SUPPLEMENTARY DATA

Imino proton NMR guides the reprogramming of A•T specific minor groove binders for mixed base pair recognition

Narinder K. Harika[1], Ananya Paul[1], Ekaterina Stroeva[1], Yun Chai[1,2], David W. Boykin[1], Markus W. Germann[1,*] and W. David Wilson[1,*]

[1]Department of Chemistry, Georgia State University, Atlanta, GA 30303-3083, USA

Address correspondence to either author

***W. David Wilson Tel: 404-413-5503; Fax: 404-413-5505; Email: wdw@gsu.edu

***Markus W. Germann Tel: 404-641-4981; Fax: +1 404 413 5561; Email: mwg@gsu.edu

[2]Present address: Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China

Supplementary Material and Methods:

Sample Preparation: Sequences for selected hairpin DNAs are provided in Table 1 (Integrated DNA Technologies, Coralville, IA, USA). The heterocyclic diamidine, DB2277, was synthesized as previously described (1). ^{15}N-labeled central G in G-hp5 was synthesized with ^{15}N-labeled phosphoramidites using a 391 DNA synthesizer and then desalted by FPLC using an AKTA purifier. The purity of DNA was checked using denaturing gel electrophoresis. The DB2277 was prepared as 2 mM stock solution in double distilled water (ddH_2O) and its concentration was calibrated against a known concentration of DSS (4, 4-dimethyl-4-silapentane-1-sulfonic acid), being used as an internal standard.

NMR Experiments: 1D and 2D NOE data were collected. Water suppression in NMR spectra was obtained using 1-1 pulse sequence. In 1D-NOE difference experiment, selective imino proton resonances are pre-irradiated to observe the NOE build up at proton resonances in proximity. Phase-sensitive NOESY spectra were recorded with 2048 × 400 data points in the two dimensions at different mixing times (150 ms and 50 ms). The spectral width was set at 24 ppm and 1 s relaxation delay was used for experiments in H_2O. Both dimensions were acquired with cosine bell functions (SSB = 2). From 2D Exchange Spectroscopy (2D EXSY), total chemical exchange rate is calculated for a two-site exchange using the following equations (Table 2):

$$k = (1 / tm) \ln ((r + 1) / (r - 1))$$

$$r = 4 X_M X_m ((I_M + I_m) / (I_{M-m} + I_{m-M})) - (X_M - Xm)^2$$

The term r corresponds to unequal populations of the major and minor species. I_M and I_m are the diagonal peak intensities of major and minor exchangeable resonances in the EXSY spectrum and I_{M-m} and I_{m-M} are the intensities of the exchange cross peaks, tm is the mixing time. X_M and X_m are the mole fractions of the major and minor species (2,3).

Surface Plasmon Resonance: As the Biosensor-SPR approach responds to mass, it is an excellent method for comparative studies of dications that have very large differences in properties and equilibrium binding constants (K). The SPR experiments were performed at 25 °C in degassed and filtered Tris-HCl buffer (50 mM Tris-HCl, 100 mM NaCl. 1 mM EDTA, pH 7.4). Steady-state binding analysis was performed with multiple injections of compound concentration as previously described (4). Kinetic analysis was performed by globally fitting the binding results for the entire concentration series using a standard 1:1 kinetic model with integrated mass transport-limited binding parameters as described previously (5,6). The signal at saturation of binding sites allows direct determination of stoichiometry.

Molecular Modeling and Docking: The center of the macromolecule is the grid center with a grid size of 72 Å×68 Å×108 Å and a grid spacing of 0.375 Å. Docking runs were performed using the Lamarckian genetic algorithm (LGA) with no modifications of docking parameters. LGA was used because of the existence of rotatable bonds in the ligands and to evaluate the correct conjugate DNA conformation, as it is known to reproduce various experimental ligand−DNA complex structures. Initially, we used a population of random individuals (population size of 150), a maximal number of 2500000 energy evaluations, a maximal number of evaluations of 2700, and a mutation rate of 0.02 fs. Two hundred independent flexible docking runs were conducted for each ligand, and then the lowest-energy dock conformation obtained from the Flexible docking was resubmitted for rigid docking to remove the internal energy of the ligand (steric clashes) and retain the hydrogen bonding interaction with ds-DNA bases.

Supplementary Figures and Tables:

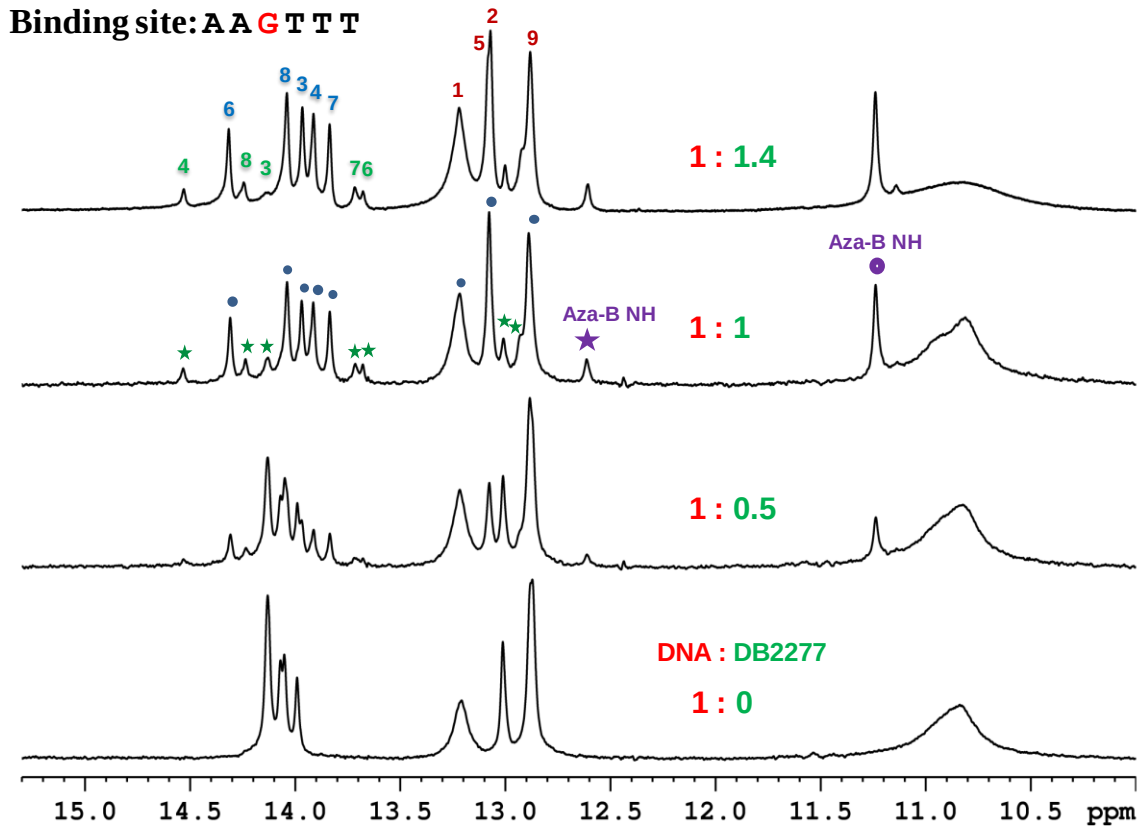

Figure S1: Shifts in the imino proton peaks at different G-hp2:DB2277 ratios at 285 K. The major and minor binding species of the DNA-DB2277 complex are represented by "●" and "★" labeled peaks respectively along with the exchangeable protons of aza-benzimidazole NH that are assigned from 1D-NOE and 2D NOESY shown in Figure S2 and S5. The NMR sample consisted of 100 μM DNA at pH 6.7.

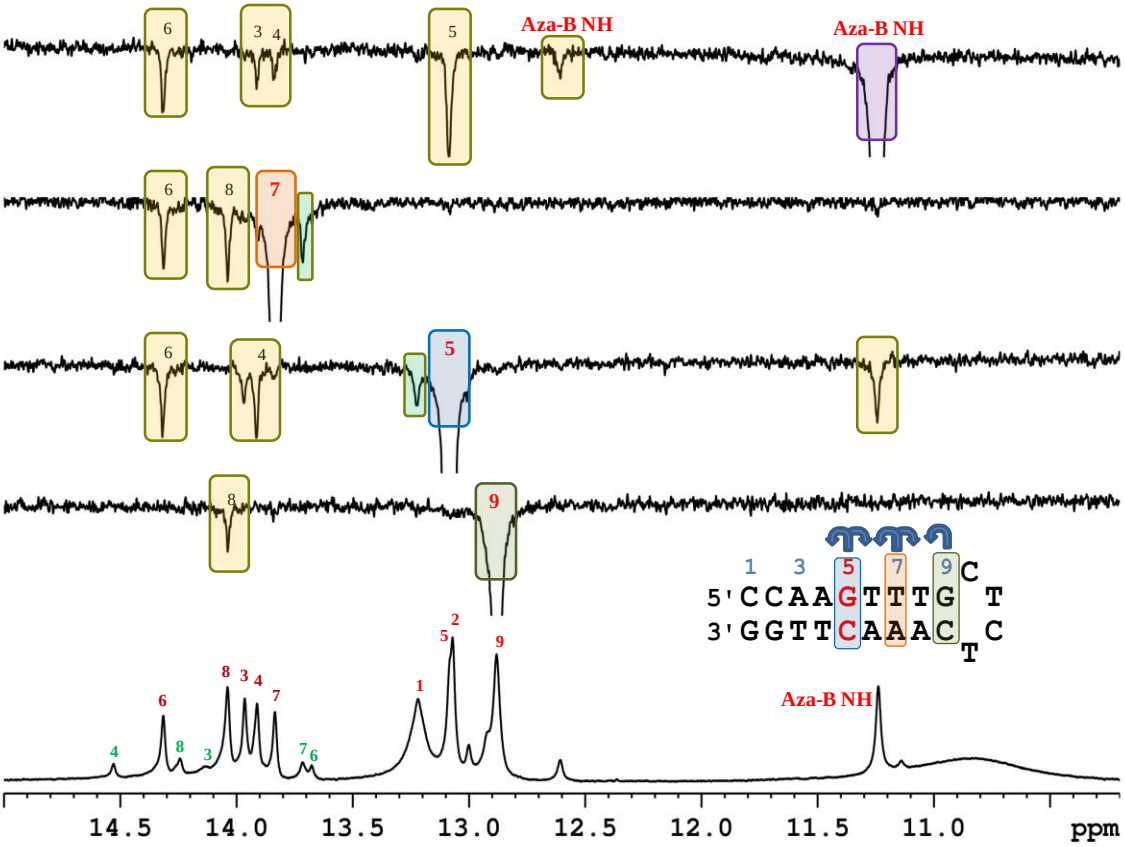

Figure S2: Sequential assignment of imino protons of G-hp2 complex with DB2277 using 1D-NOE difference spectra.

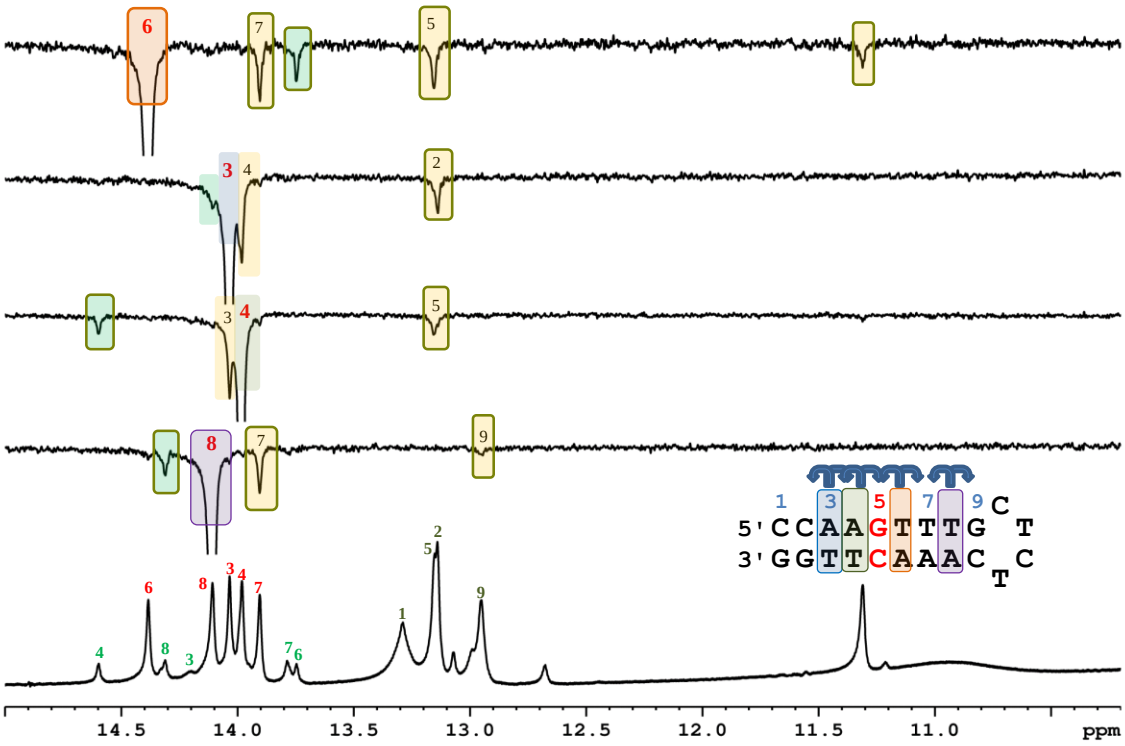

Figure S3: Sequential assignment of imino protons of G-hp2 complex with DB2277 using several 1D-NOE difference spectra in conjunction with 2D NOESY experiments (Figure S3).

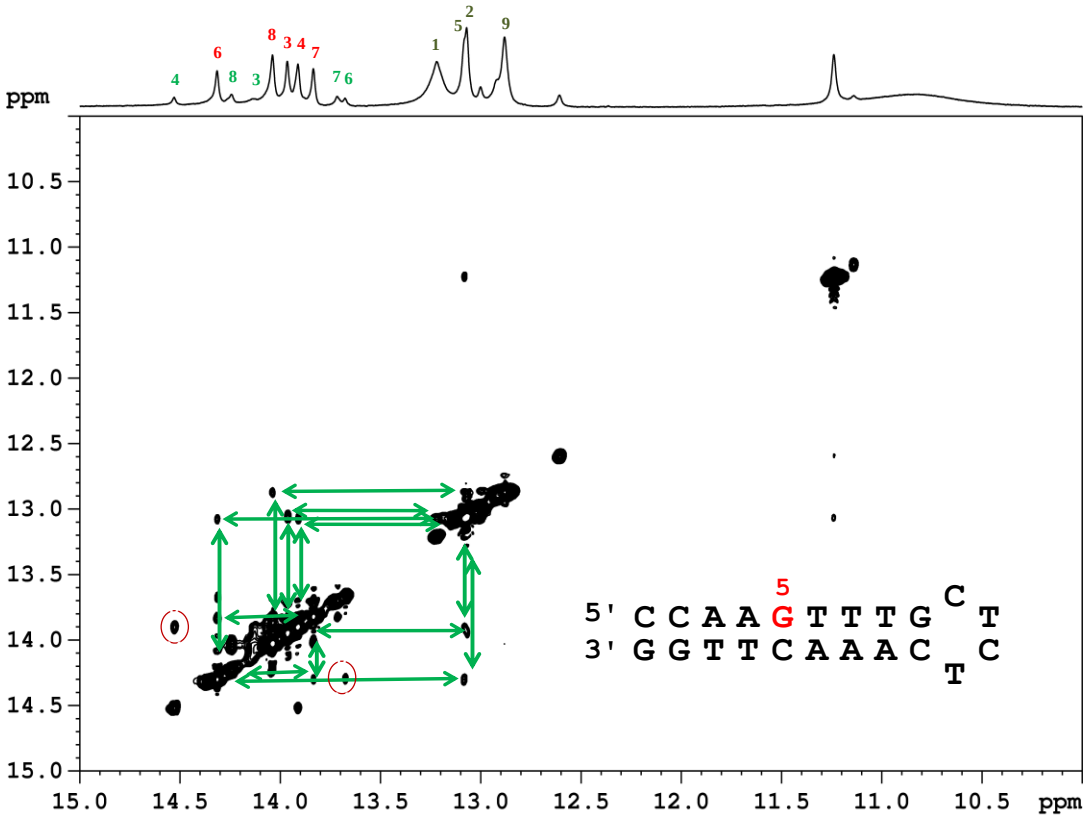

Figure S4: Imino proton region of 2D NOESY spectrum of G-hp2-DB2277 complex using a 150 ms mixing time at 285 K. Both NOE and exchange peaks (in circle) are visible.

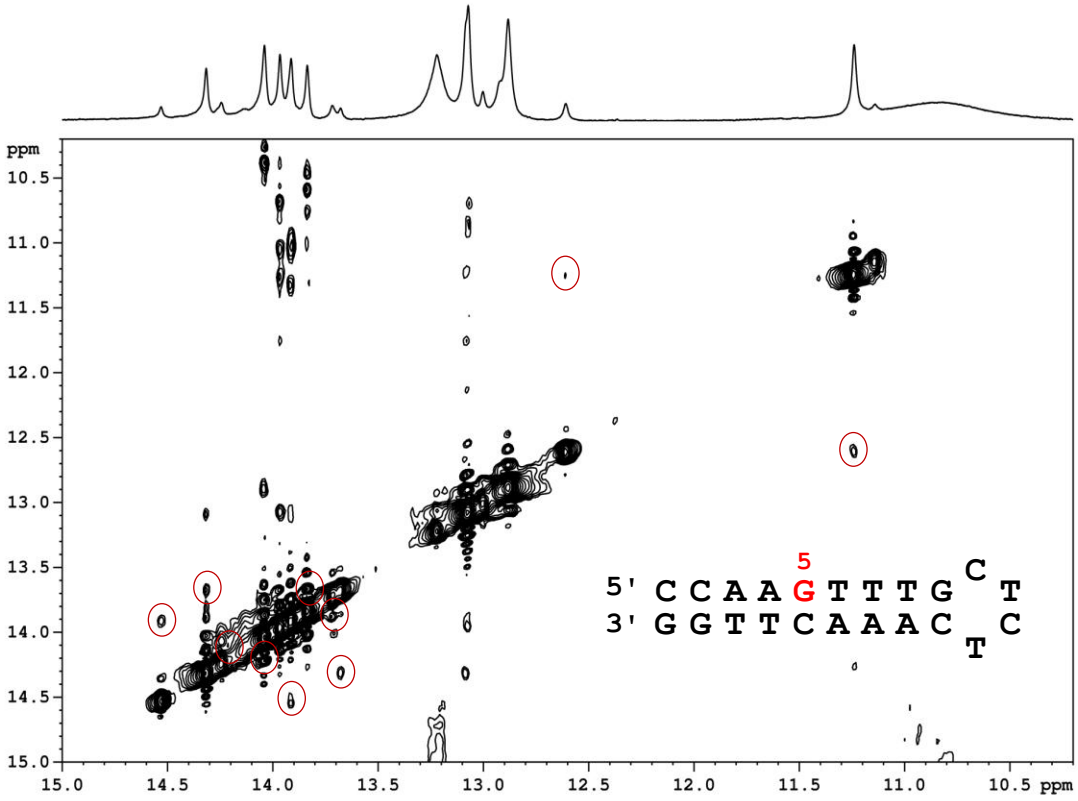

Figure S5: Imino proton region of EXSY (2D NOESY recorded at short mixing time of 50 ms) spectrum of G-hp2-DB2277 complex at 285 K. Exchange cross peaks are marked with circles.

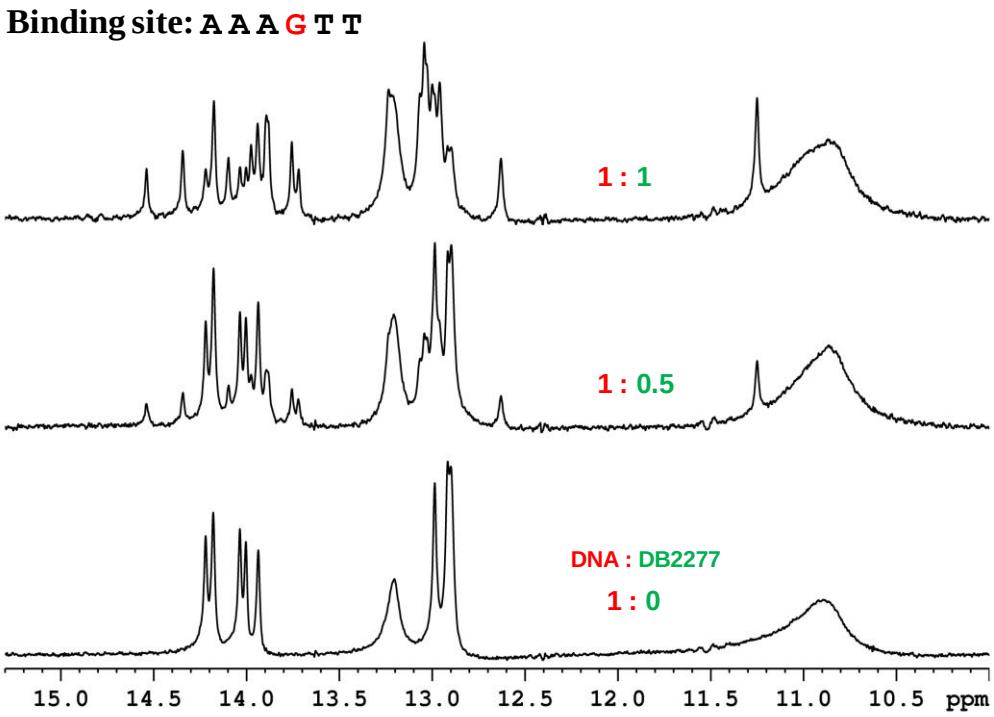

Figure S6: Shifts in the imino spectral region peaks at different G-hp3:DB2277 ratios at 285 K.

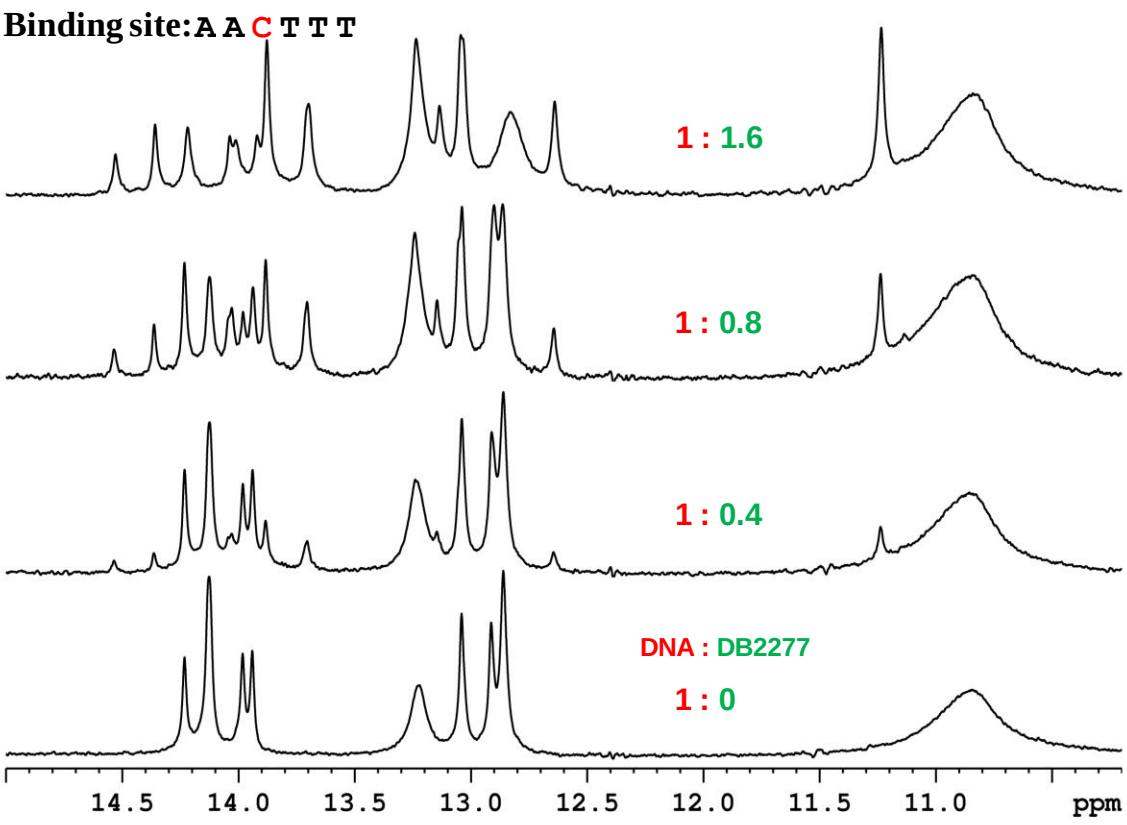

Figure S7: Shifts in the imino spectral region peaks at different G-hp4:DB2277 ratios at 285 K.

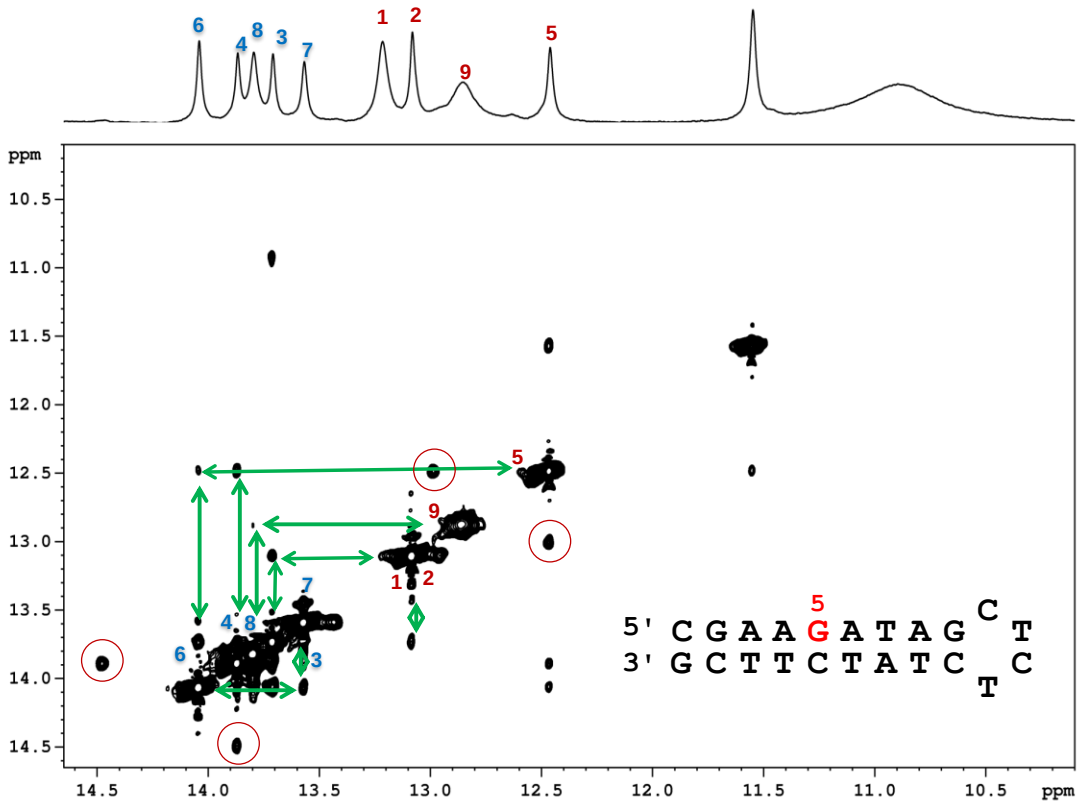

Figure S8: Imino proton assignment of G-hp5-DB2277 complex using 2D NOESY at 150 ms mixing time. Exchange cross peaks are marked in circle.

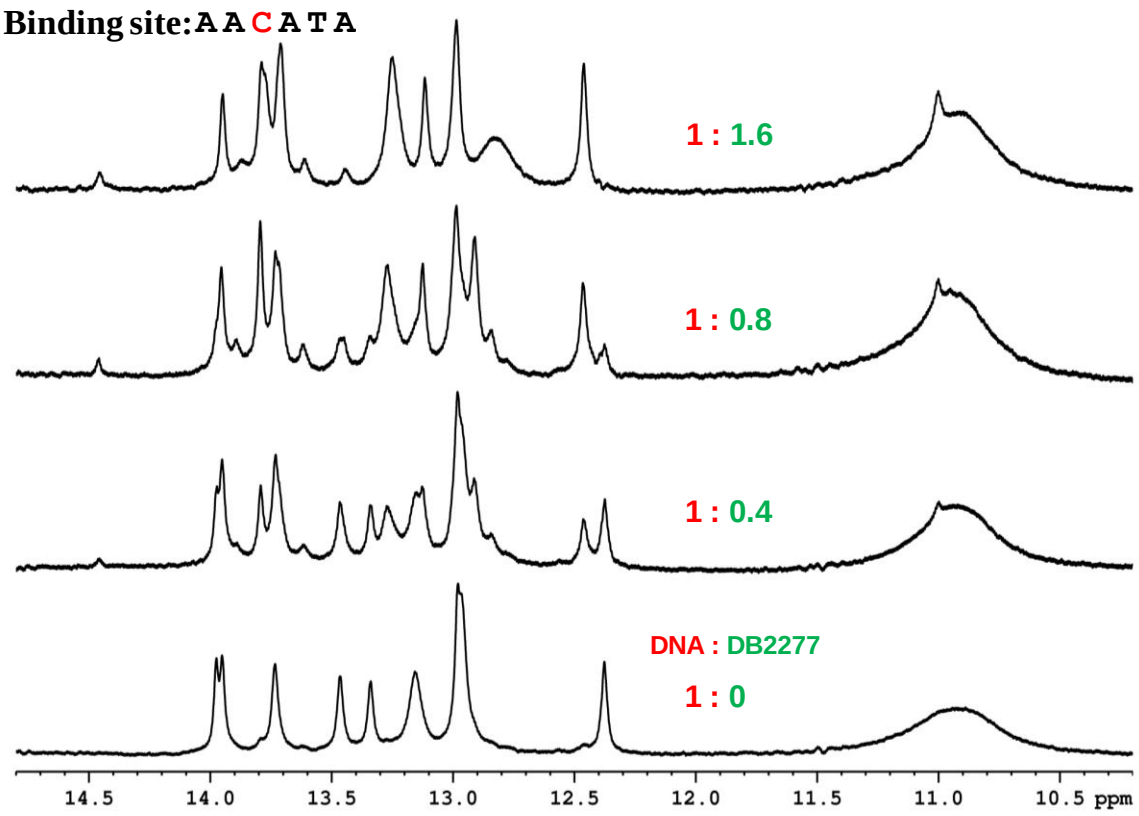

Figure S9: Shifts in the imino spectral region peaks at different G-hp6:DB2277 ratios at 285 K.

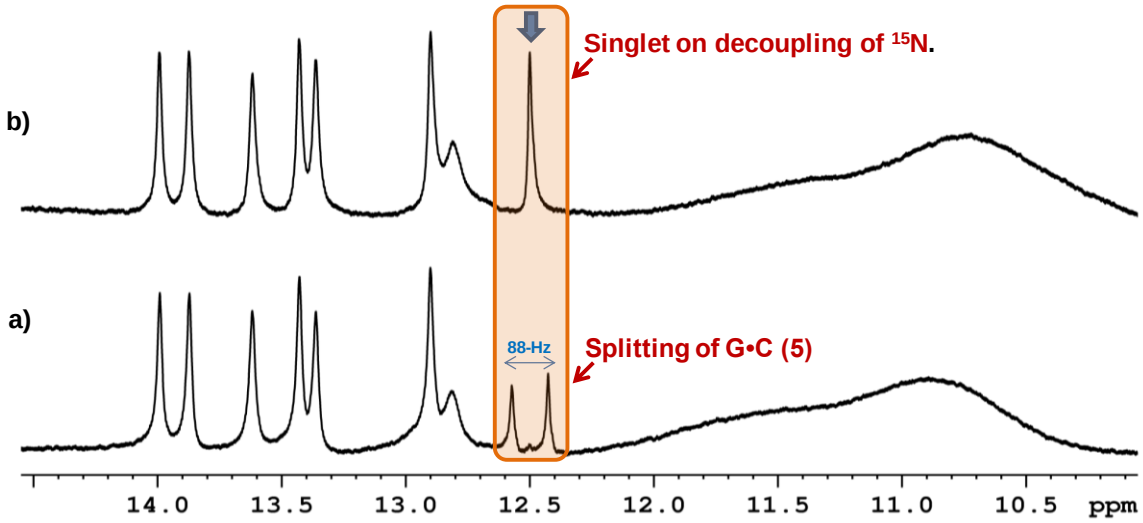

Figure S10: a) Splitting of G•C (5) imino proton peak due to ^{15}N-^{1}H coupling in the imino spectral region of ^{1}H NMR spectrum at 308 K. b) Splitting of G•C (5) imino proton peak reverts back to singlet upon decoupling of ^{15}N.

Hairpin DNA	Native DNA T_m (°C)	ΔT_m (DNA-DB2277 complex) (°C)
G-hp1	68.5	13.5
G-hp2	68.1	13.8
G-hp3	68.5	12.0
G-hp4	68.1	11.9
G-hp5	65.2	10.3
G-hp6	66.9	8.7

Table S1: T*m* results of DNA hairpin oligomers and their complexes with DB2277 are reported within ± 0.5 °C.

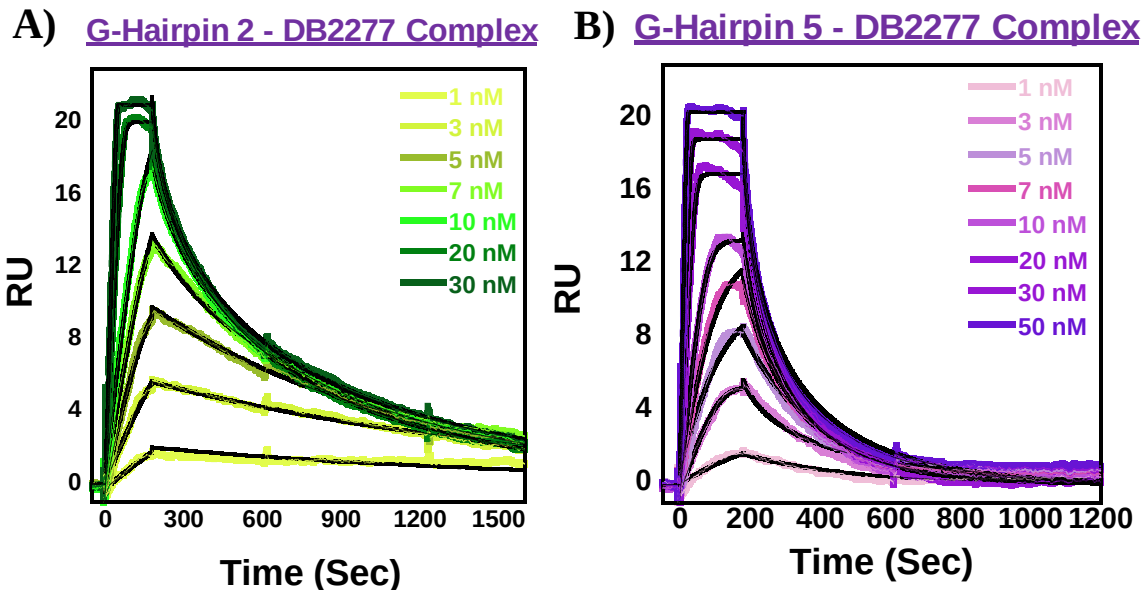

Figure S11: Representative SPR sensorgrams for DB2277 binding to G-hp2 and G-hp5. Inset, the increase in the concentration of DB2277 as injected on immobilized DNA sequence. The solid black lines are best-fit values for the global kinetic fitting of the results with a single site function.

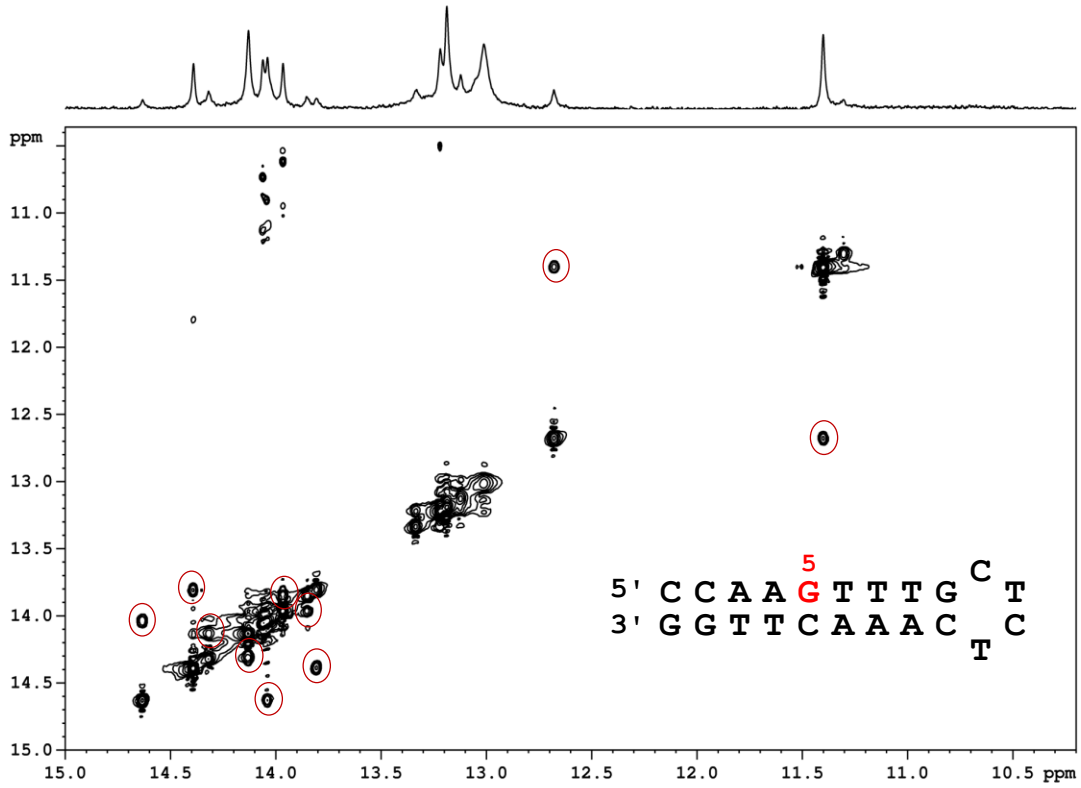

Figure S12: Imino proton region of EXSY spectrum of G-hp2-DB2277 complex (recorded at a short mixing time 50 ms) at 303 K. Exchange cross peaks are marked in circle.

Table S2 (A): Intensity and molar fraction of major and minor well-resolved peaks at 303 K and 285 K.

Peak No. A•T bp (303 K)	Intensity Major peak I_M	Intensity Minor peak I_m	Intensity Major to minor peak I_{M-m}	Intensity Minor to major peak I_{m-M}	Molar fraction X_M	Molar fraction X_m
4	2.38E+09 2.42E+09 2.29E+09	3.13E+08 3.25E+08 3.22E+08	9.02E+07 9.36E+07 8.96E+07	9.83E+07 1.09E+08 9.83E+07	0.88 0.88 0.88	0.12 0.12 0.12
6	1.97E+09 1.93E+09 1.93E+09	4.68E+08 4.09E+08 4.59E+08	9.79E+07 9.95E+07 9.95E+07	8.47E+07 8.09E+07 8.51E+07	0.81 0.81 0.81	0.19 0.19 0.19
7	2.04E+09 2.05E+09 1.91E+09	5.65E+08 5.95E+08 5.46E+08	1.43E+08 1.40E+08 1.35E+08	1.65E+08 1.67E+08 1.67E+08	0.78 0.78 0.78	0.22 0.22 0.22

Peak No. A•T bp (285 K)	Intensity Major peak I_M	Intensity Minor peak I_m	Intensity Major to minor peak I_{M-m}	Intensity Minor to major peak I_{m-M}	Molar fraction X_M	Molar fraction X_m
4	8.72E+08 8.72E+08 8.25E+08	1.16E+08 1.17E+08 1.16E+08	1.43E+07 1.33E+07 1.43E+07	9.61E+06 9.61E+06 9.61E+06	0.89 0.89 0.88	0.11 0.11 0.12
8	8.63E+08 9.86E+08 1.03E+09	2.16E+08 2.14E+08 2.52E+08	2.50E+07 2.66E+07 3.64E+07	2.38E+07 2.51E+07 2.51E+07	0.80 0.82 0.80	0.20 0.18 0.20

Table 2 (B): Exchange rate calculation of G-hp2-DB2277 complex from EXSY at 303 K and 285 K.

Peak No. AT bp	**k_{ex} (at 303 K)**	**ln(K_{eq}) (at 303 K)**	**ΔG (Kcal/mol) (at 303 K)**	**k_{ex} (at 285 K)**	**ln(K_{eq}) (at 285K)**	**ΔG (Kcal/mol) (at 285 K)**
4	7.7 ± 0.3	- 2.00	1.20 ± 0.02	2.5 ± 0.1	- 2.00	1.13 ± 0.02
6	5.3 ± 0.1	- 1.47	0.89 ± 0.04	-	-	-
7	7.4 ± 0.3	- 1.26	0.76 ± 0.02	-	-	-
8	-	-	-	3.0 ± 0.1	- 1.44	0.82 ± 0.04

[a]The reported errors are calculated using standard deviation from three repeated measurements on the same sample.

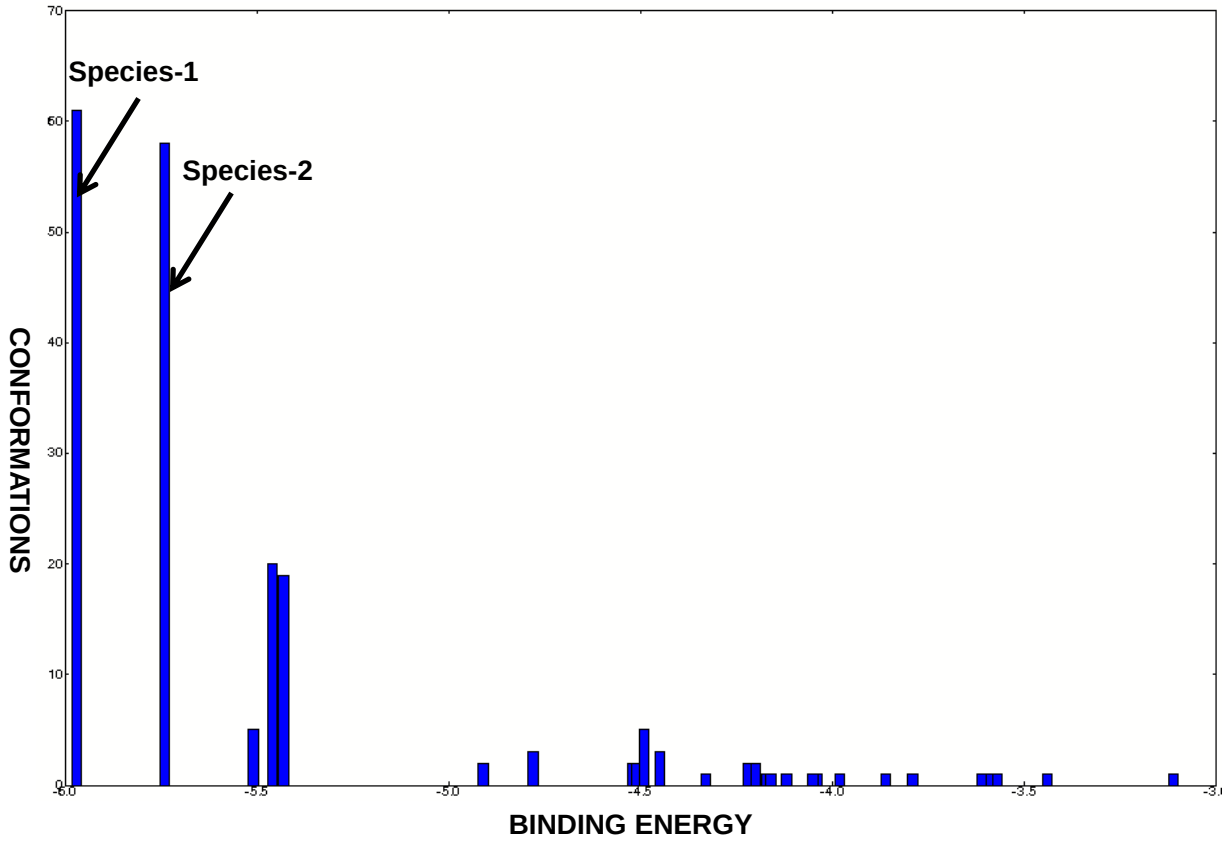

Figure S13: Clustering Histogram of multi-member conformations found out of 200 Lamarckian Genetic Algorithm docking runs.

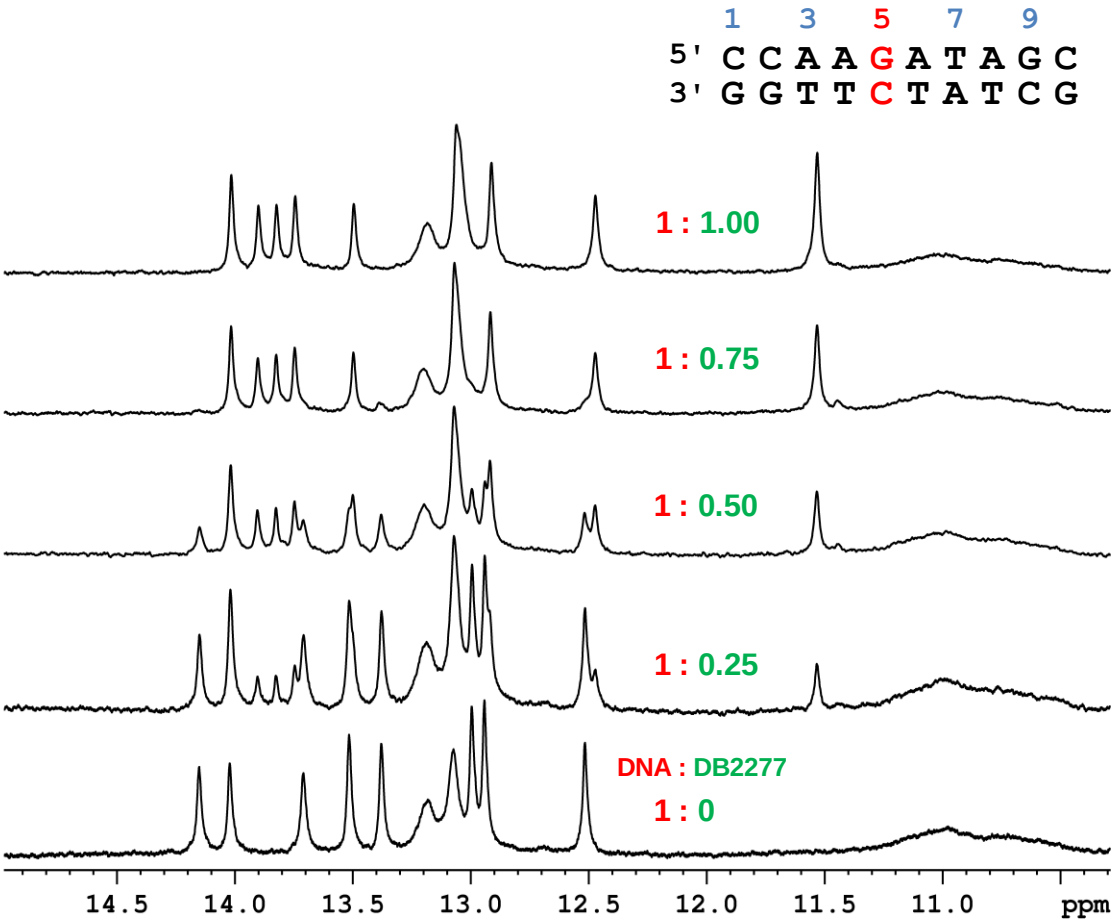

Figure S14: Shifts in the imino spectral peaks at duplex DNA:DB2277 ratios at 285 K. Duplex DNA sequence containing AAGATA binding site similar to G-hp5 is shown at the top. Well resolved five T-imino peaks are observed with no minor species. The NMR sample consisted of 100 μM DNA at pH 6.9.

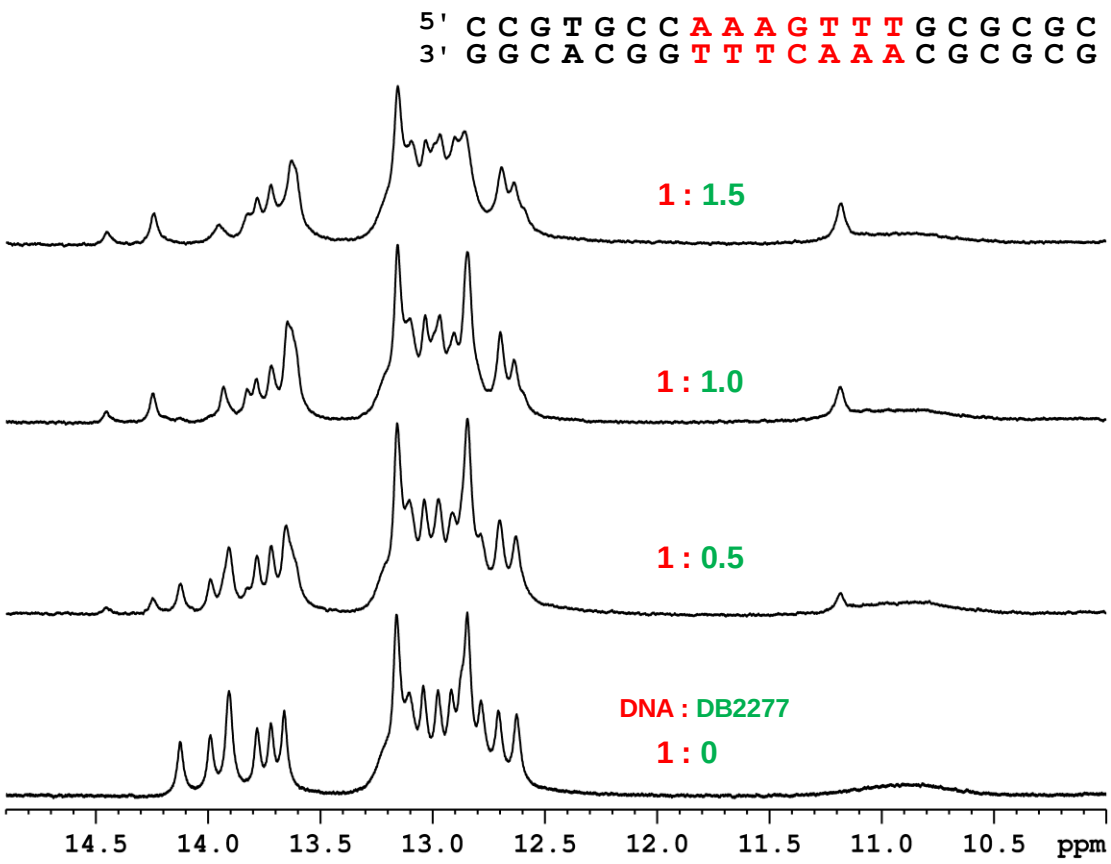

Figure S15: Shifts in the imino spectral region peaks at different DNA:DB2277 ratios at 285 K. Duplex DNA sequence is shown at the top. The NMR sample consisted of 100 µM DNA at pH 6.7.

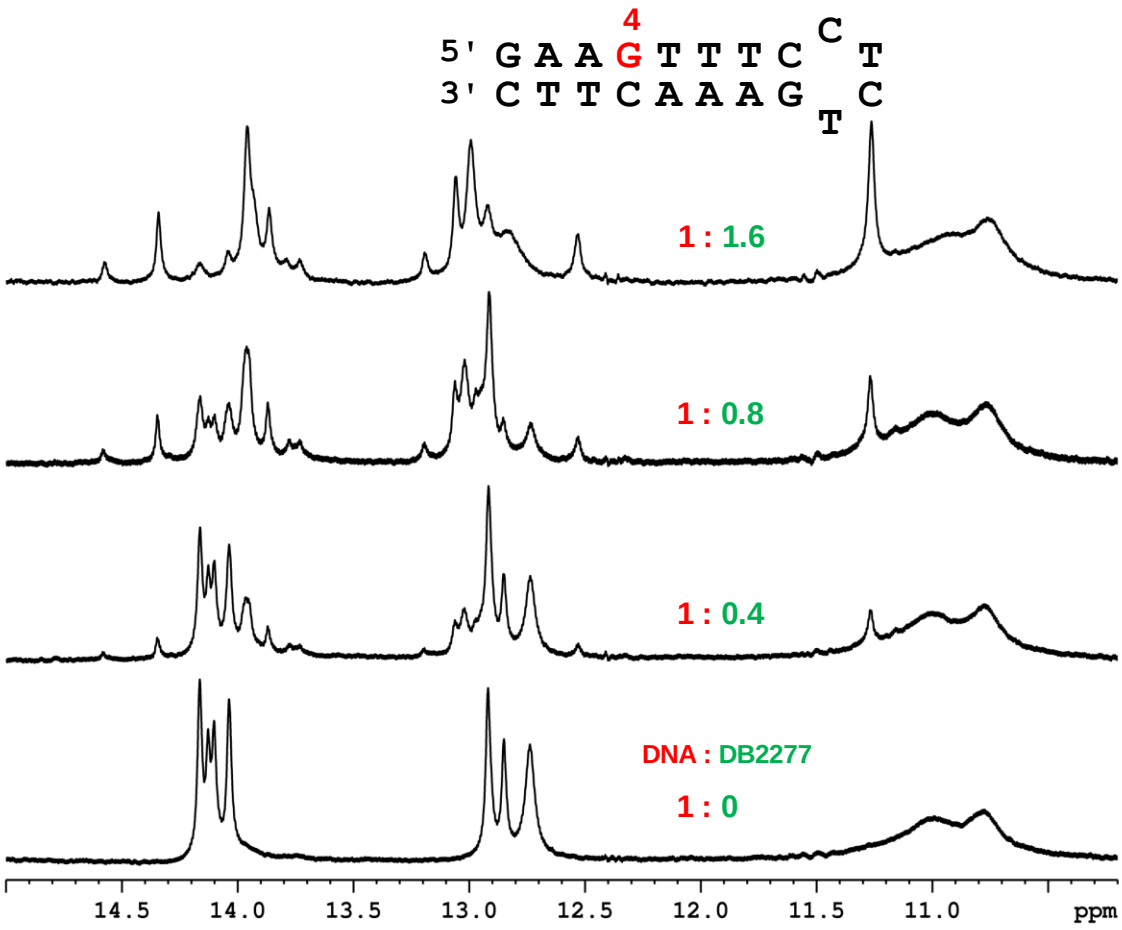

Figure S16: Shifts in peaks at different G-hp7:DB2277 ratios in imino proton spectrum at 285K. Spectra show significant minor form.

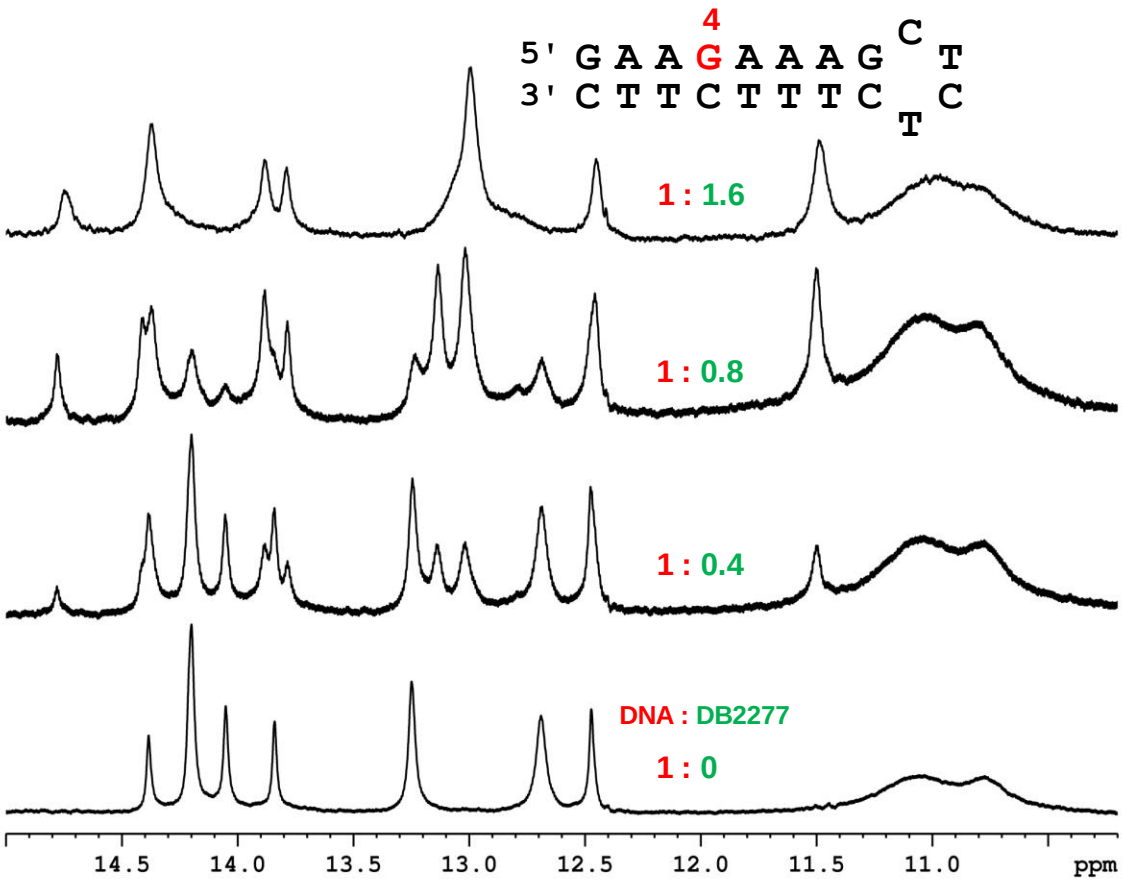

Figure S17: Shifts in peaks at different G-hp8:DB2277 ratios in imino proton spectrum at 285K. In this sequence context the minor form is suppressed.

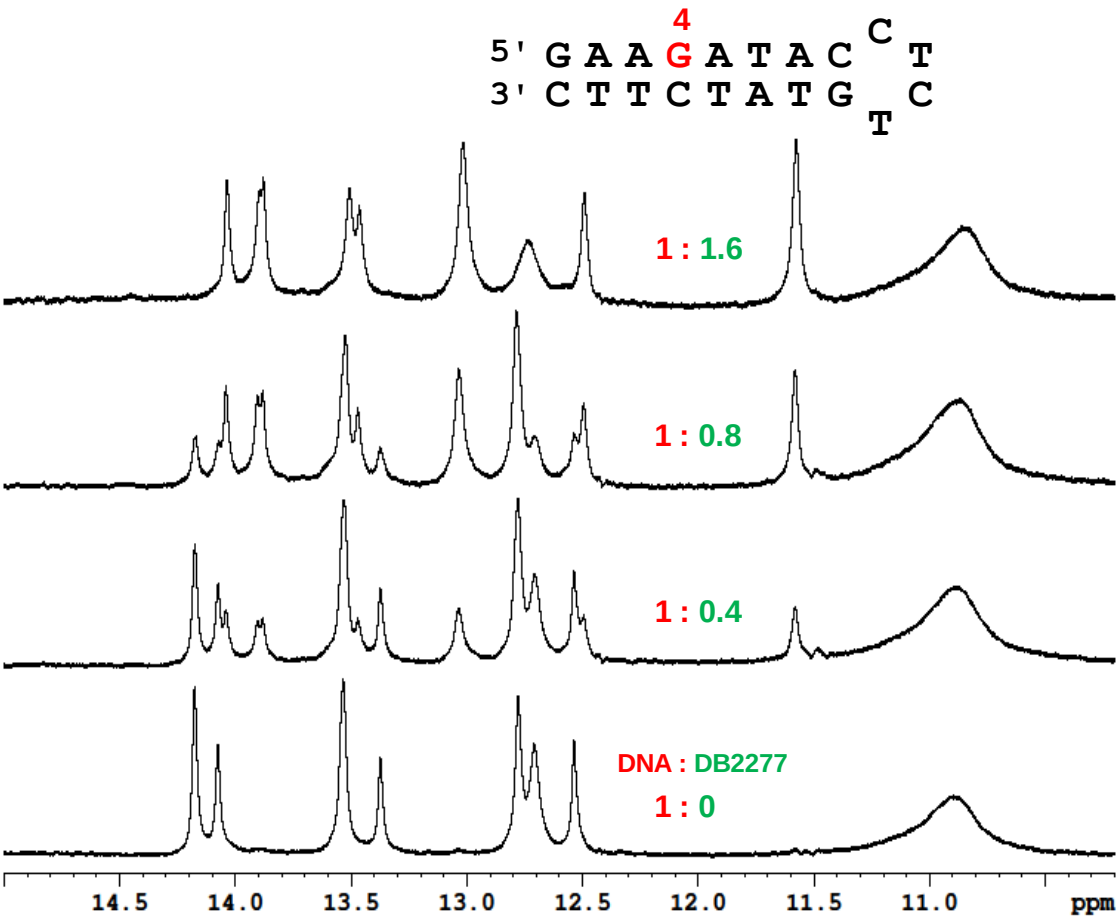

Figure S18: Shifts in peaks at different G-hp9:DB2277 ratios in imino proton spectrum at 285K. In this sequence context the minor form is suppressed.

References:

1. Chai, Y., Paul, A., Rettig, M., Wilson, W. D., and Boykin, D. W. (2014) Design and synthesis of heterocyclic cations for specific DNA recognition: From AT-rich to mixed-base-pair DNA sequences. *J. Org. Chem.,* **79**, 852-66.

2. Zhang, J., and Germann, M. W. (2011) Characterization of secondary amide peptide bond isomerization: Thermodynamics and kinetics from 2D NMR spectroscopy. *Biopolymers,* **95**, 755-762.

3. Perrin, C.L. and Dwyer, T.J. (1990) Application of two-dimensional NMR to kinetics of chemical exchange. *Chem. Rev.*, **90**, 935-967.

4. Paul, A., Chai, Y., Boykin, D. W., and Wilson, W. D. (2015) Understanding mixed sequence DNA recognition by novel designed compounds: the kinetic and thermodynamic behavior of azabenzimidazole diamidines. *Biochemistry,* **54**, 577-87.

5. Nguyen, B., Tanious, F. A., and Wilson, W. D. (2007) Biosensor-surface plasmon resonance: Quantitative analysis of small molecule–nucleic acid interactions. *Methods*, **42**, 150-161.

6. Liu, Y. and Wilson, W.D. (2010) Quantitative analysis of small molecule-nucleic acid interactions with a biosensor surface and surface plasmon resonance detection. Methods Mol. Biol. (Totowa, NJ, U. S.), 613, 1-23.