Supplementary Materials and Methods for:

Bound Compound, Interfacial Water, and Phenyl

Ring Rotation Dynamics of a Compound

in the DNA Minor Groove

Narinder K. Harika and W. David Wilson*

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

Address correspondence to either author

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

Supporting Information

Methods and Materials:

Force fields are key foundations of molecular mechanics with the fundamental assumption of Born-Oppenheimer approximation[1], where the energy of a system is described as a function of nuclear coordinates only. The accuracy of force fields further depends on the validity of additional assumptions; additivity and transferability[1]. Additivity corresponds to the expression of the potential energy of any system as a sum of different potentials and transferability refers to the development of the potential energy functions on a small set of molecules which can be applied to a much wider range of molecules with similar chemical groups. However, there is the limited availability of parameters for organic molecules in AMBER force fields[2]. The equation to represent the additive form of force fields is as follows:

$$E_{total} = \sum_{bonds} K_r (r - r_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2$$
$$+ \sum_{torsions} \frac{V_n}{2} [1 \pm \cos(n\phi - \gamma)]$$
$$+ \sum_{non-bonded} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{C_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{r_{ij}} \right] \qquad (1)$$

The force field involving separation of potential energy terms in the equation 1 is empirical and provides useful insight into the determination of molecular structure and dynamics. The energy terms used in the functional form of the force field are described below:

The first term in the equation 1 represents the bond stretching and is denoted with a harmonic potential. K_r represents the force constant of the bond (generally very high), is an indication of the amount of energy required to stretch or compress a chemical bond significantly. r_0 represents the reference bond length, defined by the value of the bond length when all other terms in the potential energy function are zero.

S2

The second term in the above equation 1 represents the change in potential energy in relation to deformation of bond angles. K_φ represents the force constant for the bond angle with values typically less than for bond stretching and is an indication that it takes less energy for a bond angle to deviate from its reference value.

The third term in the above equation 1 is referred to as the torsional term and is represented as the change in the potential energy of the molecular system as a function of the rotation about each dihedral angle. The energies involved here are significantly lower than for bond stretching and angle bending. V_n represents the barrier height, n represents the number of minima in the energy function (multiplicity) and γ is the phase factor, which determines the position of the minima.

The fourth term in the above equation 1 is represented by the non-bonded interactions involving atoms in the molecule separated by 3 or more bonds and between atoms in the different molecule. This term includes both the electrostatic interaction through the Coulombic potential, and the van der Waals interaction using the Lennard-Jones 12-6 potential. The van der Waals potential contains an attractive and a repulsive term. The attractive part represents the dispersion forces generated between instantaneous dipoles, which arise from fluctuations in electronic charge distributions in all molecules. The repulsive term reflects the observation that atoms repel each other below a certain distance (typically close to 0.3 nm). The electrostatic potential term is represented by the sum of electrostatic potentials generated by charges placed on atomic nuclei, i.e., partial atomic charges where r_{ij} is the distance between the nuclei i and j.

The foundation of classical force fields assumes that that similar chemical groups in different molecules interact in the same way[3]. Therefore, the development of a force field is based on reproducing the energy surfaces for a set of small molecules that are typically well-characterized and contain the functional groups that occur in all biomolecules. This assumption gives a force field very broad applicability. Thus, force field parametrization expansion for new atom and bond types plays a very important role in molecular mechanics. For the internal terms in the force field function, equilibrium

S3

bond length and equilibrium bond angles are mostly obtained from the experiment and high-level *ab-initio* calculations whereas force constants are obtained through an empirical approach.[2]

The first step in the generation of force field parameters includes specification of atom type of the molecule. Usage of more specific atom types helps in describing the chemical environment of molecule more accurately at the cost of burdening the parametrization.[4] Selection of the consistent or appropriate charge approach is necessary to accurately fit the conformational and non-bonded energies in a transferable way. Default charge scheme used in the AMBER force fields (or GAFF parametrization), RESP "Restrained electrostatic potential" at HF/6-31G* works well with the small molecules.[5-7].

Force Field Parameters for DB2277:

Most of the parameters in the force field parameter files for the DB2277 molecule were derived from the existing set of bonds, angles and dihedrals for the similar atom types in parm99 and GAFF force fields.[4, 8] The assignment of specifying atom types for the DB2277 molecule used in generating the force field parameters is shown in the figure S3 as previously reported NMR structure of DB2277-DNA complex.[9] Optimized geometry of DB2277 from Gaussian 09[10] was used to obtain the equilibrium bond length for CK-CA and the equilibrium bond angle for CA-CA-N2. Torsional angles of DB2277 listed below play a significant role in defining its conformation and its interaction with DNA;

1) NB-CK-CA-CA and N2-CA-CA-CA: Obtained from previously reported parameters for DAPI and DB921 molecules[11,12]

2) N*-CK-CA-CA, CT-OS-CA-CA, CT-OS-CA-NC CA-CT-OS-CA, CA-CA-CT-H1 and CA-CA-CT-OS: Obtained from previously reported parameters for DB2277 in DB2277-DNA NMR structure[9]

Since the dihedral angles, CA-CA-CT-OS, CA-CT-OS-CA, CT-OS-CA-NC and N*-CK-CA-CA are not defined in GAFF/Parm99 force field, therefore these dihedrals were re-

S4

parametrized and were reported earlier. These dihedrals define the rotation around bonds connecting the phenyl group and aza-benzimidazole moiety and are crucial in predicting the dynamic rotational movements of the phenyl and amidine moieties. Spartan16[13] was used to perform *ab-initio* calculations using DFT/6-31+G (d)[14] level of theory for defining the above mentioned torsional angle parameters. Potential energy profiles were obtained for dihedral angles varying from 0 to 180 degrees[11, 12] and were fitted with the least square fitting in kaleidaGraph software using the following cosine function equation:

$$\sum V_n/2 \left[1 + \cos\left(n\Phi - \gamma\right)\right]$$

where V_n is the torsional barrier, n is periodicity (0,1,2,3,4 or 6), and γ is phase angle.

The frcmod file (force field modification) modified with these newly parametrized dihedral angles of DB2277 can be found in the previously published NMR structure of DB2277-DNA complex.[9]
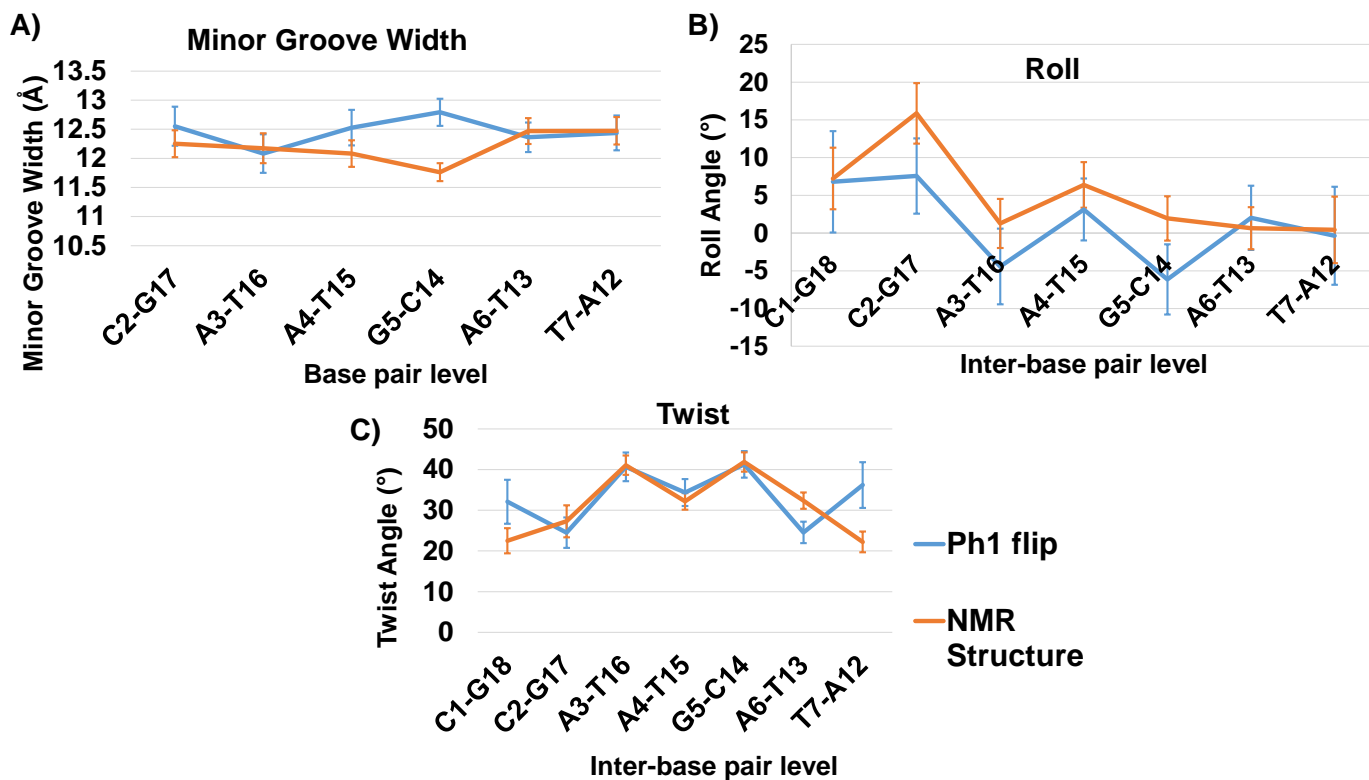
Figure S1. A) Minor groove width (in Å) of bound DNA for 1-ns trajectory (blue) in comparison to the minor groove width of DNA in the NMR structure of DB2277-DNA complex (orange). B) Roll angles and C) Twist angles at inter-base pair level for the bound DNA in 1ns Ph1-flip trajectory compared to roll and twist angles of DNA in NMR structure of DB2277-DNA complex (orange). Standard deviation calculated from all time steps in 1-ns trajectory of Ph1 flip.
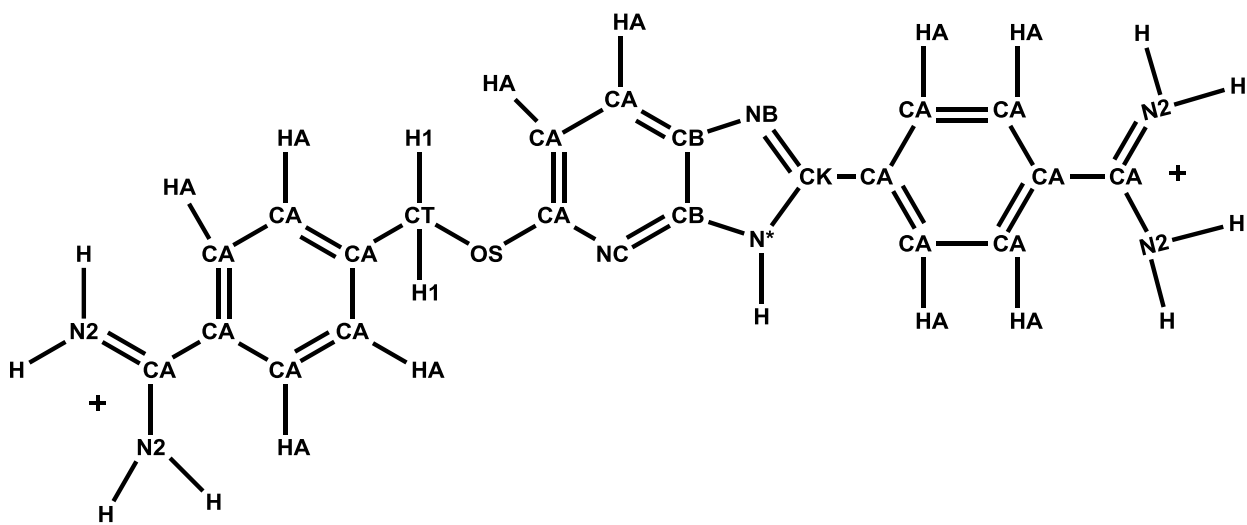
Figure S2. Molecular structure with specific atom types used for the DB2277 molecule in the DB2277-DNA NMR structure.

References:

1) Leach, A. (2009) Molecular modelling: principles and applications, Pearson Prentice Hall.

2) Wang, J., Wolf, R. M.; Caldwell, J. W.;Kollman, P. A.; Case, D. A. "Development and testing of a general AMBER force field". Journal of Computational Chemistry, 25, 2004, 1157-1174.

3) L. Monticelli and D. Peter Tieleman, Force Fields for Classical Molecular Dynamics, 2013, p. 197-213.

4) Wang, J., Wang, W., Kollman P. A.; Case, D. A. "Automatic atom type and bond type perception in molecular mechanical calculations". Journal of Molecular Graphics and Modelling, 25, 2006, 247260.

5) Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, The Journal of Physical Chemistry 97, 10269-10280.

6) Cornell, W.D., Cieplak, P., Bayly, C.I. and Kollmann, P.A. (1993) Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. Journal of the American Chemical Society, 115, 9620-9631.

7) Singh UC, Kollman PA (1984) An approach to computing electrostatic charges for molecules. J Comput Chem 5(2):129–145.

8) Pérez, A., Marchán, I., Svozil, D., Sponer, J., Cheatham, T.E., Laughton, C.A. and Orozco, M. (2007) Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. Biophysical Journal, 92, 3817-3829.

9) Harika, N. K., Germann, M. W., and Wilson, W. D. (2017) First Structure of a Designed Minor Groove Binding Heterocyclic Cation that Specifically Recognizes Mixed DNA Base Pair Sequences, Chemistry – A European Journal 23, 17612-17620.

10) Gaussian 09, Revision B.01 Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT (2009) by M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al.

S8

11) Athri, P. and W. D. Wilson (2009). "Molecular Dynamics of Water-Mediated Interactions of a Linear Benzimidazole−Biphenyl Diamidine with the DNA Minor Groove." Journal of the American Chemical Society 131(22): 7618-7625.

12) Špačková, N.a., Cheatham, T.E., Ryjáček, F., Lankaš, F., van Meervelt, L., Hobza, P. and Šponer, J. (2003) Molecular Dynamics Simulations and Thermodynamics Analysis of DNA−Drug Complexes. Minor Groove Binding between 4', 6-Diamidino-2-phenylindole and DNA Duplexes in Solution. Journal of the American Chemical Society, 125, 1759-1769.

13) Computational Chemistry, David Young, Wiley-Interscience, 2001. Appendix A. A.1.6 pg 330, Spartan.

14) Orio, M., Pantazis, D.A. and Neese, F. (2009) Density functional theory. Photosynthesis Research, 102, 443-453.