J Mol Model (2011) 17:383–389
DOI 10.1007/s00894-010-0728-2

ORIGINAL PAPER

Docking study and binding free energy calculation of poly (ADP-ribose) polymerase inhibitors

Kazuki Ohno · Takashi Mitsui · Yoshiaki Tanida · Azuma Matsuura · Hideaki Fujitani · Tatsuya Niimi · Masaya Orita

Received: 26 February 2010 / Accepted: 21 April 2010 / Published online: 18 May 2010

Abstract Recently, the massively parallel computation of absolute binding free energy with a well-equilibrated system (MP-CAFEE) has been developed. The present study aimed to determine whether the MP-CAFEE method is useful for drug discovery research. In the drug discovery process, it is important for computational chemists to predict the binding affinity accurately without detailed structural information for protein / ligand complex. We investigated the absolute binding free energies for Poly (ADP-ribose) polymerase-1 (PARP-1) / inhibitor complexes, using the MP-CAFEE method. Although each docking model was used as an input structure, it was found that the absolute binding free energies calculated by MP-CAFEE are well consistent with the experimental ones. The accuracy of this method is much higher than that using molecular mechanics Poisson-Boltzmann / surface area (MM / PBSA). Although the simulation time is quite extensive, the reliable predictor of binding free energies would be a useful tool for drug discovery projects.

Kazuki Ohno and Takashi Mitsui equally contribute to this work.

K. Ohno · T. Niimi · M. Orita (✉)
Drug Discovery Research, Astellas Pharma Inc.,
21 Miyukigaoka, Tsukuba,
Ibaraki 305-8585, Japan
e-mail: masaya.orita@jp.astellas.com

T. Mitsui
Fujitsu Ltd.,
1-9-3 Nakase, Mihama,
Chiba 261-8588, Japan

Y. Tanida · A. Matsuura · H. Fujitani
Fujitsu Laboratories Ltd.,
10-1 Morinosato-Wakamiya,
Atsugi 243-0197, Japan

Keywords Binding free energy calculation · MM-PBSA · MP-CAFEE · Poly (ADP-ribose) polymerase inhibitors

Introduction

Poly (ADP-ribose) polymerase-1 (PARP-1, EC 2.4.2.30) is a chromatin-bound nuclear enzyme protein involved in a number of cellular processes, including mainly DNA repair and programmed cell death [1]. Inhibition of PARP-1 has been shown to be effective in animal models of stroke, traumatic brain injury, Parkinson's disease and cancer [2–4]. Therefore, PARP-1 inhibitors may be useful in therapies of neurodegenerative disorders, cancers, and several other diseases involving PARP-1 activation. Up to now, several potent PARP-1 inhibitors have been described in structure-based discovery [4, 5]. Because much information regarding SAR and crystal structure is available, PARP-1 is an ideal system to evaluate the binding affinity prediction method.

In the area of computational drug design, the accurate prediction of protein-ligand binding affinities is one of the most important problems [6, 7]. Many methods have been presented to calculate the binding free energies of protein-ligand complexes. MM-PBSA (molecular mechanics Poisson-Boltzmann surface area) and MM-GBSA (molecular mechanics generalized Born surface area) are commonly used methodologies to obtain the binding energies in this field [8, 9]. These methods have several appealing features: First, these are significantly faster than the other methods such as free energy perturbation (FEP) and thermodynamics integration (TI). Second, the protein flexibility is partially taken into account; however, accuracy is limited because many important contributions are not well considered. For example, Friesner's group showed that

the entropy contribution of water molecules in active site pockets is essential to explain the structure-activity relationship of Factor Xa [10]. On the other hand, FEP and TI may be the most rigorous and strict techniques in use [11, 12]. These methods have been used frequently to calculate the free energy difference between similar molecules. However, they are very inefficient for obtaining absolute binding energies.

Recently, Jarzynski proved that the distribution of nonequilibrium work values could yield an equilibrium free energy, based on the exponential average of the set of nonequilibrium work values [13]. However, it is difficult to obtain accurate free energy by exponential averaging, because the results strongly depend on the behavior at the tails of the work value distribution. Shirts *et al*. demonstrated that the Bennett acceptance ratio (BAR) method can be interpreted in terms of the maximum likelihood estimate of the free-energy difference, given a set of nonequilibrium work values in the forward and reverse directions [14]. The MP-CAFEE [15] method was developed as a combination of the above procedure without the auxiliary restraint of keeping ligands in the binding pockets.

In the previous studies, the MP-CAFEE method successfully reproduced the rank order of binding affinities of ligands [15, 16]. However, the computation has not been performed for cases where the atomic coordinates of a complex system are somewhat ambiguous. From the viewpoint of application to drug discovery research, it is very important to discern whether this method can be applied to protein/ligand complexes for which the binding modes are unclear.

In this study, we examined the performance of MP-CAFEE to predict the free energy difference induced by ligand binding to a protein without detailed crystal structure information for the protein/ligand complex. Using a docking program, we first built PARP-1/ligand complex structures, and then investigated the binding free energies, using the MP-CAFEE method. Although we did not use crystal structures but structures built by the docking program, the binding free energies calculated by MP-CAFEE are consistent with the experimental values. In contrast, the calculated values by using the MM-PBSA method are inconsistent with the experimental values. The results indicate that MP-CAFEE is a promising binding energy calculation method for the field of drug discovery.

Methods

Data preparation

Figure 1 shows six PARP-1 inhibitors used in this study. Experimental IC50 values range from 16 nM to 1200 nM [17]. The initial atomic coordinates were taken from the Protein Data Bank (PDB). To build the free form of PARP-1, we removed the ligand from the PARP-1/C1 complex (1UK0) [18]. Hydrogen atoms were added by using a molecular modeling tool, the Molecular Operating Environment (MOE 2006.08) [19]. All the ionizable residues (arginine, lysine, asparatic acid, and glutamic acid) were charged. The ligand structures were built and optimized using MOE. Each ligand was docked with the free form of PARP-1 in a docking program, Genetic Optimization for Ligand Docking (GOLD version 3.2) [20].

Molecular dynamics simulations

Molecular dynamics (MD) simulations were performed using the GROningen MAchine for Chemical Simulations (GROMACS) package of version 3.3.1 for equilibration, version 3.1.4 with modification for free energy calculation [21–23]. The Amber force field parameters (ff99) were used to describe the PARP-1, including 1–4 interactions between hydrogen atoms. The atomic structures of ligands were optimized in a vacuum, by means of the General Atomic and Molecular Electronic Structure System (GAMESS, released Mar 24, 2007) [24, 25] with the HF/6-31 G* energy calculation. The atomic partial charges were calculated by the restrained electrostatic potential (RESP) method [26] with the RESP module of ANTECHAMBER (1.27) [27]. The general Amber force field (GAFF) parameters [28] were used to describe the ligand, and the atom types were assigned by the FF-FOM program (FUJITSU Ltd.) [29]. The chemical structures of the ligands investigated here are shown in Fig. 2. The TIP3P water model was used to describe the solvent [30].

All MD simulations were carried out at 298 K with Nose-Hoover temperature control [31, 32] with a time constant 0.5 ps, and at 1 atm with Berendsen pressure control [33] with a time constant 1.0 ps and a compressibility of 4.5×10^{-5}. All simulations were run in a truncated octahedron simulation box. The solvated ligand system had 566-854 water molecules and the solvated PARP-1-ligand complex system had 12020-12037 water molecules. The

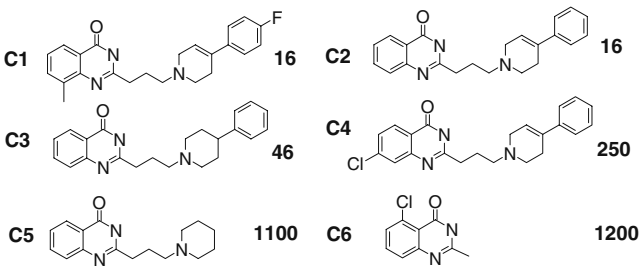

Fig. 1 Structures of six ligands investigated. Experimental IC50 values in nM on the right of the structure [17]

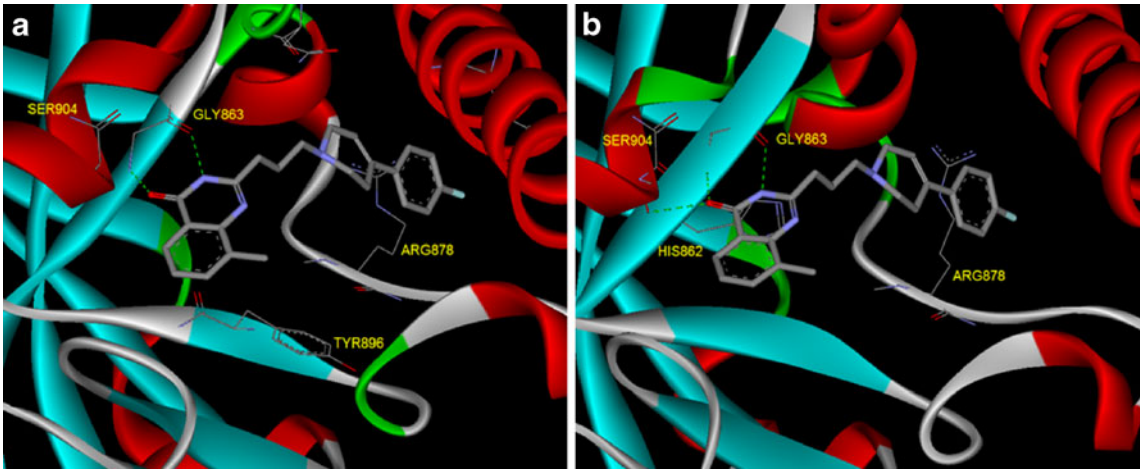

Fig. 2 Complex structure of PARP-1/C1 (**a**) before 20 ns simulation, and (**b**) after 20 ns simulation

complex system had a chlorine ion to neutralize the charge of PARP-1.

A neighbor list of 1.1 nm was utilized with an update frequency of 10 fs, and van der Waals interactions were 0.9 nm and 1.0 nm distance. Particle mesh Ewald (PME) with an interpolation order of 4 was used for long-range Coulomb interaction. The Fourier spacing was approximately 0.12 nm. A long-range correction for the finite cutoff of the Lennard-Jones potential was taken into account for energy and pressure corrections. To constrain all the bonds, the LINCS algorithm [34] with order 8 and 2 fs time step was used.

After energy minimization using steepest descent method, we performed a molecular dynamics simulation for 200 ps with the solute position restraint. We performed multiple equilibration with different initial momentum for 20 ns while monitoring the intermolecular potential energy between PARP-1 and a ligand, and carefully checked whether the solvated system reached a stable state.

Binding free energy calculations

With the well-equilibrated structure, we performed massively parallel calculation of binding free energy. Only the intermolecular interactions between the ligand and other molecules were annihilated using the soft-core potential. The optimal non-uniform λ−spacing was used. We first turned off Coulomb interaction at 12 λ−points, and then van der Waals interaction at 21 λ−points. For each λ−point, we performed 12 MD simulation with different initial momentum distribution.

We also performed MM-PBSA calculation using AMBER 9 [35] in order to compare our results with the conventional approach. The atomic coordinates of PARP-1 and each ligand were extracted from a single trajectory of molecular dynamics simulation with explicit water molecules for equilibration, and each energy term of MM-PBSA was evaluated for 4 ns of the trajectory from 16 ns to 20 ns. The dielectric constant for the surrounding solvent was 80.0. The entropy term usually estimated by normal mode analysis was omitted.

Computational cost

To calculate the absolute binding energy values of six inhibitors, we conducted 2376 molecular dynamics simulation jobs (2376=33 λ×12 samples×6 compounds). The simulation time per one job ranged from 110 hours to 140 hours using an Intel Xeon5160 processor at 3 GHz. In total, the simulation time reached 260,000 – 330,000 hours, which corresponds to 50-70 days of simulations with an Intel 100 CPU (200 core) Xeon5160 (dual core) 3 GHz cluster system.

Results and discussion

Equilibrium MD calculation

Figure 2a shows the PARP-1/C1 complex structure built using the docking program, GOLD. In this docking model, the backbone of Gly 863 has two hydrogen bonds with C1, which is consistent with the crystal structure (1UK0). Although the docking mode of the model of PARP-1/C1 complex is almost the same as that of the crystal structure, the RMSD value of C1 between the docking model and the crystal structure is not negligible (0.11 nM). This docking model was used as an initial coordinate for MD calculation. Figure 2b shows the PARP-1/C1 complex structure after 20-ns MD simulation. Again, there are two hydrogen bonds between Gly 863 and C1, which means the docking mode

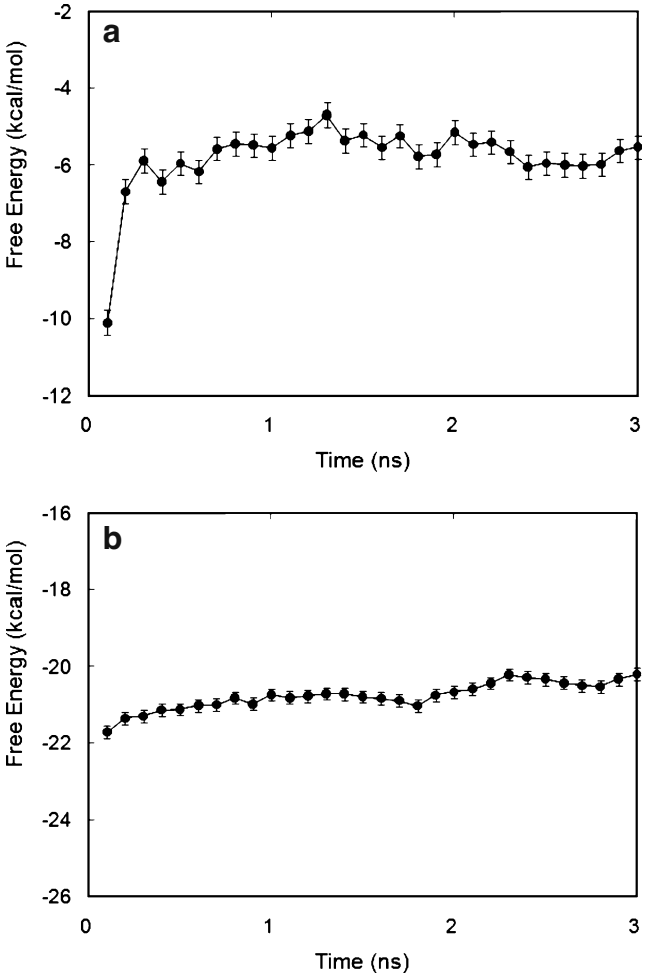

Fig. 3 Time dependency of calculated free energies for C1 (**a**) van der Waals interaction and (**b**) Coulomb interaction

before and after MD simulation is almost conserved. The total RMSD value of all heavy atoms is 0.11 nM, which indicates that there is not a large conformational change. The same situations were observed for the other ligands.

Absolute binding free energy

Figure 3 shows the convergence of the free energy of PARP-C1 complex over time. Figure 3a and b show the van der Waals and the Coulomb interaction, respectively. Both interactions reached convergence after 2 ns calculation. Each value of interaction was obtained as the average value after convergence. The absolute binding energy was obtained as the sum of each interaction. Figure 4 and Table 1 show the calculated absolute binding free energies of inhibitors and the experimental values that were obtained using $\Delta G=-RT\ln IC50$ [17]. In the case of MP-CAFEE, there is an effective linear relationship between the calculated values and the experimental ones. The square of the correlation coefficient (r^2) is 0.91 and the slope is 1.8. Furthermore, the calculation was able to reproduce the correct rank order. In contrast, in the case of MM-PBSA, the linear relationship between the calculated and experimental values is poor: the square of the correlation coefficient (r^2) is 0.74 and the slope is 2.0. Also, as shown in Fig. 4, the calculation could not reproduce the correct rank order. The prediction accuracy of MP-CAFEE is markedly higher than that of MM-PBSA.

It is notable that the computed binding energies by MP-CAFEE are approximately 5 kcal mol^{-1} smaller than the experimental values [15, 16]. This phenomenon was also observed in previous studies. Recently, Fujitani *et al.* showed that the force field selection affected the absolute binding energy values [16]. The underestimation of binding energies may arise from the force field. In particular, the binding energy expresses the change in Gibbs free energy, and thus the ligand has to be placed at a fixed position. However, the spring constraining the ligand to be bound within the binding site was not introduced. The so-called quasiharmonic approximation induces an overestimation of entropy when multiple energy valleys sampled [36]. Furthermore, our main interest was to examine the predicting

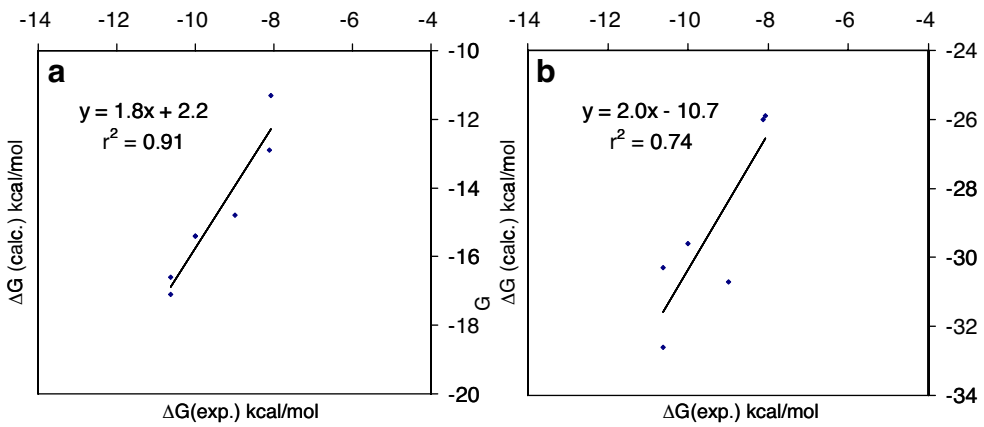

Fig. 4 Comparison of calculated absolute binding free energies with the experimental values (**a**) MP-CAFEE and (**b**) MM-PBSA

Table 1 Calculated absolute binding free energy

ligand	Exp.		MP-CAFEE			MM-PBSA		
	IC50 nM	ΔG kcal/mol	complex kcal/mol	solvation kcal/mol	ΔG kcal/mol	MM kcal/mol	PBSA kcal/mol	ΔG kcal/mol
C1	16	-10.6	-26.4	-9.3	-17.1	-79.6	49.4	-30.3
C2	16	-10.6	-29.1	-12.5	-16.6	-77.1	44.5	-32.6
C3	46	-10.0	-27.9	-12.5	-15.4	-83.7	54.2	-29.6
C4	250	-9.0	-26.7	-11.9	-14.8	-68.9	38.2	-30.7
C5	1100	-8.1	-23.4	-10.5	-12.9	-69.9	43.9	-26.0
C6	1200	-8.1	-22.0	-10.7	-11.3	-52.4	26.5	-25.9

performance of our methodology, and therefore we adopted the calculation model used here without any constraint to the ligand.

Relative binding free energy

In the lead optimization process, the relative binding free energy between two compounds has to be predicted. To evaluate the prediction accuracy of the relative binding free energies using MP-CAFEE, the relative binding free energies between inhibitors were calculated (Table 2). The calculated values using MP-CAFEE are consistent with the experimental ones. The root mean square error is 1.6 kcal mol^{-1}. On the other hand, the prediction accuracy using the MM/PBSA method is quite poor, with a root mean square error of 2.5 kcal mol^{-1}. From the viewpoint of relative binding free energy, the prediction accuracy of MP-CAFEE is higher than that of MM-PBSA.

Table 2 Calculated relative binding free energy

Pair		Exp. kcal/mol	MP-CAFEE kcal/mol	MM-PBSA kcal/mol
C1	C2	0.0	0.5	-2.3
	C3	0.6	1.7	0.7
	C4	1.6	2.3	-0.4
	C5	2.5	4.2	4.3
	C6	2.6	5.8	4.4
C2	C3	0.6	1.2	3.0
	C4	1.6	1.8	1.9
	C5	2.5	3.7	6.6
	C6	2.6	5.3	6.7
C3	C4	1.0	0.6	-1.1
	C5	1.9	2.5	3.6
	C6	1.9	4.1	3.7
C4	C5	0.9	1.9	4.7
	C6	0.9	3.5	4.8
C5	C6	0.0	1.6	0.1
RMSE			1.6	2.5

Comparison between C1 and C2

The values calculated using MP-CAFEE are almost the same (-17.1 kcal mol^{-1} for C1 and -16.6 kcal mol^{-1} for C2, respectively), which is consistent with the experimental results (-10.6 kcal mol^{-1} for C1 and C2). The substitution of a H atom to the methyl group (from C2 to C1) makes the hydrophobic surface area high, indicating that the compound becomes more hydrophobic. Thus, this change has a good affect on the complex term because of hydrophobic interaction, and a bad affect on the solvation term. In fact, the complex term of C1 is higher than that of C2 by 2.7 kcal mol^{-1} and the solvation term of C1 is lower than that of C2 by 3.2 kcal mol^{-1} (Table 1). MP-CAFEE gives us information not only about binding energies but also solvation, which may be beneficial in a lead optimization process.

Comparison between C2 and C3

Interestingly, although there is little difference between the chemical structure of C2 and that of C3, the IC50 value of C2 is three times smaller than that of C3. The calculation revealed that the energy difference arose not from the solvation energy term but from the interaction energy term. The conformation of C3 is more restricted owing to the double bond. The difference in the interaction term may arise from the entropic contribution. MP-CAFEE successfully reproduces this order, whereas the MM-PBSA method fails to reproduce it.

Comparison between C2 and C4

The experimental binding free energy for C4 is -9.0 kcal mol^{-1}, which is larger than that for C2 by 1.6 kcal mol^{-1}. The substitution of a H atom for a Cl atom

(from C2 to C4) makes a compound more hydrophobic, which is predicted to have a positive effect on the complex term and a negative effect on the solvation term. In fact, the solvation term of C4 is lower than that of C2 by 0.6 kcal mol^{-1}. However, the complex term of C4 is lower than that of C2 by 2.4 kcal mol^{-1}, which means that the contrary occurs. The substitution of a H atom for a Cl atom may cause steric conflict between ligand and protein, which induces the lowering of the complex term.

Application of MP-CAFEE to drug discovery research

As mentioned above, the binding free energies calculated by MP-CAFEE without complex crystal structure are well consistent with the experimental values. The prediction accuracy of MP-CAFEE is higher than that of MM-PBSA. Furthermore, the detailed information about solvation energy, which can help us to explain structure-activity relationship (SAR) and design novel compound, is obtained using MP-CAFEE. In total, MP-CAFEE would be an useful tool for drug discovery projects.

To put this method in practice, powerful hardware (*e.g.*, many core architecture) is required. Recently, some peta-flops super computers and specialized machines for MD simulation, and grid computing systems have been developed. These powerful systems would change the situation in terms of computation time and cost.

Conclusions

In summary, we calculated the absolute binding energies for PARP-1/inhibitor complexes, using MP-CAFEE. Although we used docking models as input structures, there is a good linear relationship between the calculated values and the experimental ones. The accuracy of this method is markedly higher than that using MM/PBSA, which is often used in the drug discovery field. Although the simulation time is currently long, and therefore the method is quite expensive, this reliable predictor of binding free energies would be a useful tool for drug discovery projects.

Acknowledgments The authors thank E. Lindahl for help in modifying GROMACS. The authors thank Ms. Mariko Katsuyama, Mr. Makoto Takeuchi, Dr. Yuzo Matsumoto, Ms. Naoko Katayama, Dr. Makoto Oku, Ms. Ayako Moritomo, Dr. Kenichi Mori, and Dr. Hideyoshi Fuji for helpful discussions.

References

1. Nguewa PA, Fuertes MA, Alonso C, Petéz JM (2003) Mol Pharmacol 64:1007–1014
2. LaPlaca MC, Zhang J, Raghupathi R, Li JH, Smith F, Bareyre FM, Snyder SH, Graham DI, Mclntosh TK (2001) J Neurotrauma 18:369–376
3. Mandir AS, Przedborski S, Jackson-Lewis V, Wang Z, Simbulan-Rosenthal CM, Smulson ME, Hoffman BE, Guastella DB, Dawson VL, Dawson TM (1999) Proc Natl Acad Sci USA 96:5774–5779
4. Menear KA, Adcock C, Boulter R, Cockcroft XL, Cranston CL, Dillon KJ A, Drzewiecki J, Garman S, Gomez S, Javaid H, Kerrigan F, Knights C, Lau A, Loh VM Jr, Matthews IT, Moore S, O'Connor MJ, Smith GC, Martin NM (2008) J Med Chem 51:6581–6591
5. Southan GJ, Szabó C (2003) Curr Med Chem 10:321–340
6. Meirovitch H, Cheluvaraja S, White RP (2009) Curr Protein Pept Sci 10:229–243
7. Fukunishi Y, Mitomo D, Nakamura H (2009) J Chem Inf Model 49:1944–1951
8. Card GL, Blasdel L, England BP, Zhang C, Suzuki Y, Gillette S, Fong D, Ibrahim PN, Artis DR, Bollag G, Milburn MV, Kim SH, Schlessinger J, Zhang KY (2005) Nat Biotechnol 23:201–207
9. Lyne PD, Lamb ML, Saeh JC (2006) J Med Chem 49:4805–4808
10. Abel R, Young T, Farid R, Berne BJ, Friesner RA (2008) J Am Chem Soc 130:2817–2831
11. Beveridge DL, DiCapua FM, Beveridge DL, DiCapua FM (1989) Annu Rev Biophys Biophys Chem 18:431–492
12. Kollman P (1993) Chem Rev 93:2395–2417
13. Jarzynski C (1997) Phys Rev Lett 78:2690–2693
14. Shirts MR, Bair E, Hooker G, Pande VS (2003) Phys Rev Lett 91:140601
15. Fujitani H, Tanida Y, Matsuura A (2009) Phys Rev E 79:021914
16. Tanida Y, Ito M, Fujitani H (2007) Chem Phys 337:135–143
17. Hattori K, Kido Y, Yamamoto H, Ishida J, Kamijo K, Murano K, Ohkubo M, Kinoshita T, Iwashita A, Mihara K, Yamazaki S, Matsuoka N, Teramura Y, Miyake H (2004) J Med Chem 47:4151–4154
18. Kinoshita T, Nakanishi I, Warizuka M, Iwashita Y, Kido Y, Hattori K, Fujii T (2004) FEBS Lett 556:43–46
19. MOE is a product of Chemical Computing Group, Inc, 1010 Sherbrooke Street West, Suite 910, Montreal, Quebec, Canada
20. Gold is a product of a collaboration between the University of Sheffield, GlaxoSmithKline plc and Cambridge Structural Database System (CCDC). It is available from CCDC (12 Union Road, Cambridge, CB2 1EZ, UK)
21. Berendsen HJC, van der Spoel D, van Drunen R (1995) Comput Phys Commun 91:43–46
22. Lindahl E, Hess B, van der Spoel D (2001) J Mol Model 7:306–317
23. van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) J Comput Chem 26:1701–1718
24. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347–1363
25. Gordon MS, Schmidt MW (2005) Theory and Applications of Computational Chemistry: the first forty years. Elsevier, Amsterdam
26. Bayly CI, Cieplak P, Cornell W, Kollman PA (1993) J Phys Chem 97:10269–10280
27. Wang J, Wang W, Kollman PA, Case DA (2006) J Mol Graph Model 25:247–260
28. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) J Comput Chem 25:1157–1174
29. Fujitani H, Matsuura A, Sakai S, Sato H, Tanida Y (2009) J Chem Theor Comput 5:1155–1156
30. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926–935
31. Nosé S (1984) Mol Phys 52:255–268
32. Hoover WG (1985) Phys Rev A 31:1695–1697

33. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690
34. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) J Comput Chem 18:1463–1472
35. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Pearlman DA, Crowley M, Walker RC, Zhang W, Wang B, Hayik S, Roitberg A, Seabra G, Wong KF, Paesani F, Wu X, Brozell S, Tsui V, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Beroza P, Mathews DH, Schafmeister C, Ross WS, Kollman PA (2006) AMBER 9. University of California, San Francisco
36. Chang C, Chen W, Gilson MK (2005) J Chem Theor Comput 1:1017–1028