CHAPTER 5

STRONG AND WEAK HYDROGEN BONDS IN DRUG–DNA COMPLEXES: A STATISTICAL ANALYSIS

5.1 Introduction

DNA sequences, both Watson-Crick and mismatched base pairs, can be recognized with a variety of compounds that bind in the DNA minor groove [5.1–5.5]. These compounds can be broadly classified as those that bind without a significant increase in the groove width of the free DNA, and those that bind with a broadening of the groove. The former class is exemplified by ligands containing aromatic rings and charged end-groups while the latter is typified by selective polyamide hairpin and related compounds like distamycin and netropsin [5.6]. The forces that dominate small molecule minor groove binding interactions are electrostatic, van der Waals, hydrophobic and hydrogen bonding. A crucial requirement for the ligand is a crescent shape that is complementary to the curvature of the minor groove.

Various aspects of minor groove drug–DNA recognition are revealed with the help of deposited X-ray structures in the protein databank (PDB) and the nucleic acid databank (NDB) [5.7]. Experimental, comparative structural and molecular modeling studies suggest that sequence specificity is often linked to key hydrogen bonds between the base pairs and the small molecule [5.8– 5.12].

The importance of weak C−H···O hydrogen bonds in macromolecules is a well-established phenomenon [5.13]. Their significance, as supporting interactions of stronger N−H···O and O−H···O bonds, in protein–ligand complexation have been described earlier [5.14, 5.15]. Based on the assumption that strong hydrogen bonding in drug–receptor interactions are thus inherently assisted by weak hydrogen bond [5.16], the aim of the present study is to analyze the importance of strong and weak hydrogen bonds in drug–DNA complexes.

The first section of the present study deals with a comparative analysis of strong and weak hydrogen bonds in a database of 70 drug–DNA complex crystal structures. The second part deals with molecular modeling applications of these results in a particular system. For the latter exercise, I have chosen a set of 26 furan derivatives targeted against Human

African Trypanosomes [5.17], *Trypanosoma brucei rhodesiense* (TBR) and *Trypanosoma brucei gambiense* (TBG). These organisms are responsible for Human African Trypanosomiasis (HAT) or sleeping sickness. The current drugs in the treatment of HAT are either toxic or difficult to use [5.18]. Only one drug for treating HAT is currently undergoing clinical trials [5.4, 5.19]. The orally available prodrug DB289 is converted systemically into another diamidine (DB75) that is active against the early-stage disease.

5.2 Materials and methods

5.2.1 *Drug–DNA complexes from PDB*

A set of 70 unique minor groove drug–DNA complexes was obtained from the PDB (Table 5.1). The drug molecules present in these cases are berenil (3 complexes), furans and thiophenes (8), pentamidines (5), netropsin (12), distamycin (5), Hoechst drugs (16), benzimidazoles (14), DAPI (2), polyamide (3), pyridines (2). The therapeutic uses of the above mentioned drugs are listed in Table 5.2. H–atoms were added to the drug–DNA complexes, and then subjected to minimization keeping the heavy atoms rigid. This was carried out in MOE with the MMFFx force field [5.20, 5.21]. The MOE optimized geometries were analyzed with the hydrogen bond analysis program, HBAT.

Table 5.1: Minor groove drug–DNA complexes in this study: (a) berenil 1-3, (b) furans 4-11, (c) pentamidines 12-16, (d) netropsin 17-28, (e) distamycin 29-33, (f) Hoechst drugs 34-49, (g) benzimidazoles 50-63, (h) DAPI 64-65, (i) polyamides 66-68, (j) pyridines 69-70.

Sl. No.	PDB ID	NDB ID	Resolution	PUBMED ID	Hetero group
1	1D63	GDL016	2.0	1640462	BRN
2	2DBE	GDL009	2.5	2323343	BRN
3	268D	GDLB42	2.0	n/a	BRN
4	289D	GDL045	2.2	8917643	D19
5	1VZK	-	1.7	n/a	D1B
6	1EEL	-	2.4	11128631	D24
7	1FMQ	DD0034	2.0	11128631	D34
8	1FMS	DD0035	1.9	11128631	D35
9	360D	GDL056	1.8	9611230	BPF
10	227D	GDL036	2.2	8639524	BGF
11	298D	GDL044	2.2	8917643	D18
12	1M6F	DD0052	1.7	12431090	CGQ
13	166D	GDL027	2.2	7813486	PET
14	1D64	GDL015	2.1	1643044	PNT

15	102D	GDL032	2.2	7608897	TNT
16	1PRP	GDL023	2.1	8268158	TNT
17	101D	GDLB31	2.2	7711020	NT
18	121D	GDL014	2.2	8395202	NT
19	195D	GDL030	2.3	n/a	NT
20	1D85	GDLB17	2.5	1332773	NT
21	1D86	GDL018	2.2	1332773	NT
22	1DNE	GDL004	2.4	2539859	NT
23	1DVL	DD0024	2.4	11914483	NT
24	261D	GDJ046	2.4	9125500	NT
25	358D	GDJB55	2.5	9826773	NT
26	375D	GDJ059	2.4	9826773	NT
27	474D	GDJB58	2.4	9826773	NT
28	6BNA	GDLB05	2.2	2991536	NT
29	267D	GDLB41	2.0	n/a	DMY
30	1JTL	DD0042	1.9	12071949	DMY
31	1K2Z	DD0046	2.4	12071949	DMY
32	2DND	GDL003	2.2	3479798	DMY
33	378D	GDH060	2.4	10089456	DMY
34	127D	GDL022	2.0	1371249	HT
35	128D	GDLB19	2.5	1371249	HT
36	129D	GDL021	2.2	1371249	HT
37	130D	GDLB20	2.5	1371249	HT
38	1D43	GDL010	2.0	1718416	HT
39	1D44	GDL011	2.0	1718416	HT
40	1D45	GDL012	1.9	1718416	HT
41	1D46	GDL013	2.0	1718416	HT
42	1DNH	GDL002	2.2	2452403	HT
43	264D	GDL026	2.4	7517864	HT
44	296D	GDL028	2.2	7515488	HT
45	8BNA	GDL006	2.2	2445998	HT
46	311D	GDL052	2.2	9162901	HT2
47	303D	GDL048	2.2	9017011	RO2
48	302D	GDL047	2.2	9017011	RO2
49	447D	DD0007	2.2	10666470	BBZ
50	263D	GDL039	2.2	8901516	TBZ
51	459D	DD0014	2.3	10373586	TBZ
52	1FTD	-	2.0	11170623	E97
53	403D	BDD001	1.4	9692982	HT1
54	443D	DD0004	1.6	10666470	IA
55	445D	DD0006	2.6	10666470	IA
56	449D	DD0009	2.1	10666470	IA
57	442D	DD0003	1.6	10666470	IB
58	444D	DD0005	2.4	10666470	IB
59	448D	DD0008	2.2	10666470	IB
60	109D	GDL033	2.0	n/a	IBB
61	1LEX	GDL037	2.2	8527438	IPL
62	1LEY	GDL038	2.2	8527438	IPL
63	269D	GDLB43	2.1	n/a	HT
64	1D30	GDL008	2.4	2627296	DAP
65	432D	DD0002	1.9	10600105	DAP
66	1CVX	DD0020	2.2	10623546	HP2
67	1CVY	DD0021	2.1	10623546	IPY
68	408D	BDD003	2.1	9756473	IPY
69	144D	GDLB24	2.2	8373768	SN6
70	328D	GDL053	2.6	9321660	SN7

Table 5.2: Minor groove binders (MGB) and their respective therapeutic use.

Sl. No.	Drug	Therapeutic uses
1	Berenil	Antiprotozoal agents, Intercalating agents Trypanocidal agents
2	Furans	Antitrypanosomal, *Pneumocystis carinii* and *Ctyptosporidium parvum*
3	Pentamidines	African trypanosomiasis and leishmaniasis, AIDS-related *Pneumocystis carinii* pneumonia
4	Netropsin	Filaricides, Antiviral agents, Anticancer agents
5	Distamycin	Anticancer and Antiviral agents
6	Hoechst drugs	Filaricides and Antifungal agent
7	Benzimidazoles	Antiviral
8	DAPI	Trypanocidal Agents
9	Polyamide	Anticancer
10	Pyridines	Anticancer

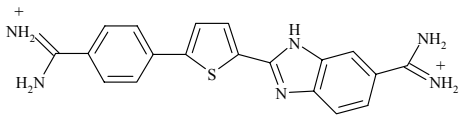
5.2.2 *Hydrogen bond analysis tool (HBAT)*

Strong and weak hydrogen bonds were analyzed with the HBAT software [5.22]. HBAT, analyzes and tabulates all hydrogen bonds present in a PDB file. The output file provides distance-angle distributions across various geometry ranges while tabulation of frequencies for each residue, ligand, water, and nucleic acids is done easily for any kind of interaction. It is a user-friendly desktop tool, which operates both with default and user-selected parameters. The standard H-bonding criteria were set as d (H···A) ≤ 2.8 Å and θ (X−H···A) ≥ 90°.

5.2.3 *Docking of HAT inhibitors*

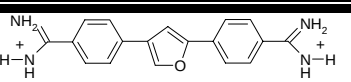
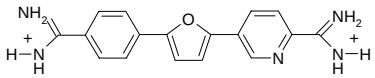
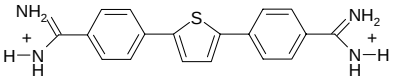
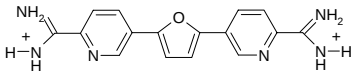
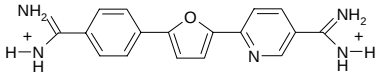
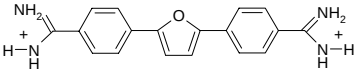
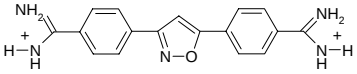
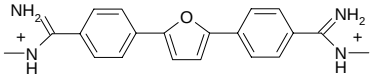
Molecule building, geometry optimizations and minimizations were carried out using the MMFFx force field in the MOE software. Docking was carried out with GOLD 3.0 [5.24]. The amidinium inhibitors (26 compounds, Table 5.3) were selected from the literature [5.17]. Of these, seven (furan-based compounds) are present in the database of 70 experimental crystal structures analyzed in the first section of this chapter. Docking was performed for each of these 26 inhibitors into the DNA in eight experimental crystal structures, in other words a total of 208 docking experiments were performed. Seven of these structures correspond to the seven furans mentioned above; the eighth corresponds to a thiophene-based inhibitor in structure PDB ID-1VZK (Scheme 5.1), which has been shown to be a good choice for docking for minor groove binders [5.23]. This eighth structure is also contained in the earlier database of 70 structures. Significantly, the correlations

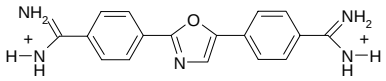
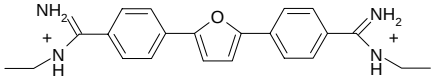
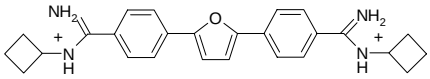
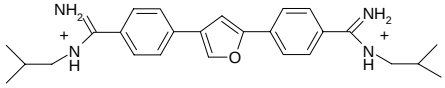
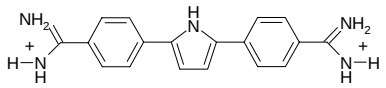
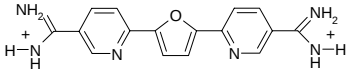
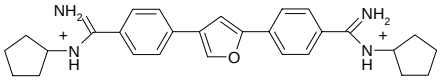
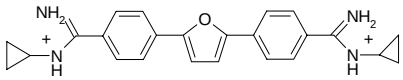
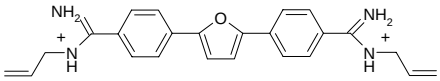
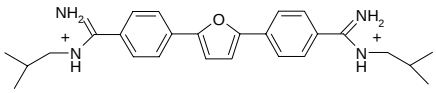
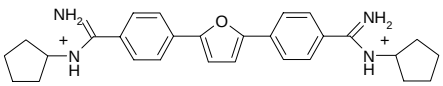
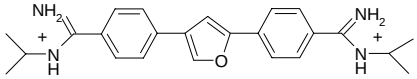
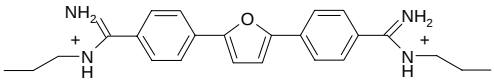
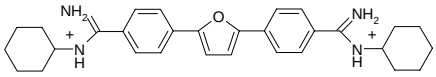
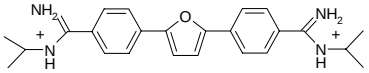
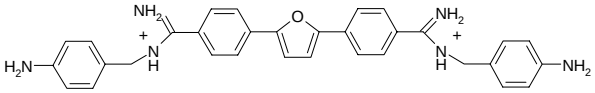
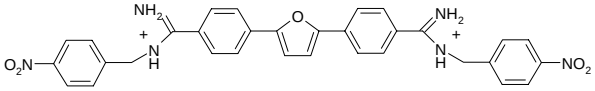
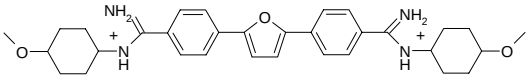
observed for the docking geometries obtained from this eighth (thiophene-based) structure were significantly better than the others, and accordingly only results from this set of 26 docking experiments are discussed further. A region of 7.0 Å radius around the ligand was defined as the active site for each drug–DNA complex. Default-set parameters were used in the docking. For each of the 10 independent Genetic Algorithm (GA) runs, with a selection pressure 1.1, 100,000 GA operations were performed on a set of 5 islands. The population size of 200 individuals was specified. Default operator weights were used for crossover, mutation, and migration of 95, 95 and 10 respectively. To further speed up the calculation, the GA docking was stopped when the top three solutions were within 1.5 Å RMSD of each other. The interaction analysis was carried out with HBAT as described above.



Scheme 5.1

Table 5.3: Amidinium based HAT inhibitors used in docking and their activities.

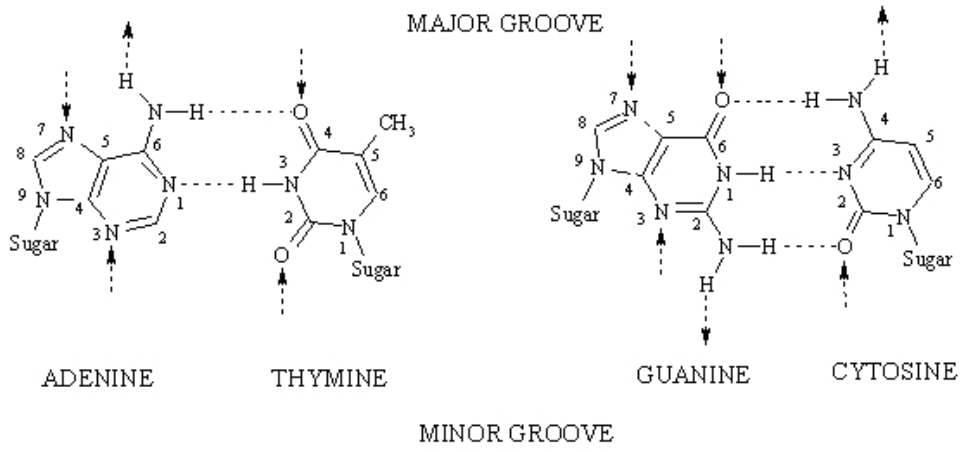
Sl. No.	Code	Structure	pIC_{50}	IC_{50} (in μM)
1	DB690		8.79	0.0016
2	DB867		8.69	0.0020
3	DB351		8.51	0.00302
4	DB994		8.39	0.0040
5	DB820		8.31	0.0048
6	DB75		8.29	0.0051
7	1RJL164		8.28	0.0052
8	DB417		8.18	0.0066

9	DB484		8.15	0.0070
10	DB427		8	0.0100
11	DB313		7.97	0.0107
12	DB518		7.86	0.0138
13	DB262		7.85	0.0141
14	DB829		7.77	0.0169
15	DB481		7.72	0.0190
16	DB193		7.63	0.0234
17	DB312		7.63	0.0234
18	DB235		7.45	0.0354
19	DB244		7.45	0.0354
20	DB480		7.41	0.0389
21	DB240		7.35	0.0446
22	DB249		7.17	0.0676
23	DB181		7.16	0.0691
24	DB422		7.01	0.0977
25	DB421		6.99	0.1023
26	DB568		6.77	0.1698

5.3 Results and discussion

5.3.1 *Overview of strong and weak hydrogen bond in drug–DNA complexes*

DNA contains a variety of strong and weak hydrogen bond functional groups (Scheme 5.2). These groups are evenly exposed in the minor and major grooves. The ratio of the number of exposed donors to acceptors for AT and GC pairs in the major groove is 1:2, while it is 1:2 for GC and 0:2 for AT in the minor groove. The possible acceptors present in the major groove are [A{N(7)}], [G{N(7)}], [G{N(6)}] while donors belong to the amino groups in the [A{N(6)}] and [C{N(4)}] positions. In the minor groove, the acceptors are [A{N(3)}], [G{N(3)}], [T{O(2)}], [C{O(2)}] and the donor is the amino group of [G{N(2)}]. Several proteins of functional importance bind to the major groove, whereas non-covalently binding drugs bind to the minor groove. As discussed earlier, the minor groove binders (MGB) selectively bind to the AT rich region. Therefore acceptors present in the minor groove, notably N(3) of purine and C(2)=O of pyrimidine, are important sites for drug–DNA interaction. Apart from nucleotides, phosphate groups and the ribose O-atoms are also acceptors.



Scheme 5.2: Strong and weak hydrogen bond functional groups in DNA.

5.3.2 *Hydrogen bond analysis*

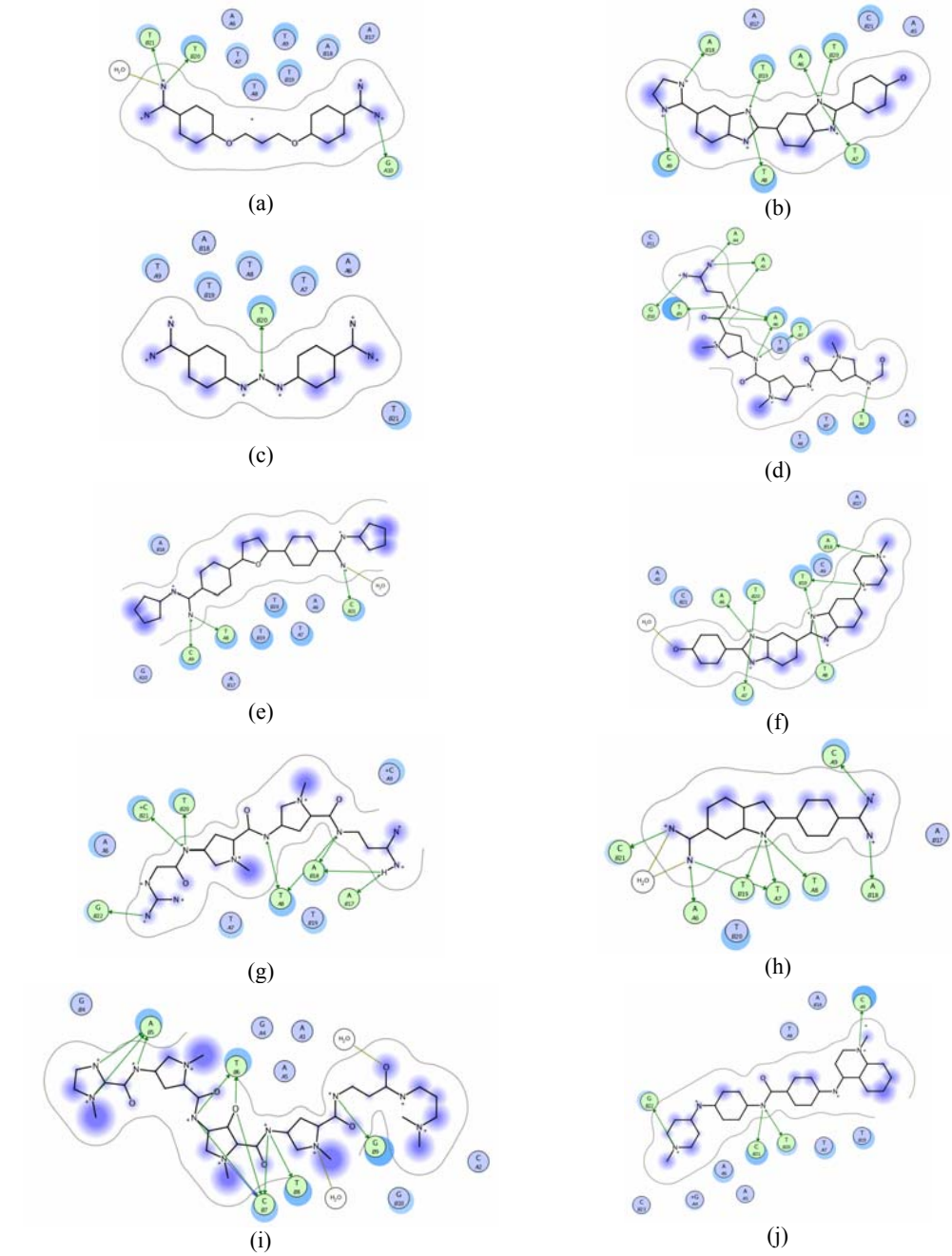
Both strong and weak hydrogen bonds are observed in the minor groove of drug–DNA complexes. As expected N(3) of purine and O(2) of pyrimidine are important (Table 5.4). The exceptions are polyamides and DAPI drugs, which interact with adenine N(3), guanine N(3), thymine O(2) and cytosine O(2). A schematic diagram for a typical inhibitor

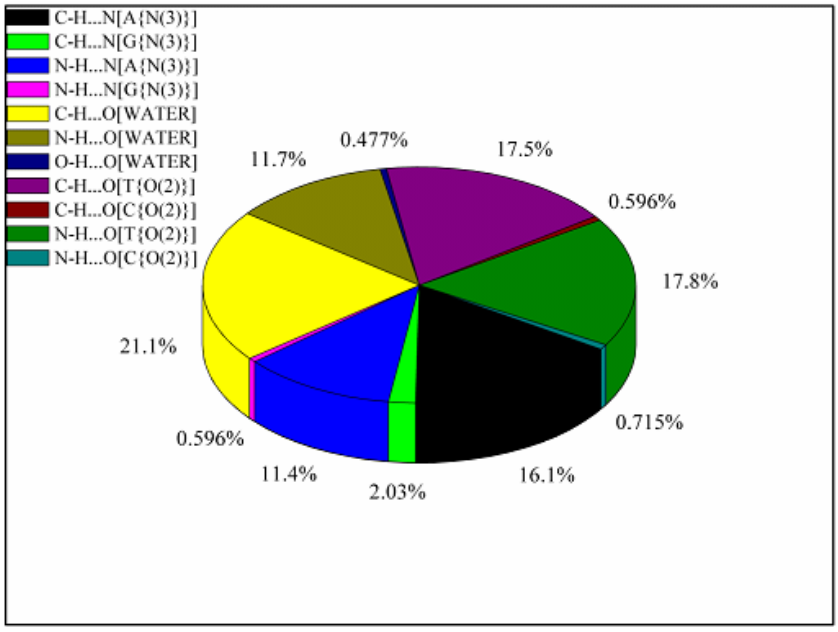
for each class is shown in Figure 5.1. The types of strong and weak hydrogen bonds observed in the case of purine is N−H···N(3) and C−H···N(3), and for pyrimidine it is N−H···O(2) and C−H···O(2). Also observed clearly are N−H···Ow and C−H···Ow interactions between the drug and water. The distribution of strong and weak hydrogen bond types in a total of 835 interactions are: (a) purine C−H···N(3) 18%, N−H···N(3) 12%, (b) pyrimidine C−H···O(2) 18%, N−H···O(2) 18%, (c) C−H···Ow 21%, N−H···Ow 12% (Figure 5.2*a*). The total number of C−H···N and C−H···O hydrogen bonds taken together is 481, while for N−H···N and N−H···O the combined total is 354. The present dataset contains 70 drug molecules. Thus on average, each drug molecule interacts with DNA through seven weak interactions and five strong interactions in the minor groove, which is effectively 1.4 weak hydrogen bonds for each strong hydrogen bond. Apart from nucleotides, drug molecules also interact with the sugar moiety, but they rarely interact with the phosphate group. The number of interactions observed between drug molecules and the deoxyribose sugar and phosphates taken together are 118. Among these interactions, deoxyribose sugar O4’ of adenine and thymine have more hydrogen bonds (Figure 5.2*b*).

Table 5.4: Strong and weak hydrogen bonds in drug–DNA complexes.

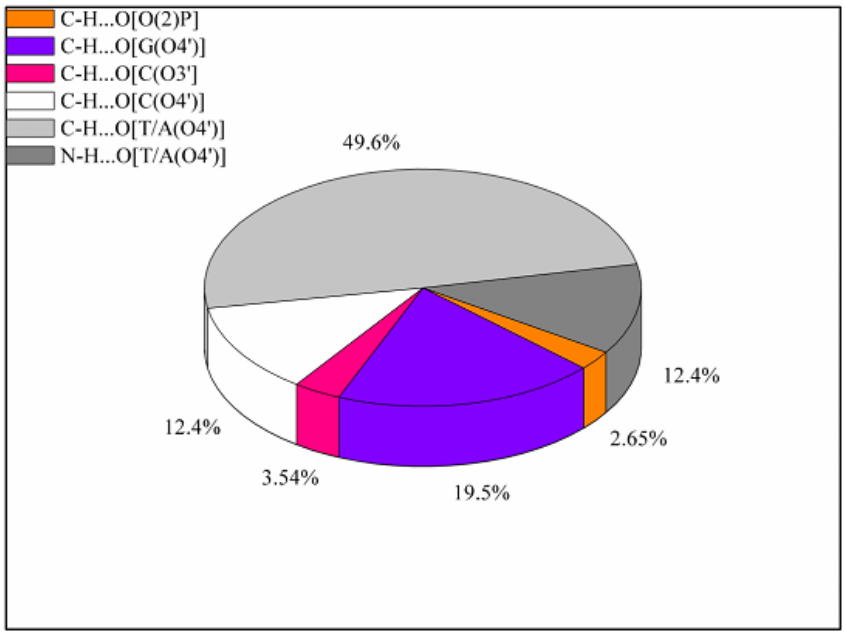
	C−H···N [A{N(3)}]	C−H···N [G{N(3)}]	N−H···N [A{N(3)}]	N−H···N [G{N(3)}]	C−H···O (WATER)	N−H···O (WATER)	O−H···O (WATER)	C−H···O [T{O(2)}]	C−H···O [C{O(2)}]	N−H···O [T{O(2)}]	N−H···O [C{O(2)}]
Berenil	5		6		3	10		5		4	
Furans	15		3		44	18		14		14	
Pentamidines	14		6		8	11		18		3	
Netropsin	12		15	2	12	20		17		18	
Distamycin	14		19		10	8		11		16	
Hoechst drugs	33		15		45	13	3	39		34	
Benzimidazoles	29	2	20		19	8		33		40	
DAPI	1	1	2	1	1	5		3		5	
Polyamides	3	14	10	2	28	4	1		5	13	6
Pyridines	9				7	1		7		2	
TOTAL	135	17	96	5	177	98	4	147	5	149	6

Figure 5.1: Schematic representation of MGBs in the minor groove. Strong hydrogen bonds are shown with green arrows, exposed ligand atoms in blue contours, and the nearest nucleotides as circles: (a) pentamidine, (b) benzimidazoles, (c) berenil, (d) distamycin, (e) furan, (f) Hoechst drugs, (g) netropsin, (h) DAPI, (i) polyamide, (j) pyridine.





(a)



(b)

Figure 5.2: Distribution of hydrogen bonds in drug–DNA complexes. Interactions of drug with (a) nucleotides and water, (b) sugar and phosphate.

5.3.3 *Donor furcation*

The specificity and affinity of minor groove-binding agents are sensitive to the local width of the groove. In AT-rich sequences, the minor groove is unusually narrow (0.3-0.4 nm). Due to the narrow width, a donor present in a drug molecule possibly is hydrogen bonded to more than one acceptor on the wall of the groove. This type of situation, where a donor can interact with several acceptors simultaneously or an acceptor can interact simultaneously with many donors is termed furcation. The most frequently observed furcated interactions, in this study, involve bifurcated and trifurcated donors. Figure 5.3 shows examples of bi- and trifurcated donors in the DB244–DNA complex in PDB ID 1EEL. Typically, a single donor from the ligand is shared by two or three acceptors of the minor groove. Table 5.5 shows a distribution of strong and weak hydrogen bonds at various levels of furcation. With an increase in furcation level, the numbers of weak hydrogen bonds increase. The numbers of donor-furcated cases are: (a) bifurcated, 205; (b) trifurcated, 41; (c) tetrafurcated, 7; (d) pentafurcated, 2.

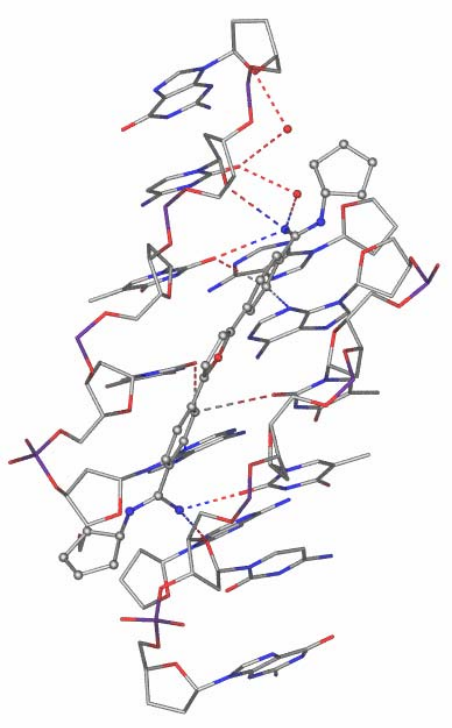


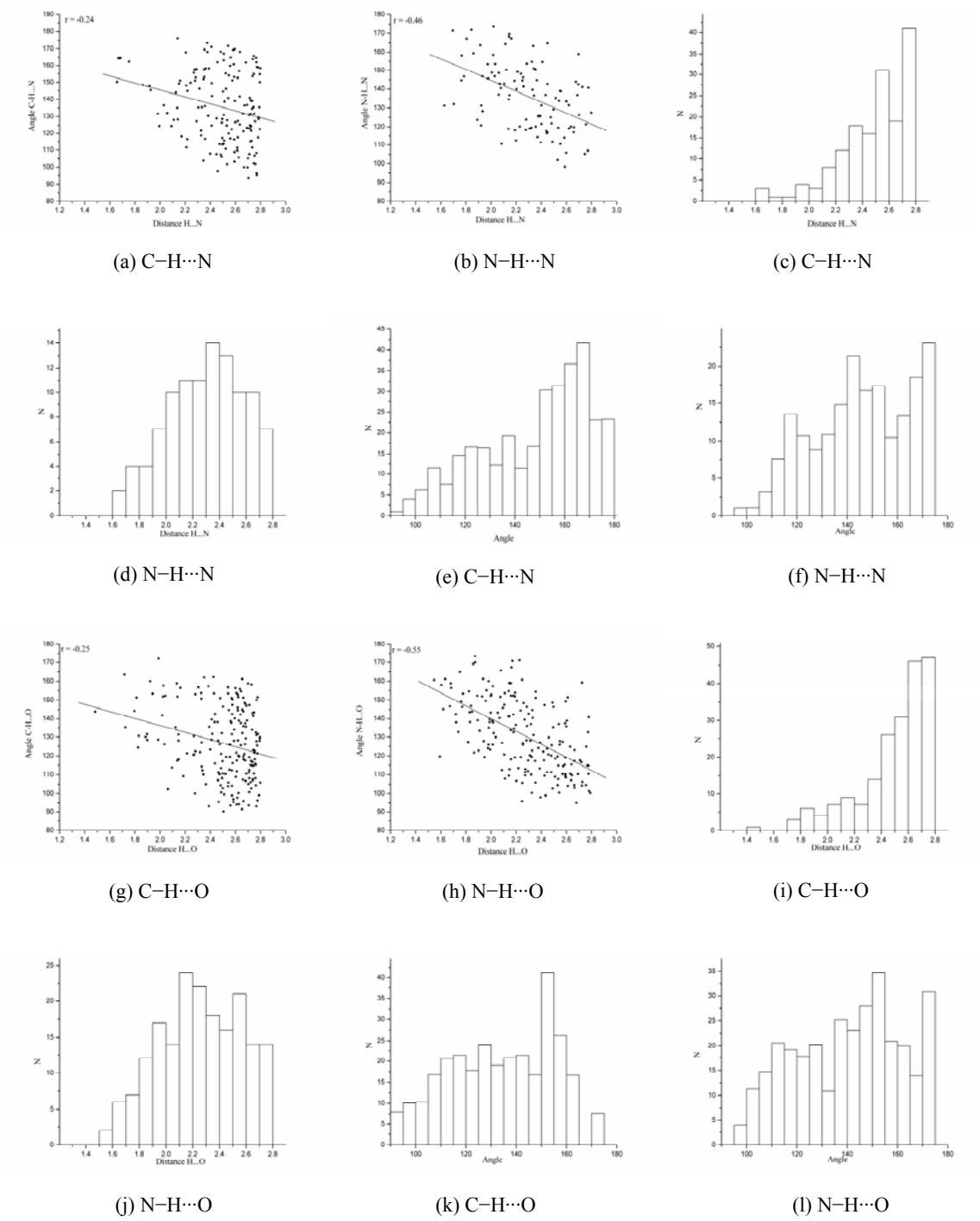
Figure 5.3: Bifurcated and trifurcated hydrogen bonds in DB244–DNA complex, PDB ID 1EEL.

Table 5.5: Strong and weak hydrogen bonds at various levels of donor furcation.

	C−H···N	N−H···N	C−H···O	N−H···O
Furcation level				
Bifurcated	84	69	216	141
Trifurcated	20	20	52	31
Tetrafurcated	6	5	10	7
Pentafurcated	6	-	4	-

5.3.4 *Hydrogen bond geometry*

For strong hydrogen bonds of the type N−H···N and N−H···O, the median H···N/H···O distances, *d*, are less than 2.4 Å and 2.2 Å respectively (Figure 5.4). The cone-corrected angular distributions for N−H···N and N−H···O are similar with maxima in the range 170-175^{o}. The inverse length-angle correlations are also well-behaved in these cases. Strong N−H···N and N−H···O hydrogen bonds shows better linearity compared to weak C−H···N and C−H···O hydrogen bonds. The weak C−H···N, C−H···O hydrogen bonds have variable geometry. The cone-corrected angular distributions of C−H···N and C−H···O interactions, have maxima at 165-170º and 150-155º respectively. To summarize, the geometries for strong and weak hydrogen bonds observed in the drug–DNA complexes are consistent and fall within acceptable limits.



(a) C−H···N (b) N−H···N (c) C−H···N

(d) N−H···N (e) C−H···N (f) N−H···N

(g) C−H···O (h) N−H···O (i) C−H···O

(j) N−H···O (k) C−H···O (l) N−H···O

Figure 5.4: Hydrogen bond geometry for purine (a)-(f) and pyrimidine (g)-(l) acceptors. In each case, the inverse length-angle scatterplot is followed by histograms of distances and cone-corrected angular distributions.

5.3.5 *Human African Trypanosomiasis (Docking)*

Molecular docking is a very useful technique in rational drug design. However, the purpose of using docking in the present context is different. Molecular docking was carried out to estimate the best drug–DNA geometries in cases where the crystal structures of the complex is unknown, assuming that strong and weak hydrogen bonds are optimized. Accordingly, the best geometry was selected (poses) for such complexes to obtain virtual crystal geometries for the 26 selected HAT inhibitors. The inhibitors were docked separately to each of the 8 drug–DNA complexes of the amidinium category. Each docking run was evaluated through regression analysis of experimental activities *versus* docking scores. The best correlation r = 0.83 was obtained in the case of PDB ID 1VZK, where molecules with higher activity show greater docking scores (Figure 5.5) showing that the assumption regarding optimization of hydrogen bonds is a valid one. The best poses for the inhibitors were exported to HBAT for hydrogen bond analysis. The docking solutions of molecule **DB690** and **DB568** are shown in Figure 5.6.

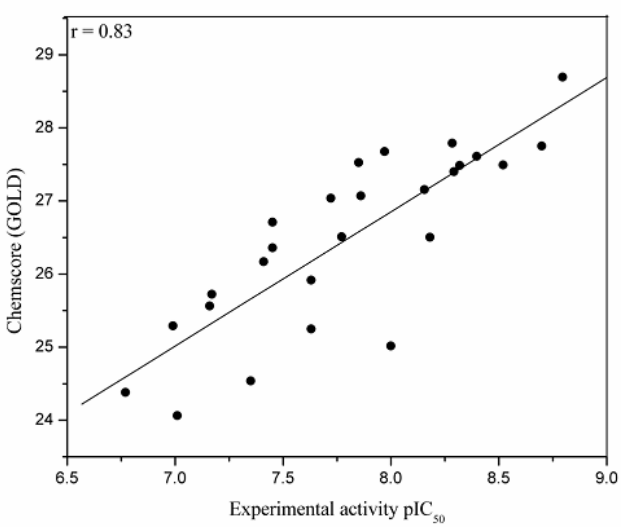
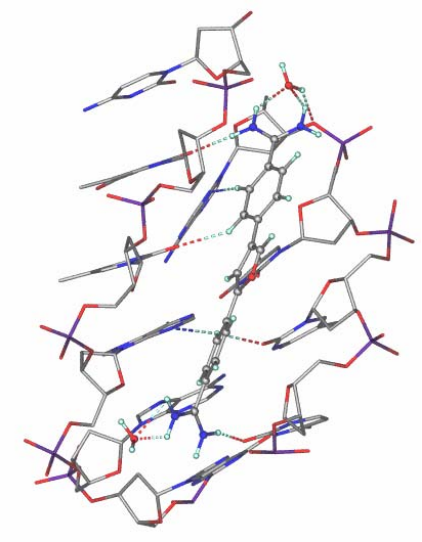


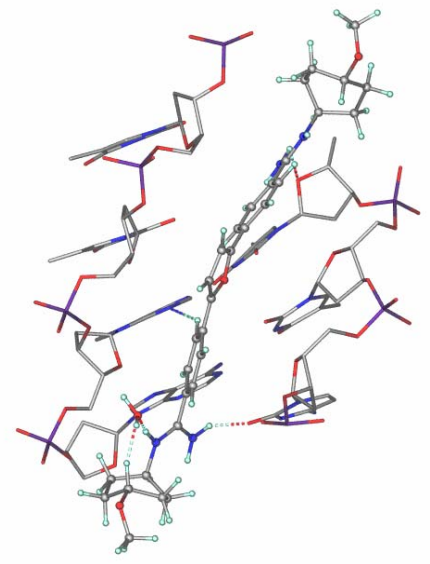
Figure 5.5: Fitness scores versus biological activities of HAT inhibitors.

Molecule **DB690** interacts with DNA through six weak and five strong hydrogen bonds. At one end, one amide group of the drug molecule is hydrogen bonded to T(O2). Further, this amide is also hydrogen bonded to a water molecule. The water molecule is again hydrogen bonded to the phosphate group present at the opposite strand. The other

amide group is hydrogen bonded to C(O2). Between these two ends, the weak C−H···N and C−H···O hydrogen bonds are present (Figure 6*a*). The inactive molecule, **DB568** binds to DNA with six weak and four strong interactions (Figure 5.6*b*). The interactions observed between molecule **DB690**, **DB568** and DNA are listed in Table 5.6.



(a)



(b)

Figure 5.6: Best docking solutions of (a) molecule **DB690** and (b) molecule **DB568** in the minor grove.

Table 5.6: (a) Hydrogen bonds in the molecule DB690–DNA complex.

Type	Acceptor	Acceptor atom	Residue number	*d* (H···A)	*D* (X···A)	θ (X−H···A)
C−H···N	A	N3	6	2.303	3.382	171.8
C−H···N	A	N3	18	2.602	3.378	127.5
C−H···O	T	O2	19	2.718	3.28	111.9
C−H···O	T	O2	19	2.449	3.284	132.4
C−H···O	T	O2	20	2.693	3.098	101.4
C−H···O	T	O2	8	2.43	3.053	114.8
N−H···O	C	O2	9	2.117	3.097	162.7
N−H···O	HOH	O	24	2.155	2.994	139.1
N−H···O	HOH	O	24	2.056	2.925	142.6
N−H···O	HOH	O	25	1.51	2.163	116.7
N−H···O	T	O2	20	2.135	3.085	155.7

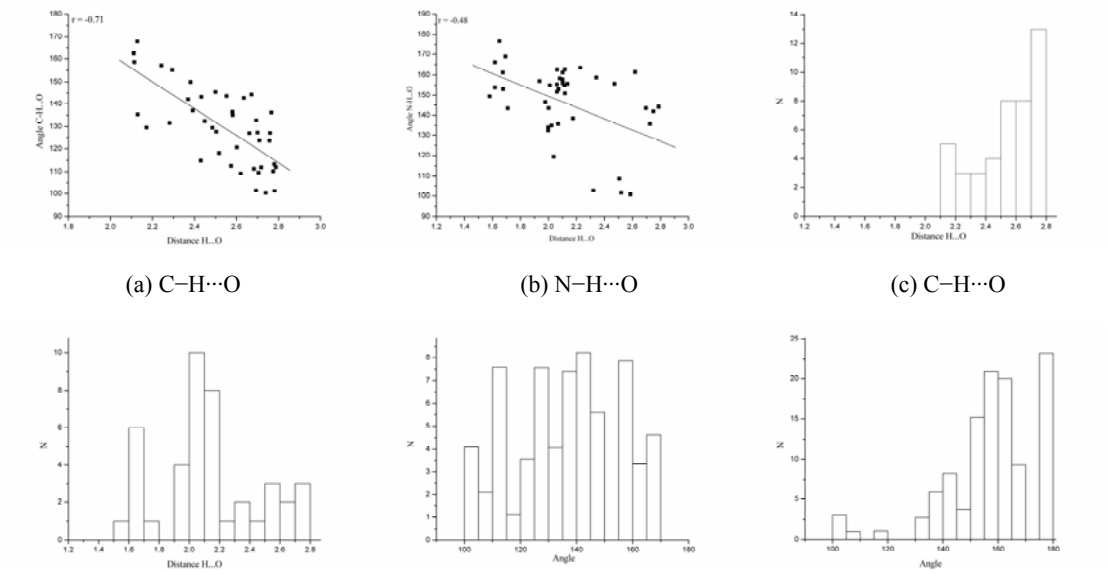
Table 5.6: (b) Hydrogen bonds in the molecule DB568–DNA complex.

Type	Acceptor	Acceptor atom	Residue number	*d* (H···A)	*D* (X···A)	θ (X−H···A)
C−H···N	A	N3	18	2.775	3.654	137.5
C−H···O	A	O1P	18	2.767	3.299	109.4
C−H···O	A	O4’	18	2.694	3.071	99.67
C−H···O	HOH	O	25	2.076	3.122	158.1
C−H···O	T	O1P	8	2.599	3.222	115.1
C−H···O	T	O2	19	2.683	3.241	111.2
N−H···N	G	N2	16	2.667	3.146	108.8
N−H···O	C	O2	9	1.709	2.592	143.4
N−H···O	HOH	O	25	1.82	2.807	163.5
N−H···O	T	O4’	7	2.403	2.836	104.6

5.3.6 *Hydrogen bonds in HAT inhibitors*

The hydrogen bond geometry for C−H···O(2) and N−H···O (2) pyrimidine acceptors are discussed here as a representatives of strong and weak hydrogen bonds. Figure 5.7 shows the C−H···O and N−H···O hydrogen bond geometries which involve DNA pyrimidine for the 26 selected HAT inhibitors. For N−H···O and C−H···O hydrogen bonds, the median H···O distances, *d* are less than 2.1 Å and 2.8 Å respectively. This is normal. The inverse length-angle correlations are also consistent for both types of hydrogen bonds. The cone-corrected angular distribution for N−H···O has a maximum at 175-180º. For C−H···O interactions, the maximum lies at 140-145º. This too is as expected. Similarly, the hydrogen

bond geometries obtained for both X-ray structures and the virtual complexes are consistent. Docking has produced a set of reasonable drug–DNA recognition geometries in the 19 cases where no crystal structure is available (virtual crystal structures).



(a) C−H···O (b) N−H···O (c) C−H···O

(d) N−H···O (e) C−H···O (f) N−H···O

Figure 5.7: C−H···O and N−H···O hydrogen bond geometries for 26 amidinium derivatives. In all cases, pyrimidine O(2) is the acceptor.

5.4 Conclusions

A statistical analysis of strong and weak hydrogen bonds in the minor groove of DNA is presented for a set of 70 drug–DNA complexes. The dataset was extracted from the PDB. The analysis was performed with the HBAT software. Both strong and weak hydrogen bonds are implicated in molecular recognition. On an average, there are 1.4 weak hydrogen bonds for every strong hydrogen bond. For both categories of interaction, the N(3) of purine and the O(2) of pyrimidine are favoured acceptors. Donor multifurcation is common with the donors generally present in the drug molecules, and being shared by hydrogen bond acceptors in the minor groove. Bifurcation and trifurcation are most commonly observed. The metrics for strong hydrogen bonds are consistent with established trends. For weak hydrogen bonds the geometries are variable. A database of recognition geometries for 26

literature amidinium-based inhibitors of HAT was generated with a docking study using seven inhibitors which occur in published crystal structures included in the list of 70 complexes mentioned above, and 19 inhibitors for which the drug–DNA complex crystal structures are unknown. The virtual geometries so generated correlate well with published activities for these 26 inhibitors.