Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters 18 (2008) 542–545
Bioorganic & Medicinal Chemistry Letters

Binding free energy calculation for duocarmycin/DNA complex based on the QPLD-derived partial charge model

Haizhen Zhong,[a] Karl N. Kirschner,[b] Moses Lee[c] and J. Phillip Bowen[a,*]

[a]*Center for Drug Discovery, Department of Chemistry and Biochemistry, The University of North Carolina at Greensboro, Greensboro, NC 27402, USA*
[b]*Department of Chemistry, Center for Molecular Design, Hamilton College, Clinton, NY 13323, USA*
[c]*Division of Natural Sciences and Department of Chemistry, Hope College, Holland, MI 49423, USA*

Received 7 August 2007; revised 20 November 2007; accepted 21 November 2007
Available online 28 November 2007

Abstract—The 3 ns unrestrained MD simulations were carried out on the DNA/duocarmycin complex based on (1) the classic RESP charge model, and (2) the QM-polarized ligand docking (QPLD)-based charge model. The RMSDs of the trajectories and the ΔG_{bind} of the QPLD model perform much better than the RESP model, with the ΔG_{bind} estimation for QPLD model (−16.11 kcal/mol) versus ΔG_{bind} estimation for RESP model (−10.05 kcal/mol).

It is now well recognized that electrostatic interactions play a critical role in noncovalent interactions between DNA/ligand, protein/ligand, and DNA/protein complexes. A combination of structural feature studies and binding affinity calculations makes it possible to design better ligands with high affinity. The development of computational approaches to predict accurately the binding free energy, therefore, is of great significance. Molecular mechanics/Poisson–Boltzmann surface area (MM/PBSA) and molecular mechanics/generalized Born surface area (MM/GBSA) have been widely employed for the binding affinity calculations for the protein/ligand and/or protein/peptide interactions.[1,2] In addition, the thermodynamic integration techniques[3] have been used by Spiegel et al.[4] to predict the binding free energy (ΔG_{bind}) for (+)-duocarmycin SA (Fig. 1) binding to the DNA duplex [*d*(G_1ACTAATTGAC$_{11}$)–*d*(G_{12}TCA ATTAGTC$_{22}$)]. The ΔG_{bind} was estimated to be −10.5 kcal/mol, which is a severe underestimation of the experimental value of −22.7 kcal/mol.[4] Spiegel et al. attributed the large discrepancy to the level of accuracy of DFT used in their hybrid Car-Parinello QM/MM simulations.[4] Due to the highly charged nature of DNA, we assumed that the electrostatic interactions between DNA and its ligand might play a critical role in the accurate estimation of the binding free energy. With an aim to develop a method that can estimate the binding free energy with improved accuracy, we present here the molecular dynamics simulations and the MM/GBSA calculations that were carried out based on the quantum mechanics (QM)-polarized ligand docking (QPLD)-derived partial charges. To our best knowledge, this is the first QPLD-based MM/GBSA calculation. We are very excited to observe that our QPLD-based MM/GBSA calculations have shown a significant increase in the accuracy of ΔG_{bind} estimation.

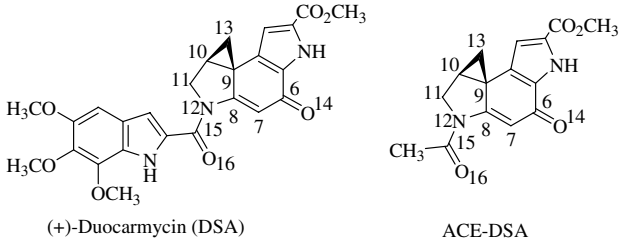

Figure 1. Chemical structures of (+)-duocarmycin (DSA) and *N*-acetyl-DSA (ACE-DSA).

The classic MD simulations use the RESP (restrained electrostatic potential) charges, which are routinely developed from an isolated ligand through the quantum mechanics calculations at the 6-31G* level using Gauss-

Keywords: MD simulations; DNA/duocarmycin complex; RESP charge model; QM-polarized ligand docking (QPLD); MM/GBSA.
* Corresponding author. Tel.: +1 336 334 4257; fax: +1 336 334 5402; e-mail: jpbowen@uncg.edu

0960-894X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2007.11.090

ian 94/98/03. The minimized structure is then used to calculate a molecular electrostatic potential (MEP) on a three-dimensional grid, and the final step is to derive the 'RESP' and 'ESP' charges for the ligand.[5] The RESP charges developed from an isolated ligand might be problematic for the DNA/ligand interactions due to the extent of electrostatic contributions. DNA is an important target for developing anticancer/antibiotic compounds, including duocarmycin SA (DSA, Fig. 1),[6] polyamides,[7] and biarylpyrimidines.[8] An accurate estimation of the ΔG_{bind} for DNA/ligand and DNA/protein complexes, therefore, will greatly benefit our understanding of the interactions and result in a better design of ligands as medical agents.

The importance of accurate charges in molecular docking has been documented in a series of docking studies using the QM dock method.[9] The QM dock (with the QM/MM charges) appears to have an advantage over other docking methods. MM/PBSA and MM/GBSA have been acclaimed to provide the absolute binding free energy for protein/ligand complexes. Will the accuracy of ΔG estimation be improved with the QM/MM charges? To answer this question, we carried out unrestrained molecular dynamics simulations of a selected DNA duplex of oligomers complexed with duocarmycin SA (DSA), using the classical RESP charges (called the RESP model), or using the QPLD-derived charges (called the QPLD model). We then analyzed the structural fluctuations in terms of distances and dihedral angles that are involved in the binding. The binding free energies for the two models were estimated with the MM/GBSA method.

The original structure of the DNA/DSA complex was taken from the first of the 20 NMR structures in 1DSA.[10] The reversible, stereoelectronically controlled N3′ of adenine 19 (A19) covalently binds to the least substituted cyclopropane carbon (C13 in DSA, Fig. 1).[11] For the MD simulation purpose, the C13—N3′ covalent bond was modified to a non-bonded interaction. The DSA was extracted from 1DSA and optimized using the ab initio density functional theory (DFT) with the 6-31G* basis set developed by Pople and coworkers,[12] and the hybrid B3LYP method.[13,14] To observe the effect of the substituted indole ring, DSA was modified to an acetylated structure (ACE-DSA) and optimized at the same B3LYP/6-31G* level of theory using Gaussian 03.[15] The RESP charges were developed based on the optimized DSA using HF/6-31G* level of theory, following the established RESP protocol in the AMBER program. The QPLD charges were derived from the QM-polarized ligand docking protocol in the Schrödinger software suite.[16] The QPLD protocol involves the usage of the Glide and the QSite programs in the Schrödinger suite. DSA was docked to DNA using the Glide program. Each configuration of the docked DNA/DSA was evaluated using the mixed quantum mechanical/molecular mechanics (QM/MM) computational routine, where DSA was treated by the QM approach and DNA was treated by the MM method. The docked pose with the lowest root mean square deviation (RMSD) from the reference ligand DSA structure was adopted and the original DSA was replaced with this newly obtained conformation. The partial charges associated with the adopted pose were used for the MD simulations.

MD simulations were carried out on two DNA/DSA systems: one with the RESP charges for DSA (RESP model), the other with the QPLD-derived charges for DSA (QPLD model). The AMBER force fields parm99[17] within the AMBER 9 package[18] were used for the simulation. The SANDER module of AMBER 9 was used for the initial minimizations and the subsequent MD simulations. All systems were solvated with neutralizing Na^+ counterions by a rectangular box of TIP3P water,[19] which extended at least 10.5 Å away from any given DNA or DSA atom. The minimizations were carried out first on the water molecules for 1000 steps while holding the solute and the counterions fixed, followed by minimizations on the complex for 1000 steps while water molecules were held fixed. The whole system was then relaxed for 2000 steps of minimization. The resulting system was heated from 10 K to 300 K in two intervals over 30 ps, followed by a 100 ps equilibration run for the whole system. During the heating process and the equilibration phase, the DNA/DSA/Na^+ were applied with a force constant constraint of 10 kcal/mol Å^2. During the first ns of MD production run, counterions Na^+ were restrained with a force constant of 10 kcal/mol Å^2, while the water, DNA, and DSA were allowed to relax. After the first ns, the whole system including the Na^+ ions was carried out without any constraint. During the MD simulations, the periodic boundary conditions were applied. The particle mesh Ewald method[20] was used to treat the long-range electrostatics. The van der Waals (vdw) interactions were treated with a 9.0-Å cutoff. The SHAKE algorithm was used to constrain all bonds involving hydrogen atoms. The trajectories of the last 3 ns snapshots were processed with the PTRAJ program in AMBER 9, and compared to the NMR structure by the root-mean-square deviations (RMSD) of the main chain atoms (Fig. 2). The graphics of model structures were made with the PYMOL program.[21]

The RMS deviations from the reference 1DSA, on average, are larger and experience more deviations in the RESP model than those in the QPLD model. The average RMSD for the RESP model is 2.43 , with standard

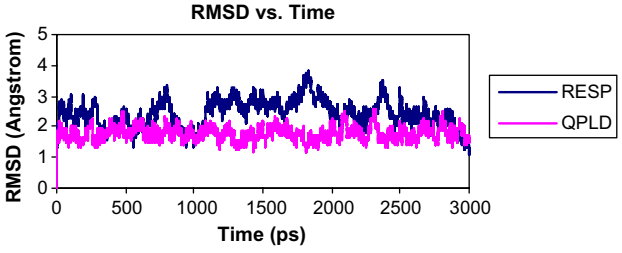

Figure 2. The root-mean-square deviations for the trajectories in comparison of the NMR structure 1DSA based on the DNA backbone atoms. The RMSD trajectories are colored in blue for the traditional partial charge (RESP) model, in magenta for the QPLD model.

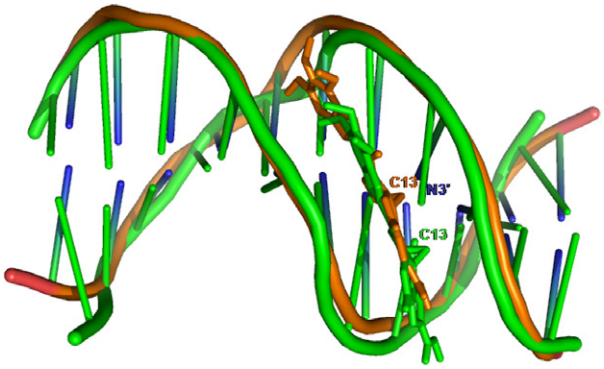

Figure 3. The cartoon structures for the DNA/DSA complexes. The complex structures were taken from the 4000th ps snapshot. The structure from the RESP model is colored in green, while the QPLD model in orange.

deviation of 0.45 . The average RMSD for the QPLD model, on the other hand, is 1.78 , with the standard deviation 0.22 (Fig. 2). The MD simulations clearly show the improved stability of the simulations and better convergence toward the NMR structure (1DSA) in the QPLD model. The C13—N3′ distance for the RESP model is 3.58 ± 0.32 , for the QPLD model, 3.32 ± 0.20 (Fig. 3). The closer distance between C13 and N3′ in the QPLD model can be rationalized by the fact that the covalent bond that is formed in the NMR structures necessitates a close interaction between these two atoms.

The probability of bonding depends on the stability of the cyclopropyl unit in the DSA, which can be measured by torsional angles χ1(C7—C8—N12—C15) and χ2(C8—N12—C15—O16). The more distorted χ1 and χ2 indicate the less stability of the cyclopropyl ring and the higher reactivity of DSA. Boger and Garbaccio have suggested that the larger distortion in DSA leads to the higher reactivity over (+)-duocarmycin SI (DSI), in which the three methoxyl groups of DSA are absent.[22] Our QM calculations reveal that torsions χ1 and χ2 in isolated DSA are 19.6° and 7.9°, respectively. The corresponding torsions in the *N*-acetyl substituted DSA (ACE-DSA) are 13.5° and 4.8°, respectively. The decreased π-electron conjugation in the DSA, caused by the sterically bulky substituted indole ring, reduces the stability of the cyclopropyl unit and enhances its reactivity. The average torsions χ1and χ2 in the DNA/DSA complex for the RESP model are 37.8 ± 11.2° and 3.7 ± 8.3°, for the QPLD models, 33.5 ± 10.4° and 11.7 ± 7.7°, for the NMR structure (1DSA), 22.2° and 11.7°, respectively. Our simulations show that binding of DSA to DNA is accompanied by the reduction in the vinylogous amide conjugation. The χ1 torsion increased from 19.6° to 33.5° (in QPLD model) or 37.8° (in RESP model) for DSA during the binding to DNA. This ground state destabilization of the ligand upon binding to DNA has been observed by other scientists.[10,21] Our QPLD model simulations for χ2 (11.7°) are in better agreement with the NMR data than the RESP model (3.7°). Therefore, structurally, the QPLD method clearly shows greater advantage over the RESP model.

The MM/GBSA method was used to estimate the ΔG_{bind} for the two studied models. In the QPLD model, the ΔG_{bind} is estimated to be −16.11 kcal/mol (Table 1), while the ΔG_{bind} of the RESP model is estimated to be −10.05 kcal/mol (Table 2). The experimental ΔG_{bind} of the DNA/DSA was estimated to be −22.70 kcal/mol. The thermodynamic integration (TI) calculation by Spiegel et al. for the DNA/DSA with the RESP charges was estimated to be −10.50 kcal/mol.[4] Our estimation on the ΔG_{bind} for the RESP charge model based on the MM/GBSA method is almost identical to Spiegel's calculations, which were based on the TI approach. The RESP charge model in both our calculations and Spiegel's calculations gave the large discrepancy to the experimental data. The QPLD-based charge model, however, performs much better (−16.11 kcal/mol), which corresponds to a 60% increase of the accuracy relative to −10.50 kcal/mol for the RESP model! The main contributor that leads to such an increase is the electrostatic contribution (Tables 1 and 2). The ΔE (electrostatic) for the QPLD model is −38.57 kcal/mol. The corresponding ΔE (electrostatic) for the RESP model is −26.21 kcal/mol. Obviously, drastic improvements in accuracy of ΔG estimation have been achieved using the QPLD based charges. The improved performance in the QPLD model was also substantiated by the less RMS deviations of the trajectories and closer distance between C13 and N3′. The improvement can be attributed to the electrostatic partial charges in the QPLD

Table 1. Energy components and free energy of binding for DNA/DSA complex using the QPLD charge model

Contributions (kcal/mol)	Complex		DNA		DSA		Δ	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
E (electrostatic)	693.03	31.64	694.23	31.68	37.37	1.22	−38.57	3.90
E (vdw)	−245.97	9.82	−192.79	9.02	8.41	2.17	−61.60	2.70
E (internal)	1318.75	20.75	1036.55	20.09	282.20	5.54	0.00	0.00
E (gas)	1765.81	34.28	1538.00	34.53	327.98	5.42	−100.17	4.04
G (GBsur)	30.92	0.29	32.02	0.22	5.49	0.03	−6.60	0.17
G (GB)	−5273.19	27.47	−5304.68	27.92	−33.51	0.86	65.00	3.45
G (GBsol)	−5242.27	27.60	−5272.65	28.00	−28.02	0.85	58.40	3.45
G (GBele)	−4580.16	11.10	−4610.45	10.82	3.85	1.06	26.43	1.81
G (GBtot)	−3476.46	20.18	−3734.65	19.53	299.96	5.36	−41.76	2.41
−TS	−590.27		−554.58		−61.34		25.65	
ΔG_{bind}							−16.11	

Table 2. Energy components and free energy of binding for DNA/DSA complex using the RESP charge model

Contributions (kcal/mol)	Complex		DNA		DSA		Δ	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
E (electrostatic)	713.21	36.84	671.60	36.15	67.82	1.19	−26.21	3.92
E (vdw)	−240.71	11.12	−188.73	9.00	8.37	2.26	−60.35	5.16
E (internal)	1317.94	21.32	1034.22	20.56	283.72	5.57	0.00	0.00
E (gas)	1790.44	38.57	1517.08	39.13	359.92	5.44	−86.56	5.45
G (GBsur)	31.53	0.51	32.31	0.23	5.48	0.04	−6.26	0.40
G (GB)	−5261.95	31.68	−5285.53	31.92	−30.50	0.85	54.08	3.51
G (GBsol)	−5230.42	31.95	−5253.22	32.02	−25.01	0.85	47.82	3.46
G (GBele)	−4548.74	11.84	−4613.93	11.08	37.33	0.97	27.87	2.42
G (GBtot)	−3439.97	20.65	−3736.14	20.04	334.91	5.37	−38.74	4.07
−TS	−588.43		−555.59		−61.53		28.69	
ΔG_{bind}							−10.05	

model. The electrostatic partial charges in the QPLD model were derived within the binding pocket, and therefore contain the environmental polarization effects, which are normally ignored in an isolated ligand system.

In summary, the QPLD-based charge model provided a better performance in the MD simulations. The QPLD-based ΔG_{bind} demonstrated a significant increase in accuracy. Application of this new method will greatly enhance the design of ligands, such as polyamides, with more potency. The authors predict that the application of the QPLD-based charge model and the QPLD-based MM/GBSA method will greatly promote our understanding of the protein/DNA and protein/ligand interactions, and will greatly assist structure-based drug design. We are investigating the application of the QPLD-based charge model in the protein/ligand complex, and protein/DNA complexes and will report the results elsewhere.

References and notes

1. Massova, I.; Kollman, P. A. *J. Am. Chem. Soc.* **1999**, *121*, 8133.
2. Zhong, H.; Carlson, H. A. *Proteins: Struct., Func., Bioinfo.* **2005**, *58*, 222.
3. Ivanov, I.; Klein, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 13380.
4. Spiegel, K.; Rothlisberger, U.; Carloni, P. *J. Phys. Chem. B* **2006**, *110*, 3647.
5. Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1357.
6. Schnell, J. R.; Ketchem, R. R.; Boger, D. L.; Chazin, W. J. *J. Am. Chem. Soc.* **1999**, *121*, 5645.
7. Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horick, S. M.; Uthe, P. B.; Le, N. M.; Cox, K. K.; Nguyen, B.; Pacheco, K. A. O.; Wilson, W. D.; Lee, M. *J. Am. Chem. Soc.* **2005**, *127*, 742.
8. Wheelhouse, R. T.; Jennings, S. A.; Phillips, V. A.; Pletsas, D.; Murphy, P. M.; Garbett, N. C.; Chaires, J. B.; Jenkins, T. C. *J. Med. Chem.* **2006**, *49*, 5187.
9. Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. D. *J. Comput. Chem.* **2005**, *26*, 915.
10. Eis, P. S.; Smith, J. A.; Rydzewski, J. M.; Case, D. A.; Boger, D. L.; Chazin, W. J. *J. Mol. Biol.* **1997**, *272*, 237.
11. Boger, D. L.; Johnson, D. S.; Yun, W. *J. Am. Chem. Soc.* **1994**, *116*, 1635.
12. Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys.* **1993**, *98*, 5612.
13. Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
14. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
15. Gaussian 03, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Pittsburgh, PA, 2003.
16. QPLD, Schrödinger, LLC: Portland, OR, 2006.
17. Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
18. AMBER 9, Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A.; AMBER 9, University of California, San Francisco, URL: http://amber.scripps.edu, 2006.
19. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J. *Chem. Phys.* **1983**, *79*, 926.
20. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
21. PYMOL, DeLano Scientific LLC. http://pymol.sourceforge.net/.
22. Boger, D. L.; Garbaccio, R. M. *Acc. Chem. Res.* **1999**, *32*, 1043.