8) torsional angles for 7G corresponding to QM calculations (black), MM calculations with the FF99 parameter sets keeping the glycosidic torsion parameters zero (red) and MM calculations with the FF99 parameter sets combined with the newly derived **χ** torsional parameters and the newly developed partial atomic charges (FF99_χ_{KOL0}) (green) by fitting the difference between the QM and MM energies for (a) SC1, (b) SC2, (c) SC3 and (d) SC4 respectively. The minimum energies were set to zero for convenience.

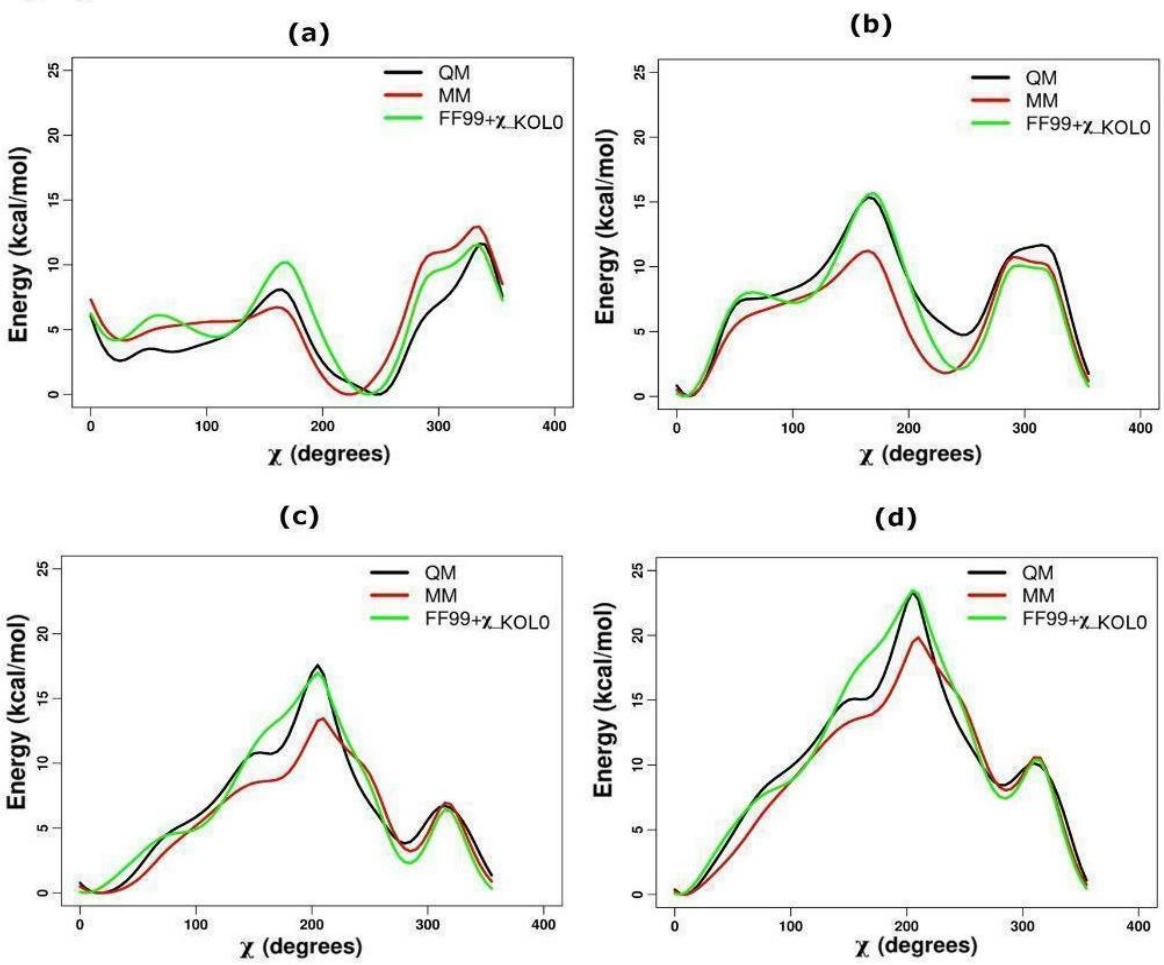
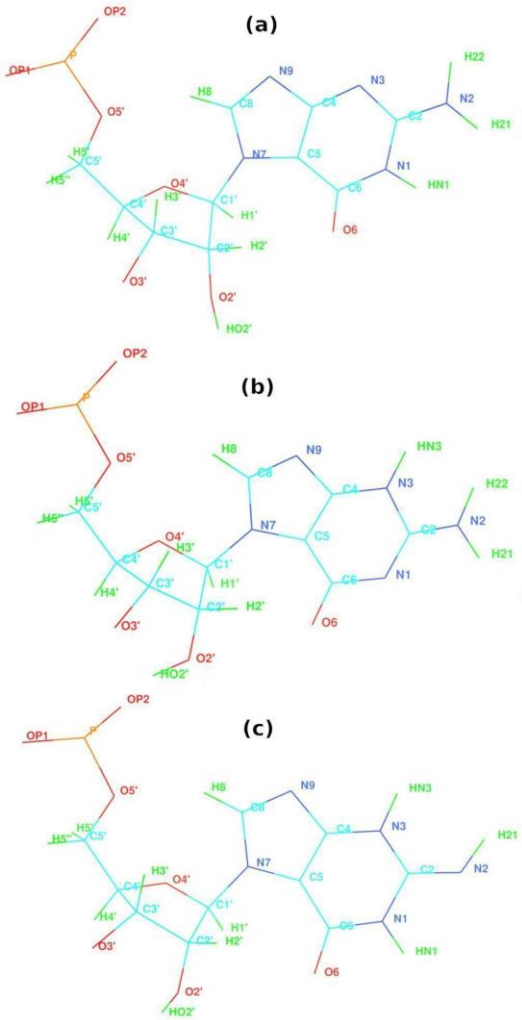


Figure S17. Representative diagrams of the 7G tautomers (a) 7G-N1H; (b) 7G-N3H and (c) 7G-N1H,N3H,N2-imino in their nucleotide forms with the atom names.



APPENDIX 1

AMBER preparatory file for 7G-N1H

```
  0    0    2

Developed at Institute of Bioorganic Chemistry PAS by Nivedita Dutta
7G.res
7G   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0       .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0       .0      .00000
   3  DUMM  DU    M    2   1   0     1.523   111.21      .0      .00000
   4  P     P     M    3   2   1     1.609   119.956  -133.992  1.1662
   5  OP1   O     E    4   3   2     1.485   107.081   145.987 -0.7760
   6  OP2   O     E    4   3   2     1.486   109.428    15.165 -0.7760
   7  O5'   OS    M    4   3   2     1.593   103.856   -99.934 -0.4989
   8  C5'   CI    M    7   4   3     1.426   120.975   -62.582  0.0558
   9  H5'   H1    E    8   7   4     1.091   109.725    30.107  0.0679
  10  H5"   H1    E    8   7   4     1.090   109.765   -89.566  0.0679
  11  C4'   CT    M    8   7   4     1.511   110.226   150.387  0.1065
  12  H4'   H1    E   11   8   7     1.091   107.708   175.369  0.1174
  13  O4'   OS    M   11   8   7     1.454   109.053   -60.938 -0.3548
  14  C1'   CT    M   13  11   8     1.418   109.791   140.832  0.1358
  15  N7    N*    S   14  13  11     1.463   108.323  -115.257 -0.1054
  16  C5    CB    S   15  14  13     1.374   126.614  -139.229 -0.1525
  17  C6    C     B   16  15  14     1.344   127.448    -0.092  0.4915
  18  N1    NA    B   17  16  15     1.332   110.623  -179.820 -0.5096
  19  C2    CA    B   18  17  16     1.338   129.242    -0.105  0.6987
  20  N3    NC    S   19  18  17     1.350   118.594     0.063 -0.6296
  21  C4    CB    S   20  19  18     1.405   117.753     0.077  0.5754
  22  N9    NB    S   21  20  19     1.388   132.274   179.802 -0.6217
  23  C8    CK    S   22  21  20     1.311   103.867  -179.994  0.2691
  24  H8    H5    E   23  22  21     1.080   123.107   179.801  0.1270
  25  N2    N2    B   19  18  17     1.332   120.712  -179.953 -0.8829
  26  H21   H     E   25  18  17     1.010   119.949     0.027  0.3901
  27  H22   H     E   25  18  17     1.009   119.979   179.985  0.3901
```

```
  28  HN1   H     E   18  17  16     1.091   115.414   179.954  0.3546
  29  O6    O     E   17  16  15     1.239   124.675     0.199 -0.5356
  30  H1'   H2    E   14  13  11     1.090   107.041   128.337  0.1210
  31  C2'   CT    M   14  13  11     1.529   107.230     7.740  0.0670
  32  O2'   OH    S   31  14  13     1.415   108.728    88.062 -0.6139
  33  HO2'  HO    E   32  31  14     0.957   106.681  -130.227  0.4186
  34  H2'   H1    E   31  14  13     1.092   111.690  -146.692  0.0972
  35  C3'   CT    M   31  14  13     1.523   100.960   -28.912  0.2022
  36  H3'   H1    E   35  31  14     1.089   112.602   -81.930  0.0615
  37  O3'   OS    M   35  31  14     1.425   112.670   157.951 -0.5246

LOOP
   C4   C5
   N7   C8
  C4'  C3'

IMPROPER
   C6   C4   C5   N7
   C5   N1   C6   O6
   C6   C2   N1  HN1
   N1   N3   C2   N2
   C2  H21   N2  H22
   C5   N3   C4   N9
   H8   N7   C8   N9
  C1'   C8   N7   C5

DONE
STOP
```

AMBER preparatory files for 7G-N3H

```
   0    0    2

Developed at Institute of Bioorganic Chemistry PAS by Nivedita Dutta
7G.res
7G  INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0       .0      .00000
```

```
 2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
 3  DUMM  DU    M    2   1   0     1.523   111.21       .0      .00000
 4  P     P     M    3   2   1     1.609   119.956  -133.992   1.1662
 5  OP1   O2    E    4   3   2     1.485   107.081   145.987  -0.7760
 6  OP2   O2    E    4   3   2     1.486   109.428    15.165  -0.7760
 7  O5'   OS    M    4   3   2     1.593   103.856   -99.934  -0.4989
 8  C5'   CI    M    7   4   3     1.426   120.975   -62.582   0.0558
 9  H5'   H1    E    8   7   4     1.091   109.725    30.107   0.0679
10  H5"   H1    E    8   7   4     1.090   109.765   -89.566   0.0679
11  C4'   CT    M    8   7   4     1.511   110.226   150.387   0.1065
12  H4'   H1    E   11   8   7     1.091   107.708   175.369   0.1174
13  O4'   OS    M   11   8   7     1.454   109.053   -60.938  -0.3548
14  C1'   CT    M   13  11   8     1.418   109.791   140.832   0.0896
15  N7    N*    S   14  13  11     1.463   108.323  -115.257  -0.0442
16  C5    CB    S   15  14  13     1.374   126.614  -139.229  -0.1063
17  C6    Q2    B   16  15  14     1.344   127.448    -0.092   0.6844
18  N1    NC    S   17  16  15     1.332   110.623  -179.820  -0.7313
19  C2    CA    B   18  17  16     1.338   129.242    -0.105   0.7943
20  N3    NA    B   19  18  17     1.350   118.594     0.063  -0.5465
21  C4    CB    S   20  19  18     1.405   117.753     0.077   0.3786
22  N9    NB    S   21  20  19     1.388   132.274   179.802  -0.5942
23  C8    Q1    S   22  21  20     1.311   103.867  -179.994   0.2237
24  H8    H5    E   23  22  21     1.080   123.107   179.801   0.1430
25  HN3   HQ    E   20  19  18     1.010   121.123  -179.896   0.3903
26  N2    N2    B   20  19  18     1.332   120.712  -179.953  -0.9297
27  H21   H     E   26  20  19     1.010   119.949     0.027   0.4024
28  H22   H     E   26  20  19     1.009   119.979   179.985   0.4024
29  O6    O     E   17  16  15     1.239   124.675     0.199  -0.5527
30  H1'   H2    E   14  13  11     1.090   107.041   128.337   0.1122
31  C2'   CT    M   14  13  11     1.529   107.230     7.740   0.0670
32  O2'   OH    S   31  14  13     1.415   108.728    88.062  -0.6139
33  HO2'  HO    E   32  31  14     0.957   106.681  -130.227   0.4186
34  H2'   H1    E   31  14  13     1.092   111.690  -146.692   0.0972
35  C3'   CT    M   31  14  13     1.523   100.960   -28.912   0.2022
36  H3'   H1    E   35  31  14     1.089   112.602   -81.930   0.0615
37  O3'   OS    M   35  31  14     1.425   112.670   157.951  -0.5246
```

AMBER preparatory files for 7G-N1H,N3H,N2-imino

```
  0    0    2

This is a remark line
molecule.res
7G   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0       .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0       .0      .00000
   3  DUMM  DU    M    2   1   0     1.523   111.21       .0      .00000
   4  P     P     M    3   2   1     1.609   119.956  -133.992  1.1662
   5  OP1   O2    E    4   3   2     1.485   107.081   145.987 -0.7760
   6  OP2   O2    E    4   3   2     1.486   109.428    15.165 -0.7760
   7  O5'   OS    M    4   3   2     1.593   103.856   -99.934 -0.4989
   8  C5'   CI    M    7   4   3     1.426   120.975   -62.582  0.0558
   9  H5'   H1    E    8   7   4     1.091   109.725    30.107  0.0679
  10  H5"   H1    E    8   7   4     1.090   109.765   -89.566  0.0679
  11  C4'   CT    M    8   7   4     1.511   110.226   150.387  0.1065
  12  H4'   H1    E   11   8   7     1.091   107.708   175.369  0.1174
  13  O4'   OS    M   11   8   7     1.454   109.053   -60.938 -0.3548
  14  C1'   CT    M   13  11   8     1.418   109.791   140.832  0.1638
  15  N7    N*    S   14  13  11     1.463   108.323  -115.257 -0.1049
  16  C5    CB    S   15  14  13     1.374   126.614  -139.229  0.0143
  17  C6    Q2    B   16  15  14     1.339   126.854    -0.089  0.3474
  18  N1    NA    B   17  16  15     1.362   110.006  -179.821 -0.2614
  19  C2    CA    B   18  17  16     1.319   129.535    -0.114  0.4061
  20  N3    NP    B   19  18  17     1.390   118.313     0.082 -0.2727
  21  C4    CB    S   20  19  18     1.395   117.046     0.057  0.2187
  22  N9    NB    S   21  20  19     1.388   131.556   179.812 -0.5708
  23  C8    Q1    S   22  21  20     1.311   103.867  -179.992  0.2667
  24  H8    H5    E   23  22  21     1.080   123.107   179.801  0.1256
  25  HN3   HQ    E   20  19  18     1.000   120.135  -179.895  0.3181
  26  N2    ND    S   19  18  17     1.251   121.753  -179.948 -0.8397
  27  H21   H     E   25  19  18     1.009   120.750   179.991  0.3788
  28  HN1   H     E   18  17  16     1.004   114.611   179.895  0.3077
  29  O6    O     E   17  16  15     1.229   125.845     0.219 -0.4958
  30  H1'   H2    E   14  13  11     1.090   107.041   128.337  0.1141
```

```
 31  C2'   CT    M   14  13  11     1.529   107.230     7.740  0.0670
 32  O2'   OH    S   31  14  13     1.415   108.728    88.062 -0.6139
 33  HO2'  HO    E   32  31  14     0.957   106.681  -130.227  0.4186
 34  H2'   H1    E   31  14  13     1.092   111.690  -146.692  0.0972
 35  C3'   CT    M   31  14  13     1.523   100.960   -28.912  0.2022
 36  H3'   H1    E   35  31  14     1.089   112.602   -81.930  0.0615
 37  O3'   OS    M   35  31  14     1.425   112.670   157.951 -0.5246

LOOP
  C4   C5
  N7   C8
 C4'  C3'

IMPROPER
  C6   C4   C5   N7
  C5   N1   C6   O6
  C2   C4   N3  HN3
  N1   N3   C2   N2
  C5   N3   C4   N9
  H8   N7   C8   N9
 C1'   C8   N7   C5

DONE
STOP
```

AMBER parameter files for 7G-N1H

```
Developed at Institute of Bioorganic Chemistry PAS by Nivedita Dutta
MASS
Q1 12.01         0.360

BOND
Q1-H5  367.0    1.080
Q1-N*  440.0    1.371
Q1-NB  529.0    1.304

ANGLE
H5-Q1-N*    50.0      123.05
```

H5-Q1-NB 50.0 123.05
N*-Q1-NB 70.0 113.90
CB-N*-Q1 70.0 105.40
CT-N*-Q1 70.0 128.80
CB-NB-Q1 70.0 103.80
CT-OS-HO 55.0 108.50 manually added from parm99.dat
CI-OS-HO 55.0 108.50 manually modified from parm99.dat & frcmod.parmbsc0
N*-CB-C 70.0000 134.500
NB-CB-NC 70.0000 122.000

DIHE
X -Q1-N*-X 4 6.80 180.0 2.
X -Q1-NB-X 2 20.00 180.0 2.
OS-CT-N*-Q1 1 0.0649123 180.0 -4. chi KOL0
OS-CT-N*-Q1 1 0.434356 180.0 -3. chi KOL0
OS-CT-N*-Q1 1 0.952733 0.0 -2. chi KOL0
OS-CT-N*-Q1 1 1.12182 180.0 1. chi KOL0
CT-CT-N*-Q1 1 0.0797247 180.0 -4. chi KOL0
CT-CT-N*-Q1 1 0.683783 180.0 -3. chi KOL0
CT-CT-N*-Q1 1 0.686632 180.0 -2. chi KOL0
CT-CT-N*-Q1 1 0.808183 0.0 1. chi KOL0

IMPROPER
X -X -Q1-H5 1.1 180. 2.
CB-CT-N*-Q1 1.0 180. 2.

NONBON
Q1 1.9080 0.0860

AMBER parameter files for 7G-N3H

Developed at Institute of Bioorganic Chemistry PAS by Nivedita Dutta
MASS
Q1 12.01 0.360
Q2 12.01 0.616
HQ 1.008 0.161

BOND
Q1-H5 367.0 1.080

```
Q1-N*  440.0    1.371
Q1-NB  529.0    1.304
NA-CB  461.0    1.354
NA-HQ  434.10   1.010
Q2-O   570.0    1.229
Q2-NC  457.0    1.358
Q2-CB  447.0    1.419

ANGLE
H5-Q1-N*    50.0      123.05
H5-Q1-NB    50.0      123.05
N*-Q1-NB    70.0      113.90
CB-N*-Q1    70.0      105.40
CT-N*-Q1    70.0      128.80
CB-NB-Q1    70.0      103.80
CT-OS-HO    55.0      108.50    manually added from parm99.dat
CI-OS-HO    55.0      108.50    manually modified from parm99.dat & frcmod.parmbsc0
N*-CB-Q2    70.0000   134.500   QM_KOL0
NB-CB-NA    70.0000   126.900   QM_KOL0
CB-NA-HQ    50.0000   120.100   QM_KOL0
CA-NA-HQ    50.0000   122.200   QM_KOL0
CA-NA-CB    70.0000   117.900   QM_KOL0
CB-Q2-NC    70.0000   114.300   QM_KOL0
CB-CB-NA    70.0000   122.700   QM_KOL0
CA-NC-Q2    70.0000   122.600   QM_KOL0
O- Q2-NC    80.0000   121.600   QM_KOL0
O- Q2-CB    80.0000   124.100   QM_KOL0
Q2-CB-CB    63.0      119.20    parm10.dat

DIHE
X -Q1-N*-X    4    6.80        180.0            2.
X -Q1-NB-X    2   20.00        180.0            2.
OS-CT-N*-Q1   1    0.0649123   180.0           -4.  chi KOL0
OS-CT-N*-Q1   1    0.434356    180.0           -3.  chi KOL0
OS-CT-N*-Q1   1    0.952733      0.0           -2.  chi KOL0
OS-CT-N*-Q1   1    1.12182     180.0            1.  chi KOL0
CT-CT-N*-Q1   1    0.0797247   180.0           -4.  chi KOL0
CT-CT-N*-Q1   1    0.683783    180.0           -3.  chi KOL0
CT-CT-N*-Q1   1    0.686632    180.0           -2.  chi KOL0
```

CT-CT-N*-Q1 1 0.808183 0.0 1. chi KOL0
NB-CB-NA-HQ 1 4.150 180.0 2.
CA-NA-CB-NB 1 4.150 180.0 2.
CB-CB-NA-CA 1 4.150 180.0 2.
CB-CB-NA-HQ 1 4.150 180.0 2.
X -Q2-NC-X 2 8.00 180.0 2. Parm10.dat
X -Q2-O -X 4 11.20 180.0 2. Parm10.dat
X -Q2-CB-X 4 12.00 180.0 2. Parm10.dat

IMPROPER
X -X -Q1-H5 1.1 180. 2.
CB-CT-N*-Q1 1.0 180. 2.

NONBON
Q1 1.9080 0.0860
Q2 1.9080 0.0860
HQ 0.6000 0.0157

AMBER parameter files for 7G-N1H,N3H,N2-imino

Developed at Institute of Bioorganic Chemistry PAS by Nivedita Dutta
MASS
Q1 12.01 0.360
Q2 12.01 0.616
HQ 1.008 0.161
ND 14.01 0.530
NP 14.01 0.530

BOND
Q1-H5 367.0 1.080
Q1-N* 440.0 1.371
Q1-NB 529.0 1.304
CQ-ND 481.0 1.340
NA-HQ 434.10 1.010
NP-HQ 434.10 1.010
Q2-O 570.0 1.229
Q2-NA 457.0 1.358
Q2-CB 447.0 1.419

NA-CB 461.0 1.3540
NP-CB 461.0 1.3540
H -ND 434.0 1.010
CA-ND 481.0 1.340
NP-CA 483.0 1.3390

MASS
Q1 12.01 0.360
Q2 12.01 0.616
HQ 1.008 0.161
ND 14.01 0.530
NP 14.01 0.530

BOND
Q1-H5 367.0 1.080
Q1-N* 440.0 1.371
Q1-NB 529.0 1.304
CQ-ND 481.0 1.340
NA-HQ 434.10 1.010
NP-HQ 434.10 1.010
Q2-O 570.0 1.229
Q2-NA 457.0 1.358
Q2-CB 447.0 1.419
NA-CB 461.0 1.3540
NP-CB 461.0 1.3540
H -ND 434.0 1.010
CA-ND 481.0 1.340
NP-CA 483.0 1.3390

DIHE
X -Q1-N*-X 4 6.80 180.0 2.
X -Q1-NB-X 2 20.00 180.0 2.
OS-CT-N*-Q1 1 0.0649123 180.0 -4. chi KOL0
OS-CT-N*-Q1 1 0.434356 180.0 -3. chi KOL0
OS-CT-N*-Q1 1 0.952733 0.0 -2. chi KOL0
OS-CT-N*-Q1 1 1.12182 180.0 1. chi KOL0
CT-CT-N*-Q1 1 0.0797247 180.0 -4. chi KOL0
CT-CT-N*-Q1 1 0.683783 180.0 -3. chi KOL0
CT-CT-N*-Q1 1 0.686632 180.0 -2. chi KOL0

```
CT-CT-N*-Q1   1    0.808183      0.0            1.  chi KOL0
X -CA-ND-X    4    9.60        180.0            2.        reinterpolated 93'
CA-NA-CB-NB   1    4.150       180.0            2.
CA-NP-CB-NB   1    4.150       180.0            2.
NB-CB-NA-HQ   1    4.150       180.0            2.
NB-CB-NP-HQ   1    4.150       180.0            2.
CB-CB-NA-CA   1    4.150       180.0            2.
CB-CB-NP-CA   1    4.150       180.0            2.
CB-CB-NA-HQ   1    4.150       180.0            2.
CB-CB-NP-HQ   1    4.150       180.0            2.
X -Q2-NA-X    2    8.00        180.0            2.        Parm10.dat
X -Q2-O -X    4   11.20        180.0            2.        Parm10.dat
X -Q2-CB-X    4   12.00        180.0            2.        Parm10.dat
X -NP-CA-X    4    1.5         180.0            2.        Parm10.dat

IMPROPER
X -X -Q1-H5         1.1         180.         2.
CB-CT-N*-Q1         1.0         180.         2.
X -X -ND-H          1.0         180.         2.          JCC,7,(1986),230
ND-NA-CA-NA         1.1         180.         2.          dac, 10/94
ND-NA-CA-NP         1.1         180.         2.          dac, 10/94

NONBON
Q1          1.9080  0.0860
Q2          1.9080  0.0860
HQ          0.6000  0.0157
ND          1.8240  0.1700
NP          1.8240  0.1700
```

References

Baranowski D, Boryski J, Ostrowski T. 2022. 7-(β-D-Ribofuranosyl)-3-methylguanine: Synthesis from guanine and comparative multinuclear NMR studies. *Journal of Molecular Structure* **1250:**131911. doi:10.1016/j.molstruc.2021.131911

Dutta N, Deb I, Sarzynska J, Lahiri A. 2023. Conformational preferences of inosine and its methyl derivatives: Comparison of the AMBER derived force field parameters and reparameterization of the glycosidic torsion parameters. doi:10.21203/rs.3.rs-2486374/v1

Schaftenaar G, Noordik JH. 2000. Molden: a pre-and post-processing program for molecular and electronic structures. *Journal of computer-aided molecular design* **14**: 123-134. doi:10.1023/a:1008193805436

Vanquelef E, Simon S, Marquant G, Garcia E, Klimerak G, Delepine JC, Cieplak P, Dupradeau FY, 2011. R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Research* 39:W51137. doi:10.1093/nar/gkr288

Wang J, Cieplak P, Kollman PA. 2000, How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry* 21: 1049–1074

Wawrzyniak D, Framski G, Januszczyk P, Ostrowski T, Baraniak D, Jahnz-Wechmann Z, Fogt J, Manikowski A, Baranowski D, Rolle K. 2020. 7-(β-d-Ribofuranosyl) guanine and its analogues modified in the sugar portion: synthesis and antiglioma properties. *ChemistrySelect* **5**: 13370-13375. doi:10.1002/slct.202003393

Yildirim I, Stern HA, Kennedy SD, Tubbs JD, Turner DH. 2010. Reparameterization of RNA χ torsion parameters for the AMBER force field and comparison to NMR spectra for cytidine and uridine. *Journal of chemical theory and computation* **6**: 1520-1531. doi:10.1021/ct900604a