Brief Articles

A Swift All-Atom Energy-Based Computational Protocol to Predict DNA−Ligand Binding Affinity and ΔT_m

Saher Afshan Shaikh and B. Jayaram*

Department of Chemistry, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi 110 016, India

Received May 10, 2006

A hybrid molecular mechanics−statistical mechanics−solvent accessibility-based computational protocol is developed to calculate DNA−ligand binding affinity without any database training and is validated on 50 DNA−ligand complexes. The calculated binding energies yield high correlation coefficients of 0.95 (R^2 = 0.90) and 0.96 (R^2 = 0.93) in linear plots against experimental binding free energies (ΔG^o) and ΔT_m, respectively. The protocol is translated into a swift, web-enabled, freely accessible computational tool, http://www.scfbio-iitd.res.in/preddicta, for ΔG^o and ΔT_m prediction for DNA−ligand complexes to aid and expedite rational drug design attempts.

Introduction

Energetics provides a vital link between structure and function of molecular systems. Correlating structural features with binding affinities allows the design of better binders via structure based drug design methods.[1] Thus, a computational tool that can determine the binding affinity accurately and expeditiously is of immediate value in drug design endeavors.

DNA is an important anticancer/antibiotic drug target, and recent years have seen an increasing number of ligands being designed to target specific DNA sequences, notable among which are polyamides, bi-benzimidazole derivatives, bisanthracyclines, bisfuramidines, and the more recent biarylpyrimidines (citations to original literature in Supporting Information). Several energy/scoring functions[2] based on force field, empirical, and knowledge based methods have been reported in literature for protein−ligand, protein−DNA, and protein−protein interactions and widely employed for docking and virtual screening and also for attempting binding affinity prediction,[3] but parallel efforts for DNA−ligand systems have begun only recently.[4] Approaches like MMPBSA and MMGBSA have also been widely employed for binding affinity calculation, but show limited success for DNA−ligand complexes.[5,6] Thus, for binding affinity predictions and for scoring DNA−ligand complexes, the need for fast and accurate methods persists.

Addressing this concern, we present here a computational methodology and web tool for generating swift estimates of DNA−ligand affinities while maintaining a phenomenological physicochemical approach. The energy function is validated on 50 DNA−ligand complexes containing 16 crystal and 34 model-built structures via linear correlations between theoretically calculated binding energies and experimentally determined binding free energies and ΔT_m (changes in melting, i.e., helix-to-coil transition temperature of DNA on drug binding) values reported in literature.

* To whom correspondence should be addressed. Tel.: +91-11-2659 1505 (office), +91-11-2659 6786 (lab). Fax: +91- 11-2658 2037. E-mail: bjayaram@chemistry.iitd.ac.in.

Theory

Accurate determination of absolute binding free energies in biomolecular systems,[7] especially in DNA−ligand complexes, has been a daunting task for computational chemists due to major challenges. First, modeling and simulation[8] of DNA requires careful consideration of solvent and counterion atmosphere to ensure system stability. Second, evolving a computational technique based on classical and statistical mechanical principles to obtain estimates of binding affinities that correlate well with experiment requires a thorough accounting of all possible energy components and their accurate estimation.[9] To this end, we propose an energy function, which represents the calculated DNA−ligand binding energy as a combination of electrostatic and steric complementarities as well as entropic and solvent effects, including hydrophobicity, and is expressed as

$$\Delta G^\circ_{cbe} = \Delta H^\circ_{el} + \Delta H^\circ_{vdw} - T\Delta S^\circ_{rt} + \Delta G^\circ_{w} \qquad (1)$$

Each component in eq 1 is described below.

ΔH°_{el}, Electrostatic Term. ΔH°_{el} includes interactions between partial atomic charges on the DNA and ligand atoms calculated from Coulomb's law, employing a sigmoidal dielectric function[10] for solvent screening and scaled phosphate charges according to Manning theory,[11] thus implicitly incorporating desolvation electrostatics and counterion effects. For a single time-averaged crystal/minimized/model-built structure, $\Delta H^\circ_{el} \approx \Delta E^\circ_{el}$. For atom i of DNA and j of the ligand

$$E_{el}\ (\text{kcal/mol}) = \frac{\sum_i \sum_j 332 q_i q_j}{D(r) r_{ij}}$$

where q_i and q_j are the partial atomic charges on atoms i and j, r_{ij} is the $i-j$ distance, $D(r)$ is a sigmoidal dielectric function taken as $D(r) = D - [(\alpha^2 + 2\alpha + 2)e^{-\alpha}(D - D_i)/2]$; D = 78; D_i = 4; $\alpha = sr$; and s = 0.395. Finally, $\Delta H^\circ_{el} \approx \Delta E^\circ_{el} = E^\circ_{el(complex)} - \{E^\circ_{el(DNA)} + E^\circ_{el(ligand)}\}$.

10.1021/jm060542c CCC: $37.00
Published on Web 04/10/2007

ΔH°_{vdw}, van der Waals Term. ΔH°_{vdw} includes direct van der Waals contributions from DNA−ligand interactions as well as desolvation van der Waals due to the interaction of solvent with complex, free DNA and ligand. For a single time-averaged crystal/minimized/model built structure, $\Delta H^\circ_{vdw} \approx \Delta E^\circ_{vdw}$ and is represented as, $E_{vdw} = E_{vdwdir} + E_{vdwdsl}$.

Direct van der Waals interactions were modeled using a (12, 6) Lennard−Jones potential between the atom i of DNA and j of ligand

$$E_{vdwdir}\ (\text{kcal/mol}) = \sum_i \sum_j \left[\frac{C_{12}^{ij}}{r_{ij}^{12}} - \frac{C_6^{ij}}{r_{ij}^{6}}\right]$$

where $C_{12}^{ij} = \epsilon_{ij}(R^*_{ij})^{12}$; $C_6^{ij} = 2\epsilon_{ij}(R^*_{ij})^6$; $R^*_{ij} = R^*_i + R^*_j$; $\epsilon_{ij} = (\epsilon_i\epsilon_j)^{1/2}$. ϵ_i is the well-depth parameter, and R^*_{ij} is the $i-j$ separation when energy is minimum. R_i^* and R_j^* values are adopted from the Cornell et al. force field.[12]

The desolvation van der Waals component is calculated[13] as

$$E_{vdwdsl}\ (\text{kcal/mol}) = -0.0398 \times \Delta A$$

where ΔA is the accessible surface area loss (in Å^2) of the DNA and ligand upon binding. Finally,

$$\Delta H^\circ_{vdw} \approx \Delta E^\circ_{vdw} = E^\circ_{vdw(complex)} - \{E^\circ_{vdw(DNA)} + E^\circ_{vdw(ligand)}\}$$

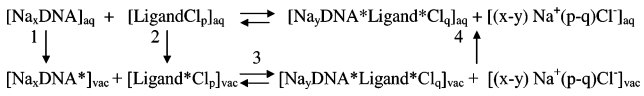
$T\Delta S^\circ_{rt}$ Term. $T\Delta S^\circ_{rt}$ represents the rotational and translational entropy changes on complex formation, calculated from the ideal gas partition function Q using statistical mechanics where

$$S_o = k \ln Q - (1/T)\,(\partial \ln Q/\partial\beta)_V$$

The translational partition function is computed as $Q_{trans} = (V/h^3)(2\pi m/\beta)^{3/2}$, where V is the volume, referring to a standard state of 1 M, $\beta = 1/kT$, k is the Boltzmann constant, T is the temperature, and m is the mass. The rotational partition function is computed as $Q_{rot} = \sigma^{-1}(1/\beta hc)^{3/2}(\pi/ABC)^{1/2}$, where σ is the symmetry number and A, B, and C are the rotational constants of the molecule.

The Sackur−Tetrode equation has been employed here for ΔS°_{rt} estimation. Though controversy surrounds the use of this equation in solution,[14] it is applicable in vacuum with reasonable accuracy on adopting the following thermodynamic cycle, which is theoretically consistent.

Thermodynamic cycle adopted to compute DNA−ligand binding affinity in aqueous media:[13] Here, * denotes the



structural adaptation/deformation as a consequence of binding. Steps 1 and 2 account for desolvation of DNA and ligand, along with counter and coions, step 3 involves DNA−ligand binding in vacuum and counterion release ($T\Delta S^\circ_{rt}$ is determined at this stage), and step 4 represents solvation of complex and counter and coions.

ΔG°_w, Hydration Term. ΔG°_w accounts for the energy associated with reordering of waters around DNA and ligand upon binding and is calculated from the change in hydration heat capacity (ΔCp) as[15]

$$\Delta G_w\ (\text{kcal/mol}) \approx T\Delta S = 0.080 \times \Delta\text{Cp}\ (\text{cal mol}^{-1}\ \text{K}^{-1})$$

ΔCp is calculated from the relation,[16]

$$\Delta\text{Cp}\ (\text{cal mol}^{-1}\ \text{K}^{-1}) = 0.382\sum_{np}\Delta A_{np} - 0.121\sum_{p}\Delta A_p$$

where ΔA_{np} and ΔA_p represent the changes in the accessible molecular surface area (ASA, in Å^2 units) for nonpolar (C, P, and H attached to C) and polar (N, O, and H attached to N and O) atoms, respectively.

The above relation is known in literature[15,16] and yields a high correlation between calculated and experimental ΔCp for DNA−ligand systems. Exceptions reported[17] for some complexes were attributed to the electrostatic contribution to hydration ΔCp, which does not scale with ASA. Also, a study[18] on nucleic acid bases shows that the polar contribution to desolvation is not accounted well by ASA-based ΔCp calculations and that ΔCp may be significantly more negative than estimated. Hence, we have employed a dielectric function in ΔH°_{el} and the ΔE°_{vdwdsl} term to calculate desolvation electrostatics and van der Waals, respectively, as already discussed.

Thus, the energetics attendant upon DNA−ligand binding is captured in eq 1, except for structural deformations and vibrational/configurational entropy changes, which require further methodological extensions. The energy function is phenomenological in nature, does not necessitate any training, and enables direct evaluation of the nonbonded energy of DNA−ligand interactions in an aqueous environment from the Cartesian coordinates of all atoms. Furthermore, it accommodates systematic theoretical improvements within the framework of additivity approximation.

Methods

DNA−ligand crystal structures present in RCSB (Protein Data Bank)[19] for which experimental data on ΔT_m and/or ΔG^o were available in literature were collected for this study. These comprise 16 minor groove binder−DNA complexes. Additionally, to augment the data set, models were built for 34 DNA−ligand complexes for which schematic structures and experimental ΔG^o or ΔT_m values were reported in literature, but crystal/NMR structures were not available.

The 16 crystal structure PDB files were prepared for analysis as follows. Hydrogen atoms were added explicitly and partial atomic charges were calculated using AM1BCC[21] methodology. GAFF[22] and Cornell et al.[12] force field parameters were assigned to ligand and DNA atoms, respectively. Waters of crystallization and ions were removed and counterions added to ensure electroneutrality along with an 8 Å box of TIP3P waters.[23] The systems were subjected to a series of minimizations (500 steps hydrogen minimization for complex, 5000 steps water minimization with 25 kcal/mol restraint on complex and ions, 5000 steps all-atom minimization, where the restraint was relieved more gradually on the complex than on the ions, and finally, 5000 steps free minimization) to alleviate any steric clashes in the crystal structures and to achieve the nearest stable low-energy conformations. These 16 minimized crystal structures were adopted for energy analysis.

Model building was carried out for 34 complexes, and these structures have been deposited at the RCSB (Protein Data Bank, www.rcsb.org) site in the theoretical models section. These can be accessed using the RCSB-assigned codes listed in Table S1 in Supporting Information. High-resolution crystal structures of DNA−ligand complexes were used as templates for building the starting structures. The DNA structures for all these complexes were available in the PDB, and only ligand structure generation/modification was required, which was carried out using the XLEAP module of AMBER8.[20] The ligands were modified/extended within the minor groove, incorporating binding site information wherever available in literature or considering the binding site to be similar to that in structurally related DNA−ligand crystal structures. These structures were then subjected to minimization with explicit solvent and counterions, following the protocol described above. Though low energy conformations and preferred orientations of functional

Table 1. Comparison of PreDDICTA with Commonly Employed Methods for Scoring/Binding Affinity Prediction

method	brief description	advantages of PreDDICTA	disadvantages of PreDDICTA
Force field *Scoring*	Comprises van der Waals and electrostatics terms	Includes entropic, solvent, and counterion effects	Computationally slower
Empirical *Scoring*	Constituent terms, e.g., hydrogen bonding, hydrophobic contacts, rotor terms, desolvation etc. are trained against experimental data; the resultant equation is adopted for scoring	Based on binding phenomenon, thus includes enthalpy and entropy Not trained, hence extensible to other systems	Computationally slower
Knowledge Based *Scoring*	Frequencies of occurrence of properties like interatomic contacts, pairwise potentials, etc., are determined across large data sets and adopted for scoring	Specific enthalpic and entropic contributions are determined Not trained/derived from data sets, hence extensible to other systems	Computationally slower
MMPBSA/ MMGBSA *Binding Affinity*	Sum of molecular mechanical (MM) terms, solvation electrostatics from Poisson−Boltzmann (PB) or generalized Born (GB) method, solvation van der Waals, and hydrophobicity from solvent accessibility (SA)	Involves a sigmoidal dielectric to account for solvation electrostatics, hence, is computationally swifter	PB or GB methods treat desolvation more rigorously
Linear Interaction Energy, LIE *Binding Affinity*	Binding affinity calculated from two parameterized terms, where parameters are derived from experimental data and applied on ensemble-averaged energy terms obtained from simulations	Computationally swifter because time intensive simulations are not required May be applied to single structures as well as MD trajectories	Explicit solvent effects not included
Free energy Simulations *Binding Affinity*	Binding affinity is directly determined via simulation employing techniques, e.g., free energy perturbation and thermodynamic integration	Computationally swifter because time intensive simulations are not required	Lower theoretical rigor Explicit solvent/ion effects not included

groups were considered, simple model building followed by minimization is not expected to reflect the conformation adopted by the ligand in a dynamic environment. Hence, the 34 minimized model-built structures were subjected to nanosecond scale explicit solvent molecular dynamics (MD) simulations. The final structures obtained at the end of each MD simulation were collected and examined for any major structural distortions and deviations from ideal bond, angle, or torsional geometries using NUcheck,[25] available at the RCSB (Protein Data Bank) site.

Thus, a total of 50 structures, comprising 16 energy-minimized crystal structures and 34 processed (minimization and MD) model-built structures were obtained. These structures were evaluated using CURVES[24] for DNA structural analysis, and no major deviations from similar DNA−ligand crystal structures were observed. Local inter-base pair parameters for the DNA for all complexes are reported in the Supporting Information. The DNA−ligand binding energy was calculated for each structure using eq 1.

Results and Discussion

The scoring function for DNA−ligand complexes proposed here was guided by our earlier experiences in development of MMGBSA[6,13,26] and MMPBSA[27] and empirical[28] approaches for scoring/binding affinity prediction for protein−DNA,[26,27] protein−ligand,[26,28] and DNA−ligand[6] systems, and we have aimed to capture the theoretical rigor of MMGBSA/MMPBSA methods, as well as the swiftness of simpler empirical/knowledge-based approaches. The former quality is necessary for good binding affinity prediction, while the latter is essential for a scoring function. The resultant of these efforts is an online tool named PreDDICTA (http://www.scfbio-iitd.res.in/preddicta), which is not aimed to replace any of these approaches, yet presents an attractive alternative in terms of computational speed, binding affinity prediction accuracy, and reliability of underlying theory (Table 1).

Experimentally determined binding free energies and ΔT_m (change in melting temperature of DNA upon drug binding) values are employed for a validation of the proposed computational protocol. The former are a direct measure of binding affinities, while the latter are closely related to binding affinities of DNA−ligand complexes.[29] Figure 1a illustrates the linear correlation obtained between the calculated DNA−ligand binding energy and the experimental ΔG^o (correlation coefficient,

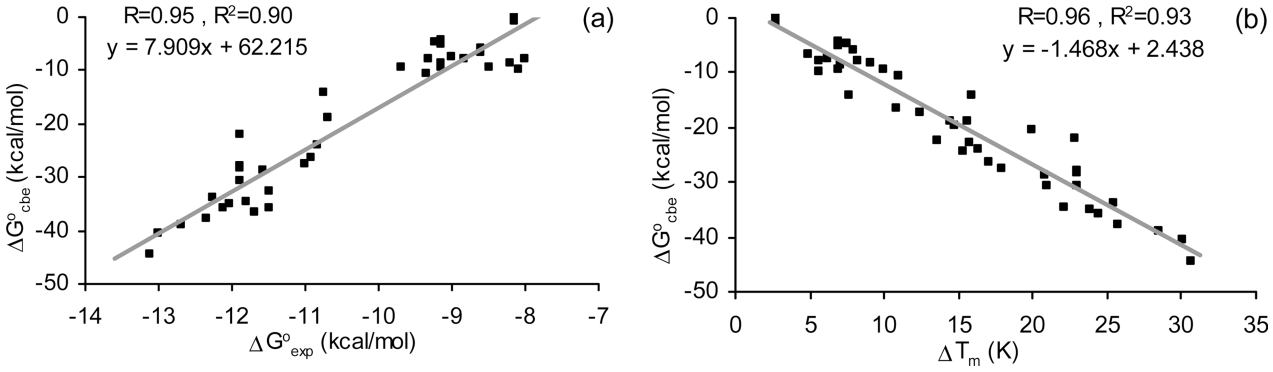


Figure 1. Calculated DNA−ligand binding energy (ΔG^o_{cbe}) versus experimentally determined (a) standard free energies of binding, ΔG^o_{exp} for 39 complexes, and (b) changes in the melting temperatures of DNA oligomers on ligand binding, ΔT_m for 44 complexes.

$R = 0.95$, $R^2 = 0.90$) for 39 complexes, and Figure 1b shows the correlation ($R = 0.96$, $R^2 = 0.93$) with experimental ΔT_m values for 44 complexes. The high correlations observed despite statistical noise associated with both the theoretical and the experimentally determined quantities indicate the overall stability and viability of the protocol adopted and the utility of this computational tool in a predictive mode.

The energy components, ΔH°_{el}, ΔH°_{vdw}, $T\Delta S^\circ_{rt}$, and ΔG°_w computed for each complex are reported in Table S2 in Supporting Information. Direct electrostatics and van der Waals interactions are highly favorable and, though largely compensated by desolvation effects, the overall terms, ΔH°_{el} and ΔH°_{vdw}, remain favorable in almost all cases, in conformity with earlier observations[6] in DNA−ligand systems. Rotational and translational entropy changes are expectedly unfavorable due to entropy losses on DNA−ligand binding. Release of ordered waters from free DNA and the drug appears to compensate for the entropy loss in reordering of waters around the complex, resulting in overall favorable hydration terms in all cases.

On consideration of each term individually, a moderate correlation of the experimental data is observed with ΔH°_{el} (with ΔG^o_{exp}, $R^2 = 0.70$, and ΔT_m, $R^2 = 0.51$), a strong correlation is observed with ΔH°_{vdw} (ΔG^o_{exp}, $R^2 = 0.80$; ΔT_m, $R^2 = 0.85$), and no correlation is observed with $T\Delta S^\circ_{rt}$ or ΔG°_w. From these observations, one may expect that ΔH°_{vdw} alone or alongwith ΔH°_{el} would suffice for a binding affinity prediction. However, we consider the inclusion of all four terms in the function necessary for two reasons: first, to maintain theoretical consistency, because they are based on the binding phenomenon, and second, for novel ligands, where entropy could play a stronger role than enthalpy in binding, $T\Delta S^\circ_{rt}$ and ΔG°_w would be more significant and their exclusion could lead to erroneous affinity predictions for such ligands.

Further statistical analyses of the energy terms using multiple linear regression indicate the relative weights of each term in the function in the resultant equation,

$$\Delta G^\circ = (0.19\Delta H^\circ_{el}) + (0.10\Delta H^\circ_{vdw}) - (0.19T\Delta S^\circ_{rt}) + (0.10\Delta G^\circ_w) - 10.08$$

The p values of <0.0001 for ΔH°_{el} and ΔH°_{vdw}, 0.2095 for $T\Delta S^\circ_{rt}$, and 0.0052 for ΔG°_w indicate that all four terms affect the correlation with high statistical significance, further supporting the inclusion of all the terms in the energy function. ANOVA analysis results show a very high F value of 106.34 and a correspondingly low p value of <0.0001, implying that the overall correlation also has high statistical significance (detailed discussion in Supporting Information).

The linear regression equations obtained from the correlations (Figure 1a,b) can be used to predict ΔT_m and ΔG°. This is shown in Table S2 in the Supporting Information, wherein the predicted values are reported alongside the experimental data for comparison. The equations are

$$\Delta G^\circ_{pred}(\text{predicted}) = (\Delta G^\circ_{cbe} - 62.215)/7.909 \quad (2)$$

$$\Delta T_m(\text{predicted}) = (\Delta G^\circ_{cbe} - 2.438)/(-1.468) \quad (3)$$

for binding affinity and ΔT_m prediction, respectively. It may be noted here that this approach differs from using a trained function, because the latter contains coefficients for each term derived from the training data set, while here only the final calculated binding energy is proportionately scaled/converted using the linear regression equation.

The energy function (1) and relations (2) and (3) have been incorporated in PreDDICTA. The tool takes a geometry optimized structure of the DNA−ligand complex in PDB format as input and proceeds with force field parameter and partial atomic charge assignment, followed by calculation of binding energies (ΔG°_{cbe}) and prediction of ΔT_m and ΔG°. The time taken for energy calculation is 0.05 to 0.20 min, while the complete process, which includes charge derivation (AM1BCC) and parameter assignment, takes 0.08 to 9.00 min on an AMD Opteron processor, depending on the ligand structure, size, and conformation in the complexes in the data set.

It was observed during this study that a few bis-amidinium ligand (i.e., positively charged NH_2CHNH_2 group at both terminals of ligand) DNA complexes do not show conformity to the observed correlation. Experimental data for three complexes of $d(CGCGAATTCGCG)_2$ bound to DAPI (PDB ID: 1D30), Hoechst derivative (311D), and a model-built symmetric bi-benzimidazole derivative (2GXV) indicates very high binding affinity (ΔT_m is 17.1, 32.8, and 22.8 K, respectively), which is underestimated in the calculated binding energies (ΔG°_{cbe} is −6.8, −23.6, and −22.1 kcal/mol, respectively). The reason for this is not clear, but it may be due to very strong electrostatic interactions at the amidinium ends, which are not adequately accounted for by the general route of partial atomic charge assignment and electrostatic calculations adopted here and needs system-specific attention. It may also be due to certain phenomena under experimental conditions not reflected in the structure, hence not visible from the calculation. However, it is of interest to note that mono-amidinium containing ligands (2GXJ, 2GXK, 2IRJ, 2GY6, 2GY8, and 2GYE in Table S1 in Supporting Information) bind with affinities that correlate well with those predicted by this energy function.

Independence of the energy function from training on any particular system or structure and the observed strong correlation obtained with experimental data on a dataset of 50 complexes leads to the expectation that not many exceptions may be anticipated when this computational model is employed on any structure outside the data set studied. Also, this data set is fairly diverse structurally, containing mono- as well as dicationic ligands, including seven types of molecules (Supporting Information): bisfuramidines, phenylbenzimidazoles, symmetric and asymmetric substituted bi-benzimidazoles, polyamides, substituted triazene (berenil), and propamidine. The data set also includes six different oligomeric sequences, and though the sequence $d(CGCGAATTCGCG)_2$ dominates, the binding site varies from three to eight base pairs within the set of ligands. The energy function performs well when applied separately on each type of drug in the data set, indicating its extensibility, which is discussed in greater detail in the Supporting Information. Additionally, preliminary results (work in progress) emerging from the application of this tool on 130 protein−ligand complexes indicate a high correlation ($r = 0.88$) between the calculated and experimental binding free energies, strongly supporting the extensibility of the method to systems differing significantly from the data set considered. Further validation and systematic improvements can be envisioned with more diverse ligand structures and DNA sequences as and when the corresponding experimental data is available. As for now, the energy function and the associated computational protocol are expected to perform reasonably well on DNA−ligand complexes, especially when a series of probable drug candidates are to be ranked in terms of their binding affinities.

Conclusions

A computational tool that accounts for diverse energy contributors to DNA−ligand binding affinities at the atomic

level has been developed and tested on 50 DNA−ligand complexes. The correlations obtained between predicted and experimental ΔG^{o} and ΔT_m data are quite satisfactory for the protocol to be harnessed in a predictive mode. The suite of programs for predicting binding affinity and ΔT_m from any input structure of a DNA−ligand complex has been organized as an online tool named PreDDICTA and made freely accessible at http://www.scfbio-iitd.res.in/preddicta. The tool has potential value in scoring/binding affinity prediction and assessment of binding properties of newly designed ligands targeted to DNA.

Acknowledgment. Funding from the Department of Science and Technology, Government of India, is gratefully acknowledged. S.A.S. is a recipient of the Senior Research Fellow award from the Council of Scientific and Industrial Research, Govt. of India. The authors also thank Shailesh Tripathi, Anuj Gupta, and Vidhu S. Pandey for help in web-enabling the tool.

Supporting Information Available: Citations to original literature on DNA-directed ligands, complex crystal structures, and experimental binding affinity data; calculated energy components, ΔG^{o}_{cbe} and ΔT_m values; chemical structures of the ligands studied; DNA structural analysis data for the 50 complexes; detailed discussion on extensibility/generality of the method; and results of the multiple regression analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Brooijmans, N.; Kuntz, I. D. Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335−373.
(2) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discovery* **2004**, *3*, 935−949. Schulz-Gasch, T; Stahl, M. Scoring functions for protein-ligand interactions: A critical perspective. *Drug Discovery Today* **2004**, *1*, 231−239.
(3) Warren, G. L.; Andrews, C. W.; Capelli, A; Clarke, B; LaLonde, J; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M; Peishoff, C. E; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49*, 5912−5931.
(4) Ge, W.; Schneider, B.; Olson, W. K. Knowledge-based elastic potentials for docking drugs or proteins with nucleic acids. *Biophys. J.* **2005**, *88*, 1166−1190.
(5) Spackova, N.; Cheatham, T. E.; Ryjacek, F.; Lankas, F.; Van Meervelt, L.; Hobza, P.; Sponer, J. Molecular dynamics simulations and thermodynamics analysis of DNA−drug complexes: Minor groove binding between 4′,6-diamidino-2-phenylindole and DNA duplexes in solution. *J. Am. Chem. Soc.* **2003**, *125*, 1759−1769.
(6) Shaikh, S. A.; Ahmed, S. R.; Jayaram, B. A molecular thermodynamic view of DNA−drug interactions: A case study of 25 minor-groove binders. *Arch. Biochem. Biophys.* **2004**, *429*, 81−99.
(7) Lazaridis, T. Binding affinity and specificity from computational studies. *Curr. Org. Chem.* **2002**, *6*, 1319−1332. Cooper, A. Thermodynamic analysis of biomolecular interactions. *Curr. Opin. Chem. Biol.* **1999**, *3*, 557−563.
(8) Jayaram, B.; Beveridge, D. L. Modeling DNA in aqueous solutions: Theoretical and computer simulation studies on the ion atmosphere of DNA. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 367−394. Beveridge, D. L.; McConnell, K. J. Nucleic acids: Theory and computer simulation, Y2K. *Curr. Opin. Struct. Biol.* **2000**, *10*, 182−196. Orozco, M.; Perez, A.; Agnes, N.; Javier, L. F. Theoretical methods for the simulation of nucleic acids. *Chem Soc. Rev.* **2003**, *6*, 350. Cheatham, T. E., III. Molecular dynamics simulation of nucleic acids. *Curr. Opin. Struct. Biol.* **2004**, *14*, 360−367.
(9) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The statistical-thermodynamic basis for computation of binding affinities. A critical review. *Biophys. J.* **1997**, *72*, 1047−1069.
(10) Arora, N.; Jayaram, B. Energetics of base pairs in B-DNA in solution: An appraisal of potential functions and dielectric treatments. *J. Phys. Chem. B.* **1998**, *102*, 6139−6144. Young, M. A.; Jayaram, B.; Beveridge, D. L. Local dielectric environment of B-DNA in solution: Results from a 14 ns molecular dynamics trajectory. *J. Phys. Chem. B.* **1998**, *102*, 7666−7669.
(11) Manning, G. S. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* **1978**, *11*, 179−246.
(12) Cornell, W. D.; Cieplak, P.; Bayly, C. F.; Gould, I. R.; Kenneth, M. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second-generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(13) Jayaram, B.; McConnell, K. J.; Dixit, S. B.; Beveridge, D. L. Free energy analysis of protein−DNA binding: The EcoRI endonuclease−DNA complex. *J. Comput. Phys.* **1999**, *151*, 333−357.
(14) Yu, Y. B.; Privalov, P. L.; Hodges, R. B. Contribution of translational and rotational motions to molecular association in aqueous solution. *Biophys. J.* **2001**, *81*, 1632−1642.
(15) Spolar, R. S.; Ha, J. H.; Record, M. T., Jr. Hydrophobic effect in protein folding and other noncovalent processes involving proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8382−8385. Lane, A. N.; Jenkins, T. C. Thermodynamics of nucleic acids and their interactions with ligands. *Q. Rev. Biophys.* **2000**, *33*, 255−306.
(16) Ren, J.; Jenkins, T. C.; Chaires, J. B. Energetics of intercalation reactions. *Biochemistry* **2000**, *39*, 8439−8447.
(17) Gallagher, K.; Sharp, K. A. Electrostatic contributions to heat capacity changes of DNA−ligand binding. *Biophys. J.* **1998**, *75*, 769−776.
(18) Madan, B.; Sharp, K. A. Hydration heat capacity of nucleic acid constituents determined from the random network model. *Biophys. J.* **2001**, *81*, 1881−1887.
(19) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(20) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California, San Francisco: San Francisco, CA, 2004.
(21) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC Model: I. Method *J. Comput. Chem.* **2000**, *21*, 132−146.
(22) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, A. P.; Case, D. A. Development and testing of a general Amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(24) Lavery, R.; Sklenar, H. Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn.* **1989**, *6*, 655−667.
(25) Feng, Z; Westbrook, J; Berman, H. M. *NUCheck*; Rutgers University: New Brunswick, NJ, 1998; Report No. NDB-407.
(26) Kalra, P.; Reddy, T. V.; Jayaram, B. Free energy component analysis for drug design: A case study of HIV-1 protease-inhibitor binding. *J. Med. Chem.* **2001**, *44*, 4325−4338. Jayaram, B.; McConnell, K.; Dixit, S. B.; Das, A.; Beveridge, D. L. Free energy component analysis of 40 protein−DNA complexes: A consensus view on the thermodynamics of binding at the molecular level. *J. Comput. Chem.* **2002**, *23*, 1−14.
(27) Jayaram, B.; Rajasekaran, E. Counterion condensation in DNA systems: The cylindrical Poisson−Boltzmann model revisited. *Biopolymers* **1994**, *34*, 443−445. Dixit, S. B.; Jayaram, B. Role of hydrogen bonds in protein−DNA recognition: A comparison of generalized born and finite difference Poisson−Boltzmann solvation treatments. *J. Biomol. Struct. Dyn.* **1998**, *16*, 237−242.
(28) Jain, T.; Jayaram, B. An all atom energy based computational protocol for predicting binding affinities of protein-ligand complexes. *FEBS Lett.* **2005**, *579*, 6659−6666.
(29) Crothers, D. M. Statistical thermodynamics of nucleic acid melting transitions with coupled binding equilibria. *Biopolymers* **1971**, *10*, 2147−2160.

JM060542C