Communication

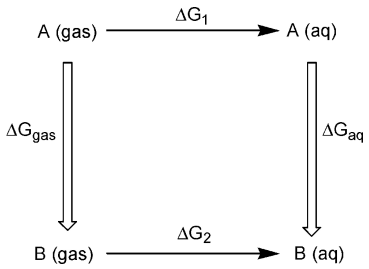
Development of a Quantum Mechanics-Based Free-Energy Perturbation Method: Use in the Calculation of Relative Solvation Free Energies

M. Rami Reddy, U. C. Singh, and Mark D. Erion

J. Am. Chem. Soc., **2004**, 126 (20), 6224-6225• DOI: 10.1021/ja049281r • Publication Date (Web): 22 April 2004

$$E_{tot} = E_{QM} + E_{MM} + \sum_{i=1}^{M}\sum_{j=1}^{L} E_{QM/MM}^{ij}$$

$$\Delta\Delta G_{sol} = \Delta G_{aq} - \Delta G_{gas}$$



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 4 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 04/22/2004

Development of a Quantum Mechanics-Based Free-Energy Perturbation Method: Use in the Calculation of Relative Solvation Free Energies

M. Rami Reddy,*,† U. C. Singh,‡ and Mark D. Erion†

Metabasis Therapeutics, Inc., 9390 Towne Centre Drive, Building 300, San Diego, California 92121, and Azaya Therapeutics, Inc., 130 West Rhapsody Drive, San Antonio, Texas 78216

Received February 9, 2004; E-mail: reddy@mbasis.com

Free-energy perturbation (FEP) is considered the most accurate computational method[1] for calculating relative solvation[2] and binding[3] free-energy differences. Despite impressive results generated over the past two decades and some successes applying FEP methods to both drug design[4] and lead optimization,[5,6] FEP calculations are rarely used in the pharmaceutical industry. The primary reason is frequently attributed to the inability of current methods to feasibly evaluate large numbers of compounds over a relatively short period of time. While the high CPU demand of these calculations is one major factor limiting throughput, another important factor relates to the availability of validated molecular mechanics force field parameters. Since most drug candidates contain substructures not fully described by existing parameters, the user must develop and input parameters prior to initiating the calculation. The process is time-consuming and often limited by the absence of relevant experimental data. Moreover, the process is difficult to automate since it is highly dependent on the structure and requires considerable user expertise and judgment.

A strategy with the potential to generate accurate force field parameters, possibly in an automated manner, entails use of quantum mechanics (QM) to describe the ligand structure and molecular mechanics (MM) to describe the surrounding environment (solvent, protein). Coupling of the QM and MM subsystems, QM/MM, has been used successfully to characterize transition-state structures.[7] In this communication, we describe the integration of QM/MM with FEP (QM/MM-based FEP) within the program Galaxy (AM Technologies, Inc.)[8] and the relative solvation free energies calculated using this method for a series of structurally diverse molecules.

As in conventional FEP methods, relative solvation free energies were calculated by transforming solute A into solute B using the λ coupling method.[1] The thread technique,[4−5,9] which is used frequently in conventional FEP calculations for mapping structurally dissimilar molecules, was used in all transformations in this work and proved to be essential for our success. As illustrated in Figure 1, "threading" together phenylalanine and isoleucine results in a dual topology in which the portions of the solutes being transformed are described by topologies that for one solute start and the other solute end the simulation entirely as dummy atoms.[6a] Dummy atoms are identical to real atoms except that their Lennard-Jones parameters and charges are set to zero. At intermediate points during the transformation, all atoms in both topologies have fractional Lennard-Jones parameters and charges.

Using the thread technique with the conventional FEP method entails scaling the MM parameters according to λ and calculating the corresponding MM energies. In contrast, the QM/MM-based FEP method uses either ab initio or semiempirical QM methods to calculate the energies and forces for the solute in the system and

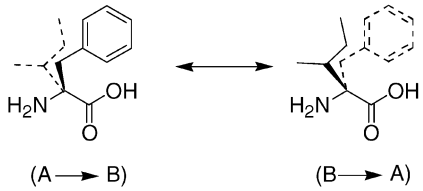


Figure 1. Dual topology definition for phenylalanine (A) to isoleucine (B). Common atoms are represented by a single topology (amino acid backbone). For $\lambda = 0$, the noncommon atoms for solute B are "dummy atoms" as represented by the dashed structure. Hydrogen atoms are removed for clarity.

MM methods to calculate the energies of the solvent (environment). To calculate the QM energy, we implemented a procedure that separated the threaded molecule into two molecules (A and B) at each dynamic step. QM forces and energies were then computed and scaled based on λ using eqs 1 and 2, respectively.

$$f_\lambda^i = \lambda f_A^i + (1-\lambda) f_B^i \tag{1}$$

$$E_\lambda^{QM} = \lambda E_A^{QM} + (1-\lambda) E_B^{QM} \tag{2}$$

The total energy for the system was determined using eq 3 wherein the term $E_{QM/MM}$ represents the interaction energy for an atom i in the MM part of the system and an atom j in the QM part of the system. The free-energy change (eq 4) is decomposed into the free-energy contribution from the subsystem treated by QM and the free-energy contribution from the surroundings, i.e., the subsystem not treated by QM (non-QM or NQM).

$$E_{tot} = E_{QM} + E_{MM} + \sum_{i=1}^{M}\sum_{j=1}^{L} E_{QM/MM}^{ij} \tag{3}$$

$$\Delta G_{tot} = \Delta G_{QM} + \Delta G_{NQM} \tag{4}$$

Calculations using the conventional and QM/MM-based FEP methods were performed using procedures previously described.[5,6a] The solutes were immersed in a 13 Å box of equilibrated SPC/E water[10] and energy minimized. A molecular dynamics (MD) time step of 1 fs was used to minimize the incidence of premature termination of the simulation due to poor convergence of the QM energies. The energy-minimized system was equilibrated with 20 ps of MD at constant temperature (300 °K) and pressure (1 atm), periodic boundary conditions in all directions, SHAKE to constrain all bond lengths and a nonbonded interaction cutoff of 12 Å. Fifty-one windows were used for each electrostatic and van der Waals free-energy difference. Each window consisted of 1.5 ps of equilibration and 3 ps of data collection. Accordingly, each calculation required 479 ps to complete the transformation. Errors were estimated for each window by dividing the window statistics

† Metabasis Therapeutics.
‡ Azaya Therapeutics.

10.1021/ja049281r CCC: $27.50

Table 1. Relative Solvation Free Energies (kJ/mol)

transformation[a]	$\Delta\Delta G$(AM)[b]	$\Delta\Delta G$(QM)[c]	$\Delta\Delta G$(FP)[d]	$\Delta\Delta G$(E)[e]
$CH_3OH \rightarrow EtH$	18.9 ± 1.9	29.7 ± 1.8	31.5 ± 1.9	29.0
$AcCH_3 \rightarrow AcNH_2$	-19.6 ± 2.2	-27.9 ± 2.3	-28.8 ± 2.2	-24.8
$C_6H_6 \rightarrow C_6H_5OH$	-18.6 ± 1.6	-22.2 ± 1.7	-21.6 ± 1.7	-23.5
$C_6H_6 \rightarrow C_6H_5NH_2$	-15.0 ± 1.5	-17.8 ± 1.4	-18.2 ± 1.4	-16.7
$CH_3CCl_3 \rightarrow EtH$	9.5 ± 2.2	9.1 ± 2.1	9.3 ± 2.2	8.4
$C_6H_6 \rightarrow C_5H_5N$	-9.9 ± 1.6	-13.2 ± 1.5	-12.6 ± 1.5	-14.7
Ser → Cys	13.8 ± 1.8	18.8 ± 1.8	18.5 ± 1.8	16.0
Phe → Ile	8.8 ± 3.0	10.7 ± 3.2	10.5 ± 3.1	12.1
Cyt → Thy	20.2 ± 2.0	25.6 ± 2.3	24.1 ± 2.2	
Ade → Gua	-24.5 ± 1.8	-38.1 ± 2.6	-36.3 ± 2.6	

[a] Abbreviations: Cyt = cytidine, Thy = thymidine, Ade = adenine, Gua = guanine. [b] Calculated using AM1 for both the gradients and ESP partial atomic charges. [c] Calculated using AM1 for gradients and ab initio (HF/6-31G*)/ESP for partial atomic charges. [d] Calculated using a conventional FEP method and HF/6-31G*/ESP for partial atomic charges. [e] Values obtained from experimental data reported in the literature.[11]

into four groups and computing the standard deviation.[4b] The reported standard deviation is the root-mean-square of the window errors.

Solvation free-energy differences for 10 molecular pairs were calculated and compared to results from experimental data ($\Delta\Delta G$(E))[11] and/or results using conventional FEP methods ($\Delta\Delta G$(FP)) (Table 1). Two methods were used in the QM/MM-based FEP calculations. Both relied on AM1 semiempirical quantum mechanics for calculating energies and gradients at each MD step. One method used AM1/ESP derived partial atomic charges ($\Delta\Delta G$(AM)), whereas the other method used HF/6-31G*/ESP derived partial atomic charges ($\Delta\Delta G$(QM)).[12]

A comparison of relative solvation free energies computed using the QM/MM-based FEP method ($\Delta\Delta G$(QM)) with the experimental values suggested or values obtained from conventional methods showed that the QM/MM-based FEP method was at least as accurate, if not more accurate. In addition, a comparison of the standard deviations for the calculated results revealed no differences in calculation variability. In contrast, results obtained using the QM/MM-based FEP method ($\Delta\Delta G$(AM)) generally underestimated the experimental value presumably due to an underestimation of the electrostatic contribution to the solvation free energy. The underestimation occurs particularly with polar solutes and is attributed to an underestimation of the solute dipole moment.

Accurate results were obtained using $\Delta\Delta G$(QM) for molecular pairs with large differences in structure and aromaticity (Phe → Ile), polarity (methanol → ethane), hydrogen bonding (acetone → acetamide; adenine → guanine), and the total number of electrons (1,1,1-trichloroethane → ethane). Moreover, the results were accurate across a relative solvation free-energy difference range of 9.1 to 38.1 kJ/mol.

Not surprising, the QM/MM-based FEP method required 3- to 5-fold more CPU than the conventional FEP method to complete the calculations on the molecular pairs chosen for this study. Further increases in the CPU demand are expected for calculations involving molecules with a greater number of atoms and/or electron-rich atoms and for calculations using ab initio QM for determining both the gradients and energies. Accordingly, CPU remains a potential concern associated with the QM/MM-based FEP method but one that may be minimized in the future if the exponential increases in CPU power observed over the past two decades continues. Moreover, additional CPU power will be gained following parallelization of the code to enable simultaneous use of multiple processors.

In summary, we developed a FEP method that uses QM for treating the solute, MM for treating the solute surroundings, and the FEP method for computing free-energy differences. Relative solvation free energies for 10 molecular pairs were calculated, and the results were in close agreement with both the calculated results generated by conventional FEP methods and the experimentally derived values. While considerably more CPU demanding than conventional FEP methods, QM/MM-based FEP alleviates the need for development of molecule-specific MM force field parameters and therefore may enable future automation of FEP-based calculations. Moreover, calculation accuracy should be improved over conventional methods especially for calculations reliant on MM parameters derived in the absence of experimental data.[13,14] These advances could further the use of FEP methodology in drug discovery programs and thereby aid in more accurate drug candidate prioritization.

Supporting Information Available: Dipole moments and electrostatic and van der Waals free energies (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Zwanzig, R. J. *J. Chem. Phys.* **1954**, *22*, 1420. (b) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281. (c) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395.

(2) (a) Rao, B. G.; Singh, U. C. *J. Am. Chem. Soc.* **1989**, *111*, 3125. (b) Jorgensen, W. L.; Nguyen, T. B. *J. Comput. Chem.* **1993**, *14*, 195.

(3) (a) Bash, P.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574. (b) Brooks, C. L., III; Fleischman, S. H. *J. Am. Chem. Soc.* **1990**, *112*, 3307. (c) Rao, B. G.; Tilton, R. F.; Singh, U. C. *J. Am. Chem. Soc.* **1992**, *114*, 4447. (d) Reddy, M. R.; Erion, M. D.; Agarwal, A. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 2000; Vol. 16, pp 217−304.

(4) (a) Erion, M. D.; van Poelje, P. D.; Reddy, M. R. *J. Am. Chem. Soc.* **2000**, *122*, 6114. (b) Erion, M. D.; Reddy, M. R. *J. Am. Chem. Soc.* **1998**, *120*, 3295.

(5) For reviews, see: *Free Energy Calculations in Rational Drug Design*; Reddy, M. R., Erion, M. D., Eds.; Kluwer/Plenum Press: New York, 2001.

(6) (a) Reddy, M. R.; Erion, M. D. *J. Am. Chem. Soc.* **2001**, *123*, 6246. (b) Plout Price, M. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 9455. (c) Reddy, M. R.; Varney, M. D.; Kalish, V. N.; Viswanadhan, V. N.; Appelt, K. *J. Med. Chem.* **1994**, *37*, 1145. (d) Reddy, M. R.; Viswanadhan, V. N.; Weinstein, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10287.

(7) (a) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718. (b) Harrison, M. J.; Burton, N. A.; Hillier, I. H. *J. Am. Chem. Soc.* **1997**, *119*, 122.

(8) Galaxy Molecular Modeling Software and AM2000 Macromolecular Simulation package; AM Technologies: San Antonio, Texas, 1995.

(9) (a) Singh, U. C.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 9518. (b) Reddy, M. R.; Bacquet, R. J.; Zichi, D.; Mathews, D. A.; Welsh, K. M.; Jones, T. R.; Freer, S. *J. Am. Chem. Soc.* **1992**, *114*, 10117.

(10) (a) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269. (b) Reddy, M. R.; Berkowitz, M. *Chem. Phys. Lett.* **1989**, *155*, 173.

(11) (a) Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 292. (b) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Sol. Chem.* **1981**, *10*, 563. (c) Wolfenden, R. *Biochemistry* **1981**, *20*, 849.

(12) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 43.

(13) Efforts are underway to compare solvation free energies calculated using partial atomic charges derived from AMSOL. (a) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 87. (b) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2003**, *24*, 1291.

(14) We have also modified the QM/MM-based FEP method so that both gradients as well as partial atomic charges are obtained using ab initio methods.

JA049281R