Electronic Supplementary Material (ESI) for Molecular BioSystems.

Supporting information for:

Validation of a computational docking methodology to identify the non-covalent binding site of ligands to DNA

Christos Deligkaris,[*,†] Anthony Thomas Ascone,[†,‡] Kevin Joseph Sweeney,[†] and Alan Jonathan Quentin Greene[†]

Department of Physics, Drury University, Springfield MO 65802, USA

E-mail: cdeligkaris@drury.edu

Abstract

Ligand residue numbers (Table S1), RMSD between original and flipped DNA structures (Table S2), ΔG_{lowest} as a function of the number of free energy evaluations (Figures S1-S9), 10 lowest free energy binding sites for all ligand-DNA systems (Tables S3-S30), distance-based success rates (Table S31), predicted and experimental binding geometries (Figures S10 -S54), hydrogen bonding information for predicted ligand geometries (Table S32).

[*]To whom correspondence should be addressed

[†]Department of Physics, Drury University, Springfield MO 65802, USA

[‡]Present address: Department of Physics, Astronomy, and Materials Science, Missouri State University, Springfield MO 65897, USA

Table S1: PDB file residue numbers used as ligands in the docking studies.

PDB ID	LIGAND RESIDUE	PDB ID	LIGAND RESIDUE	PDB ID	LIGAND RESIDUE
1Z3F	42	4E1U	103	454D	81
1D30	25	432D	21	109D	25
127D	25	128D	25	264D	25
227D	25	311D	25	2B3E	25
129D	25	130D	25	2DBE	25
1D63	25	298D	25	1FMQ	25
1EEL	25	144D	25	102D	25
1PRP	25	2GB9	302	1D32	9
334D	22	2DND	25	1JTL	31
182D	13	1Z8V	31	1DVL	25
473D	21	6BNA	25		

Table S2: RMSDs (in Å) of atoms of the four central base pairs in the DNA molecule in its original orientation and the flipped orientation.

PDB ID	RMSD
1D63	0.306
227D	0.380
298D	0.315
2DBE	0.332
1D30	0.458
432D	0.683
2B3E	0.209
1JTL	0.549
2DND	0.702
1EEL	0.305
1FMQ	0.209
109D	0.331
127D	0.764
128D	1.048
129D	0.889
130D	1.048
264D	0.362
311D	0.248
1DVL	0.576
1Z8V	0.535
473D	0.328
6BNA	0.380
102D	0.311
1PRP	0.334
144D	1.146

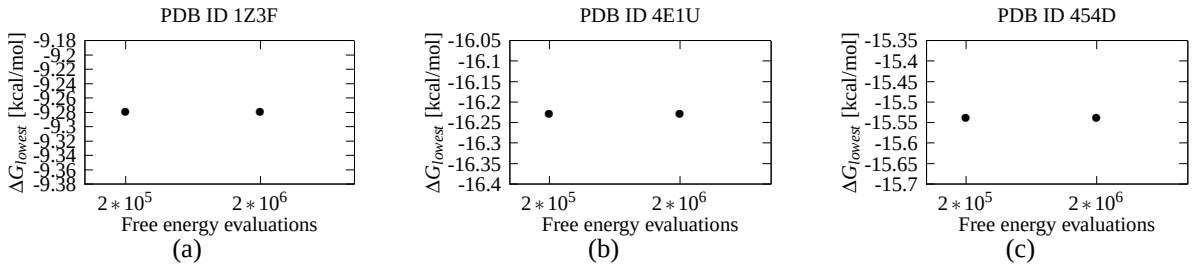

Figure S1: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with no rotatable bonds.

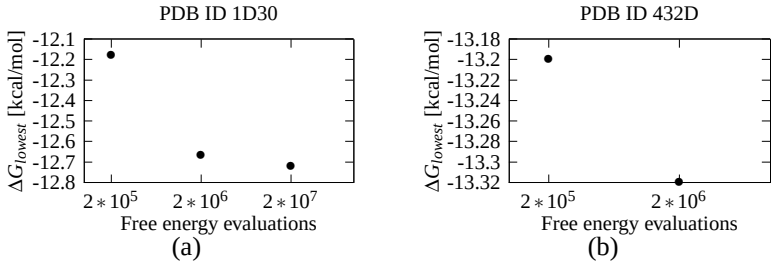

Figure S2: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with three rotatable bonds.

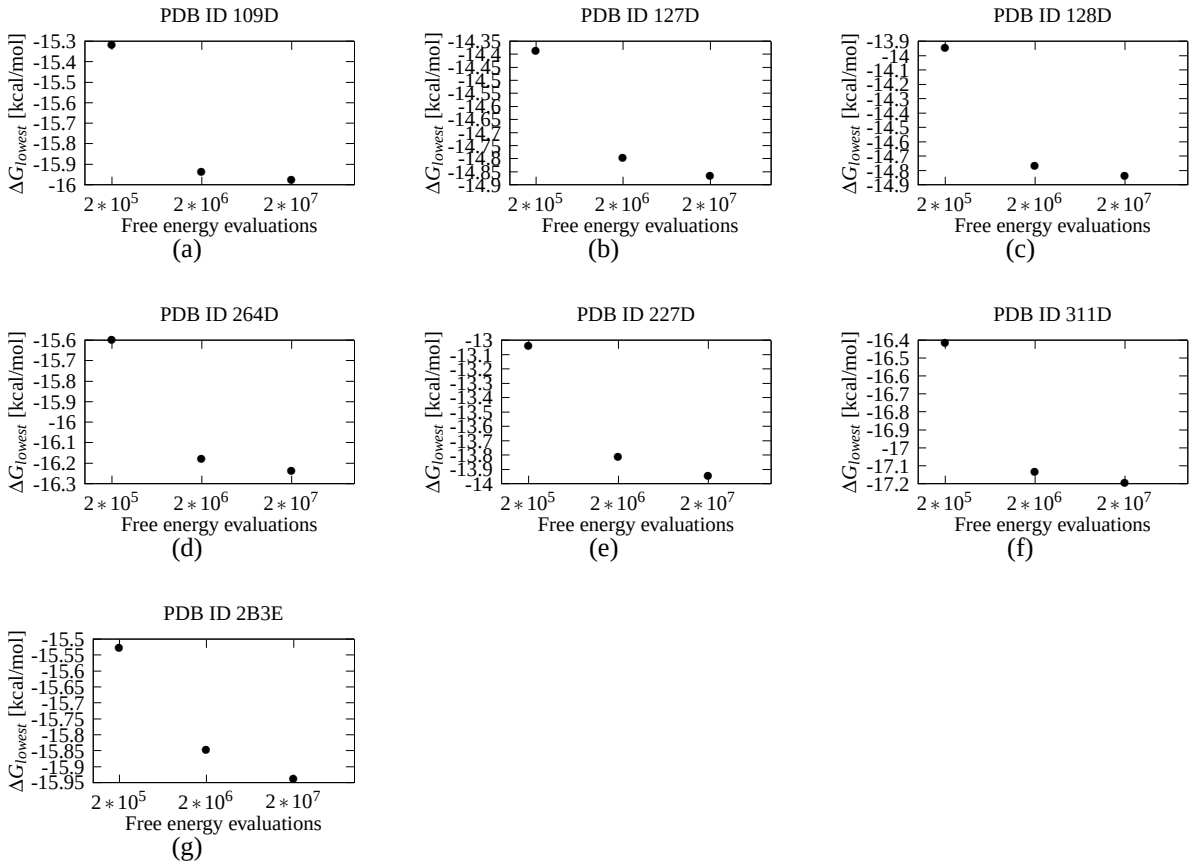

Figure S3: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with four rotatable bonds.

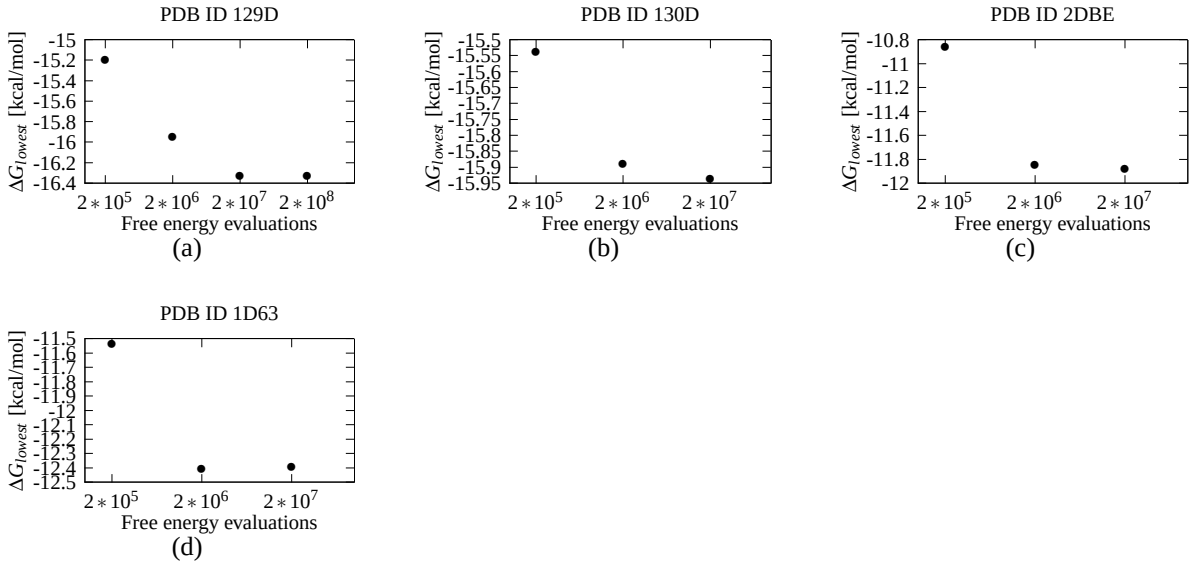

Figure S4: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with five rotatable bonds.

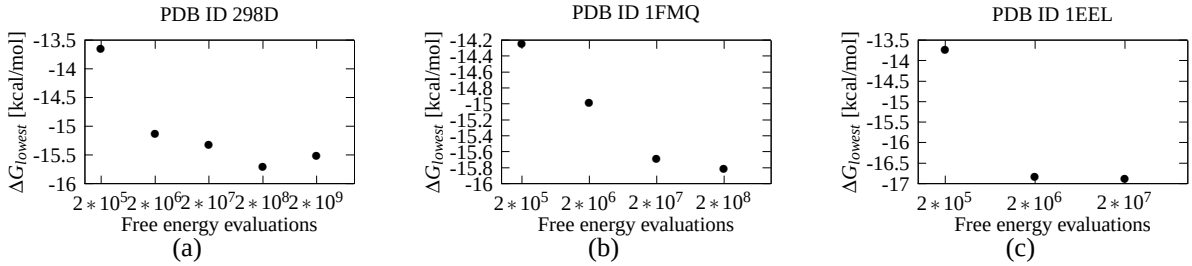

Figure S5: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with six rotatable bonds.

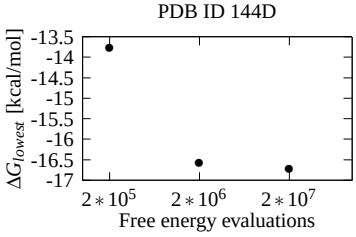

Figure S6: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with seven rotatable bonds.

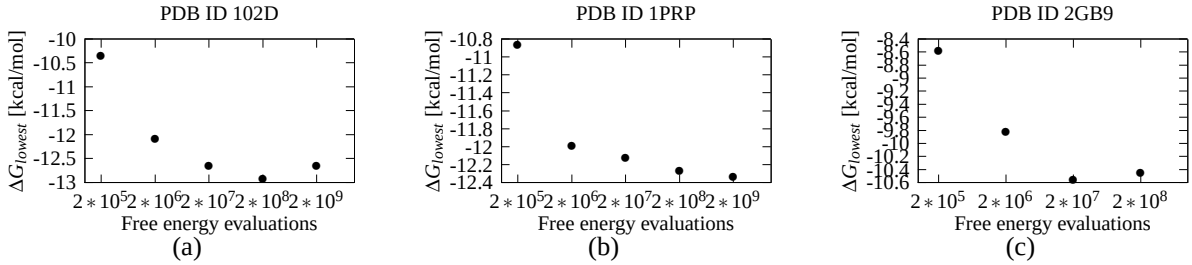

Figure S7: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with eight rotatable bonds.

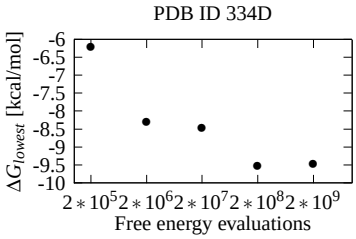

Figure S8: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with 10 rotatable bonds.

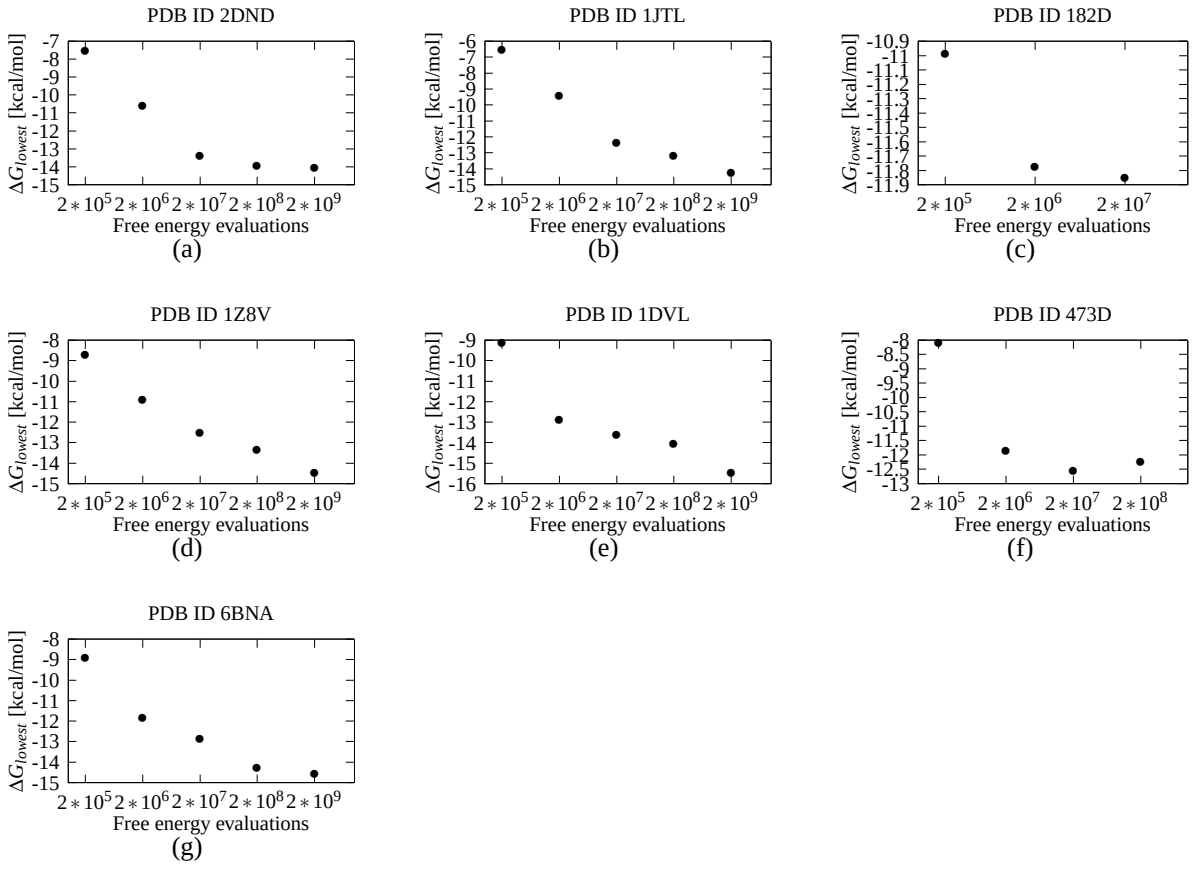

Figure S9: ΔG_{lowest} as a function of the number of free energy evaluations for ligands with 13 rotatable bonds.

Table S3: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 1Z3F. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^5$

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−9.28	25	5.67
2	−9.23	40	1.37
3	−9.13	17	6.27
4	−9.13	15	2.90
5	−8.84	1	3.72
6	−8.69	1	3.61
7	−8.43	1	5.73

Table S4: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 4E1U. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^5$

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−16.23	28	0.19
2	−16.23	33	0.20
3	−12.00	11	7.85
4	−11.93	10	7.86
5	−11.75	3	7.87
6	−11.68	1	7.78
7	−11.67	4	7.80
8	−11.66	1	7.88
9	−10.88	4	11.86
10	−10.84	3	11.75

Table S5: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 454D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^5$

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−15.54	11	8.40
2	−15.27	18	8.25
3	−15.10	3	8.38
4	−15.04	8	8.27
5	−13.14	17	0.98
6	−13.08	12	0.98
7	−12.16	2	1.87
8	−12.03	6	12.01
9	−12.02	1	9.67
10	−11.88	4	12.03

Table S6: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 1D30. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−12.67	26	2.19
2	−12.23	1	1.77
3	−12.22	10	2.34
4	−12.14	9	0.87
5	−12.11	5	10.62
6	−12.10	3	5.97
7	−11.96	3	5.94
8	−11.95	14	2.19
9	−11.86	2	4.17
10	−11.82	5	6.63

Table S7: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 432D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^5$

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−13.20	13	3.46
2	−12.92	12	8.21
3	−12.56	1	8.05
4	−11.96	12	7.90
5	−11.86	2	7.58
6	−11.59	1	6.06
7	−11.47	2	0.53
8	−11.37	3	3.55
9	−11.22	4	2.82
10	−11.20	1	8.28

Table S8: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 109D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−15.94	26	1.26
2	−15.74	16	1.15
3	−15.24	8	5.34
4	−15.09	11	6.58
5	−15.09	7	5.21
6	−15.03	2	5.80
7	−14.30	6	2.52
8	−14.02	12	2.33
9	−13.89	1	2.97
10	−13.68	6	2.91

Table S9: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 127D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−14.80	21	1.16
2	−14.12	6	3.81
3	−14.07	3	3.02
4	−14.01	9	5.82
5	−13.81	5	0.57
6	−13.56	1	2.46
7	−13.38	4	5.66
8	−13.34	3	6.10
9	−13.24	1	9.26
10	−13.20	2	14.34

Table S10: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 128D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−14.77	10	6.67
2	−14.48	8	9.15
3	−14.38	10	2.17
4	−14.18	8	3.66
5	−14.13	5	5.58
6	−13.85	6	6.08
7	−13.82	14	6.40
8	−13.79	1	8.73
9	−13.77	2	8.40
10	−13.75	2	7.21

Table S11: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 264D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−16.18	25	5.03
2	−15.94	44	0.73
3	−15.70	12	5.02
4	−15.15	5	1.08
5	−14.55	3	2.44
6	−14.39	2	9.65
7	−14.10	1	5.14
8	−14.00	1	3.31
9	−13.86	1	2.64
10	−13.78	3	5.25

Table S12: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 227D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−13.82	15	0.65
2	−13.63	17	0.72
3	−13.21	6	0.90
4	−13.12	3	0.87
5	−12.87	1	6.17
6	−12.77	12	4.14
7	−12.68	1	4.11
8	−12.61	15	4.26
9	−12.60	4	4.16
10	−12.46	6	3.50

Table S13: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 311D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−17.14	13	0.96
2	−16.93	16	1.09
3	−16.56	2	1.57
4	−16.00	1	6.26
5	−15.87	17	2.85
6	−15.59	2	4.09
7	−15.55	2	6.11
8	−15.53	15	3.58
9	−15.50	2	3.47
10	−15.29	7	11.71

Table S14: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 2B3E. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−15.85	28	0.96
2	−15.65	3	5.01
3	−15.65	21	0.70
4	−15.61	4	1.11
5	−15.31	14	5.05
6	−15.24	2	7.24
7	−14.93	2	9.60
8	−14.93	2	5.16
9	−14.38	2	4.35
10	−14.17	1	3.53

Table S15: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 129D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^7$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−16.33	32	0.70
2	−15.98	50	0.90
3	−15.93	4	2.04
4	−15.31	3	5.40
5	−15.19	4	5.63
6	−15.15	6	5.41
7	−14.27	1	6.30

Table S16: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 130D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−15.89	7	1.41
2	−15.74	19	3.08
3	−15.27	3	2.90
4	−15.18	6	5.12
5	−15.11	3	3.76
6	−14.92	1	1.25
7	−14.87	9	8.67
8	−14.84	4	2.81
9	−14.78	7	7.34
10	−14.70	5	4.49

Table S17: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 2DBE. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−11.85	7	2.85
2	−11.69	2	3.36
3	−11.51	3	3.60
4	−11.45	9	12.54
5	−11.37	8	12.13
6	−11.16	1	8.60
7	−11.04	1	4.68
8	−11.02	1	3.18
9	−11.00	2	4.46
10	−10.95	2	6.09

Table S18: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 1D63. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−12.41	9	3.07
2	−11.73	4	4.03
3	−11.72	1	0.92
4	−11.55	1	0.73
5	−11.50	1	1.29
6	−11.29	8	12.86
7	−11.28	1	9.14
8	−11.26	6	8.41
9	−11.22	5	11.99
10	−11.09	1	1.98

Table S19: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 298D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^8$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−15.71	41	1.48
2	−15.30	9	6.21
3	−15.23	20	1.58
4	−14.85	1	1.93
5	−14.79	3	4.48
6	−14.67	5	6.86
7	−14.64	2	4.49
8	−14.58	3	1.68
9	−14.54	2	9.00
10	−14.42	4	6.21

Table S20: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 1FMQ. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^7$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−15.69	16	1.25
2	−15.65	9	1.65
3	−15.51	8	7.35
4	−15.45	2	9.10
5	−15.13	7	2.02
6	−14.89	3	2.71
7	−14.48	3	6.35
8	−14.48	4	4.57
9	−14.41	3	8.49
10	−14.30	2	4.43

Table S21: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 1EEL. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−16.85	7	2.01
2	−16.25	2	4.41
3	−15.98	1	1.85
4	−15.49	4	7.40
5	−15.23	6	1.39
6	−15.16	2	4.78
7	−14.99	2	6.23
8	−14.98	1	1.06
9	−14.77	3	2.28
10	−14.75	2	4.98

Table S22: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 144D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−16.58	4	3.49
2	−15.43	1	2.89
3	−14.96	2	3.74
4	−14.82	1	5.24
5	−14.78	4	3.08
6	−14.75	2	4.80
7	−14.51	4	4.56
8	−14.36	2	5.60
9	−14.23	3	5.18
10	−14.11	2	6.08

Table S23: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 102D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^8$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−12.94	6	2.19
2	−12.78	5	1.05
3	−12.63	10	2.02
4	−12.47	6	1.80
5	−12.41	4	2.98
6	−12.38	1	4.28
7	−12.35	4	4.89
8	−12.33	6	1.21
9	−12.15	3	1.51
10	−12.13	5	2.44

Table S24: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 1PRP. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^8$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−12.28	19	2.99
2	−12.25	22	2.94
3	−12.12	7	1.67
4	−12.05	3	2.62
5	−12.00	3	5.06
6	−11.93	1	6.87
7	−11.91	7	2.24
8	−11.88	4	2.67
9	−11.88	2	2.74
10	−11.87	4	1.68

Table S25: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 2GB9. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^7$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−10.57	5	2.69
2	−10.24	1	8.17
3	−10.19	7	8.69
4	−10.12	5	4.26
5	−10.07	3	2.41
6	−9.96	1	3.91
7	−9.93	8	7.68
8	−9.86	17	8.39
9	−9.82	3	5.14
10	−9.77	3	4.21

Table S26: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 1D32. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^8$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−20.12	1	7.09
2	−20.01	4	6.55
3	−19.79	10	1.37
4	−19.46	3	7.05
5	−19.41	1	3.31
6	−19.38	1	6.47
7	−19.36	1	7.22
8	−19.32	4	6.57
9	−19.30	7	6.56
10	−19.20	1	3.22

Table S27: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 334D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^8$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−9.54	2	9.83
2	−9.26	2	12.13
3	−9.23	1	5.11
4	−9.03	1	9.59
5	−8.97	1	9.34
6	−8.89	1	8.47
7	−8.88	1	9.51
8	−8.87	2	13.25
9	−8.76	2	13.06
10	−8.59	2	10.07

Table S28: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 2DND. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^8$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−13.95	2	13.21
2	−13.18	1	3.73
3	−12.97	1	2.94
4	−12.63	1	8.52
5	−12.51	1	11.07
6	−12.49	1	13.10
7	−12.49	1	11.74
8	−12.21	1	13.77
9	−12.19	1	5.26
10	−12.15	2	13.09

Table S29: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 182D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^6$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−11.78	9	1.67
2	−10.97	19	7.04
3	−10.81	3	10.76
4	−9.70	2	10.80
5	−8.90	6	10.36
6	−5.89	1	15.08
7	−5.65	1	9.23
8	−5.56	1	11.11
9	−5.23	1	17.67
10	−5.00	1	18.09

Table S30: Binding site information (binding free energies in kcal/mol, number of conformations in each site, and RMSDs in Å) for PDB ID: 473D. At most ten binding sites with the lowest binding free energies are shown. Number of free energy evaluations: $2 * 10^7$.

BINDING SITE	ΔG	NUMBER IN CLUSTER	RMSD
1	−12.58	1	5.44
2	−12.13	1	5.92
3	−12.11	1	5.03
4	−11.86	1	9.92
5	−11.56	1	5.74
6	−11.42	1	5.60
7	−11.30	1	8.07
8	−11.28	1	1.97
9	−11.22	1	8.15
10	−11.21	1	5.17

Table S31: Success rates using multiple thresholds in RMSDs considering the site of lowest binding free energy or all sites within some percentage from the lowest binding free energy. These results include only systems for which convergence was achieved.

	Threshold		
	1 Å	2 Å	3 Å
RMSD OF ΔG_{lowest}	19%	39%	61%
LOWEST RMSD WITHIN 5% OF ΔG_{lowest}	25%	61%	75%
LOWEST RMSD WITHIN 10% OF ΔG_{lowest}	32%	68%	86%
LOWEST RMSD WITHIN 15% OF ΔG_{lowest}	36%	79%	93%
LOWEST RMSD WITHIN 20% OF ΔG_{lowest}	39%	86%	96%

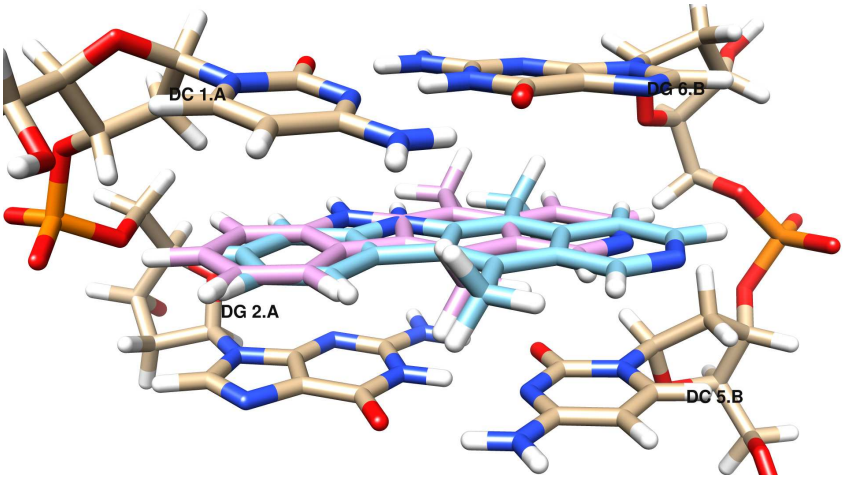

Figure S10: Binding site that corresponds to $\Delta G = -9.23$ kcal/mol for system 1Z3F. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.37 Å. The predicted geometry overlaps with the experimentally determined geometry in the intercalation site and is in the same orientation. The ligand in both geometries forms no hydrogen bonds with DNA.

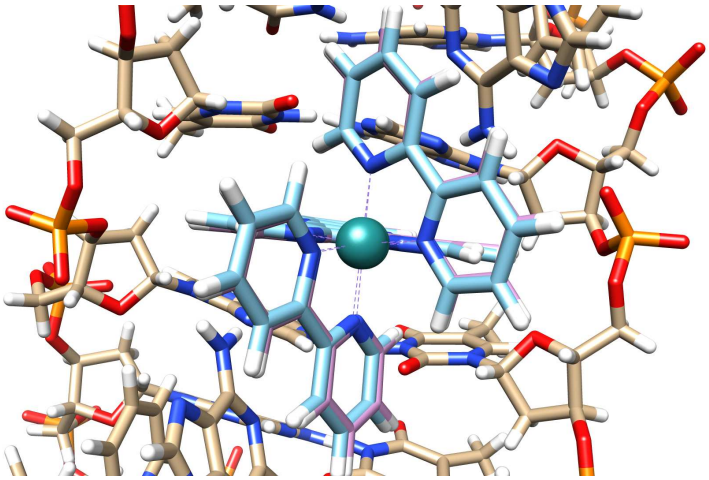

Figure S11: Binding site that corresponds to $\Delta G_{lowest} = -16.23$ kcal/mol for system 4E1U. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.19 Å. The ligand in the experimental geometry is in complete overlap with the ligand in the predicted geometry. Both ligands form no hydrogen bonds with the DNA molecule.

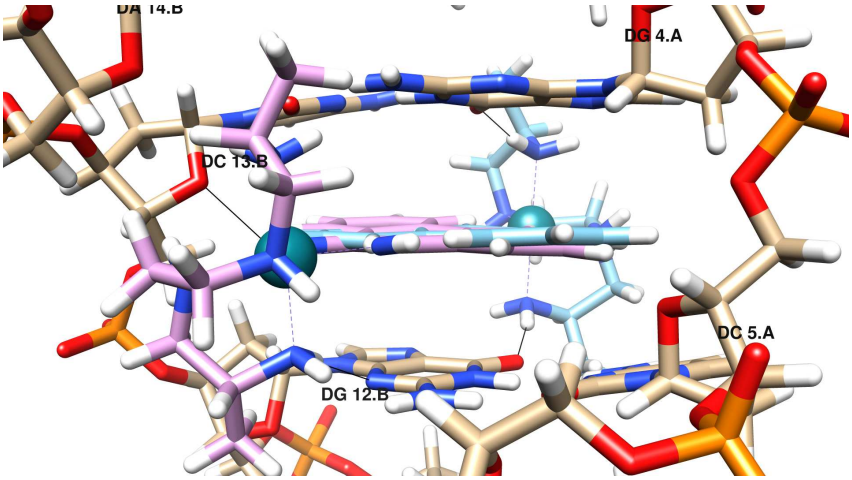

Figure S12: Binding site that corresponds to $\Delta G_{lowest} = -15.54$ kcal/mol for system 454D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 8.40 Å. The ligand in the experimental geometry intercalates from the major groove, whereas the ligand in the predicted geometry intercalates from the minor groove. In the experimentally determined geometry, the ligand forms two hydrogen bonds with two guanine bases (atom O6, residues 4 and 12). In the predicted geometry, the ligand forms a hydrogen bond with a guanine base (atom N3, residue 12) and another one with a sugar (atom O4′, residue 13).

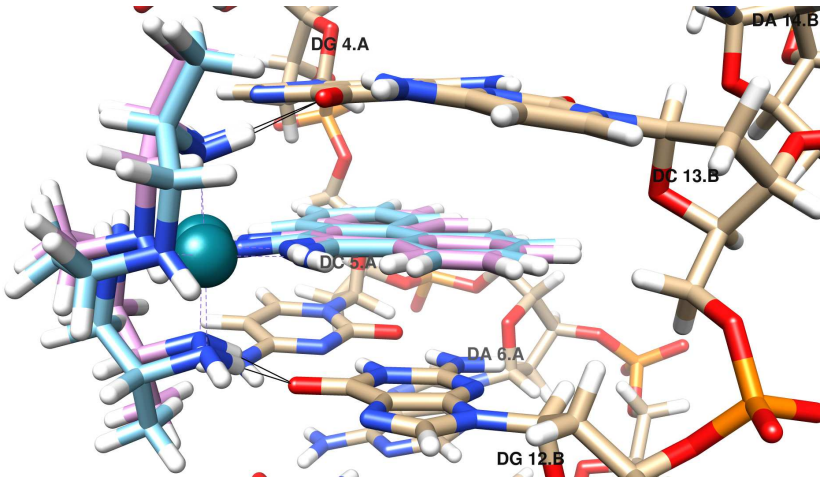

Figure S13: Binding site that corresponds to $\Delta G = -13.08$ kcal/mol for system 454D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.98 Å. The ligand in both the experimental and predicted geometries intercalates from the major groove. In the experimentally determined geometry, the ligand forms two hydrogen bonds with two guanine bases (atom O6, residues 4 and 12) with both present in the predicted geometry as well.

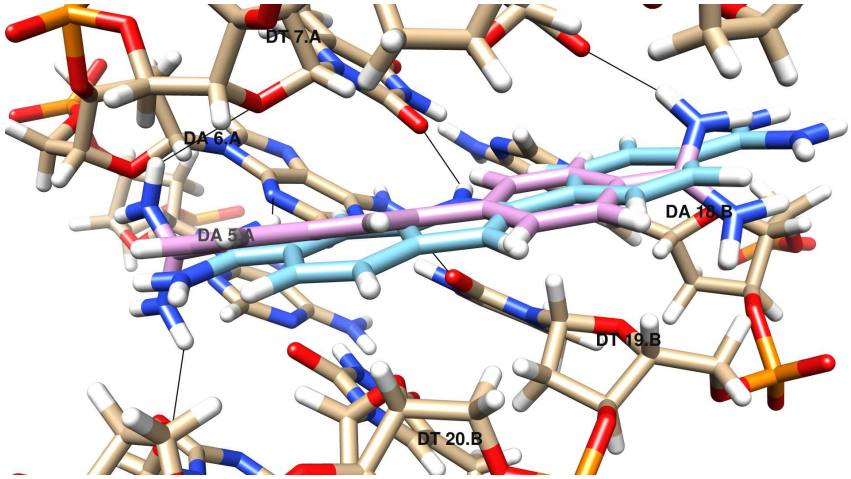

Figure S14: Binding site that corresponds to $\Delta G_{lowest} = -12.67$ kcal/mol for system 1D30. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.19 Å. The predicted geometry is displaced from the experimentally determined geometry. There are two hydrogen bonds in the experimentally determined geometry, with an adenine (atom N3, residue 6) and a thymine base (atom O2, residue 7). None of those two hydrogen bonds are present in the predicted geometry. Instead, the ligand in the predicted geometry forms hydrogen bond with two thymine bases (atom O2, residues 8 and 19), and an amidinium group forms two hydrogen bonds with two sugars on both strands (atom O4′, residues 7 and 21).

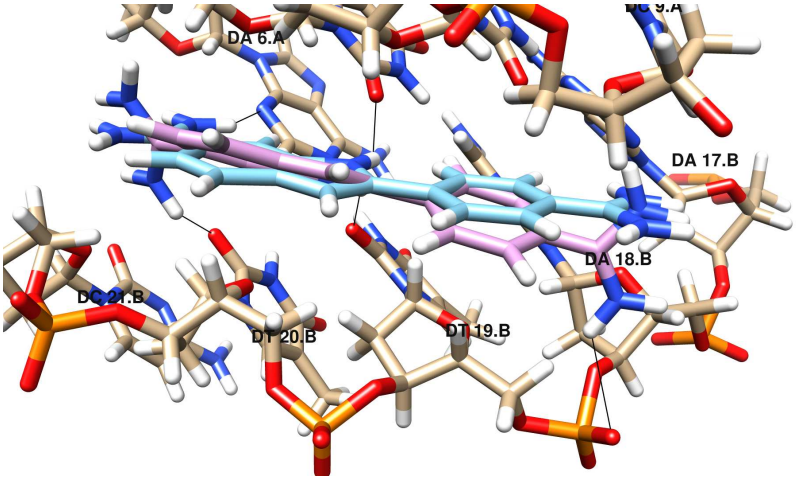

Figure S15: Binding site that corresponds to $\Delta G = -12.14$ kcal/mol for system 1D30. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.87 Å. The rings in the predicted geometry overlap significantly with the rings in the experimentally determined geometry. There are two hydrogen bonds in the experimentally determined geometry, with an adenine (atom N3, residue 6) and a thymine base (atom O2, residue 7). None of those two hydrogen bonds are present in the predicted geometry. Instead, the ligand in the predicted geometry forms hydrogen bonds with two thymine bases (atom O2, residues 19 and 20) on the same strand, and an amidinium group forms a hydrogen bonds with a phosphate group oxygen (residue 19).

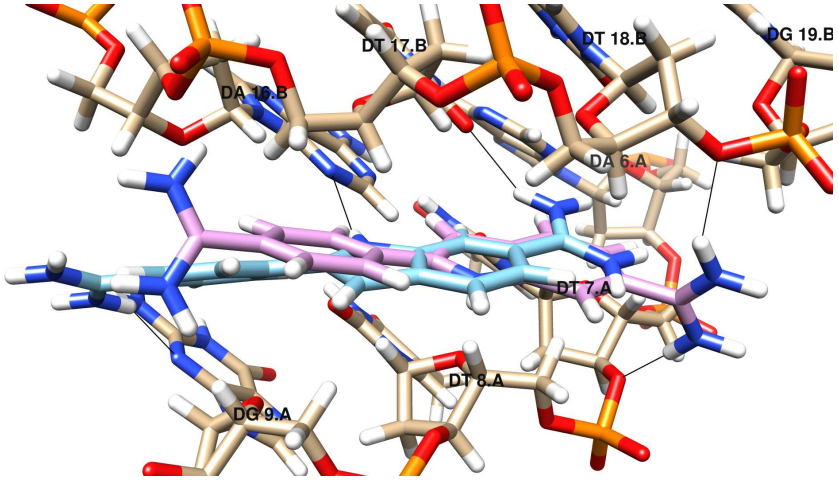

Figure S16: Binding site that corresponds to $\Delta G_{lowest} = -13.20$ kcal/mol for system 432D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 3.46 Å. The predicted geometry is displaced from the experimentally determined geometry. The ligand in the experimentally determined geometry forms three hydrogen bonds with the DNA molecule, none of which are present in the predicted geometry. In the experimentally determined geometry, the ligand forms hydrogen bonds with a thymine base (atom O2, residue 17), with an adenine base (atom N3, residue 16), and a guanine base (atom N3, residue 9). In the predicted geometry, an amidinium group forms two hydrogen bonds with two sugars on both strands (atoms O3′, residues 7 and 18).

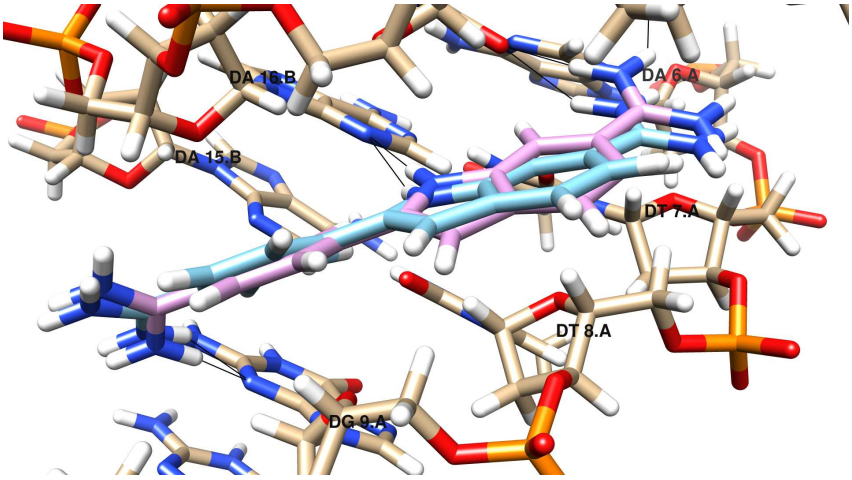

Figure S17: Binding site that corresponds to $\Delta G = -11.47$ kcal/mol for system 432D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.53 Å. The predicted geometry overlaps significantly with the experimentally determined geometry. The ligand in the experimentally determined geometry forms three hydrogen bonds with the DNA molecule, all of which are present in the predicted geometry. In the experimentally determined geometry, the ligand forms hydrogen bonds with a thymine base (atom O2, residue 17), with an adenine base (atom N3, residue 16), and a guanine base (atom N3, residue 9). In the predicted geometry, in addition to the three hydrogen bonds present in the experimentally determined geometry, there is an additional hydrogen bond with a thymine base (atom O2, residue 18).

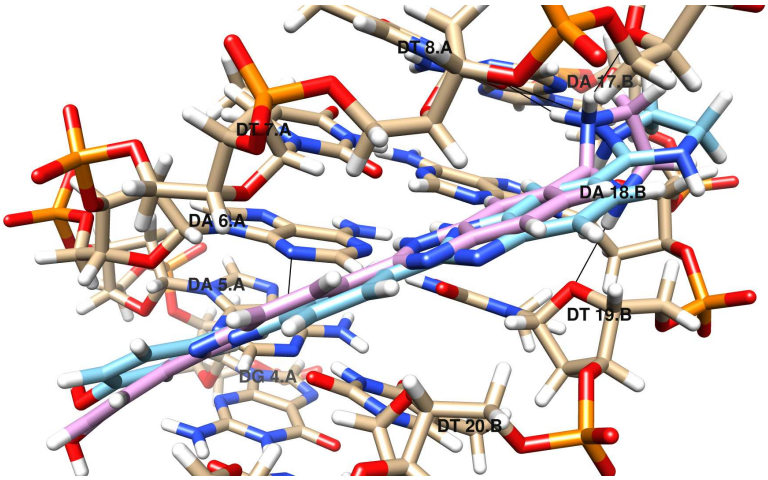

Figure S18: Binding site that corresponds to $\Delta G_{lowest} = -15.94$ kcal/mol for system 109D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.26 Å. The predicted geometry overlaps with the experimentally determined geometry. The ligand in the experimentally determined geometry forms three hydrogen bonds, one with an adenine base (atom N3, residue 6), one with a thymine base (atom O2, residue 8), and another one with a sugar (atom O4′, residue 18). The ligand in the predicted geometry forms three hydrogen bonds, one of which is also found in the experimentally predicted structure (with a thymine base, atom O2, residue 8). The other two hydrogen bonds are formed with two sugars (atom O4′, residues 9 and 19).

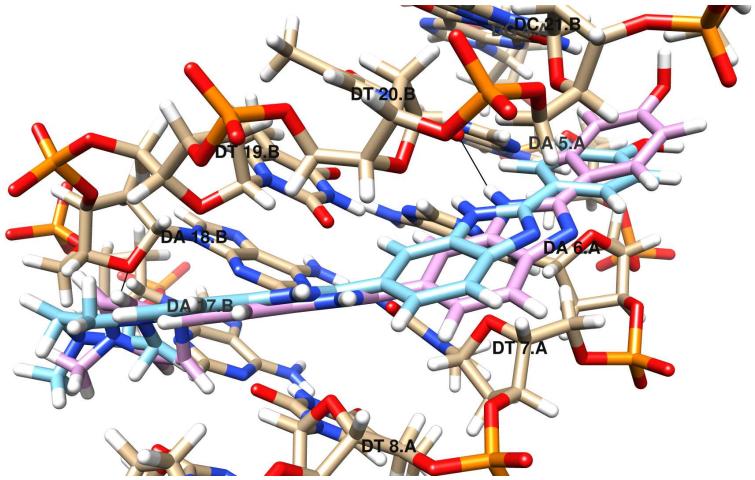

Figure S19: Binding site that corresponds to $\Delta G_{lowest} = -14.80$ kcal/mol for system 127D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.16 Å. The predicted geometry overlaps to some extent with the experimentally determined geometry but is displaced from the latter. The ligand in the experimentally determined geometry does not form any hydrogen bonds with the DNA molecule, whereas the ligand in the predicted geometry forms a hydrogen bond with a thymine base (atom O2, residue 20) and a sugar (atom O4′, residue 18).

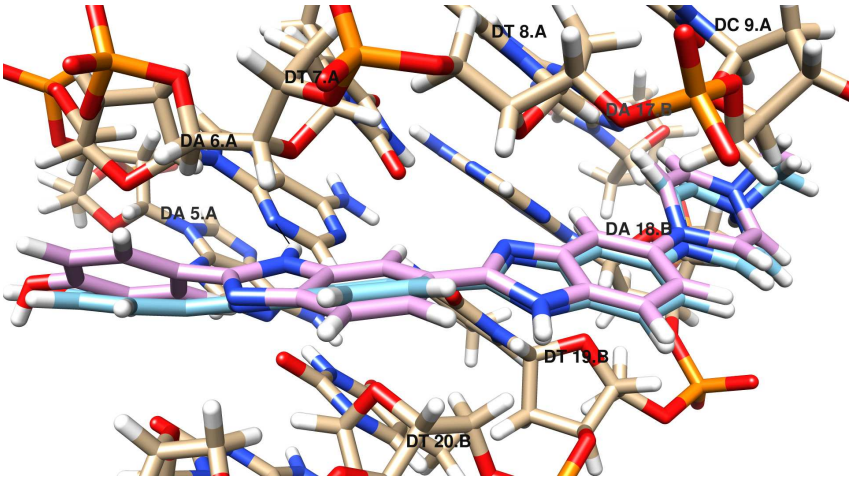

Figure S20: Binding site that corresponds to $\Delta G = -13.81$ kcal/mol for system 127D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.57 Å. The predicted geometry overlaps significantly with the experimentally determined geometry. The ligand in the experimentally determined geometry does not form any hydrogen bonds with the DNA molecule, whereas the ligand in the predicted geometry forms a hydrogen bond with an adenine base (atom N3, residue 6).

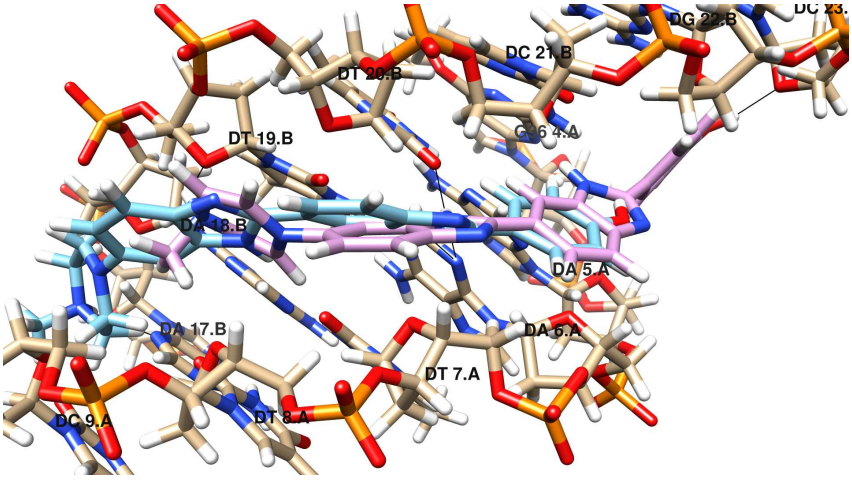

Figure S21: Binding site that corresponds to $\Delta G_{lowest} = -14.77$ kcal/mol for system 128D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 6.67 Å. The predicted geometry is significantly displaced from the experimentally determined geometry. The ligand in the experimentally determined geometry forms two hydrogen bonds with the DNA molecule, one of which is also present in the predicted geometry. In the experimentally determined geometry the ligand forms two hydrogen bonds with two adenine bases (atom N3, residues 6 and 17). In the predicted geometry, the ligand forms two hydrogen bonds with two sugars (atom O4′, residues 19 and 23), a hydrogen bond with a thymine base (atom O2, residue 20) and an adenine base (atom N3, residue 6).

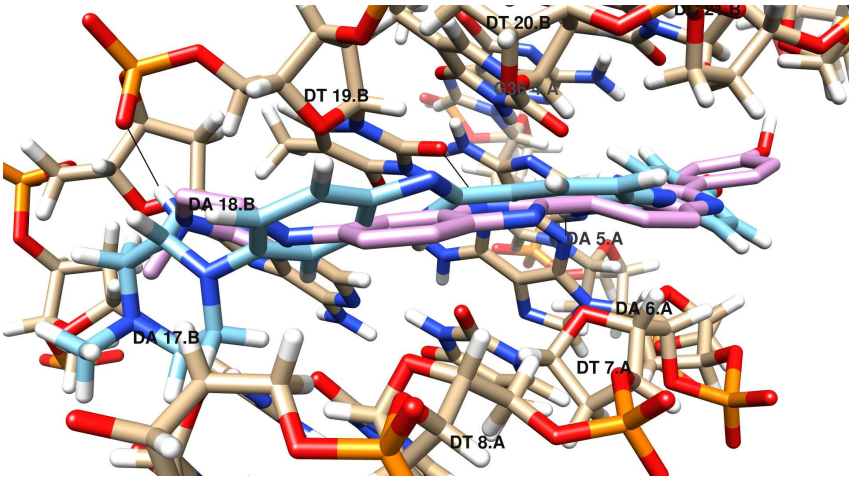

Figure S22: Binding site that corresponds to $\Delta G = -14.38$ kcal/mol for system 128D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.17 Å. The predicted geometry is displaced from the experimentally determined geometry. The ligand in the experimentally determined geometry forms two hydrogen bonds with the DNA molecule, none of which is present in the predicted geometry. In the experimentally determined geometry the ligand forms two hydrogen bonds with two adenine bases (atom N3, residues 6 and 17). In the predicted geometry, the ligand forms a hydrogen bond with a phosphate group oxygen (residue 19), and a hydrogen bond with a thymine base (atom O2, residue 19).

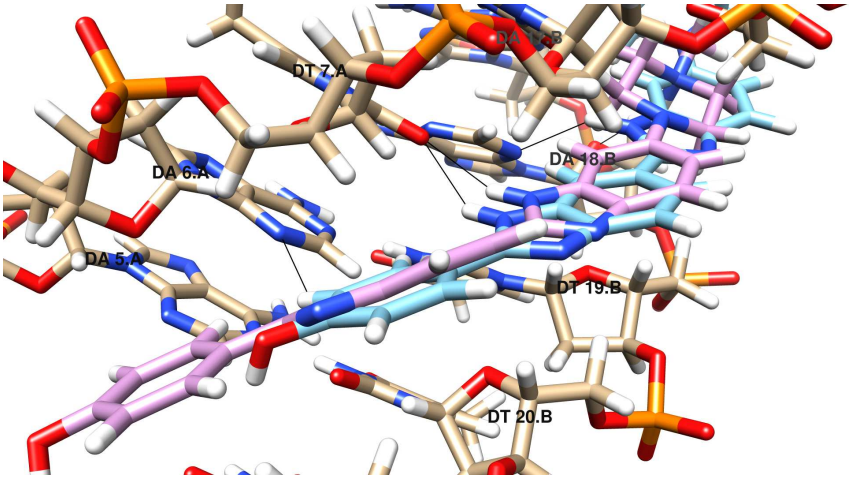

Figure S23: Binding site that corresponds to $\Delta G_{lowest} = -16.18$ kcal/mol for system 264D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 5.03 Å. The predicted geometry is displaced from the experimentally determined geometry. The ligand in the experimentally determined geometry forms two hydrogen bonds with the DNA molecule, one of which is also present in the predicted geometry. In the experimentally determined geometry the ligand forms two hydrogen bonds with an adenine base (atom N3, residue 18) and a thymine base (atom O2 residue 7). In the predicted geometry, the ligand forms a hydrogen bond with a sugar (atom O4′, residue 18), a hydrogen bond with a thymine base (atom O2, residue 7) and an adenine base (atom N3, residue 6).

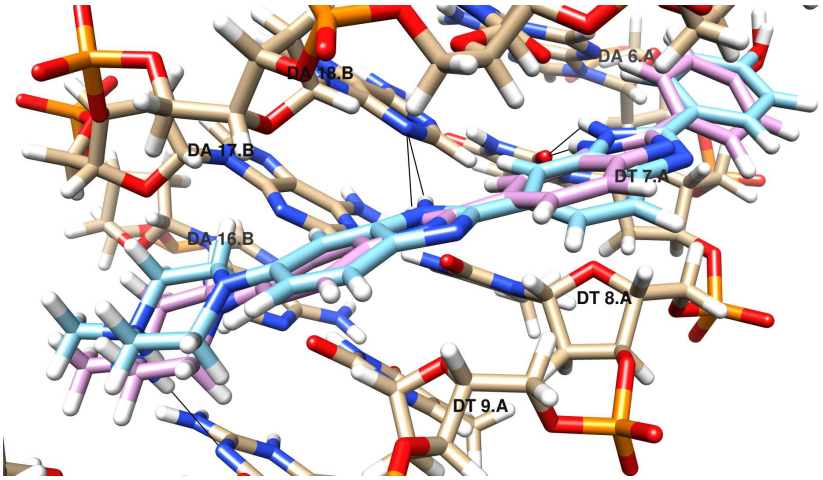

Figure S24: Binding site that corresponds to $\Delta G = -15.94$ kcal/mol for system 264D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.73 Å. The predicted geometry overlaps significantly with the experimentally determined geometry. The ligand in the experimentally determined geometry forms two hydrogen bonds with the DNA molecule, both of which are also present in the predicted geometry. In the experimentally determined geometry the ligand forms two hydrogen bonds with an adenine base (atom N3, residue 18) and a thymine base (atom O2 residue 7). In the predicted geometry, the ligand forms an additional hydrogen bond with a guanine (atom N3, residue 10).

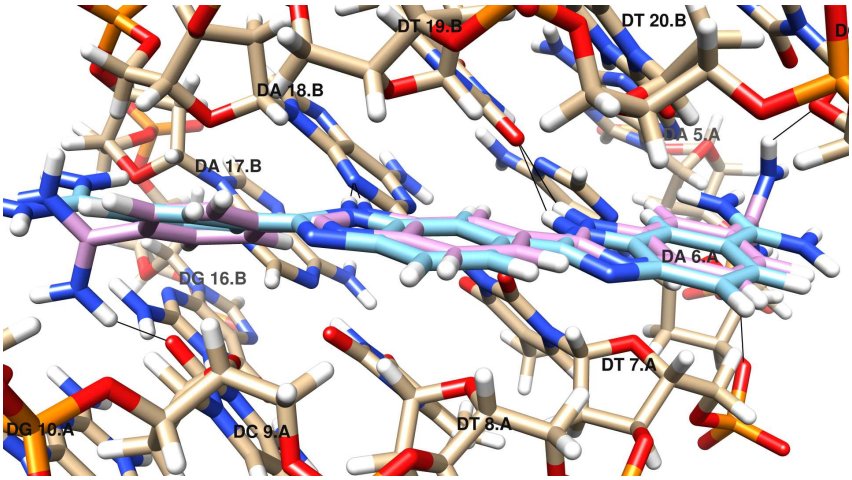

Figure S25: Binding site that corresponds to $\Delta G_{lowest} = -17.14$ kcal/mol for system 311D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.96 Å. There is significant overlap between the rings of the predicted and experimentally determined geometries. In the experimentally determined structure, the ligand forms two hydrogen bonds, with an adenine base (atom N3, residue 18) and a thymine base (atom O2, residue 19), with both of these hydrogen bonds being present in the predicted geometry. The amidinium groups of the ligand in the predicted geometry also form hydrogen bonds with a cytosine base (atom O2, residue 9) and a sugar (atom O4′, residue 21).

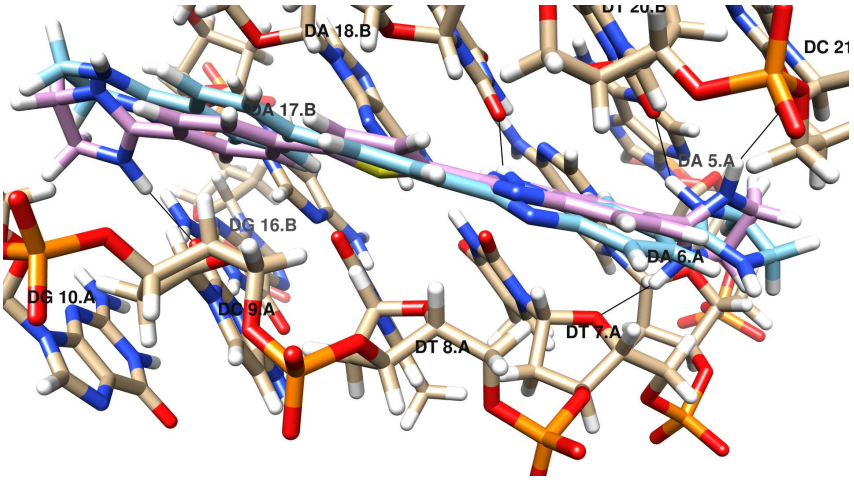

Figure S26: Binding site that corresponds to $\Delta G_{lowest} = -15.85$ kcal/mol for system 2B3E. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.96 Å. There is significant overlap between the rings of the predicted and experimentally determined geometries. In the experimentally determined structure, there are two hydrogen bonds with two thymine bases (atom O2, residues 19 and 20) and none of them appear in the predicted geometry. In the latter, there are two hydrogen bonds with two sugars (atom O4′, residues 7 and 21), and a hydrogen bond with a cytosine base (atom O2, residue 9).

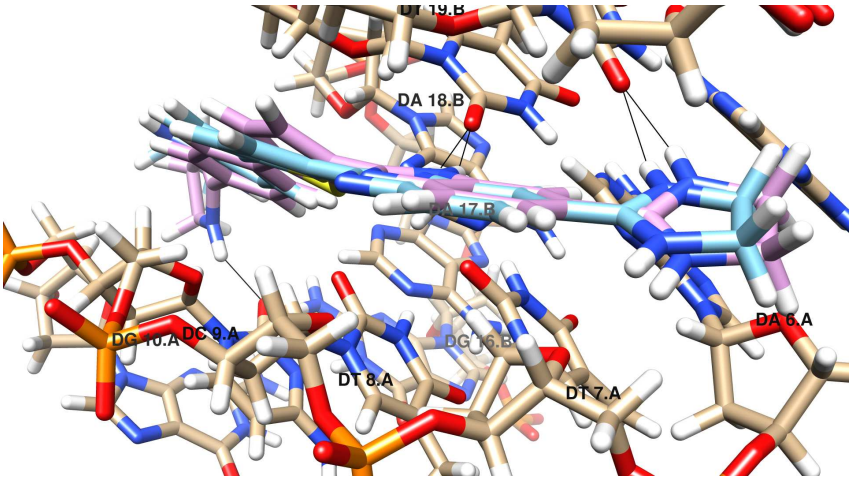

Figure S27: Binding site that corresponds to $\Delta G = -15.65$ kcal/mol for system 2B3E. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.70 Å. There is significant overlap between the rings of the predicted and experimentally determined geometries. In the experimentally determined structure, there are two hydrogen bonds with two thymine bases (atom O2, residues 19 and 20). The same two hydrogen bonds appear in the predicted structure, in addition to a third hydrogen bond with a cytosine base (atom O2, residue 9).

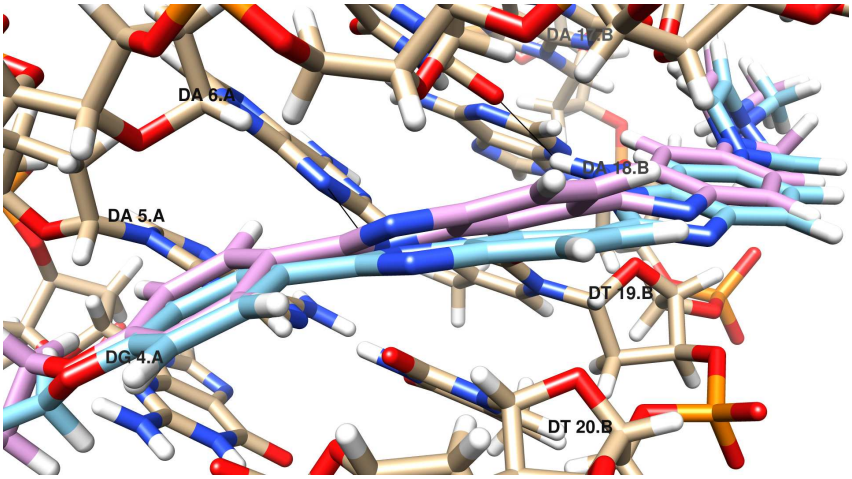

Figure S28: Binding site that corresponds to $\Delta G_{lowest} = -16.33$ kcal/mol for system 129D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.70 Å. The predicted geometry overlaps with the experimentally determined geometry. The ligand in the experimentally determined geometry forms a hydrogen bond with an adenine base (atom N3, residue 6), whereas the ligand in the predicted geometry forms a hydrogen bond with a thymine base (atom O2, residue 7).

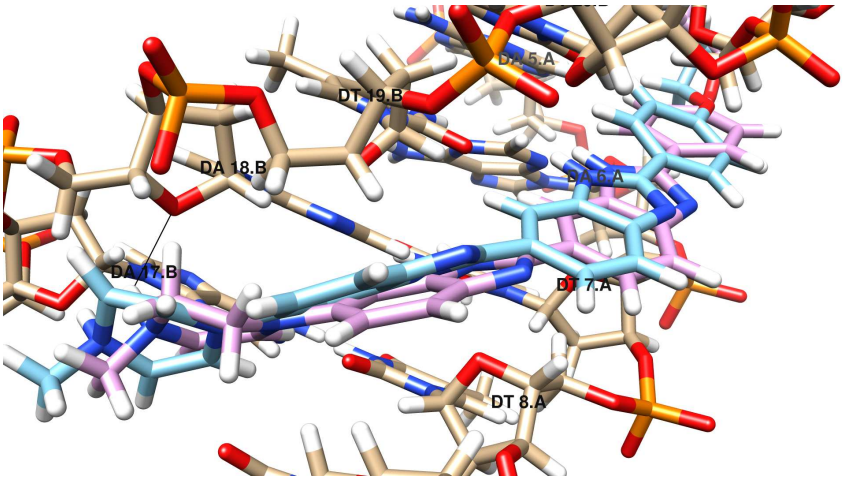

Figure S29: Binding site that corresponds to $\Delta G_{lowest} = -15.89$ kcal/mol for system 130D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.41 Å. The predicted geometry overlaps to some extent with the experimentally determined geometry. The ligand in the experimentally determined geometry forms no hydrogen bonds with the DNA molecule, whereas the ligand in the predicted geometry forms a hydrogen bond with a sugar (atom O4′, residue 18).

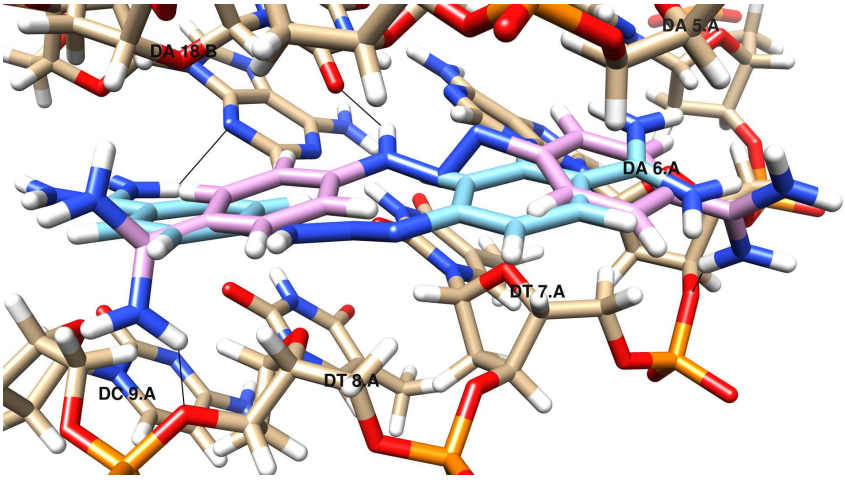

Figure S30: Binding site that corresponds to $\Delta G_{lowest} = -11.85$ kcal/mol for system 2DBE. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.85 Å. The rings of the predicted and experimentally determined geometries are displaced from each other. A secondary amine in the predicted geometry forms a hydrogen bond with a thymine base (atom O2, residue 19), whereas the same group faces the solvent in the experimentally determined structure. An amidinium group forms a hydrogen bond with an adenine (atom N3, residue 18) in the experimentally determined geometry, whereas the same group forms a hydrogen bond with a sugar (atom O3′, residue 8) located on the opposite strand in the predicted geometry. The other amidinium group forms a hydrogen bond with a sugar (atom O3′, residue 6) in the predicted geometry, whereas there are no hydrogen bonds in the experimentally determined geometry.

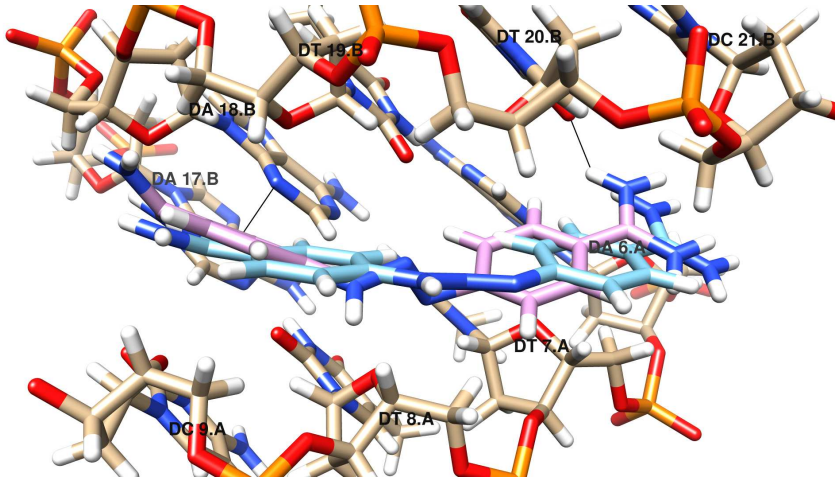

Figure S31: Binding site that corresponds to $\Delta G = -10.55$ kcal/mol for system 2DBE. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.02 Å. The rings of the predicted and experimentally determined geometries are displaced from each other. An amidinium group in the predicted geometry forms a hydrogen bond with a thymine base (atom O2, residue 20), whereas an amidinium group in the experimentally predicted geometry forms a hydrogen bond with an adenine base (atom N3, residue 18).

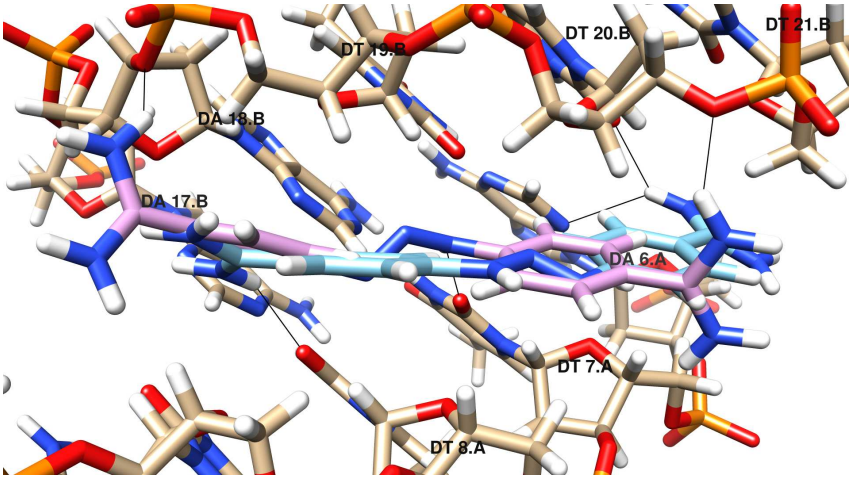

Figure S32: Binding site that corresponds to $\Delta G_{lowest} = -12.41$ kcal/mol for system 1D63. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 3.07 Å. The amidinium groups of the ligand at the predicted geometry form 2 hydrogen bonds with 2 sugars (atom O3′, residues 18 and 20) whereas the same groups at the experimental geometry form hydrogen bonds with two thymine bases (atom O2, residues 8 and 20) and an adenine base (atom N3, residue 6). In the predicted geometry there is also a hydrogen bond between the ligand's azido group and a thymine base (atom O2, residue 7), whereas no such hydrogen bond is present in the experimental structure.

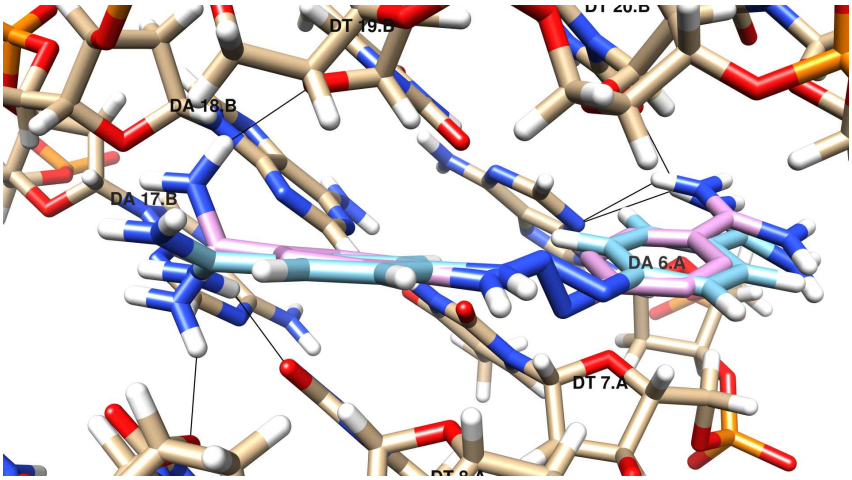

Figure S33: Binding site that corresponds to $\Delta G = -11.55$ kcal/mol for system 1D63. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 0.73 Å. In the predicted geometry one amidinium group forms two hydrogen bonds with the O4′ atoms of two DNA sugars on opposite strands (residues 9 and 19), whereas the other amidinium group forms a hydrogen bond with an adenine base (atom N3, residue 6). The latter hydrogen bond is also found in the experimentally determined geometry. The hydrogen of the ligand's azido group faces the solvent, as found in the experimentally determined structure.

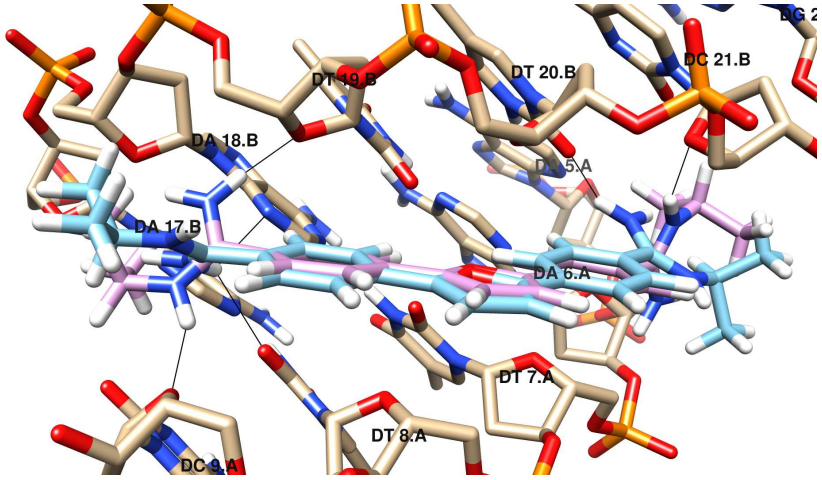

Figure S34: Binding site that corresponds to $\Delta G_{lowest} = -15.71$ kcal/mol for system 298D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.48 Å. There is significant overlap between the rings of the predicted and experimentally determined geometries. In the experimentally determined structure, an amidinium group forms a hydrogen bond with a thymine base (atom O2, residue 8), and a hydrogen bond with an adenine base from the opposite strand (atom N3, residue 18). The same group in the predicted geometry forms two hydrogen bonds with a sugar from each strand (atom O4′, residue 9 and atom O4′, residue 19). The other amidinium group forms a hydrogen bond with a thymine base (atom O2, residue 20) in the experimentally determined structure, and a hydrogen bond with a sugar (atom O4′, residue 21) in the predicted geometry.

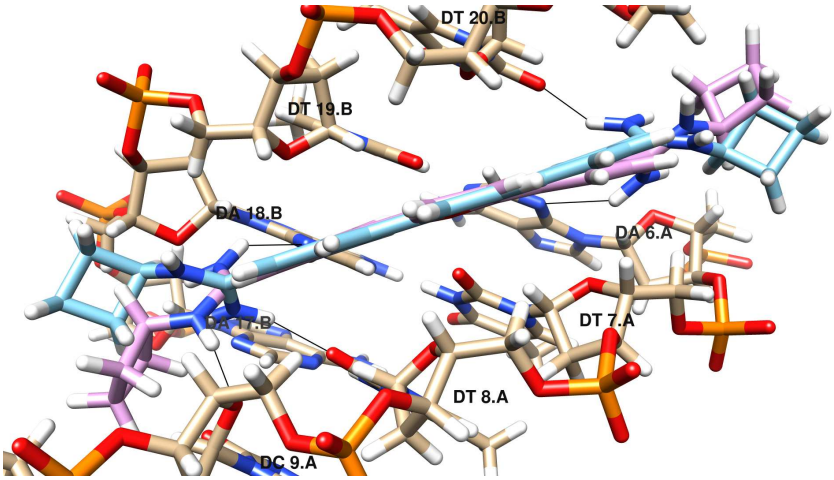

Figure S35: Binding site that corresponds to $\Delta G_{lowest} = -15.69$ kcal/mol for system 1FMQ. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.25 Å. There is significant overlap between the rings of the predicted and experimentally determined geometries. In the experimentally determined structure, the amidinium groups form two hydrogen bonds with two thymine bases (atom O2, residues 8 and 20), with none of those two hydrogen bonds shared with the predicted geometry. In the latter, the ligand forms two hydrogen bonds with two adenine bases (atom N3, residues 6 and 18) and a hydrogen bond with a sugar (atom O4′, residue 9).

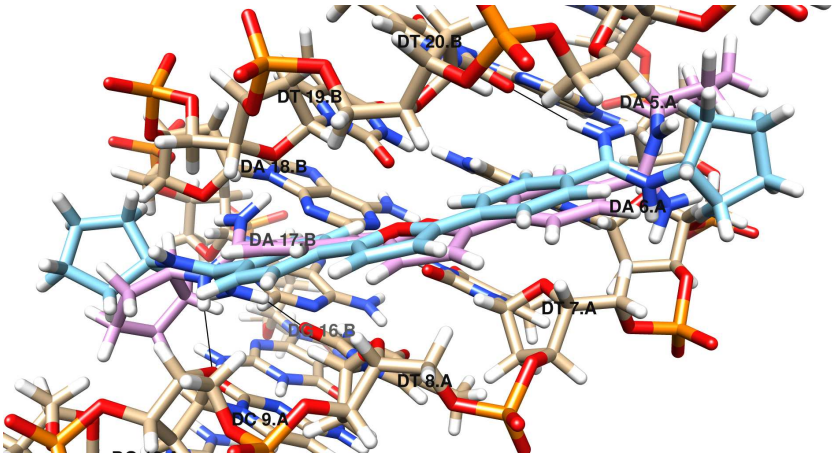

Figure S36: Binding site that corresponds to $\Delta G_{lowest} = -16.85$ kcal/mol for system 1EEL. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.01 Å. There is significant overlap between the central rings of the predicted and experimentally determined geometries. In the experimentally determined structure, an amidinium group forms two hydrogen bonds with two thymine bases (atom O2, residues 8 and 20) whereas in the predicted geometry the ligand forms a hydrogen bond with a sugar (atom O4′, residue 9).

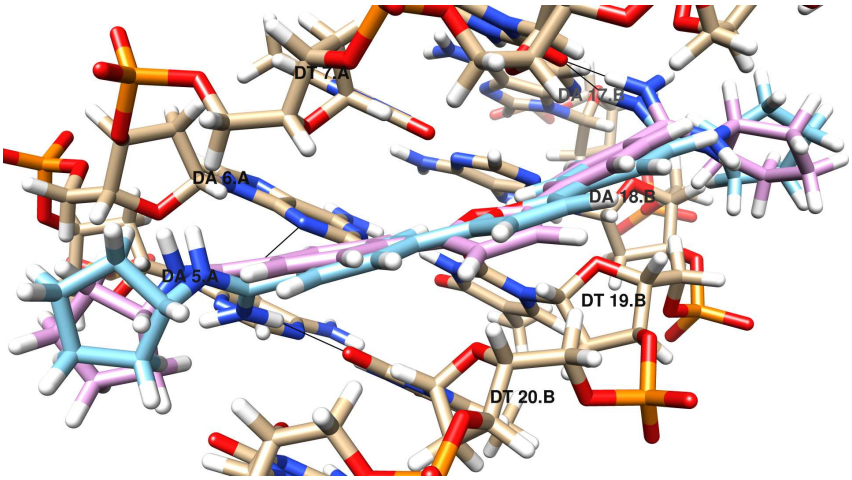

Figure S37: Binding site that corresponds to $\Delta G = -14.98$ kcal/mol for system 1EEL. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.06 Å. There is significant overlap between the rings of the predicted and experimentally determined geometries. In both structures, an amidinium group forms a hydrogen bond with a thymine base (atom O2, residue 8). In the predicted geometry the ligand forms a hydrogen bond with an adenine base (atom N3, residue 6), whereas in the experimentally determined geometry it forms a hydrogen bond with a thymine base (atom O2, residue 20).

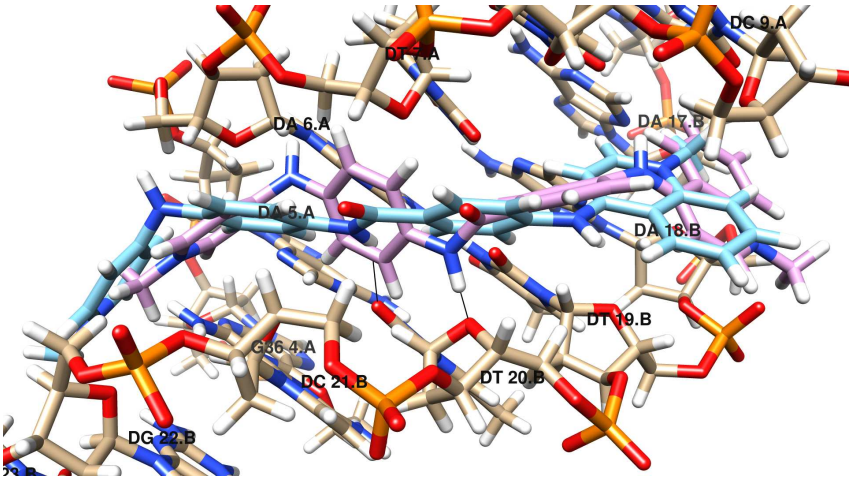

Figure S38: Binding site that corresponds to $\Delta G_{lowest} = -16.58$ kcal/mol for system 144D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 3.49 Å. There is some overlap between the predicted and experimentally determined geometries, but the two have different hydrogen bonding patterns and are displaced from each other. In the experimentally determined structure, there is a hydrogen bond with a thymine base (atom O2, residue 20), whereas in the predicted geometry the ligand forms a hydrogen bond with a sugar (atom O4′, residue 20).

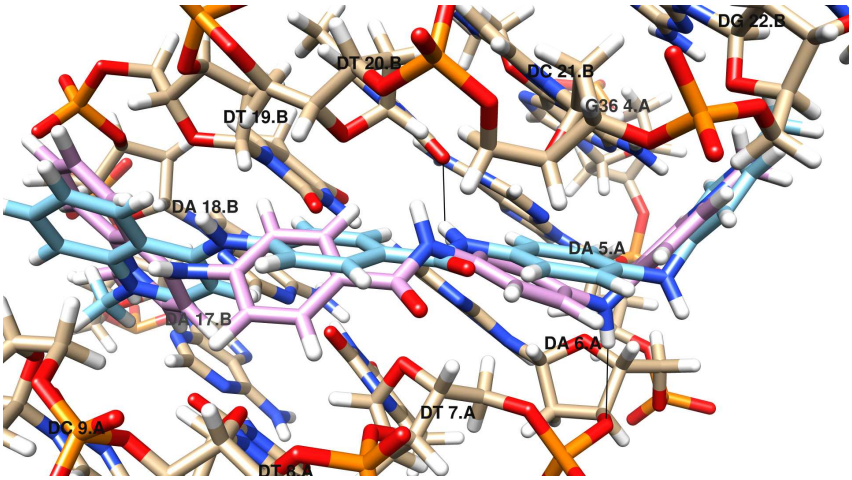

Figure S39: Binding site that corresponds to $\Delta G = -13.61$ kcal/mol for system 144D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.46 Å. There is significant overlap between the predicted and experimentally determined geometries, but the two have different hydrogen bonding patterns. In the experimentally determined structure, there is a hydrogen bond with a thymine base (atom O2, residue 20), whereas in the predicted geometry the ligand forms a hydrogen bond with a sugar (atom O3′, residue 6).

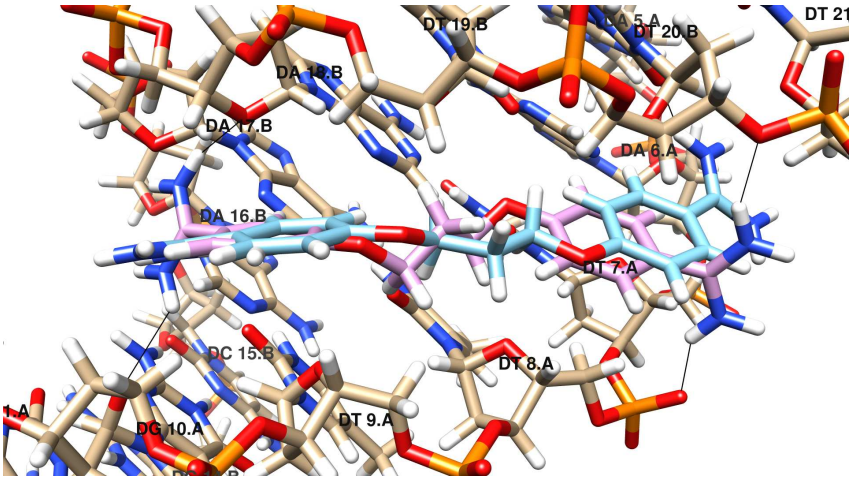

Figure S40: Binding site that corresponds to $\Delta G_{lowest} = -12.94$ kcal/mol for system 102D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.19 Å. There is some overlap between the predicted and experimentally determined geometries, but the two have different hydrogen bonding patterns. In the experimentally determined structure, an amidinium group forms a hydrogen bond with a thymine base (atom O2, residue 20) whereas in the predicted geometry the ligand forms three hydrogen bonds with three sugars (atoms O4′ and O3′, residues 10, 18, and 20), and a hydrogen bond with a phosphate group oxygen (residue 8).

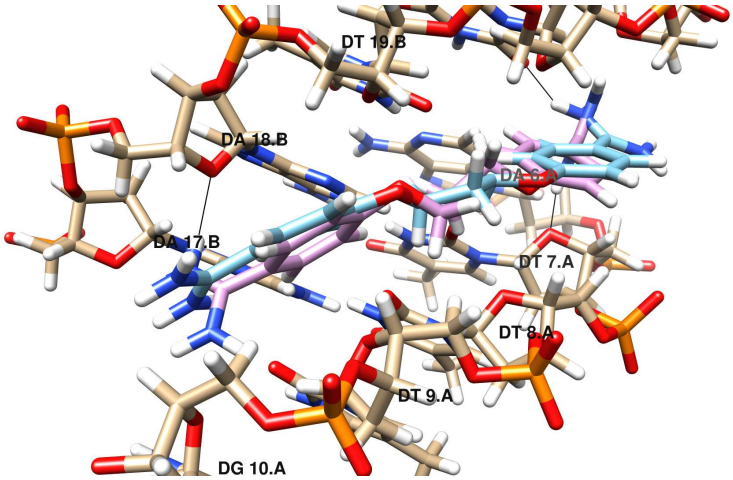

Figure S41: Binding site that corresponds to $\Delta G = -12.78$ kcal/mol for system 102D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.05 Å. There is significant overlap between the predicted and experimentally determined geometries, but the two have different hydrogen bonding patterns. In the experimentally determined structure, an amidinium group forms a hydrogen bond with a thymine base (atom O2, residue 20) whereas in the predicted geometry the ligand forms two hydrogen bonds with two sugars (atoms O4′, residues 7, 18).

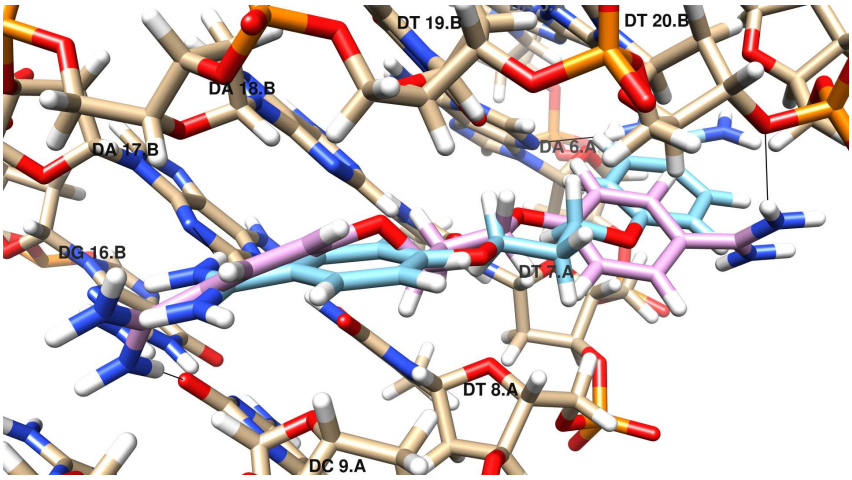

Figure S42: Binding site that corresponds to $\Delta G_{lowest} = -12.28$ kcal/mol for system 1PRP. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.99 Å. There is some overlap between the predicted and experimentally determined geometries, but the two have different hydrogen bonding patterns. In the experimentally determined structure, an amidinium group forms two hydrogen bonds with two adenine bases (atom N3, residues 6 and 17) whereas in the predicted geometry the ligand forms a hydrogen bond with a cytosine base (atom O2, residue 9) and a hydrogen bond with a sugar (atom O3′, residue 20).

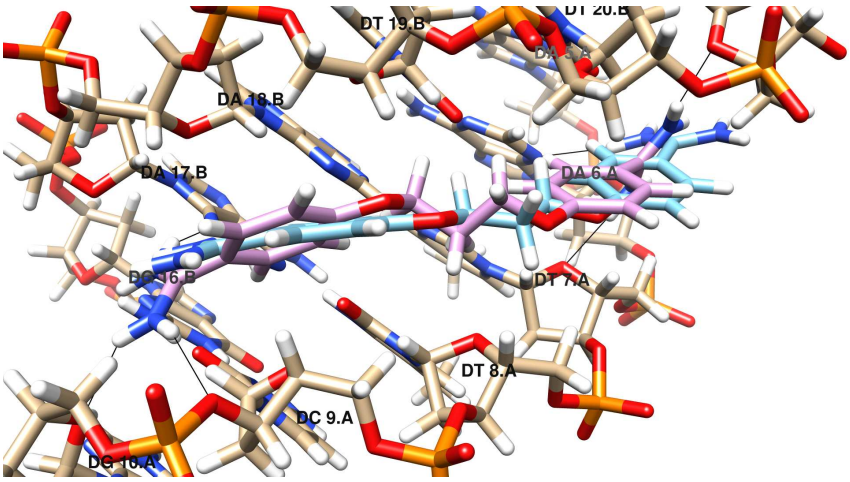

Figure S43: Binding site that corresponds to $\Delta G = -12.12$ kcal/mol for system 1PRP. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.67 Å. There is some overlap between the predicted and experimentally determined geometries, but the two have different hydrogen bonding patterns. In the experimentally determined structure, an amidinium group forms two hydrogen bonds with two adenine bases (atom N3, residues 6 and 17) whereas in the predicted geometry the ligand forms four hydrogen bonds with four sugars (atoms O3′ and O4′, residues 7, 9, 10, and 21).

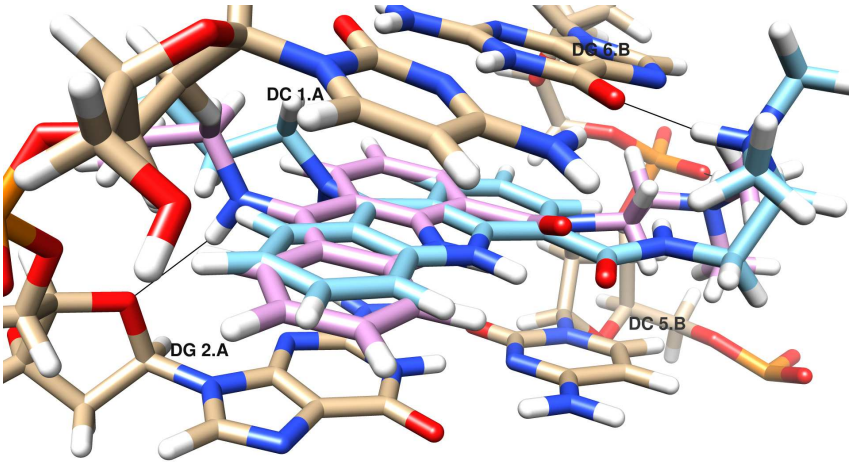

Figure S44: Binding site that corresponds to $\Delta G_{lowest} = -10.57$ kcal/mol for system 2GB9. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.69 Å. The two geometries overlap, however, the ligand in the predicted geometry is rotated with respect to the experimentally determined geometry. In the predicted geometry the ligand forms a hydrogen bond with a sugar (atom O4′, residue 2), whereas in the experimentally determined geometry the ligand forms a hydrogen bond with a guanine base (atom O6, residue 6) on the opposite strand.

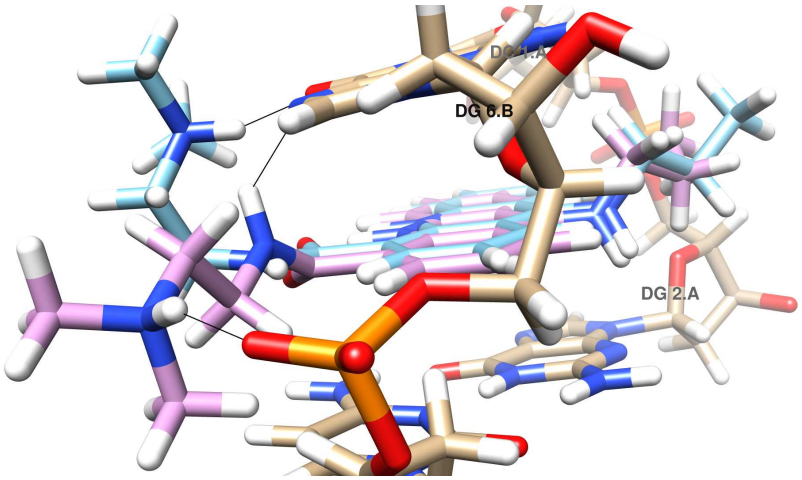

Figure S45: Binding site that corresponds to $\Delta G = -10.07$ kcal/mol for system 2GB9. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.41 Å. The rings of the two geometries overlap significantly. In the experimentally determined geometry the ligand forms a hydrogen bond with a guanine base (atom O6, residue 6), whereas in the predicted geometry, the ligand forms a hydrogen bond with an oxygen atom from a phosphate group (residue 6), and another one with a guanine base (atom N7, residue 6).

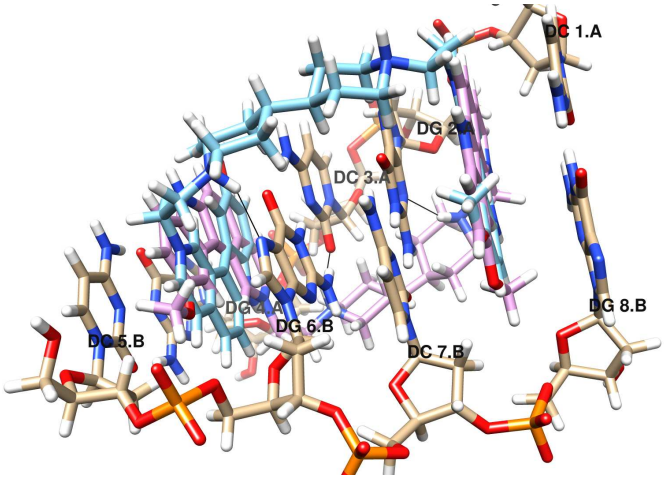

Figure S46: Binding site that corresponds to $\Delta G_{lowest} = -20.12$ kcal/mol for system 1D32. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 7.09 Å. The ligand in both geometries occupies the two intercalation sites, however, the ligand in the experimentally determined structure occupies the major groove, whereas the ligand in the predicted geometry occupies the minor groove. In the experimentally determined geometry there is a hydrogen bond with a guanine base (atom N7, residue 6), and in the predicted geometry there is a hydrogen bond with a cytosine base (atom O2, residue 3) and one with a guanine base (atom N3, residue 2).

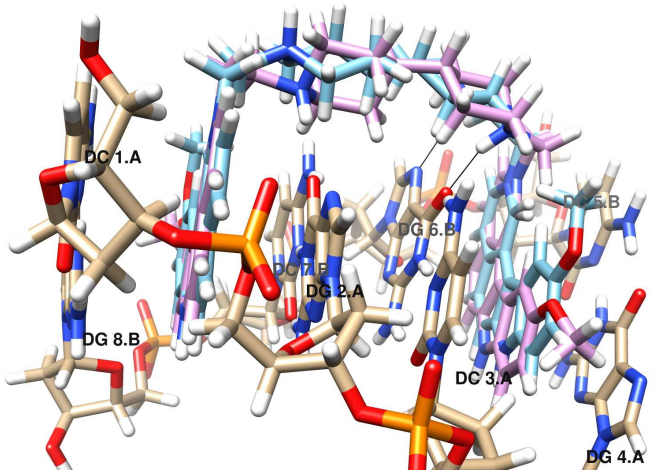

Figure S47: Binding site that corresponds to $\Delta G = -19.79$ kcal/mol for system 1D32. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.37 Å. The ligand in both geometries occupies the two intercalation sites and the major groove. In the experimentally determined geometry there is a hydrogen bond with a guanine base (atom N7, residue 6), and in the predicted geometry there is a hydrogen bond with the same guanine base but different acceptor atom (atom O6, residue 6).

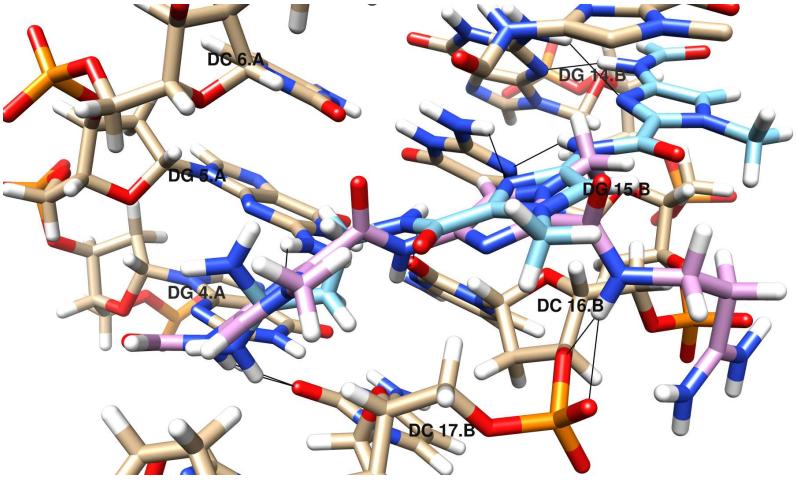

Figure S48: Binding site that corresponds to $\Delta G_{lowest} = -9.54$ kcal/mol for system 334D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 9.83 Å. There is partial overlap between the experimentally determined and the predicted geometry, however, a significant part of the ligand is not located in the minor groove. Part of the ligand in the predicted geometry is in the solvent, adjacent to two DNA phosphate groups. The ligand in the experimentally determined geometry forms forms two hydrogen bonds with two cytosine bases (atom O2, residues 16 and 17) and four hydrogen bonds with two guanine bases (atoms N2 and N3, residues 14 and 15). The ligand in the predicted geometry forms two hydrogen bonds with two guanine bases (atom N2, residues 4 and 5), two hydrogen bonds with two cytosine bases (atom O2, residues 16 and 17), a hydrogen bond with a sugar (atom O3′, residue 16) and a hydrogen bond with a phosphate group oxygen (residue 17).

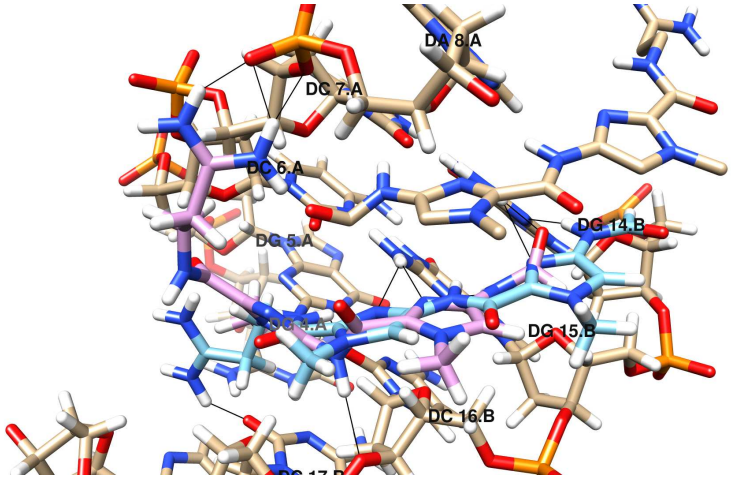

Figure S49: Binding site that corresponds to $\Delta G = -9.23$ kcal/mol for system 334D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 5.11 Å. There is partial overlap between the experimentally determined and the predicted geometry, however, a significant part of the ligand is not located in the minor groove. Part of the ligand in the predicted geometry is in the solvent, adjacent to a DNA phosphate group. The ligand in the experimentally determined geometry forms forms two hydrogen bonds with two cytosine bases (atom O2, residues 16 and 17) and four hydrogen bonds with two guanine bases (atoms N2 and N3, residues 14 and 15). The ligand in the predicted geometry forms two hydrogen bonds with two guanine bases (atom N2, residues 14 and 15), two hydrogen bonds with two sugars (atom O3′, residues 7 and 16), and an amidinium group forms two hydrogen bonds with a phosphate group oxygen (residue 8).

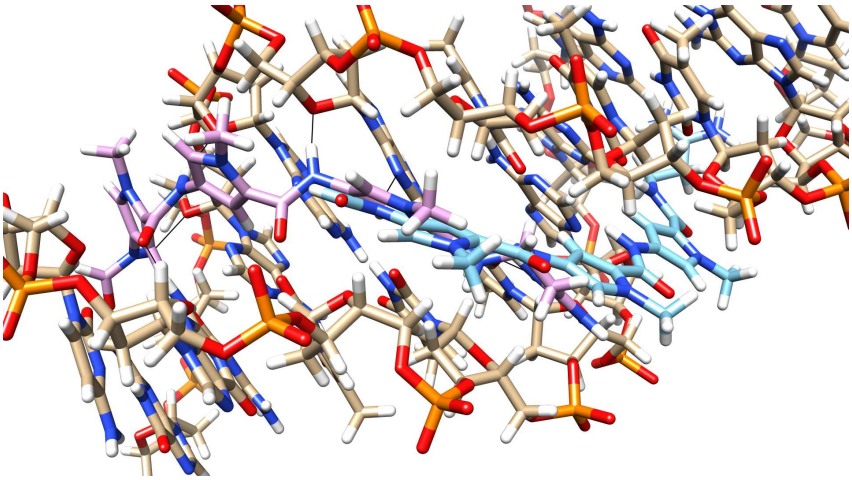

Figure S50: Binding site that corresponds to $\Delta G_{lowest} = -13.95$ kcal/mol for system 2DND. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 13.21 Å. The two ligands are both in the minor groove but in very different sites, with different hydrogen bonding patters. The ligand in the experimentally determined structure has a hydrogen bond with an adenine base (atom N3, residue 18), whereas the ligand in the predicted geometry has six hydrogen bonds. A hydrogen bond is formed with a thymine base (atom O2, residue 7), three hydrogen bonds are formed with sugars (atom O4′, residues 16,17,18), and two hydrogen bonds are formed with two guanine bases (atom N2, residues 10, 14).

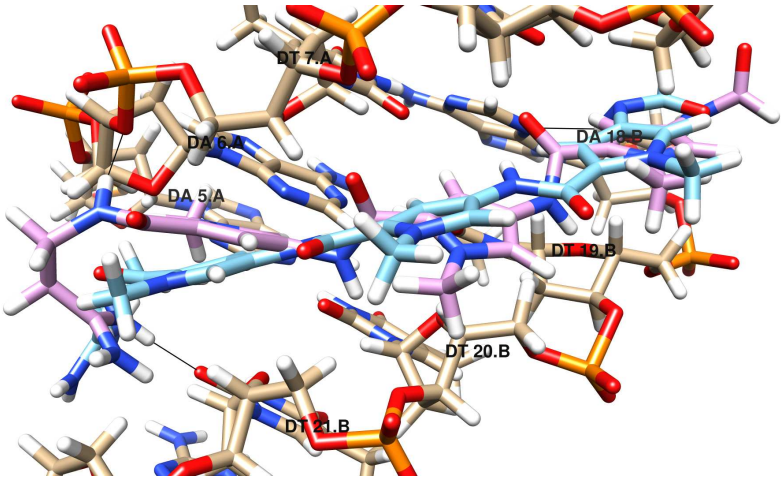

Figure S51: Binding site that corresponds to $\Delta G = -12.97$ kcal/mol for system 2DND. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 2.94 Å. The two ligands have some overlap and different hydrogen bonding patters. The ligand in the experimentally determined structure has a hydrogen bond with an adenine base (atom N3, residue 18), whereas the ligand in the predicted geometry has two hydrogen bonds. A hydrogen bond is formed with a thymine base (atom O2, residue 21), and another hydrogen bond is formed with a sugar (atom O3′, residue 6).

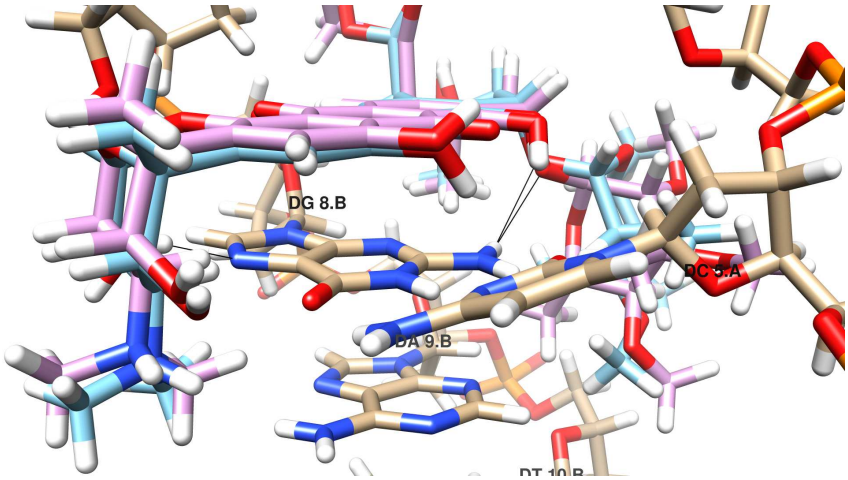

Figure S52: Binding site that corresponds to $\Delta G_{lowest} = -11.78$ kcal/mol for system 182D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.67 Å. The two ligands occupy the same intercalation site in the same orientation. The ligand in both geometries forms a hydrogen bond with a guanine base (atom N2, residue 8) in the minor groove, and the ligand in the predicted geometry forms an additional hydrogen bond with the same guanine base from the major groove (atom N7, residue 8).

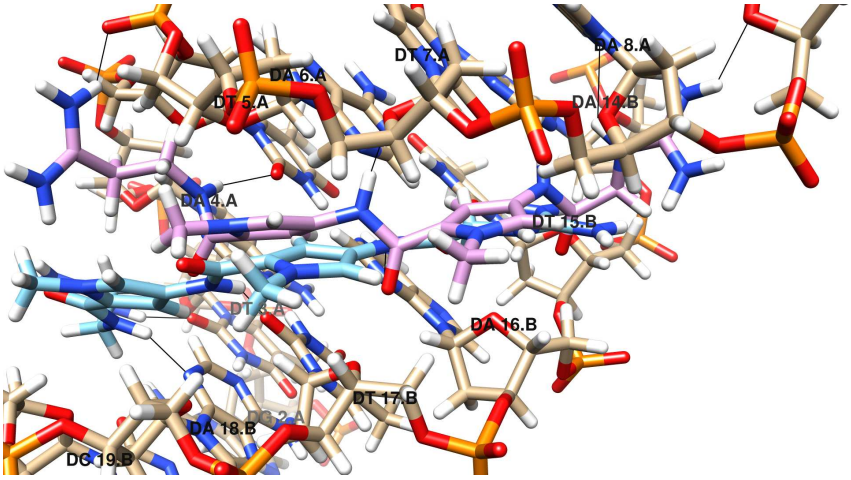

Figure S53: Binding site that corresponds to $\Delta G_{lowest} = -12.58$ kcal/mol for system 473D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 5.44 Å. The two ligands occupy the same minor groove site but with very different hydrogen bonding patterns. The ligand in the experimentally determined structure has two hydrogen bonds with two adenine bases (atom N3, residues 16 and 18), and three hydrogen bonds with three thymine bases (atom O2, residues 15, 17, 3). In contrast the ligand in the predicted geometry forms a hydrogen bond with an adenine base (atom N3, residue 8), with a thymine base (atom O2, residue 5), two hydrogen bonds with two sugars (atom O4′, residues 7 and 9), and a hydrogen bond with a phosphate group oxygen (residue 6).

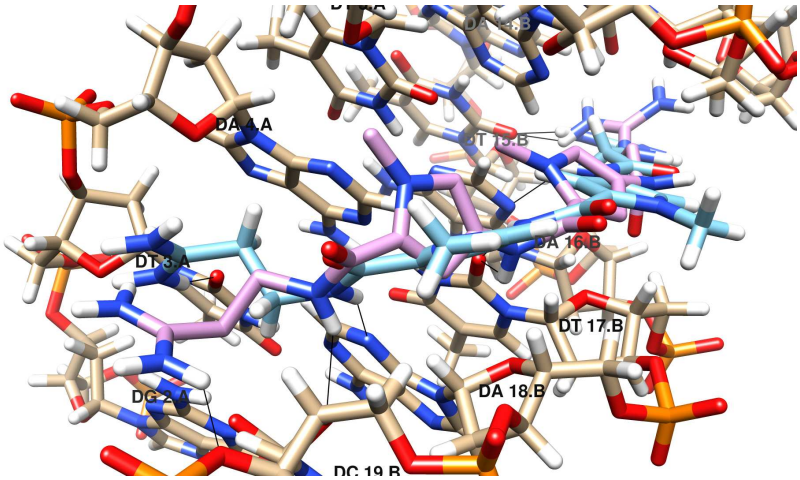

Figure S54: Binding site that corresponds to $\Delta G = -11.28$ kcal/mol for system 473D. Carbon atoms of the ligand at the experimentally determined geometry are shown with light blue, whereas carbon atoms at the predicted geometry are shown with light pink. The RMSD is 1.97 Å. The two ligands occupy the same minor groove site and have some common hydrogen bonding patters. The ligand in the experimentally determined structure has two hydrogen bonds with two adenine bases (atom N3, residues 16 and 18), and three hydrogen bonds with three thymine bases (atom O2, residues 15, 17, 3). The ligand in the predicted geometry forms two hydrogen bonds with two thymine bases (atom O2, residues 15 and 17), and three hydrogen bonds with two sugars (atoms O4′ and O3′, residues 16 and 19).

Table S32: DNA molecule residues and atoms that have hydrogen bonds with the ligand in the experimentally determined geometry and a predicted geometry with a relatively low RMSD value (other than the ΔG_{lowest} geometry). DNA atoms that participate in a hydrogen bond with the ligand in both the predicted and the experimentally determined geometry are shown in italics. If in addition, the hydrogen bond includes the same ligand atom then the DNA atoms are shown in bold. Ade: adenine, thy: thymine, gua: guanine, cyt: cytosine, dRib:deoxyribose, P: phosphate group.

	Experimental Geometry			Predicted Geometry (ΔG_{lowest})						
PDB ID	Residue	Unit	Atom(s)	Residue	Unit	Atom(s)	Residue	Unit	Atom(s)	RMSD
454D	4	Gua	**O6**	4	Gua	**O6**				0.98
	12	Gua	**O6**	12	Gua	**O6**				
1D30	6	Ade	N3	20	Thy	O2	19	*P*	O	0.87
	7	Thy	O2	19	Thy	O2				
432D	9	Gua	**N3**	9	Gua	**N3**	18	Thy	O2	0.53
	16	Ade	**N3**	16	Ade	**N3**				
	17	Thy	**O2**	17	Thy	**O2**				
127D				6	Ade	N3				0.57
128D	6	Ade	N3	19	*P*	O				2.17
	17	Ade	N3	19	Thy	O2				
264D	18	Ade	**N3**	18	Ade	**N3**	10	Gua	N3	0.73
	7	Thy	**O2**	7	Thy	**O2**				
2B3E	19	Thy	**O2**	19	Thy	**O2**	9	Cyt	O2	0.70
	20	Thy	**O2**	20	Thy	**O2**				
2DBE	18	Ade	N3	20	Thy	O2				1.02
1D63	6	Ade	**N3**	6	Ade	**N3**				0.73
	8	Thy	O2	9	dRib	O4′				
	20	Thy	O2	19	dRib	O4′				
1EEL	8	Thy	**O2**	8	Thy	**O2**				1.06
	20	Thy	O2	6	Ade	N3				
144D	20	Thy	O2	6	dRib	O3′				1.46
102D	20	Thy	O2	7	dRib	O4′	18	dRib	O4′	1.05
1PRP	6	Ade	N3	7	dRib	O4′	10	dRib	O4′	1.67
	17	Ade	N3	9	dRib	O3′	21	dRib	O4′	
2GB9	6	Gua	O6	6	Gua	N7	6	*P*	O	2.41
1D32	6	Gua	N7	6	Gua	O6				1.37
334D	16	Cyt	O2	7	dRib	O3′	8	*P*	O	5.11
	17	Cyt	O2	16	dRib	O3′	8	*P*	O	
	14	Gua	*N2-H*,N3	14	Gua	*N2-H*				
	15	Gua	*N2-H*,N3	15	Gua	*N2-H*				
2DND	18	Ade	N3	21	Thy	O2	6	dRib	O3′	2.94
473D	16	Ade	N3	19	dRib	O4′				1.97
	18	Ade	N3	19	dRib	O3′				
	3	Thy	O2	16	dRib	O4′				
	15	Thy	*O2*	15	Thy	*O2*				
	17	Thy	*O2*	17	Thy	*O2*				