Available online at www.sciencedirect.com
SCIENCE DIRECT®
Archives of Biochemistry and Biophysics 429 (2004) 81–99
ABB
www.elsevier.com/locate/yabbi

A molecular thermodynamic view of DNA–drug interactions: a case study of 25 minor-groove binders

Saher Afshan Shaikh, Sajeedha Reshmi Ahmed, and B. Jayaram*

Department of Chemistry, Indian Institute of Technology, Hauz Khas, New Delhi 110016, India

Received 23 April 2004, and in revised form 27 May 2004
Available online 26 June 2004

Abstract

Developing a molecular view of the thermodynamics of DNA recognition is essential to the design of ligands for regulating gene expression. In a first comprehensive attempt at sketching an atlas of DNA–drug energetics, we present here a detailed thermodynamic view of minor-groove recognition by small molecules via a computational study on 25 DNA–drug complexes. The studies are configured in the MMGBSA (Molecular Mechanics-Generalized Born-Solvent Accessibility) framework at the current state of the art and facilitate a structure–energy component correlation. Analyses were conducted on both energy minimized structures of DNA–drug complexes and molecular dynamics trajectories developed for the purpose of this study. While highlighting the favorable role of packing, shape complementarity, and van der Waals and hydrophobic interactions of the drugs in the minor groove in conformity with experiment, the studies reveal an interesting annihilation of favorable electrostatics by desolvation. Structural modifications attempted on the ligands point to the requisite physico-chemical factors for obtaining improved binding energies. Hydrogen bonds predicted to be important for specificity based on structural considerations do not always turn out to be significant to binding in post facto analyses of molecular dynamics trajectories, which treat thermal averaging, solvent, and counterion effects rigorously. The strength of the hydrogen bonds retained between the DNA and drug during the molecular dynamics simulations is ~1 kcal/mol. Overall, the study reveals the compensatory nature of the diverse binding free energy components, possible threshold limits for some of these properties, and the availability of a computationally viable free energy methodology which could be of value in drug-design endeavors.

Keywords: DNA–drug interactions; Free energy component analysis; A/T specificity; Shape complementarity; Hydrogen bonds; Molecular dynamics; Binding affinity

Regions of DNA involved in vital processes such as origin of replication, promotion of transcription, etc., are of particular interest as targets for a wide range of anticancer and antibiotic drugs [1–4]. Significant efforts [2,5,6] have been made to design small molecules with structural and chemical features that would allow them to recognize the properties of DNA sequences and bind to them with requisite sequence specificity and binding affinity, such that a competition with regulatory proteins could ensue. These drug-design endeavors could be greatly expedited and made more productive, if supplemented by a technique that is predictive in nature. Our aim was to develop such a theoretical model based on an atomic level description, which would eventually assist and complement the rational drug-design attempts. As a step towards this goal, we put together a computational pathway and applied it to a systematic characterization of the energetic and structural features that facilitate non-covalent binding of small molecules to DNA.

Two major problems arise in developing a theoretical account of DNA–ligand binding in particular and thermodynamics of DNA recognition in general. Modeling DNA at atomic level in aqueous solutions with counterions and added salt poses several challenges to theory, a resolution of which has become feasible only in recent years [7]. Second, computing binding free energies and their components for a large biomolecular system

* Corresponding author. Fax: +91-11-2658-2037.
E-mail address: bjayaram@chemistry.iitd.ac.in (B. Jayaram).

0003-9861/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2004.05.019

within the rigors of statistical mechanics while maintaining computational tractability is an equally formidable task [8–11]. In this paper, we present a computational protocol for a semi-quantitative estimation of DNA–drug binding free energies treating the systems as realistically as possible at the current state of art, and examine the energy components favoring binding in a number of DNA–drug complexes with a view to arriving at a set of common principles and a consensus view of DNA binding.

Energetics of DNA recognition is not fully understood at a molecular level, due mainly to the strong electrostatic interactions prevalent in the system arising due to phosphates, mobile counterions, and hydration, superposed on potential hydrogen bonds and van der Waals interactions [7,12] besides the energetic consequences of sequence dependent structural adaptation of DNA and hydrophobic effect. A detailed enumeration and quantification of these contributions can help in addressing issues of both practical and fundamental interest [13]. The fundamental interest lies in tackling challenges involved in predicting ligand binding free energies qualitatively/semi-quantitatively. The practical interest lies in designing/modifying DNA binding small molecules to achieve expected ligand binding affinities. It is thus desirable that a binding energy component database be developed for non-covalent associations in general and for DNA–drug systems in particular via experiment or theory or both. The present study is an attempt in this direction from the standpoint of theory. The task is facilitated by the fact that structures of many DNA–drug complexes are now known at high resolution [14–33].

Both from structural and energetic angles, binding of small molecules to DNA and proteins differs significantly. Protein (enzyme)–drug binding has been explained by various popular models such as the lock and key model, induced fit model [34], etc., and it is also believed that hydrophobic effects play an important role in the binding process. However, a straightforward extrapolation of these interaction models to DNA–drug systems is not feasible since there is no formal active site in DNA, unlike enzymes. Certain base sequence dependent chemical, structural, and conformational characteristics of DNA double helix, nonetheless, carry sufficient information for recognition by regulatory proteins as well as small molecules [35,36], which bind non-covalently or introduce small covalent modifications. Fairly strong and specific binding is observed between some ligands and DNA (with nanomolar dissociation constants or better) and this is attributed to various factors such as hydrogen bonding, snug fit, etc. There are two principally different modes by which drugs bind non-covalently to DNA—groove binding and intercalation. Minor-groove-binding drugs typically possess a crescent shape, which complements the shape of the groove and facilitates binding by promoting van der Waals interactions. Additionally, these drugs can form hydrogen bonds to bases, typically to N3 of adenine and O2 of thymine. Most minor-groove-binding drugs bind to A/T rich sequences. This preference in addition to the designed propensity for the electronegative pockets of AT sequences is probably due to better van der Waals contacts between the ligand and groove walls in this region, since A/T regions are narrower than G/C groove regions [37] and also because of the steric hindrance in the latter, presented by the C2 amino group of the guanine base. However, a few synthetic polyamides have been designed which have specificity for G–C and C–G regions in the grooves. Notable among these are the series of lexitropsins [38] and imidazole–pyrrole polyamides [39] where the design strategy was based on hydrogen-bonding interactions between the DNA and the drug atoms. Minor-groove binders have been used in therapeutic applications, e.g., berenil is a trypanocidal drug [21], pentamidine and bisguanylphenylfuran are effective against the AIDS-associated pathogen, *Pneumocystis carinii* [19,24] while netropsin and distamycin are known antitumor drugs [29]. Intercalators constitute the other class of non-covalently binding drugs. These drugs typically possess flat, heteroatomic ring systems that can stack between two adjacent base pairs in a helix. The complex, among other factors, is thought to be stabilized by π–π stacking interactions between the drug and DNA bases. Intercalators introduce structural perturbations in the DNA [40]. Intercalators also exhibit potential for clinical use, e.g., nogalamycin and hedamycin inhibit complex formation between DNA and transcription factors [41]. As the number of DNA–drug complex structures that have been solved by X-ray crystallography and NMR analysis is increasing, it is becoming possible to identify structural features and their energetic consequences that guide observed properties like affinity and sequence selectivity.

The role of experimental studies has been fundamental to the understanding of DNA–drug recognition. Detailed thermodynamic studies can yield information on binding constants, and corresponding free energy, enthalpy, entropy, and heat capacity changes on complex formation [3,42]. Such information is immensely helpful for validating interpretations based on structural analyses and theoretical predictions. Among the earliest studies focusing on thermodynamics of DNA–drug complex formation, Breslauer and co-workers made the remarkable observation of enthalpy–entropy compensation [43] in DNA–drug systems and also contributed significantly to developing thermodynamic profiles of these complexes [44,45]. Significant structural insights into the DNA–drug binding process were provided by Neidle and co-workers [5,16–25] by a synthesis and study of crystal structures. Chaires, Haq, and co-workers [1–3] carried out elegant investigations on thermodynamics of DNA–drug binding [46] and developed an effective protocol based on experimental studies for

parsing the binding energy into various components [47]. This methodology was extended and applied [48,49] with encouraging results, but due to problems such as compound solubility and the large binding constants of many minor-groove-binding agents with DNA, systematic studies on a large set of minor-groove complexes in AT sequences could not be reported. The key question in the molecular thermodynamics of recognition is concerned with the identification of the favorable energy components that add up to the experimentally observed binding free energies and the energetic consequences of mutation/optimization to eventually design better binding ligands. Hence, the intervention of theory becomes essential at this stage to expeditiously elicit this information. The studies reported here are intended to complement existing experimental and theoretical data and to provide additional results for steering towards the discovery of some general energetic and structural principles that characterize minor-groove interactions.

Some recent progresses in force fields, simulation techniques, free energy methodologies, and computational resources [50] have enabled us now to undertake a detailed free energy component analysis using state of the art molecular dynamics (MD)[1] simulation methods. Tracing back to the historical milestones in the field of modeling of DNA–drug complexes, some of the earliest studies in this field [51,52] focused on netropsin and proposed models to elucidate the exact roles played by formal charges and hydrogen bonds besides snug fit in the netropsin–DNA complex. Computer graphics model building techniques and molecular mechanical calculations based on energy minimization protocols were employed to widen the horizons and study different classes of DNA-binding molecules, notable among which were studies on the minor groove binding netropsin, bisintercalating drug, triostin A [53], and the covalent binder acridine [54]. Successful MD simulations of DNA [55,56] underlined the importance of providing the appropriate environment to the polyanionic DNA during simulation, bringing explicit solvent and counterions into the picture. Soon enough, for DNA–drug complexes too, the energy minimization based technique graduated to more complex MD studies [57]. Transition from structure to free energy of binding in drug–DNA complexes is a more recent development.

Though the focus in early computational studies on DNA–drug interactions lay on development in molecular mechanics protocols, other theoretical methods were also developed, notable among which were algorithms for computerized selection of potential DNA binding compounds [58] and finding potential DNA-binding candidates using molecular shape as a criterion [59]. Concurrently, the era of free energy calculations [60–62] commenced and calculation of binding free energies and various components contributing to it became a major area of interest. Importance of solvent, counterions [56], and salt effects on ligand–DNA binding [63,64] was demonstrated. The interactions between ligand and DNA were mainly classified as electrostatic, which included the Coulomb attraction of a ligand to DNA and the desolvation expense upon binding, and as hydrophobic effects and van der Waals interactions, the latter including packing considerations. A detailed study on the electrostatics in DNA–netropsin system revealed that the electrostatic binding free energy is very low, since the favorable electrostatic interactions are largely compensated for by unfavorable changes in the solvation of both the ligand and the DNA upon binding [65]. Calculation of relative binding free energies for distamycin and its analog was reported to result in good agreement with experiment [66]. Determination of absolute binding free energies is a relatively difficult task and remains semi-quantitative at the current state of the art. A few successful attempts have been reported [67–69] wherein the theoretically determined absolute binding free energy is in correspondence with experiment. These theoretical studies on DNA–drug interactions have resulted in important contributions to the understanding of binding in a few DNA–drug complexes. Since these studies typically focused on a single DNA–drug complex [67–69], the need to examine a series of complexes persists to enable extraction of a set of general principles. The necessity for a comprehensive molecular thermodynamic study on a series of DNA–drug complexes cannot be overstated, to eventually facilitate tailor-making better binders.

Rules governing DNA recognition continue to be an enigma and efforts to unravel them for protein–DNA or drug–DNA systems have not yet been fully rewarded. Our study employs a novel energy component strategy wherein it is possible to make correlations between structure and energetics, thus orienting the study towards eliciting these rules. We report here such a study on a series of minor-groove binders complexed with various DNA sequences (Table 1). All the minor-groove binders studied have A/T specificity, but a few also exhibit specific interactions with C/G base pairs. The drugs are cationic in nature, with the positive charge at one or both the ends of the molecule. Noting that DNA is polyanionic in nature and the minor groove in AT regions presenting only electronegative (N3 and O2) atoms, favorable electrostatic interactions are expected between the drug and the floor of the A/T rich minor groove which has a high negative potential [70]. The experimental binding free energies for these systems are typically in the range of −7 to −10 kcal/mol [1]. The main objective of our study was to aim towards evolving a general methodology and developing a consensus view of the energetics of recognition and its correlation to structure and not to focus on a single

[1] *Abbreviations used:* MD, molecular dynamics; BFEE, binding free energy estimates; ASA, accessible surface area; SA, solvent accessible.

Table 1
The DNA–drug complexes investigated and the calculated binding free energy components

	PDB	DNA[a]	Drug	n_a[b]	q_{dr}[c]	vdW[d]	el_{dir}[e]	solv[f]	cav[g]	ent[h]	ΔG_E^0[i]	ΔG_{ex}^0[j]	Ref.[k]
1	264D	A3T3	Hoechst 33258	32	+1	−56.5	−554.5	545.1	−5.9	25.1	−46.7	−7.7	[14]
2	127D	A2T2	Hoechst 33258	32	+1	−68.9	−587.1	575.2	−6.4	25.2	−62.0	−7.9	[15]
3	129D	A2T2	Hoechst 33342	34	+1	−69.2	−597.3	578.7	−6.8	25.5	−69.1		[14]
4	302D	A2T2	Metahydroxy Hoechst	32	+1	−69.5	−580.0	569.7	−6.2	25.2	−60.8		[16]
5	102D	A3T3	Propamidine	23	+2	−47.6	−1154.4	1140.4	−5.6	24.4	−42.9		[17]
6	1PRP	A2T2	Propamidine	23	+2	−45.8	−1133.3	1122.6	−5.3	24.4	−37.3	−8.2	[17]
7	1D64	A2T2	Pentamidine	25	+2	−50.2	−1121.8	1117.7	−5.8	24.8	−35.3	−7.0	[18]
8	166D	A2T2	γ-Oxapentamidine	25	+2	−46.0	−1139.5	1123.6	−5.7	24.8	−42.8		[19]
9	1D63	A3T3	Berenil	21	+2	−44.1	−1159.5	1126.0	−4.8	24.1	−58.2	−8.0	[20]
10	2DBE	A2T2	Berenil	21	+2	−41.2	−1160.0	1126.8	−4.8	24.1	−55.1	−8.6	[21]
11	289D	A2T2	Cyclopropyl bisfuramidine	29	+2	−52.7	−1110.0	1090.9	−6.0	24.9	−52.8		[22]
12	298D	A2T2	Isopropyl bisfuramidine	29	+2	−56.0	−1110.1	1100.1	−6.1	25.0	−47.2		[22]
13	227D	A2T2	Guanyl bisfuramidine	23	+2	−41.6	−1158.0	1143.6	−4.6	24.2	−36.4	−9.3	[23]
14	360D	A2T2	Ethyl bisfuramidine	27	+2	−49.5	−1133.3	1117.1	−5.8	24.8	−46.7		[24]
15	1FMQ	A2T2	Cyclobutyl bisfuramidine	31	+2	−56.4	−1103.6	1090.0	−6.3	25.1	−51.2		[25]
16	1FMS	A2T2	Cyclohexyl bisfuramidine	35	+2	−68.7	−1144.9	1130.9	−7.2	25.5	−64.3		[25]
17	1EEL	A2T2	Piperidinoethyl bisfuramidine	33	+2	−54.4	−1105.3	1091.2	−6.5	25.3	−49.6		
18	121D	A3T3	Netropsin	31	+2	−70.7	−1180.6	1168.5	−7.0	25.4	−64.3	−8.6	[26]
19	1D86	A2T2	Netropsin	31	+2	−58.0	−1205.6	1193.9	−6.3	25.4	−50.6	−8.8	[27]
20	195D	T2A2	Netropsin	31	+2	−66.9	−1175.7	1159.3	−6.7	25.4	−64.6	−8.7	[28]
21	1DNE	ATAT	Netropsin	31	+2	−64.5	−1198.9	1184.4	−6.6	25.4	−60.2		[29]
22	2DND	A3T3	Distamycin	35	+1	−74.0	−595.7	581.7	−7.2	25.7	−69.5		[30]
23	1LEX	A2T2	Monoimidazole lexitropsin	37	+2	−66.3	−1186.5	1168.4	−6.7	25.5	−65.6		[31]
24	1D30	A2T2	DAPI	21	+2	−35.8	−1186.0	1167.2	−4.6	23.9	−35.2	−8.0	[32]
25	328D	A2T2	SN7167	37	+2	−73.2	−1109.8	1091.6	−7.2	25.7	−72.9		[33]

[a] A2T2 refers to the DNA sequence d(CGCGAATTCGCG), A3T3 to d(CGCAAATTTCGC), T2A2 to d(CGCGTTAACGCG), and ATAT to d(CGCGATATCGCG).
[b] The number of heavy atoms in each drug molecule.
[c] Charge on the drug.
[d] The direct van der Waals component to the binding free energy in kcal/mol (1 kcal = 4.186 kJ).
[e] The direct electrostatics component of the binding free energy in kcal/mol.
[f] The solvation electrostatics component of the binding free energy in kcal/mol.
[g] The cavitation component of the binding free energy in kcal/mol.
[h] The rotational and translational entropy component, $T\Delta S$, in kcal/mol.
[i] The theoretically calculated net binding free energy in kcal/mol.
[j] Experimental binding free energy values available in the literature for a few complexes in kcal/mol.
[k] Corresponding references for the crystal/NMR structures of the complexes.

system or obtain quantitative agreement with experimental data using specialized parameters for a single system. To our knowledge, this is the first attempt based on theory at laying out a comprehensive free energy component analysis on a series of drugs.

Transiting from structure to thermodynamics of binding via computational means is a challenging task, as noted earlier—some of the difficulties being, modeling the complex and DNA at the atomic level with counterions and solvent [7,71], handling the condensation and release of ions and solvent upon drug binding [72], describing accurately the partial atomic charges and other energy parameters of drug molecules in a force field compatible manner, and finally setting up a computationally viable but expeditious free energy methodology for simulations on large biomolecular systems that is at least semi-quantitative. In this study, we address the above issues and present a computational methodology which is phenomenological in origin and in addition to yielding binding free energies reveals various components contributing to the overall free energy, which otherwise cannot be determined by experimental methods. An additional goal is to seek an atomic level correlation between structure and thermodynamics. The methodology presented here makes it possible to predict with some confidence, whether a particular structural change would have an unfavorable or favorable effect on binding. Employing the information gained from general trends observed in our results, we have successfully designed 'mutated' ligands wherein we introduced structural changes that improve DNA–ligand binding. The methodology adopted and the results obtained are presented and discussed below.

Methods

A flowchart describing the systematic analysis undertaken of the DNA–drug complexes and the protocol followed is presented in Fig. 1.

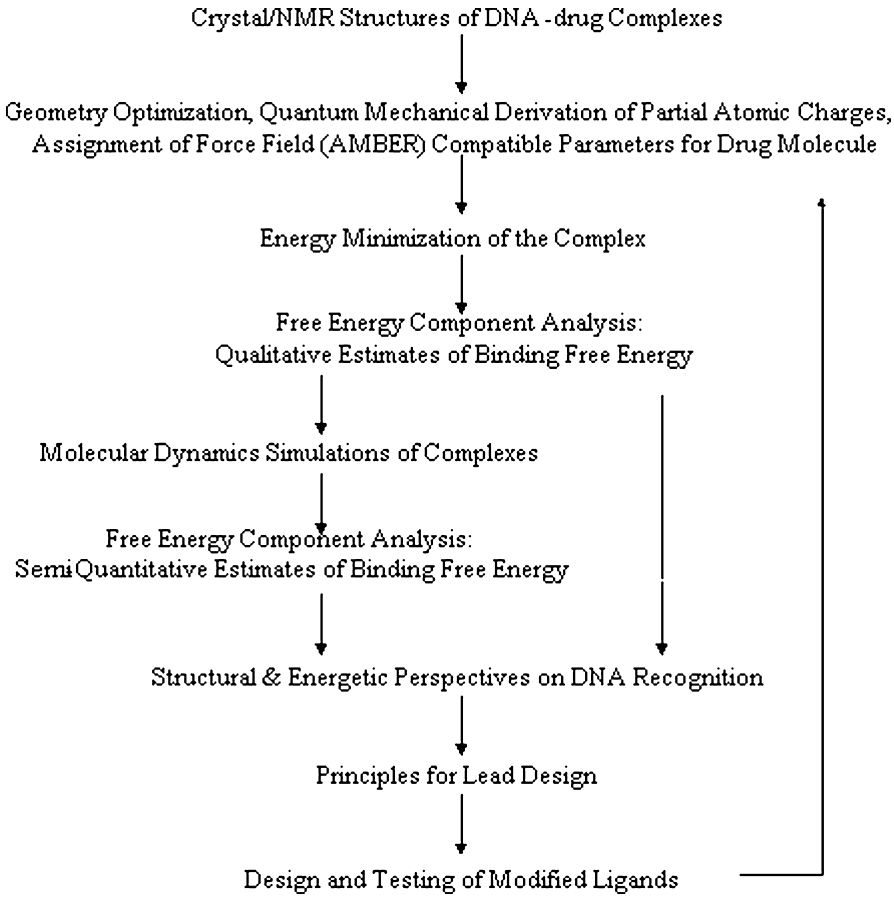

Fig. 1. Flowchart describing the computational protocol adopted.

Preparation of DNA–drug complexes for thermodynamic analysis

The atomic coordinates of the complexes were obtained from the RCSB Data Bank [73]. Waters of crystallization and ions were removed from all crystal structures and hydrogen atoms were added explicitly using the LEAP module of the AMBER [74] package; the NMR structures were used as such. Partial atomic charges for the drug atoms were derived consistent with the AMBER protocol. The electrostatic potentials were generated with the 6-31G* basis set using GAMESS [75] and were fitted by a restrained electrostatic potential method [76]. Force field parameters were assigned to the drug atoms by analogy with the original Cornell et al. [77] parameterization. Energy minimization was then performed to achieve the nearest stable low energy conformations. The structures were first subjected to 500 steps of hydrogen minimization (50 steps of steepest descent, SD, followed by 450 steps of conjugate gradient, CG) to relieve any steric clashes due to the addition of hydrogens. To the thus minimized structures, counterions were added to ensure electroneutrality and each system was surrounded by a 9 Å box of waters. This was followed by 5000 steps of water minimization (500 SD + 4500 CG) with a 25 kcal/mol restraint on the complex and ions. A restrained all atom minimization was then carried out in which the complex and ions were initially subjected to a 25 kcal/mol restraint, which was relieved more gradually on the complex than on the ions, over 5000 steps of minimization (SD:CG = 1:9). Finally, 6000 steps of free minimization were carried out. This series of minimizations undertaken is to alleviate any steric clashes in the crystal structures and to prepare them for energy analyses.

Free energy analyses

We have employed a second-generation all atom molecular mechanics force field AMBER for the determination of the intramolecular energetics while the electrostatic contribution to solvation was calculated via the AMBER-compatible modified "generalized Born solvent accessibility" [78] model, which employs effective radii parameters derived by Jayaram et al. [79]. MMGBSA [78–81] and MMPBSA [82,83] are two recent methods, which elicit free energies from structural information circumventing the computational complexity of free energy simulations. The MMGBSA approach is parameterized within the additivity

approximation [84] wherein the net free energy change is treated as a sum of a comprehensive set of individual energy components, each with a physical basis. Each component is estimated in a force field compatible manner. The modified MMGBSA method as applied to energy minimized structures is computationally rapid and fairly reliable for studying the various components contributing to the binding free energy, limited only by the semi-quantitative nature of the results obtained. The methodology is amenable to further systematic improvements. Some of the limitations inherent in single point energy calculations (studies on energy minimized structures) can be overcome by subjecting the systems to configurational averaging via molecular dynamics simulations [85]. These are also investigated in the present study.The governing equation for the estimation of free energy change upon binding is

$$\Delta G^0_{\text{net}} = \Delta G^0_{\text{tr}} + \Delta G^0_{\text{rot}} + \Delta G^0_{\text{vib.config.}} + \Delta G^0_{\text{adpt}} + \Delta G^0_{\text{inter}} + \Delta G^0_{\text{solvn.}}. \quad (1)$$

A detailed description and discussion of the above equation are presented in the Appendix. The energy components in Eq. (1) can be described as:

$$\Delta G^0_{\text{trvc}} = \Delta G^0_{\text{tr}} + \Delta G^0_{\text{rot}} + \Delta G^0_{\text{vib.config.}} = -T\,\Delta S^0_{\text{trvc}}. \quad (2a)$$

$$\Delta G^0_{\text{adpt}} = \Delta G^0_{\text{intra}} = \Delta H^0_{\text{intra}}. \quad (2b)$$

$$\Delta G^0_{\text{inter}} = \Delta G^0_{\text{el}} + \Delta G^0_{\text{vdW}} = \Delta H^0_{\text{inter}}. \quad (2c)$$

$$\Delta G^0_{\text{solvn.}} = \Delta G^0_{\text{solvn,el}} + \Delta G^0_{\text{solvn,nel}}. \quad (2d)$$

$$\Delta G^0_{\text{net el}} = \Delta G^0_{\text{el}} + \Delta G^0_{\text{solvn,el}}. \quad (2e)$$

$$\Delta G^0_{\text{cav}} = \Delta G^0_{\text{solvn,nel}}. \quad (2f)$$

These components are computed via the MMGBSA methodology adopted here. The thermodynamic cycle employed to construct the standard free energies of DNA–drug binding in solution is illustrated in Fig. 2. Building on Eqs. (1) and (2a)–(2f), the net binding process is decomposed into six steps and the corresponding binding free energy is calculated as a sum of five components:

$$\Delta G^0_{\text{net}} = \Delta G^0_{\text{trvc}} + \Delta G^0_{\text{adpt}} + \Delta G^0_{\text{vdw}} + \Delta G^0_{\text{net el}} + \Delta G^0_{\text{cav}}. \quad (3)$$

The net binding free energy is considered to be a sum of the free energy changes due to translational, rotational, vibrational, and configurational entropy losses, deformation or adaptation expense, van der Waals interactions between DNA and drug, net electrostatics, and cavitation effects [86]. The net electrostatics includes Coulomb interactions between DNA and the drug, explicit counterion contributions, added salt

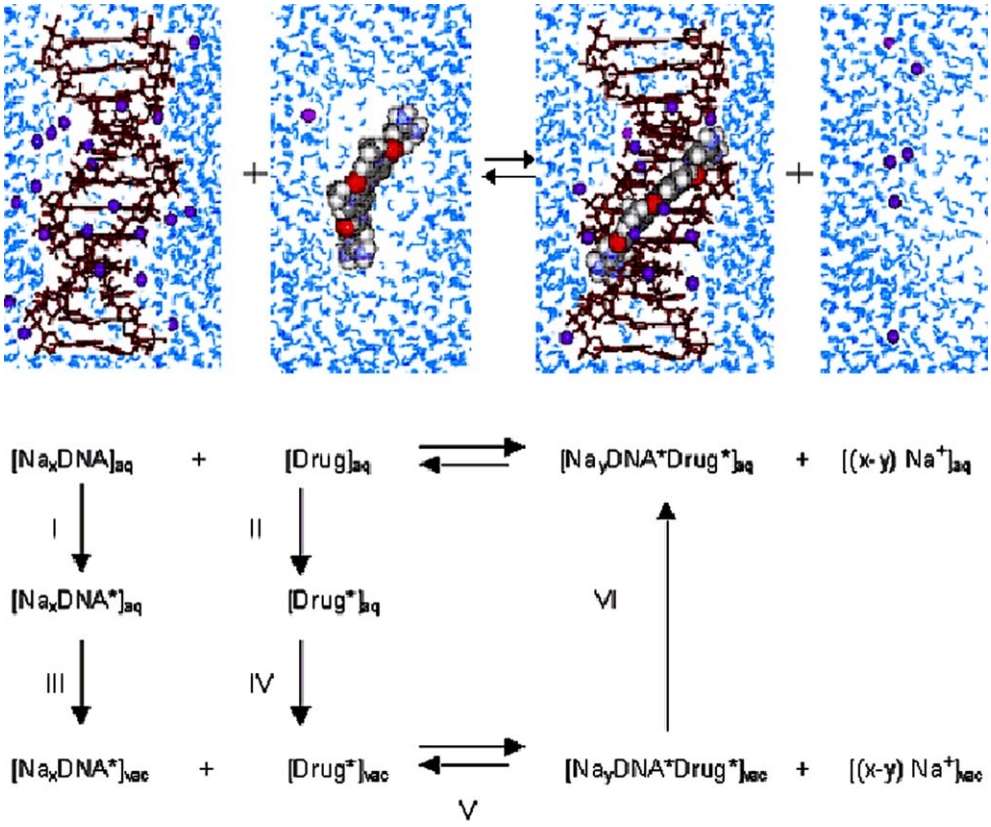

Fig. 2. The thermodynamic cycle adopted to construct the standard free energies of DNA–drug binding in aqueous medium.

(0.18 M NaCl) effects at the Debye–Huckel level, and electrostatic component of solvation. The change in size and shape of solvent cavity on complexation gives rise to water reorganization, a component of which, originating from non-polar sources, is the hydrophobic effect. Here the non-electrostatics of desolvation of both polar and non-polar atoms is accounted for in the cavitation term [78]. The molecular surface area calculations, required in computing the cavitation contributions [78,80], were performed using ACCESS program [87] but with AMBER force field compatible van der Waals radii.

Free energy analyses were performed on the energy minimized structures (single points in configuration space), using the MMGBSA methodology detailed above. The systems investigated include 25 DNA–drug complexes and a few modified complexes that contained drugs in which mutations were introduced in silico for the purposes of this study. It may be noted at this juncture that the sum of the various components obtained by the application of the MMGBSA method is expected to result in the net binding free energy of the complex, but since we have excluded the contribution due to counterions, the deformation expense, and the vibrational entropy contribution, we refer to this sum as the binding free energy estimates (ΔG^0_E or BFEE). We have adequately confirmed from molecular dynamics simulations, which include all these contributions, that the conclusions we make about general trends in the Results and discussion are unaffected by these omissions. The reasons for these exclusions are as follows: an assessment of the counterion contribution was attempted using the bisector model for initial placement of ions around DNA and considering the displacement of ions by drug followed by Monte Carlo optimization of the ion locations, but due to convergence problems accurate estimates of the contribution were unobtainable for the energy minimized structures. Thus, a rigorous treatment of these effects is undertaken within the molecular dynamics study with explicit solvent. Calculation of adaptation expense, which is a measure of the energetics of deformation in the structures of the drug and the DNA to accommodate each other in forming the complex, requires a study of the structures before and after binding preferably with an ensemble averaging [85], hence this term is included only when analyzing the results from the molecular dynamics simulations. In the energy analysis, each component is calculated separately for the minimized complex, and the DNA and drug extracted from the minimized complex. The overall contribution is determined by subtracting the separate contributions of the DNA and drug from those of the complex. The GB/SA and entropy calculations were performed using programs developed in-house [86] based on the theory discussed in the Appendix.

Molecular dynamics simulations

Additionally, we carried out extensive molecular dynamics simulations on the complex, the drug, and corresponding canonical DNA with explicit solvent and counterions on two systems, to obtain a more comprehensive quantitative picture and determine the factors that limit the single point analyses. For the molecular dynamics simulations, the input structures were prepared as follows. Waters of crystallization and ions were removed from all crystal structures and hydrogen atoms were added explicitly. The structures were first subjected to 500 steps of hydrogen minimization (50 SD + 450 CG) to relieve any steric clashes due to the addition of hydrogens. To these structures, counterions were added to ensure electroneutrality and each system was surrounded by a 9 Å box of waters. This was followed by 500 steps of water minimization with a 25 kcal/mol restraint on the complex and ions. A restrained all atom minimization was then carried out in which the complex and ions were initially subjected to a 25 kcal/mol restraint, which was relieved more gradually on the complex than on the ions, over 500 steps of minimization. Finally, 100 steps of free minimization were carried out. The structures thus prepared were used as input for the MD simulation. A 1 femtosecond (fs, 10^{-15} s) time step was used for integrating the equations of motion in all the MD studies. Periodic boundary conditions were applied throughout the MD simulations, along with PME summation method [88] for treating the electrostatics. Here, 10 picoseconds (ps, 10^{-12} s) of heating phase was first carried out in which the system was heated to 300 K using a constraint of 25 kcal on all the atoms except waters. This was followed by a 20 ps period of equilibration in which the constraints were removed more gradually from the complex than the ions, in a stepwise manner. Further equilibration was carried out for 10 ps with no constraints on the coordinates and finally a 4 nanosecond (ns, 10^{-9} s) data collection phase was carried out under constant pressure periodic boundary conditions. Analysis of the molecular dynamics trajectories was then undertaken. About 300 structures at intervals of 10 ps were collected from each trajectory followed by a computation of the averages for each component in the thermodynamic cycle. The net contributions of all components except the deformation (adaptation) expense were determined from the trajectory of the complex by extracting the DNA and the drug from the complex structure, calculating the individual energetic contributions and then subtracting their sum from the contribution of the complex, i.e., Net energy = [Complex – (DNA + Drug)]. The deformation expense was calculated as the energy difference between the bound and unbound states of the DNA and the drug.

Binding free energy analyses were performed on 25 minor-groove binders based on energy minimized

structures of the complexes. Additionally, post facto free energy analyses were carried out from the molecular dynamics simulations on the bound and unbound complexes in two cases. A critical appraisal of the results obtained and a perspective on the nature of minor-groove binding is presented underneath. A discussion of some novel attempts to make better binders by mutation of ligands in silico to improve binding affinity is presented subsequently.

Results and discussion

A qualitative view of the energetics of DNA–drug binding

A summary of the calculated binding free energy estimates from energy minimized structures of 25 DNA–drug complexes together with an energy component-wise partitioning into the van der Waals, electrostatics, desolvation, cavitation, and entropy contributions is presented in Table 1. Note that the binding free energy estimates are based on single structures and do not include contributions due to thermal averaging, ion effects, vibrational entropy, and deformation expense, hence a large negative value is expected and observed. Accurate evaluation of binding free energy would involve extracting these from molecular dynamics simulations, which treat thermal averaging, solvent, and counterion effects with rigor. We have carried out such simulations for two representative systems and these are discussed in ''Molecular dynamics simulation studies on DNA–drug systems: a semi-quantitative view of the energetics of DNA–drug binding.'' However, the free energy estimates based on single structures are computationally fast and extremely useful in building a qualitative picture of binding and in determining relative contributions of the different computed energy components. Taking a consensus view (Fig. 3) of all the 25 minor-groove binders studied, we observe that van der Waals and cavitation effects are favorable while entropy effects are unfavorable to binding in the complexes, as seen earlier in protein–DNA [12,86] and protein–drug [85] systems. Figs. 3A and B illustrate that the operational separation of the cavitation component into hydrophobic and desolvation van der Waals does not affect the overall binding free energy. Fig. 3A suggests that the largest favorable contribution comes from the direct van der Waals component in all the cases, indicating the importance of packing and shape complementarity and size of the drug. Fig. 3B suggests that the largest favorable contributor to binding is the hydrophobic component, as suggested by experimental studies [48,49]. While both direct van der Waals and hydrophobic contributions are favorable to complexation as to which one is larger depends on the formation of compound subsets from the individual phenomenological energy components. The overall electrostatics (with desolvation effect included) is marginally favorable in all the DNA–drug complexes as seen earlier in some protein–drug systems, while it contributes unfavorably in protein–DNA complexes in a consensus view with case-specific exceptions. Thus, we observe that both, steric as well as electrostatic complementarities, are important for minor-groove binding in DNA–drug systems.

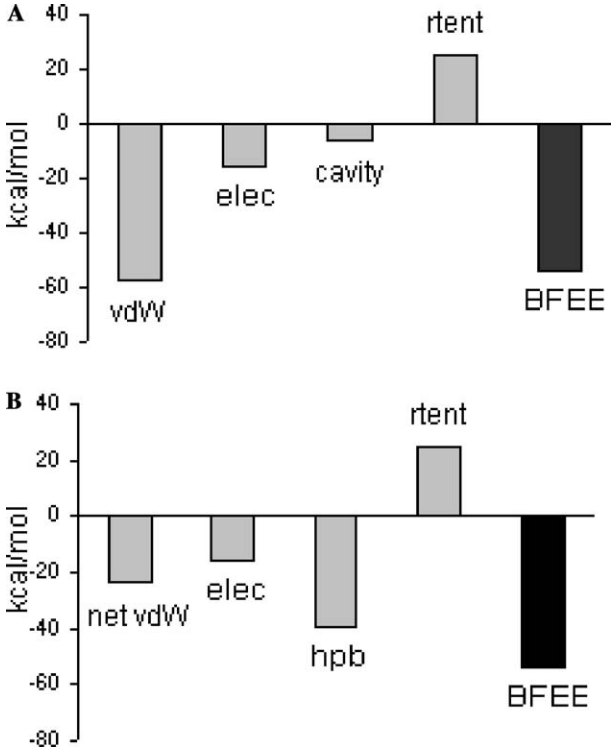

Fig. 3. Histograms depicting a consensus view of the various energy components contributing to the binding free energy estimates (BFEE). (A) Direct van der Waals (vdW) component, overall electrostatics (Elec) component (direct electrostatics + desolvation), overall cavitation (hydrophobic contribution + desolvation vdW), and rotational and translational entropy (rtent). (B) Overall van der Waals (net vdW) interactions (direct vdW + desolvation vdW), overall electrostatics component, hydrophobic contribution (Hpb), and entropy.

Mining the free energy component data

Role of van der Waals contributions in DNA–drug binding

Conclusive evidence appears on the important role played by van der Waals interactions in DNA–drug binding as observed in the plot of the direct van der Waals contribution versus the number of heavy atoms in the drug (Fig. 4A). The plot reveals a linear correlation between the two, indicating that favorable van der Waals contacts increase as the size of the drug is increased in the synthetic ligands. This result directly suggests that larger drugs compensate for the drawbacks related with their bulky presence such as steric hindrance, higher desolvation expense, etc., by maximizing favorable van der Waals contacts. From the standpoint

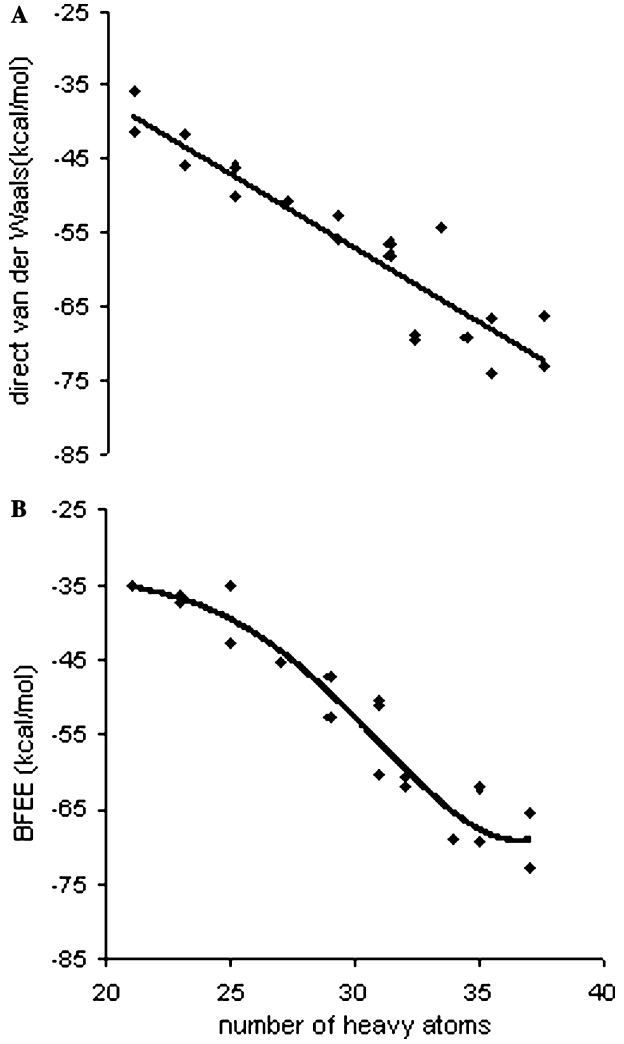

Fig. 4. (A) Variation of direct van der Waals interaction energy with number of heavy atoms in the corresponding drug ($R^2 = 0.8609$). (B) Variation of binding free energy estimate of complex with number of heavy atoms in the corresponding drug ($R^2 = 0.9258$).

of specificity, a larger molecule would be a better candidate—larger the drug, larger is the target area covered, hence more could be the specificity in targeting. However, the question that arises here is purely from a binding perspective—leaving aside drug absorption and delivery issues—whether there exists a limit on the size. Intuitively one expects a threshold that limits the size of a minor-groove binder in view of various factors like steric hindrance, lack of shape complementarity over the entire drug, etc. entropy being another issue for consideration that is accessible through molecular simulations described later, which could emerge as stronger players as the size of the drug increases. An indication to this effect appears in the correlation plot between the binding free energy estimates and the number of heavy atoms in each drug (Fig. 4B). The plot resembles a sigmoidal curve and suggests the existence of both a lower and an upper threshold limit on the size of the drug. Thus, it is only within a range (25–35 heavy atoms) that increasing the size of a drug could effectively increase the binding affinity. The existence of a possible threshold can be explained as follows: in the smaller drugs, with increase in size, a small incremental gain is achieved in shape complementarity, hence the overall energy does not change much, while among the larger drugs, gain in overall energy is very low due to loss in shape complementarity with increase in size of the drug as well as diminishing van der Waals component due to desolvation expense.

Role of shape complementarities in DNA–drug binding

An important relation emerges from the plot between the van der Waals term and loss in accessible surface area (ASA) on binding. The plot (Fig. 5) shows that greater the loss in area, larger is the van der Waals component. The loss in ASA is a good indicator of shape complementarity, thus this plot establishes a direct relation between shape complementarity and van der Waals interactions. The shape of the curve shows that ASA loss is rather less for small drugs and hardly varies for the four smallest ligands but moving towards the medium-sized and larger drugs, the ASA loss becomes sharper and for the largest drugs, the loss is highly pronounced. This indicates that shape complementarity is much more important for larger drugs than in the smaller ones. Looking at these results, it is clear that van der Waals and cavitation effects follow a pattern depending on the structure of the drug, and it is possible to predict the effect of a structural change on these properties with some degree of confidence.

Study of shape complementarity vis-à-vis van der Waals contributions through effect of mutations

Our results indicate the importance of size and packing in drug binding, and also establish a relation between these properties and shape complementarity.

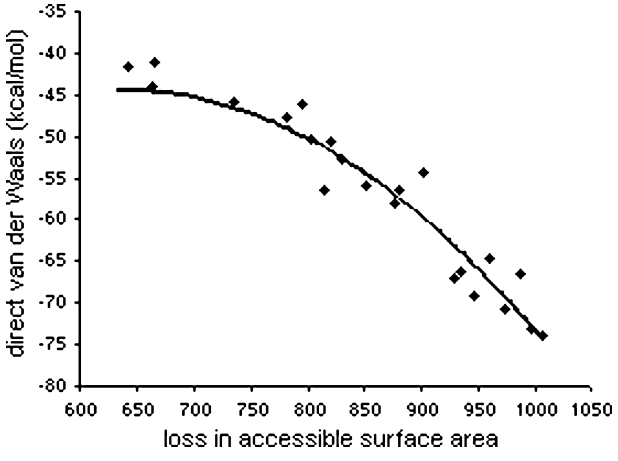

Fig. 5. Variation of direct van der Waals component with loss in accessible surface area (Å^2; $R^2 = 0.9167$).

The significance of this correlation led us to undertake another short study, with the aim of further determining the importance of shape complementarity. The working hypothesis is that a larger size for the ligand is not enough to ensure good van der Waals, and that shape complementarity was a very important factor in maximizing van der Waals contacts. For this study, we added aliphatic chains and rings to the smaller drugs—berenil, propamidine, DAPI, and bisguanylphenylfuran in silico such that the size of the newly created molecules was close to that of the larger drugs studied. The additions were made while the drugs were bound to the minor groove and the structures were checked for clashes, hence we were fairly satisfied with the placement of the new entity in the minor groove. These newly developed complexes were then subjected to the minimization protocol followed earlier and analyzed. The results (Table 2) show that when the size (number of heavy atoms) of the ligand is same, the original drugs with larger loss in solvent accessible area or better complementarity show better van der Waals than the new entities. The van der Waals interactions depend on, and improve with, shape complementarity between the drug and the DNA.

Role of electrostatics in DNA–drug binding

The modified MMGBSA (m2GB) method has been shown [79] to give improved estimates of solvation energies compatible with Coulomb interactions (direct electrostatics) described by the force field, compared to many other theoretical models. Hence, we employed this method to determine the free energy components and obtain reliable estimates of the total (Coulomb and solvation) electrostatics. It must be noted here that counterion effects have not been included in the calculation of electrostatics in the case of studies involving energy minimized structures. Building a three-dimensional model for counterions around DNA in the presence and absence of drug with continuum solvent (i.e., without explicit solvent) has been shown earlier to be thermodynamically satisfactory [89] but structurally unsatisfactory (Jayaram et al., unpublished work). These effects have been dealt with in the molecular dynamics analysis carried out with explicit solvent and are described in "Molecular dynamics simulation studies on DNA–drug systems: a semi-quantitative view of the energetics of DNA–drug binding."

Owing to the highly charged nature of polyanionic DNA, electrostatics is expected to dominate the binding thermodynamics. Interestingly however, an anti-correlation emerges between the direct electrostatic component and the desolvation expense (Fig. 6) in DNA–drug binding. In all the cases studied, we observe almost complete compensation of the gain in the net binding free energy due to the direct electrostatic (Coulomb) interactions by the loss due to desolvation. It is well known that water is essential to the stability of the highly charged double helix [7,90]. Ligand binding resulting in removal of waters from the vicinity of DNA must overcome the strong desolvation electrostatics. Thus, the contribution of overall electrostatics to the binding energy, though favorable, is very low in most cases. The fact that such compensation is observed in all the 25 complexes studied and also in some earlier theoretical work on single systems [65,67] indicates electrostatics-desolvation compensation to be a general trend in DNA–drug systems.

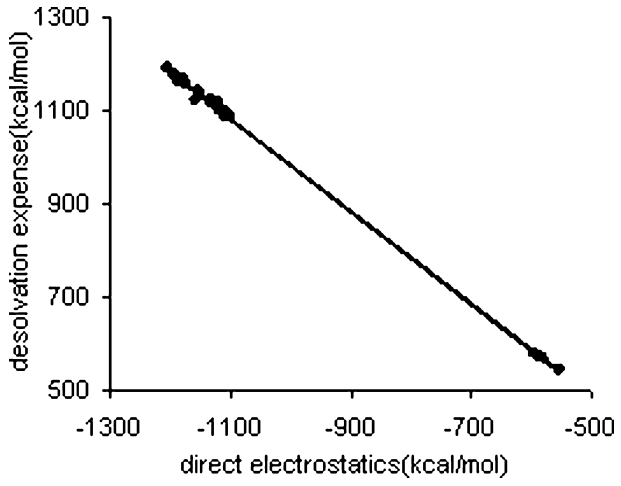

Fig. 6. Variation of direct electrostatics component with desolvation expense ($R^2 = 0.9993$).

Table 2
Comparison of van der Waals contribution in drugs and modified ligands of same size

n_{atm}[a]	Existing drug[b]	Drug modified[c]	$\Delta\Delta G_{vdW}$[d] (kcal/mol)
31	Netropsin	Guanyl bisfuramidine	9.7
33	Cyclopentadienyl bisfuramidine	Berenil	6.8
33	Cyclopentadienyl bisfuramidine	DAPI	7.2
35	Cyclohexyl bisfuramidine	Propamidine	5.6
35	Cyclohexyl bisfuramidine	Guanyl bisfuramidine	12.6

[a] Number of heavy atoms in the ligand.
[b] Existing drug containing the specified number of heavy atoms.
[c] Drug modified to form a new ligand containing the specified number of heavy atoms.
[d] Difference between direct van der Waals interaction energy of original complex and modified ligand–DNA complex.

Role of hydrogen bonding in DNA–drug binding: a study of the structure and energetics of hydrogen bonding in the energy minimized DNA–drug complexes

Hydrogen bonds are considered to make an important discriminatory contribution in the binding of ligands to the minor groove of DNA with different sequences and this forms the basis for a strategy to design new ligands [39]. Putative hydrogen bonds are reported often along with the crystal structures. However, since hydrogens are not reported in these structures, the proposed hydrogen bonds are based on heavy atom distances and the hydrogen-bond angle depends on the hybridization of the heavy atom. This approach does not take into account slightly distorted geometries, which often occur due to steric factors, hence there is a need to examine the hydrogen bonds using structures that report the positions of hydrogens and are also as close as possible to the actual DNA–drug complex structure. We considered the energy minimized structures with hydrogens added to be suitable for this purpose.

We examined the structures of the 25 DNA–drug complexes studied here for putative hydrogen bonds keeping a distance criterion of less than 2.5 Å between the hydrogen and the acceptor atom. We also classified the hydrogen bonds as 'linear' or 'bent' by applying an angle criterion of less than 30° deviation of the hydrogen atom for linear bonds and greater than 30° deviation for bent bonds. Minor-groove binders can form hydrogen bonds with functional groups on the bases exposed in the grooves via their end groups and also from their amide or other linker groups located in the middle of the drug. It may be noted that in all the complexes studied, the hydrogen atom belongs to the drug and the electron rich heavy atom (oxygen/nitrogen) to the DNA, i.e., essentially the drug behaves as a donor and the DNA as an acceptor, as expected for A/T sequences in the minor groove. We also observed that the charged end groups of the drugs formed most of the 'linear' bonds while the 'bent' bonds were formed with the linker groups. We noticed that almost 75% of the hydrogen bonds in the 25 DNA–drug complexes studied were 'linear' by our definition. The majority of the hydrogen bonds were between the drug and base pairs but a few bonds with the backbone were also seen.

The set of complexes studied here comprises a large variety of drugs and DNA sequences thus presenting an opportunity to address the role of hydrogen bonds in DNA recognition. However, a direct quantification of the contribution of hydrogen bonds is not feasible, since it requires comparison of a large number of closely related structures with and without a functional group involved in hydrogen bond. Hence, we undertook further studies in which we calculated the change in binding energy that occurred when a hydrogen bond was switched off in a particular complex. This was achieved by restraining the charge on the H-bonding hydrogen to zero and compensating for this in the charges of the neighboring groups, thus maintaining the same overall charge. The binding energy calculations were repeated on minimized structures separately for each hydrogen bond made by the drug in each of the complexes. The results of this study are summarized in Table 3. The method adopted for switching off hydrogen bonds in this study is swift and gives a qualitative picture of the energetic perspective of hydrogen-bonding interactions. A quantitative approach would be to modify the hydrogen-bonding group, re-derive partial atomic charges, and then carry out energy analyses preferably using molecular dynamics trajectories.

We observe that in every complex, switching off a hydrogen bond resulted in a loss in the net binding free energy. The energy loss, averaged over 60 hydrogen bonds, was found to be 3.2 kcal/mol. In most cases, loss of a linear H-bond causes a slightly greater loss in energy than loss of a bent bond, as expected. Our observations suggest that the hydrogen-bond network between the DNA and the drug could probably 'sew' the two together such that the ligand stays at just the right place in the minor groove for optimal binding. This would ensure better interaction and better binding between the drug and DNA besides packing/snug fit of the drug in the minor groove. Hydrogen bonds also ensure better sequence specificity, which in itself, is a very important issue apart from good binding. While these results are intuitive, qualitative, and justify drug-design attempts based on a structural appraisal of the hydrogen-bonding functional groups afforded by a sequence, the hydrogen bonds introduced may contribute to interaction energy favorably but may not always contribute to binding energy if explicit solvent and thermal averaging is taken into account. These issues are addressed in the molecular dynamics studies reported in

Table 3
Hydrogen-bond analysis of 25 DNA–drug complexes[a]

Property	Calculated averages for proposed H-bonds
Distance (Å)[b]	2.2
Angle—linear (°)[c]	19.5
Angle—bent (°)[d]	38.8
Energy loss—linear (kcal/mol)[e]	3.9
Energy loss—bent (kcal/mol)[f]	2.3
Energy loss/H-bond (kcal/mol)[g]	3.2

[a] Analysis based on energy minimized complexes (1 kcal = 4.186 kJ; 1 Å = 0.1 nm).
[b] Average distance of hydrogen from heavy atom with which hydrogen bonding occurs.
[c] Average angle of linear hydrogen bonds.
[d] Average angle of bent hydrogen bonds.
[e] Average energy loss on removal of a linear hydrogen bond.
[f] Average energy loss on removal of a bent hydrogen bond.
[g] Average energy loss on removal of any hydrogen bond.

"Molecular dynamics simulation studies on DNA–drug systems: a semi-quantitative view of the energetics of DNA–drug binding."

Towards a computational design of better binders through a qualitative appraisal of free energy components: studies on mutated drugs

One of the aims of our study was to set up a direct correlation between structure and thermodynamic properties, to predict as to which structural changes would improve or hinder drug binding and to what extent. Building on the studies reported in this paper, we have introduced structural changes in berenil in silico, with an expectation that the mutation would have the desired effect on binding. The results of this study are shown in Table 4. Berenil (Fig. 7) was chosen because of the simplicity in its structure due to symmetry and the unusual and interesting presence of a triazene group in it.

The structural changes considered in designing mutated drug berenil are as follows. (1) brne- N1′ replaced by C to reduce negative potential on the drug surface in this area and H added to N to make the overall charge on the molecule +3. (2) brne1- N and N1′ replaced by a C atom each to decrease negative potential on the drug surface in the area. (3) brnh- NH_2 groups added to C3′ and C5 since there was hydrogen-bonding potential in these regions. (4) brnv- Ethyl groups added on terminal N to increase van der Waals interactions.

Partial atomic charges were then derived for the ligand atoms as described earlier in the Methods and AMBER-compatible force field parameters were assigned to the ligand atoms. The mutations were applied to the berenil–CGCGAATTCGCG complex and these were subjected to the minimization protocol as described earlier, followed by energy analysis. The results show that the binding energy improves in all cases. The modifications in the ligands, brne and brne1, were expected to help in increasing favorable electrostatics, and this was observed too, indicating that decreasing the negative potential over the drug surface close to the groove floor results in better binding. As mentioned earlier, the A/T region has a high negative potential at the floor of the groove, hence the ideal drug should possess a complementary positive potential in this region. In brnh, the amino groups were introduced in such a manner that they were geometrically well placed to form hydrogen bonds with nearby bases in the DNA, and after subjecting the complex to the minimization protocol, the position of these groups indicated formation of hydrogen bonds. The increase in favorable electrostatics indicates that better charge relay was possible between the drug and DNA due to the intro-

Table 4
Estimated binding free energy components[a] for modified berenil ligands relative to berenil

PDB	DNA sequence	Ligand	Charge	elec[b]	vdW[c]	cav[d]	ent[e]	$\Delta\Delta G_E^{0\,f}$
2DBE	CGCGAATTCGCG	brne	3	−11.6	4.0	0.0	0.0	−7.5
	CGCGAATTCGCG	brne1	2	−8.0	−2.4	0.0	0.2	−10.2
	CGCGAATTCGCG	brnh	2	−6.8	−3.2	−0.3	0.0	−10.3
	CGCGAATTCGCG	brnv	2	0.9	−8.6	−0.8	−0.5	−7.9

[a] All energies are reported in units of kcal/mol (1 kcal = 4.186 kJ).
[b] Change in overall electrostatic component (direct electrostatics + desolvation expense) in modified ligand–DNA complex compared to original berenil–DNA complex.
[c] Change in direct van der Waals component.
[d] Change in cavitation component.
[e] Change in the rotational and translational entropy component.
[f] Change in binding free energy estimates.

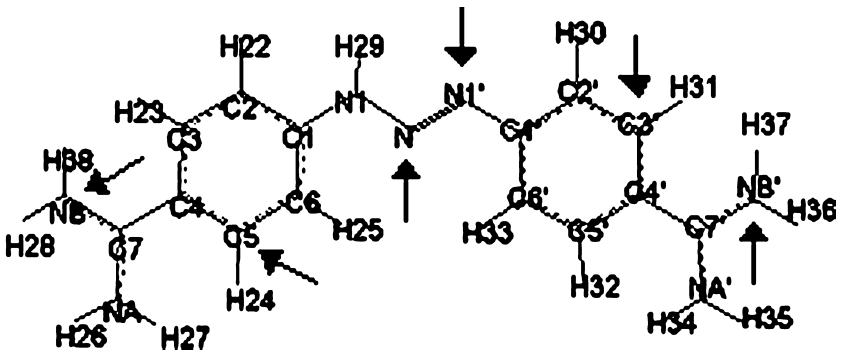

Fig. 7. A representation of the molecular structure of Berenil. Groups subjected to in silico modifications are indicated by arrows.

duction of these groups. The brnv–DNA complex shows better van der Waals energy due to the additional ethyl groups. Overall, the results show that these alterations bring about the expected and desirable changes in the thermodynamic properties, thus indicating the potential of the method employed and further providing assurance that errors and statistical uncertainties in the estimations are under control, especially when differential energetics are considered.

Molecular dynamics simulation studies on DNA–drug systems: a semi-quantitative view of the energetics of DNA–drug binding

Molecular dynamics simulations introduce theoretical rigor into binding free energy estimates and eliminate some of the limitations inherent in the single point free energy calculations. Also dynamics simulations with explicit solvent and small ions yield an enhanced view of recognition from structural and dynamic perspectives. The results from MD simulations turn out to be semi-quantitative in nature and are comparable to experimentally observed quantities, thus validating the theoretical protocol adopted throughout our study.

Energetics of DNA–drug binding based on analysis of MD simulations

Molecular dynamics simulations were performed on two systems, which required five simulations in all (two complexes + one canonical DNA + two drugs), and included counterions and explicit solvent. The chosen systems were DAPI-d(CGCGAATTCGCG) and propamidine-d(CGCGAATTCGCG). The binding free energy (BFE) was computed via the thermodynamic cycle described in Fig. 2. Further details are provided in the Methods. Fig. 8 indicates the level of convergence in the binding free energies as the run proceeded. The results obtained from analysis are listed in Table 5. A comparison of the interaction energies and BFE estimates developed from energy minimized structures with the experimentally determined BFEs and theoretical BFEs calculated from MD demonstrates the significance of the analysis based on MD simulations and the current state of free energy theory. Comparison of the consensus values (Fig. 9A) of the energy components obtained from MD and the EM method reveals the importance of the contribution from counterion effects, adaptation expense, and vibrational entropy component which are not included in the single point analysis but are present

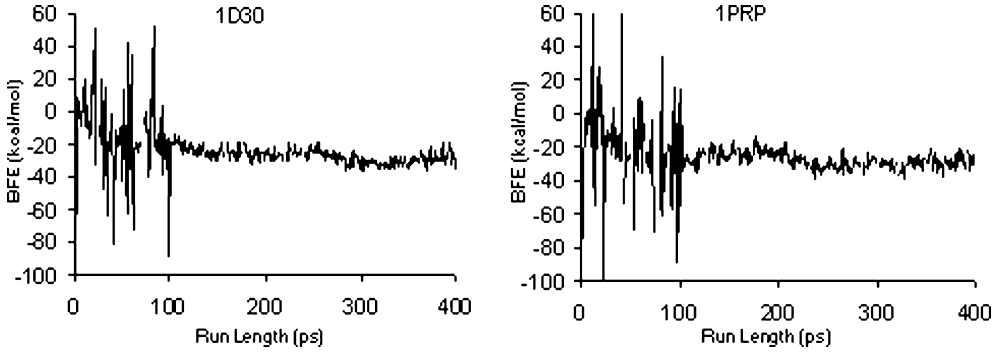

Fig. 8. Convergence plots for variation of binding free energy as a function of molecular dynamics run length.

Table 5
Calculated net binding free energies and components[a] based on post facto analyses of molecular dynamics simulations

PDB	Drug	ΔG^{0}_{ex} [b]	ΔE_{int} [c]	ΔG^{0}_{E} [d]	ΔG^{0} [e]	vdW[f]	elec[g]	cavity[h]	rtent[i]	vibent[j]	ci[k]	adapt[l]
1D30	DAPI	−8	−71.9	−35.2	−9.1	−41.8	−15.8	−4.5	23.9	1.2	18.2	9.7
1PRP	Propamidine	−8.2	−76.7	−37.3	−10.1	−46.8	−9.2	−5.3	24.4	5.2	7.7	13.8
Average		−8.1	−74.3	−36.3	−9.6	−44.3	−12.4	−4.9	24.2	3.2	13.0	11.7

[a] All energies are reported in units of kcal/mol. (1 kcal = 4.1868 kJ).
[b] Experimentally determined net binding free energy (literature values).
[c] Overall interaction energy between DNA and drug in the complex.
[d] Theoretically estimated net binding free energy from energy minimized structures.
[e] Theoretically calculated net binding free energy, from MD trajectories.
[f] Direct van der Waals component.
[g] Overall electrostatics component of the binding free energy.
[h] Cavitation component of the binding free energy.
[i] Rotational and translational entropy component.
[j] Vibrational and configurational entropy component.
[k] Contribution from counterion effects.
[l] Structural adaptation (deformation) expense during the binding process.

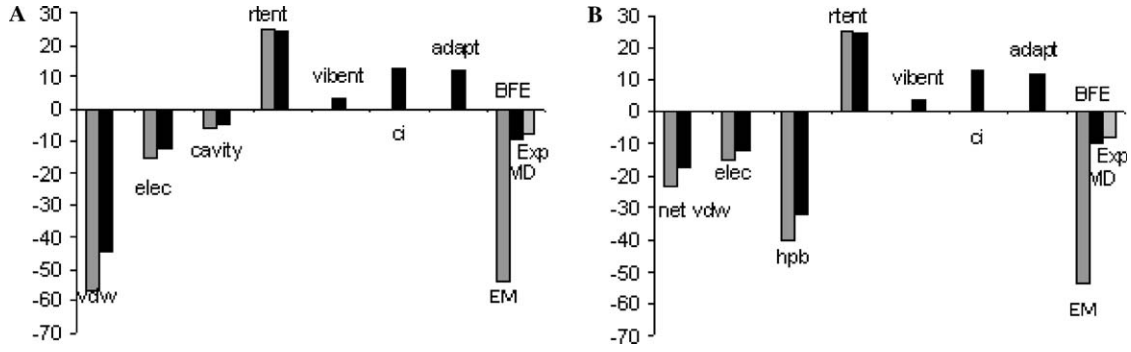

Fig. 9. Comparison of consensus energy values obtained from molecular dynamics analysis and minimization studies. Two alternative ways of separating the components are shown (A) and (B). Gray, minimization (EM) results; black, MD results; 'Ex' denotes the experimental value for binding free energy, 'vdW' denotes the direct van der Waals interactions, 'elec' refers to overall electrostatics, 'cavity' denotes the cavitation component, 'rtent' refers to rotational and translational entropy, 'vibent' refers to vibrational and configurational entropy, 'ci' refers to contribution from counterion effects, 'adapt' is the configurational adaptation expense, 'BFE' is the net binding free energy, 'netvdw' denotes the sum of desolvation van der Waals and direct van der Waals, and 'hpb' indicates the hydrophobic contribution.

in the MD results. An alternative decomposition of the consensus binding free energy components is presented in Fig. 9B for the purpose of correlation with experiment and discussed in detail in ''Correlation between experimental and theoretical data.'' Looking at Fig. 9A, displacement of counterions from the vicinity of DNA due to drug binding causes an energetic loss, which shows up as a positive contribution (+13 kcal/mol) as observed. The van der Waals term calculated from the MD simulation is lower in magnitude than that obtained from minimization as expected, since packing is not as tight during the dynamic simulation relative to the minimized structures especially in the presence of explicit solvent. The desolvation expense is almost equal to the direct electrostatic contribution, hence the overall electrostatics is low in magnitude. The cavitation term and rotational and translational entropy terms remain almost the same in the minimized and MD structures. The adaptation expense can be determined rigorously only from the MD simulations on bound and unbound molecules, as explained earlier. The binding free energies obtained from the MD simulations are −9.1 kcal/mol for the DAPI–DNA complex and −10.1 kcal/mol for the propamidine–DNA complex, which are very close to the experimentally determined values.

Role of hydrogen bonds in DNA–drug binding: an analysis of MD simulations

Introduction of hydrogen bonds between DNA and drug to promote binding is an appealing strategy, but it is important to know the dynamics of these hydrogen bonds and the thermodynamic consequences, especially in the presence of solvent and counterions. The key question for drug design is whether the hydrogen bond is retained when the DNA–drug complex is allowed dynamic behavior in a more realistic environment. We carried out an analysis of the hydrogen bonds in the structures emerging from MD simulations. The distance between hydrogen-bonding atoms was averaged over the MD trajectories of both complexes. The results (Table 6) indicate that only a few hydrogen bonds that were present in the original (energy minimized) structures are retained during the simulations and the distance between hydrogen-bonding atoms was observed to increase from that in the minimized structures.

The MD structures were analyzed to monitor the hydrogen bonds, which are retained throughout the MD simulation. Putative hydrogen-bonding groups on the drug while exhibiting fluctuations in their distance from DNA were observed often to come within the bond-forming distance. In the DAPI–DNA complex, two hydrogen bonds between the DNA and the drug are retained throughout the simulation and we considered them suitable for further analysis. The N3–H36 bond is a 'linear' bond, with H36 belonging to the charged end

Table 6
Comparison of H-bond distances (Å) in minimized and molecular dynamics structures

PDB	Drug	H-bond[a]	Minimization[b]	Molecular dynamics[c]
1D30	DAPI	O2–H35	2.5	7.8
		O2–H38[d]	**1.9**	**2.3**
		O4′–H34	1.9	5.5
		O4′–H26	2.1	4.2
		N3–H33	2.4	6.7
		N3–H36	**2.0**	**2.4**
		N3–H37	2.0	3.9
1PRP	Propamidine	N3–H44	2.0	4.2
		N3–H45	2.1	4.4

[a] The atoms involved in hydrogen bond in the energy minimized structure.
[b] Distance between hydrogen-bonded atoms in the minimized structure.
[c] Distance between hydrogen-bonded atoms averaged over 400 structures obtained from MD.
[d] H-bonds retained during MD are in bold.

group while the O2–H38 bond is 'bent,' with H38 belonging to a group in the middle part of the drug's structure. This conforms to our earlier observation that most linear bonds belong to the charged end groups while bent bonds belong to groups lying in the middle, usually linker groups. We calculated the energy associated with each of these H-bonds in the manner described earlier (''Mining the free energy component data'') and averaged it over all the structures obtained from MD. It was observed that loss of the linear bond results in an energy loss of 1.1 kcal/mol and loss of the bent bond results in an energy loss of 0.6 kcal/mol (compared to that of 3.2 and 2.4 kcal/mol, respectively, in the minimized structures). This implies a favorable contribution to the binding free energy from hydrogen-bond formation in this system. This method can be extended to other DNA–drug systems to provide quantitative estimates of the energetic contribution of all putative hydrogen bonds, which from a design perspective is extremely useful information.

Role of water in DNA–drug binding: a structural and energetic perspective

DNA hydration is a well-discussed subject [7,90–93] since DNA being highly polyanionic is stable only in the presence of water and counterions. Presence of ordered waters along the DNA grooves, termed as the spine of hydration, was first observed in the minor groove of the d(CGCGAATTCGCG) dodecamer, and seems to be a common feature to the AT rich regions [94]. The spine of hydration is presumed to stabilize the DNA conformation. Intuitively one expects that binding of a drug in the minor groove would displace waters present in the region, resulting in a net release of waters in the binding process. However, contrary to intuition, in an experimental study on a netropsin analog–DNA complex [95], net uptake of water was observed to occur on groove binding. Thus, we carried out an investigation on the DAPI–DNA and propamidine–DNA systems to establish whether the binding process in these two systems is associated with water release or uptake.

The volume occupied by propamidine and DAPI was calculated using a fairly accurate grid method and determined to be 178.5 and 155.5 $Å^3$, respectively. The volume occupied by a water molecule in a solution of density 1 g/ml is 29.8 $Å^3$. Thus, six waters were expected to be displaced by propamidine and five waters by DAPI. We then carried out a detailed analysis of the number of waters released/taken up by the DNA in these complexes on binding. This was done by counting the number of waters surrounding the free DNA and the complexed DNA structures obtained from MD trajectories, within distance cutoffs ranging from 3 to 7 Å. The difference between the numbers obtained for canonical DNA and complexed DNA would be a measure of the waters released/taken up from DNA upon binding. We observed that the ensemble averaged number of waters released by DNA in the propamidine–DNA system is around 7–8 and for the DAPI–DNA system, it is around 5–6, and these numbers remain the same irrespective of the distance cutoff (within 3–7 Å) or the run length of the simulation. These numbers correspond well with the expectation from the volume calculation mentioned above.

The waters released during binding of a drug to DNA mostly belong to the spine of hydration. Thus, this release would be favorable since the ordered spine of hydration is disrupted, resulting in a positive entropy contribution. It may be recalled that the unfavorable contribution from the electrostatics of desolvation is compensated by the gain in electrostatics between drug and DNA (Fig. 6). Also, the cavitation component, which is a measure of the hydrophobic contribution and desolvation van der Waals component, is observed to be favorable, as discussed earlier in ''Energetics of DNA–drug binding based on analysis of MD simulations,'' indicating that water release favors DNA–drug binding. In the light of these observations, the phenomenon of water uptake observed in a particular complex of DNA and a netropsin analog [95] is an interesting counter-intuitive observation, which requires further investigations in which a close mimicking of the experimental method in the computational setup may be required.

Correlation between experimental and theoretical data

Experimental studies directed at a parsing of the energy components of DNA–drug binding [48,49,96–98] suggest that the hydrophobic contribution is the largest and factors such as van der Waals, electrostatics, hydrogen-bonding contributions, etc., (ΔG_{mol} or ΔG_{int}), are less significant. Results based on MD trajectories come closest to experiment at the current state of the art and thus provide a natural point of comparison with experiment. Figs. 9A and B show two different representations of the consensus results obtained from post facto analysis of molecular dynamics trajectories, as mentioned earlier in ''Molecular dynamics simulation studies on DNA–drug systems: a semi-quantitative view of the energetics of DNA–drug binding.'' Fig. 9B clearly indicates the hydrophobic term to be the largest favorable contributor to the binding free energy, as indicated by experiment [48,49,96–98]. The highly favorable direct van der Waals contribution is canceled to a large extent by the loss in van der Waals interactions due to desolvation. The direct electrostatics component is almost completely compensated by the desolvation electrostatics and counterion effects. Hence, a sum of these components is a relatively smaller contribution in relation to the hydrophobic contribution, as indicated by experiment.

We have also determined the structural adaptation expense, which approximates the ΔG_{conf} term used in experiment, denoting the loss in free energy due to conformational changes in the DNA and drug upon binding. This term is assigned a value of zero in experimental parsing.

Adding up the theoretically obtained adaptation contribution, overall van der Waals term and overall electrostatics results in an energy term of low magnitude closely matching with the experimentally observed ΔG_{mol} or ΔG_{int} term. This implies the overall nature of correspondence between experiment and theory—it is not the individual contributions but the balance between the components that leads to the agreement.

We have determined the translational and rotational entropy contribution using the Sackur–Tetrode equation in gas phase (discussed in the Appendix) and the vibrational entropy by the quasi-harmonic approach. Much controversy surrounds the use of Sackur–Tetrode equation in solution [99] where it is considered inapplicable. We have thus circumvented this objection by constructing a thermodynamic cycle (Fig. 2), which allows the calculation of entropy using the ideal gas statistical mechanics and is theoretically consistent. The gas-phase entropy values obtained from our calculations are naturally expected to differ from the solution-phase experimentally derived entropies, however, the signature of the entropic contribution remains positive and unfavorable to binding in both cases.

The above correlations reflect the overall nature of the agreement between experiment and theory. Theory facilitates a better interpretation of the experimental data at a microscopic level. It also becomes clear that as to which component contributes most or least to binding depends on the manner in which the individual components are summed, the observed and computed binding free energies reflecting a fine balance of diverse energetic contributions at a molecular level.

Conclusions

An exhaustive computational analysis of the molecular thermodynamics of DNA–drug binding performed on 25 minor-groove binders reveals the critical role played by van der Waals and hydrophobic interactions, the compensatory effects of direct electrostatics, and desolvation and the net hydrogen-bond contributions. Our analyses indicate that better van der Waals interactions lead to a better binding of the drug to DNA, and that the magnitude of the van der Waals interactions is strongly linked with shape complementarity. The study brings out the importance of favorable hydrophobic contributions to DNA–drug binding. Electrostatics does not appear to contribute much to the net binding free energy when desolvation and counterion effects are taken into consideration. Detailed analyses of hydrogen-bonding interactions indicate that a few hydrogen bonds contribute favorably to the net binding free energy. Results from the molecular dynamics simulations by and large corroborate these findings apart from introducing more rigor in the methodology for determining binding free energies and enhancing the quality of the binding free energy estimates. Further, MD simulations facilitate in evolving a detailed molecular thermodynamic view of counterion effects, water release, and conformational adaptation upon DNA–drug binding. The study establishes a direct semi-quantitative relationship between structure and thermodynamics, in conformity with experimental data where available, thus promising to provide valuable information as well as a computational pathway for lead compound design and drug improvement. Furthermore, the atomic level description of the systems and the lack of any necessity to bring in system dependent parameterization allows for transferability of the methodology to study diverse biomolecular binding problems in aqueous media.

Acknowledgments

The authors gratefully acknowledge the financial support received from the Department of Science and Technology, Government of India. The authors also wish to thank Ms. Priyanka Srivastava and Ms. S. Sivashankari for their participation in preliminary studies.

Appendix

The statistical mechanical theory adopted [100–104] to study DNA–drug binding in aqueous media is presented here.

Let D (denoting DNA) and dr (denoting drug) be the reactants and D^*dr^*, the product of binding in aqueous medium:

$$[D]_{aq} + [dr]_{aq} = [D^*dr^*]_{aq}. \tag{A.1}$$

At equilibrium

$$\Delta G^0_{aq} = -RT \ln K_{eq.aq}. \tag{A.2}$$

In terms of canonical partition functions (Q) [105]

$$\begin{aligned}\Delta G^0_{aq} &= \Delta A^0_{aq} + P\Delta V^0_{aq} \\ &= -RT \ln K_{eq.aq} = -RT \\ &\quad \times \ln[\{Q_{D^*dr^*.aq}/(N_A Q_w)\}/\{(Q_{D.aq}/(N_A Q_w)) \\ &\quad \times (Q_{dr.aq}/(N_A Q_w))\}] + P\Delta V^0_{aq}.\end{aligned} \tag{A.3}$$

Eq. (A.3) is an exact expression for non-covalent associations in aqueous medium. ΔA^0 is the standard Helmholtz free energy of the reaction and $P\Delta V^0_{aq}$ is the

pressure–volume correction to Helmholtz free energy in the solvent medium. Q_w denotes the partition function for pure solvent (water).

Assuming that translations and rotations are separable from intra-solute degrees of freedom as well as those of solvent

$$\Delta G^0_{aq} = -RT \times \ln[\{Q^{tr}_{D^*dr^*} Q^{rot}_{D^*dr^*} Z^{int}_{D^*dr^*.aq} Q^{el}_{D^*dr^*} N_A Q_w\} /\{(Q^{tr}_D Q^{rot}_D Z^{int}_{D.aq} Q^{el}_D)(Q^{tr}_{dr} Q^{rot}_{dr} Z^{int}_{dr.aq} Q^{el}_{dr})\}] + P\Delta V^0_{aq}. \quad (A.4)$$

The superscript, 'int,' denotes the internal contribution.

The electronic partition function, Q^{el}, is assumed to be unity for non-covalent associations

$$Q^{el}_D = Q^{el}_{D^*} = Q^{el}_{dr} = Q^{el}_{dr^*} = 1.$$

Thus, the standard free energy can be expressed as

$$\Delta G^0 = \Delta G^0_{tr} + \Delta G^0_{rot} + \Delta G^0_{int.aq} + P\Delta V^0_{aq}, \quad (A.5)$$

where

$$\Delta G^0_{tr} = -RT \ln[Q^{tr}_{D^*dr^*}(N_A/V)(Q^{tr}_D Q^{tr}_{dr})], \quad (A.6)$$

$$\Delta G^0_{rot} = -RT \ln[Q^{rot}_{D^*dr^*}/(Q^{rot}_D Q^{rot}_{dr})], \quad (A.7)$$

$$\Delta G^0_{int.aq} = -RT \ln[(Z^{int}_{D^*dr^*.aq} Q_w V)/(Z^{int}_{D.aq} Z^{int}_{dr.aq})], \quad (A.8)$$

where Z^{int} is the configurational partition function. It includes contributions from intermolecular interactions and internal motions as well as solvation (hydration) effects. The $\Delta G^0_{int.aq}$ term (Eq. (A.8)) is accessible to free energy molecular simulations configured in the canonical ensemble [11], albeit they are computationally expensive. The $P\Delta V^0_{aq}$ term in equation is often neglected in liquid-state work.

We consider the following work plan to make binding free energy estimations computationally feasible.

The translational part of the free energy is given by the Sackur–Tetrode equivalent as

$$\Delta G^0_{tr} = -RT \ln[(N_A/V)(\Lambda^3_D \Lambda^3_{dr}/\Lambda^3_{D^*dr^*})] = -RT \ln[(N_A/V)(h^2/2\pi k_B T)^{3/2} \times \{m_{D^*dr^*}/(m_D m_{dr})\}^{3/2}]. \quad (A.9)$$

Similarly, the rotational part of the free energy is evaluated as

$$\Delta G^0_{rot} = -RT \ln[(\sigma_D \sigma_{dr}/\sigma_{D^*dr^*})(1/(8\pi^2)) \times (h^2/2\pi k_B T)^{3/2} \times \{(I^a_{D^*dr^*} I^b_{D^*dr^*} I^c_{D^*dr^*})/(I^a_D I^b_D I^c_D I^a_{dr} I^b_{dr} I^c_{dr})\}^{1/2}]. \quad (A.10)$$

To invoke the Eqs. (A.9) and (A.10), a consideration of reactants and products in gas phase is necessary, thus requiring the construction of a thermodynamic cycle as done in this study (Fig. 2).

I^a_D, I^b_D, and I^c_D are the components of moments of inertia of species D along the principal axes a, b, and c, and σ_D its symmetry number. $k_B T$ is the product of Boltzmann constant and temperature (in Kelvin).

Z^{int} is determined via

$$Z^{int}_{D.aq} = \int \cdots \int \exp\{(-E(X^N_D, X^M_W)/k_B T)\}\, dX^N_D\, dX^M_W = \langle \exp(E(X^N_D, X^M_W)/k_B T)\rangle, \quad (A.11)$$

where X^N_D and X^M_W represent the configurational space accessible to the solute D and solvent W, respectively, in the presence of each other. $E(X^N_D, X^M_W)$ denotes the total potential energy of the system describing non-idealities.

At this stage, to make the problem computationally tractable, one may attempt separating intramolecular interactions from solvation effects

$$Z^{int}_{D.aq} = Z^{intra}_D Z^{solvn.}_D$$

$$Z^{int}_{D.aq} \simeq \int \cdots \int \exp[-\{E(X^N_D) + E(X^{Nfixed}_D; X^M_W)\}/k_B T] \times dX^N_D\, dX^M_W \simeq \left[\int \cdots \int \exp\{-E(X^N_D) + k_B T\}\, dX^N_D\right] \times \left[\int \cdots \int \exp\{(-E(X^{Nfixed}_D; X^M_W)/k_B T)\, dX^M_W\}\right]. \quad (A.12)$$

Equations similar to (A.12) can be written for dr and D^*dr^*, and converted to excess free energies. Such a separation leads to

$$\Delta G^0 = \Delta G^0_{tr} + \Delta G^0_{rot} + \Delta G^0_{inter} + \Delta G^0_{solvn.} + \Delta G^0_{intra}. \quad (A.13)$$

Eq. (A.13) forms the basis for "master equation" methods:

$$\Delta G^0_{inter} = \Delta H^0_{inter} - T\Delta S^0_{inter},$$
$$\Delta H^0_{inter} = \Delta H^0_{el} + \Delta H^0_{vdW} = \langle \Delta E^0_{inter}\rangle = \langle \Delta E^0_{el}\rangle + \langle \Delta E^0_{vdW}\rangle, \quad (A.14)$$

where ΔE^0_{el} and ΔE^0_{vdW} are computed from (12,6,1) force field for a fixed structure (from minimization studies) or for an ensemble of structures from MD simulations.

$$\Delta G^0_{intra} = \Delta G^0_{adapt}, \quad (A.15)$$

$$\Delta S^0_{intra} = \Delta S^0_{vib,config}, \quad (A.16)$$

where $\Delta S^0_{vib,config}$ can be calculated by normal mode analysis for energy minimized structures or by quasi-harmonic approximation introduced by Karplus and Kushick [106] and subsequently extended and adapted to MD simulations by van Gunsteren et al. [107].

To account for structural deformation upon binding, we include adaptation expense explicitly in the ΔG^0_{intra} term, and it is calculated as the difference in the free

energies of the bound and unbound states of the DNA and the drug (steps I and II in Fig. 2).

In the MMGBSA model, the solvation free energies are computed as

$$\Delta G^0_{\text{solvn}} = \Delta G^0_{\text{GBSA}} = \Delta G^0_{\text{GB}} + \Delta G^0_{\text{SA}}, \tag{A.17}$$

where ΔG^0_{GB} comprises of the electrostatic component of solvation while ΔG^0_{SA} is the non-electrostatic contribution, called cavitation energy in the literature [78].

The defining equation employed for evaluating the electrostatic contribution to the solvation free energy [79] with the MMGBSA model is

$$G^0_{\text{el.solvn}} = -166(1 - 1/\epsilon)\sum_{i=1}^{n}\sum_{j=1}^{n} q_i q_j / f_{\text{m2GB}}, \tag{A.18}$$

$$\Delta G^0_{\text{GB}} = G^0_{\text{el.solvn}}(\text{final state}) - G^0_{\text{el.solvn}}(\text{initial state}). \tag{A.19}$$

The non-electrostatic (nel) contributions to the solvation free energy [79,87] are computed as a function of the solvent accessible (SA) surface area

$$G^0_{\text{nel.solvn}} = \gamma_{\text{nel}}\,\Delta A, \tag{A.20}$$

$$\Delta G^0_{\text{SA}} = G^0_{\text{el.solvn}}(\text{final state}) - G^0_{\text{nel.solvn}}(\text{initial state}). \tag{A.21}$$

The quantity γ_{nel} has a value of 7.2 cal/mol/Å^2 . This may be considered as a resultant of +47 cal/mol/Å^2 from the cavity term [108] and −39.8 cal/mol/Å^2 from van der Waals interactions of the solute and the solvent. This separation is only for the purpose of interpretation and does not alter the free energy estimates.

A combination of Eqs. (A.9), (A.10), (A.14), (A.15), (A.16), and (A.17) yields the final equation used to determine the absolute binding free energies:

$$\Delta G^0_{\text{net}} = \Delta G^0_{\text{tr}} + \Delta G^0_{\text{rot}} + \Delta G^0_{\text{vib.config.}} + \Delta G^0_{\text{adpt}} + \Delta G^0_{\text{inter}} + \Delta G^0_{\text{solvn.}}.$$

Alternative formulations of this equation are discussed in the Methods.

Computation of absolute binding free energies is a formidable task and obtaining free energy estimates from a single simulation is equally challenging. The theory and methodology proposed above is an attempt to link structure(s) with thermodynamics and to elicit the binding free energies in a computationally expeditious manner. The series of results obtained highlight the merits of the methodology and also suggests areas for further systematic improvement.

References

[1] I. Haq, Arch. Biochem. Biophys. 403 (2002) 1–15.
[2] J.B. Chaires, Curr. Opin. Struct. Biol. 8 (1998) 314–320.
[3] J.B. Chaires, Biopolymers 44 (1997) 201–215.
[4] L.H. Hurley, Nat. Rev. Cancer 2 (2002) 188–200.
[5] S. Neidle, Nat. Prod. Rep. 18 (2001) 291–309.
[6] P.B. Dervan, Bioorg. Med. Chem. 9 (2001) 2215–2235.
[7] B. Jayaram, D.L. Beveridge, Annu. Rev. Biophys. Biomol. Struct. 25 (1996) 367–394.
[8] W. Jorgensen, Acc. Chem. Res. 22 (1989) 184–189.
[9] W.F. van Gunsteren, H.J.C. Berendsen, Angew. Chem. Int. Ed. Engl. 29 (1990) 992–1023.
[10] J.A. McCammon, Curr. Opin. Struct. Biol. 1 (1991) 196–200.
[11] D.L. Beveridge, F.M. DiCapua, Annu. Rev. Biophys. Biomol. Struct. 18 (1989) 431–492.
[12] B. Jayaram, K.J. McConnell, S.B. Dixit, A. Das, D.L. Beveridge, J. Comput. Chem. 23 (2002) 1–14.
[13] T. Lazaridis, Curr. Org. Chem. 6 (2002) 1319–1332.
[14] M.C. Vega, I. Garcia Saez, J. Aymami, R. Eritja, G.A. van der Marel, J.H. van Boom, A. Rich, M. Coll, Eur. J. Biochem. 222 (1994) 721–726.
[15] M. Sriram, G.A. van der Marel, H.L. Roelen, J.H. van Boom, A.H. Wang, EMBO J. 11 (1992) 225–232.
[16] G.R. Clark, C.J. Squire, E.J. Gray, W. Leupin, S. Neidle, Nucleic Acids Res. 24 (1996) 4882–4890.
[17] (a C.M. Nunn, S. Neidle, J. Med. Chem. 38 (1995) 2317–2325; (b C.M. Nunn, T.C. Jenkins, S. Neidle, Biochemistry 32 (1993) 13838–13843.
[18] K.J. Edwards, T.C. Jenkins, S. Neidle, Biochemistry 31 (1992) 7104–7109.
[19] C.M. Nunn, T.C. Jenkins, S. Neidle, Eur. J. Biochem. 226 (1994) 953–961.
[20] D.G. Brown, M.R. Sanderson, E. Garman, S. Neidle, J. Mol. Biol. 226 (1992) 481–490.
[21] D.G. Brown, M.R. Sanderson, J.V. Skelly, T.C. Jenkins, T. Brown, E. Garman, D.I. Stuart, S. Neidle, EMBO J. 9 (1990) 1329–1334.
[22] J.O. Trent, G.R. Clark, A. Kumar, W.D. Wilson, D.W. Boykin, J.E. Hall, R.R. Tidwell, B.L. Blagburn, S. Neidle, J. Med. Chem. 39 (1996) 4554–4562.
[23] C.A. Laughton, F. Tanious, C.M. Nunn, D.W. Boykin, W.D. Wilson, S. Neidle, Biochemistry 35 (1996) 5655–5661.
[24] A. Guerri, I.J. Simpson, S. Neidle, Nucleic Acids Res. 26 (1998) 2873–2878.
[25] I.J. Simpson, M. Lee, A. Kumar, D.W. Boykin, S. Neidle, Bioorg. Med. Chem. Lett. 10 (2000) 2593–2597.
[26] L. Tabernero, N. Verdaguer, M. Coll, I. Fita, G.A. van der Marel, J.H. van Boom, A. Rich, J. Aymami, Biochemistry 32 (1993) 8403–8410.
[27] M. Sriram, G.A. van der Marel, H.L.P.F. Roelen, J.H. van Boom, A.H. Wang, Biochemistry 31 (1992) 11823–11834.
[28] K. Balendiran, S.T. Rao, C.Y. Sekharudu, G. Zon, M. Sundaralingam, Acta Crystallogr. D 51 (1995) 190–198.
[29] M. Coll, J. Aymami, G.A. van der Marel, J.H. van Boom, A. Rich, A.H. Wang, Biochemistry 28 (1989) 310–320.
[30] M. Coll, C.A. Frederick, A.H. Wang, A. Rich, Proc. Natl. Acad. Sci. USA 84 (1987) 8385–8389.
[31] D.S. Goodsell, H.L. Ng, M.L. Kopka, J.W. Lown, R.E. Dickerson, Biochemistry 34 (1995) 16654–16661.
[32] T.A. Larsen, D.S. Goodsell, D. Cascio, K. Grzeskowiak, R.E. Dickerson, J. Biomol. Struct. Dyn. 7 (1989) 477–491.
[33] C.J. Squire, G.R. Clark, W.A. Denny, Nucleic Acids Res. 25 (1997) 4072–4078.
[34] A.L. Lehninger, D.L. Nelson, M.M. Cox, Principles of Biochemistry, CBS Publishers & Distributors, New Delhi, 1993.
[35] W. Saenger, Principles of Nucleic-Acid Structure, Springer-Verlag, New York, 1983.
[36] R.R. Sinden, DNA Structure and Function, Academic Press, San Diego, 1994.
[37] H.C. Nelson, J.T. Finch, F.L. Bonaventura, A. Klug, Nature 330 (1987) 221–226.

[38] D.S. Goodsell, Curr. Med. Chem. 8 (2001) 509–516.
[39] P.B. Dervan, B.S. Edelson, Curr. Opin. Struct. Biol. 13 (2003) 284–299.
[40] W.K. Olson, A.A. Gorin, Xiang-Jun Lu, L.M. Hock, V.B. Zhurkin, Proc. Natl. Acad. Sci. USA 95 (1998) 11163–11168.
[41] J.J. Welch, F.J. Rauscher, T.A. Beerman, J. Biol. Chem. 269 (1994) 31051–31058.
[42] A. Cooper, Curr. Opin. Chem. Biol. 3 (1999) 557–563.
[43] K.J. Breslauer, D.P. Remeta, W.Y. Chou, R. Ferrante, J. Curry, D. Zaunczkowski, J.G. Snyder, L.A. Marky, Proc. Natl. Acad. Sci. USA 84 (1987) 8922–8926.
[44] T.V. Chalikian, K.J. Breslauer, Curr. Opin. Struct. Biol. 8 (1998) 657–664.
[45] D.S. Pilch, N. Poklar, E.E. Baird, P.B. Dervan, K.J. Breslauer, Biochemistry 38 (1999) 2143–2151.
[46] J.B. Chaires, Anticancer Drug Des. 11 (1996) 569–580.
[47] I. Haq, T.C. Jenkins, B.Z. Chowdhry, J. Ren, J.B. Chaires, Methods Enzymol. 323 (2000) 373–405.
[48] I. Haq, J.E. Ladbury, B.Z. Chowdhry, T.C. Jenkins, J.B. Chaires, J. Mol. Biol. 271 (1997) 244–257.
[49] S. Mazur, F. Tanious, D. Ding, A. Kumar, D.W. Boykin, I.J. Simpson, S. Neidle, W.D. Wilson, J. Mol. Biol. 300 (2000) 321–337.
[50] E. Guidice, R. Lavery, Acc. Chem. Res. 35 (2002) 350–357.
[51] K. Zakrzewska, R. Lavery, B. Pullman, Nucleic Acids Res. 11 (1984) 8825–8839.
[52] J. Caldwell, P.A. Kollman, Biopolymers 25 (1986) 249.
[53] U.C. Singh, N. Pattabiraman, R. Langridge, P.A. Kollman, Proc. Natl. Acad. Sci. USA 83 (1986) 6402–6406.
[54] S.N. Rao, U.C. Singh, P.A. Kollman, J. Med. Chem. 29 (1986) 2484–2492.
[55] B. Tidor, K.K. Irukura, B.R. Brooks, M. Karplus, J. Biomol. Struct. Dyn. 1 (1983) 231–252.
[56] G.L. Seibel, U.C. Singh, P.A. Kollman, Proc. Natl. Acad. Sci. USA 82 (1985) 6537–6540.
[57] S.N. Rao, P.A. Kollman, Proc. Natl. Acad. Sci. USA 84 (1987) 5735–5739.
[58] P.D. Grootenhuis, P.A. Kollman, G.L. Seibel, R.L. DesJarlais, I.D. Kuntz, Anticancer Drug Des. 5 (1990) 237–242.
[59] P.D. Grootenhuis, D.C. Roe, P.A. Kollman, I.D. Kuntz, J. Comput. Aided Mol. Des. 8 (1994) 731–750.
[60] B.L Tembe, J.A. McCammon, Comput. Chem. 8 (1984) 281–283.
[61] W.L. Jorgensen, C. Ravimohan, J. Chem. Phys. 83 (1985) 3050–3054.
[62] D.L. Beveridge, F.M. DiCapua, Annu. Rev. Biophys. Biophys. Chem. 18 (1989) 431–492.
[63] V.K. Misra, K.A. Sharp, R.A. Friedman, B. Honig, J. Mol. Biol. 238 (1994) 245–263.
[64] K.A. Sharp, R.A. Friedman, V. Misra, J. Hecht, B. Honig, Biopolymers 36 (1995) 245–262.
[65] V.K. Misra, B. Honig, Proc. Natl. Acad. Sci. USA 92 (1995) 4691–4695.
[66] S.B. Singh, Ajay, D.E. Wemmer, P.A. Kollman, Proc. Natl. Acad. Sci. USA 91 (1994) 7673–7677.
[67] S.B. Singh, P.A. Kollman, J. Am. Chem. Soc. 121 (1999) 3267–3271.
[68] L.F. Pineda, M. Zacharias, J. Mol. Recognit. 15 (2002) 209–220.
[69] N. Spackova, T.E. Cheatham III, F. Ryjacek, F. Lankas, L. van Meervelt, P. Hobza, J. Sponer, J. Am. Chem. Soc. 125 (2003) 1759–1769.
[70] (a B. Jayaram, K.A. Sharp, B. Honig, Biopolymers 28 (1989) 975–993;
(b R. Lavery, B. Pullman, Nucleic Acids Res. 9 (1981) 3765–3777.
[71] N.L. Marky, G.S. Manning, J. Am. Chem. Soc. 122 (2000) 6057–6066.
[72] G.S. Manning, Q. Rev. Biophys. 11 (1978) 179–246.
[73] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucleic Acids Res. 28 (2000) 235–242.
[74] D.A. Pearlman, D.A. Case, J.W. Caldwell, T.E. Cheatham III, S. Debolt, D.M. Ferguson, G.L. Seibel, P.A. Kollman, Comput. Phys. Commun. 91 (1995) 1–41.
[75] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S.J. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, J. Comput. Chem. 14 (1993) 1347–1363.
[76] C.I. Bayly, P. Cieplak, W.D. Cornell, P.A. Kollman, J. Phys. Chem. 97 (1993) 10269–10280.
[77] W.D. Cornell, P. Cieplak, C.F. Bayly, I.R. Gould, M.M. Kenneth, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, J. Am. Chem. Soc. 117 (1995) 5179–5197.
[78] W.C. Still, A. Tempczyk, R.C. Hawley, T.J. Hendrickson, J. Am. Chem. Soc. 112 (1990) 6127–6129.
[79] B. Jayaram, Y. Liu, D.L. Beveridge, J. Chem. Phys. 109 (1998) 1465–1471.
[80] G.D. Hawkins, C.J. Cramer, D.G. Truhlar, J. Phys. Chem. 100 (1996) 19824–19839.
[81] B. Jayaram, D. Sprous, D.L. Beveridge, J. Phys. Chem. B 102 (1998) 9571–9576.
[82] B. Honig, A. Nicholls, Science 268 (1995) 1144–1149.
[83] J. Srinivasan, T.E. Cheatham III, P. Cieplak, P.A. Kollman, D.A. Case, J. Am. Chem. Soc. 120 (1998) 9401–9409.
[84] K.A. Dill, J. Biol. Chem. 272 (1997) 701–704.
[85] P. Kalra, T.V. Reddy, B. Jayaram, J. Med. Chem. 44 (2001) 4325–4338.
[86] B. Jayaram, K.J. McConnell, S.B. Dixit, D.L. Beveridge, J. Comput. Phys. 151 (1999) 333–357.
[87] K. Lee, F.M. Richards, J. Mol. Biol. 55 (1971) 379–400.
[88] D.M. York, W. Yang, H. Lee, T. Darden, L.G. Pedersen, J. Am. Chem. Soc. 117 (1995) 5001–5002.
[89] B. Jayaram, S. Swaminathan, D. Beveridge, K. Sharp, B. Honig, Macromolecules 23 (1990) 3156–3165.
[90] H.M. Berman, Curr. Opin. Struct. Biol. 1 (1991) 423–427.
[91] H.W. Drew, R.E. Dickerson, J. Mol. Biol. 151 (1981) 535–556.
[92] B. Jayaram, T. Jain, Annu. Rev. Biophys. Biomol. Struct. 33 (2004) 343–361.
[93] B. Jayaram, R. Fine, K. Sharp, B. Honig, J. Phys. Chem. 93 (1989) 4320–4327.
[94] M. Kochoyan, J.L. Leroy, Curr. Opin. Struct. Biol. 5 (1995) 329–333.
[95] N. Yu-Sidorova, D.C. Rau, Biopolymers 35 (1995) 377–384.
[96] D. Rentzeperis, L.A. Marky, Biochemistry 34 (1995) 2937–2945.
[97] L.A. Marky, K.J. Breslauer, Proc. Natl. Acad. Sci. USA 84 (1987) 4359–4363.
[98] L. Wang, A. Kumar, D.W. Boykin, C. Bailly, W.D. Wilson, J. Mol. Biol. 317 (2002) 361–374.
[99] Y.B. Yu, P.L. Privalov, R.B. Hodges, Biophys. J. 81 (2001) 1632–1642.
[100] A. Ben-Naim, Statistical Thermodynamics for Chemists and Biochemists, Plenum, New York, 1992.
[101] M.K. Gilson, V. Given, B.L. Bush, J.A. McCammon, Biophys. J. 72 (1997) 1047–1069.
[102] J. Janin, Proteins 21 (1995) 30–39.
[103] Ajay, M.A. Murcko, J. Med. Chem. 38 (1996) 4593–4966.
[104] P.W. Atkins, Physical Chemistry, W.H. Freeman, New York, 1990.
[105] D.A. McQuarrie, Statistical Mechanics, Harper & Row, New York, 1976.
[106] M. Karplus, J.N. Kushick, Macromolecules 14 (1981) 325–332.
[107] W.F. van Gunsteren, P.K. Weiner, A.J. Wilkinson, Computer Simulations of Biomolecular Systems Theoretical and Experimental Applications, Escom Science Publishers, Leiden, 1989.
[108] K.A. Sharp, A. Nicholls, R.F. Fine, B. Honig, Science 252 (1991) 106–109.