conformational relaxation of the binding pocket of CDK2, allowing wider protein residues to move in MD simulations and longer MD simulations are required. The former is an especially important parameter for improving the enrichment performance, although it would increase the computational cost. In our next study, which will focus on the extent of the mobility of protein residues, along with simulation time and force-field parameters for organic small molecules, we will attempt to optimize the MD setup using a recent widely used dataset of decoy compounds [49].

The computational screening of large compound libraries involves the use of hierarchical multiple filters, such as ligand- and structure-based approaches. Molecular docking plays the primary role in these filters. With advancements in computer performance and computational chemistry, docking programs have become more accurate, but their ability to enrich hit compounds remains unsatisfactory. In order to improve the enrichment performance of molecular docking, we attempted to use the MM/PB-SA method [50] as a post-molecular docking filter. The basis of our approach was to perform massive MD simulations of protein-ligand conformations obtained from molecular docking, aim at the refinement/relaxation of protein-ligand conformations after docking, and predict more accurate binding free energies using the MM/PB-SA method in a practical time for lead discovery. Combining molecular docking and MD simulations basically allows each of them to neutralize the other's defects, but certain problems remain even with MD simulations, particularly with regard to compound screening applications. The major drawback of MD simulations is insufficient sampling due to the significant computational cost involved. To solve this problem, we performed MD simulations using various docking conformations obtained by molecular docking. However, the computational cost of this technique was approximately five to six times that of MD simulations using single docking conformations, such as the top-scored docking conformation. The enormous computational time needed for MD simulations is a serious problem. Here, we solved this problem by accelerating most of the time-consuming operations of the MD simulation using a high-performance special-purpose computer for MD simulations, "MDGRAPE-3" [27,28]. Accordingly, our approach could be performed in a practical time (about a week) for lead discovery. The evaluation in this study provides valuable information on in-silico drug design. Further, a more rigorous MD-based filter is under consideration for further improving the enrichment performance. This technique will also be applied to the lead optimization stage of drug development research.

In conclusion, our approach could improve the enrichment of virtual screening by molecular docking. Among the 12 types of binding free energies, G06, which was obtained from the MD simulations using multiple poses, showed the highest and most stable ability to enrich the active compounds. The strategy of multiple poses can be used to sample the potentially correct poses of active compounds; thus, it increases the enrichment performance. Since the G06 enrichment factors for the top 100 compounds ranged from 4 to 10 (see Table 4), which indicates approximately 1.6–4.0 times higher values than the enrichment performance of molecular docking, with the exception of CDK2, it is obvious that a stable and high enrichment can be achieved after molecular docking. In addition, G06 is suitable for compound screening because its computational cost is the least among those of the other MM/PB-SA energies obtained from the MD simulations. We also confirmed that G01, which was obtained from the MM calculations, showed good enrichment ability despite its low computational cost. This result agreed with that of the previous study [25]. The ability of G01 to enrich active compounds was lower and less stable than that of G06, but we believe that G01 acted as an effective filter between molecular docking and the MD-based MM/PB-SA method. From this study, we conclude that the application of MD simulations to virtual screening for lead discovery is effective and practical, but that further optimization of the MD simulation protocols is required for the screening of various target proteins, including kinases.

Materials and Methods

Preparation of the Target Protein

We applied our approach to four target proteins: trypsin, HIV PR, AChE, and CDK2. These structures with crystallographic resolutions of less than 3.0 Å, were retrieved from the Protein Data Bank (PDB) because the conformations of residues in the binding pocket affect the molecular docking results (PDB Id: 1C5S (trypsin) [51], 1HWR (HIV PR) [52], 1E66 (AChE) [53], and 1FVV (CDK2) [46]). All of the bound crystal water molecules, ligands, and other organic compounds were removed from each protein. Hydrogen atoms were added, and energy minimizations on the hydrogen atoms were performed using the Molecular Operating Environment (MOE) program (Chemical Computing Group Inc. [54]).

Seeded Compound Library for Docking

For each target protein, we prepared a test set of compounds that included 10,000 randomly selected compounds, or decoys, from the Maybridge library of compounds and experimentally known active compounds. It was confirmed that 95.5% of the selected decoy compounds obeyed the Lipinski rule of 5 [55]. The active compounds, which had binding affinities (K_i, K_d, or IC_{50}) below 30 μm, were selected from the PDBbind database [56,57] and by referring to the literatures [26,58]. Most of the active compounds also obeyed the Lipinski rule of 5. The numbers of active compounds selected for each of the respective target proteins was as follows: 21 (trypsin), 8 (HIV PR), 14 (AChE), and 26 (CDK2) (see Figure S1, S2, S3, S4). For each compound of the test set, a 3D conformation was generated, ionized, and energy minimized using LigPrep (Schrödinger Inc. [59]), assuming a pH of 7.0.

Docking

Molecular dockings were performed using the Genetic Optimisation of Ligand Docking (GOLD) version 3.1 [9,10]. This program employs a GA to explore the possible binding modes. The standard default settings for the GA parameters were used. The binding site radius was 12 Å. We performed the docking run three or four times using the GoldScore or ChemScore function for each target protein and selected the result that showed the best enrichment. GoldScore (default settings) was used as the scoring function for trypsin and HIV PR. In contrast, ChemScore (default settings) was used for AChE and CDK2 because docking runs using GoldScore can detect few of the successfully docked active compounds for AChE and CDK2. For AChE alone, the torsional rotations of Phe-330 (chi1 and chi2) were treated as flexible in the docking process. For each docking run, the 10 highest-scoring docking poses were saved to obtain a variety of binding modes.

Post-processing of the Docking Results

First, among the 10 highest-scoring docking poses saved for each compound, those in which the compound did not occupy the binding pocket or did not interact with the important residues were removed. The latter was used only for trypsin and HIV PR. The important residues were Asp180 for trypsin and Asp24 in

each monomer for HIV PR. These treatments had the effect of reducing the false positives for molecular docking. The docked compounds were then arranged in descending order from the highest score with respect to the multiple docking poses, and the top 1,000 compounds were selected from the test set. Finally, for the top 1,000 compounds, the docking poses of each compound were clustered using the root mean square deviation of 0.9 Å (complete link method [60]). After post-processing, approximately 6,000 docking poses were selected for the 1,000 compounds, which were then used as the initial conformations for MD simulations. Some active compounds were not ranked in the top 1,000. The numbers of active compounds in the top-scoring 1,000 were 10, 6, 7, and 17 for trypsin, HIV PR, AChE, and CDK2, respectively. In addition, the compounds in the top-scoring 1,000 were rescored with ChemScore (trypsin and HIV PR) or GoldScore (AChE and CDK2) because it is known that the rescoring approach increases the enrichment performance [61]. Furthermore, we analyzed ROC curves using molecular weight as classifier (Figure S7). From statistical analysis, it is obvious that the differences in the ROC values between G06 and molecular weight were statistically significant for trypsin, HIV PR, AChE.

MD Simulation Protocols

We performed MD simulations of each complex (ligand-bound protein), protein, and ligand to obtain various types of binding free energies (see the following subsection). The active sites of the protein-ligand complexes were immersed in an approximately 28–30 Å sphere of transferable intermolecular potential 3 point (TIP3P) water [62] molecules. The radius of the water droplet was selected such that the distance of the atoms of all the docked compounds from the water wall was greater than 15 Å (see Figure 6). The total number of atoms in the respective systems was approximately 8,000–12,000. On the solvent boundary, a half-harmonic potential (1.5 kcal/mol-Å^2) was applied to prevent the evaporation of the water molecules. The ligand, water molecules, and protein residues that were approximately 12 Å of the active center were allowed to move, but other protein residues were restrained to the X-ray structure by the harmonic energy term (1.5 kcal/mol-Å^2) in all of the MM calculations, namely the MM energy-minimization, and MD simulations. For the simulations of the ligands, each ligand was immersed in a water droplet, and this structure was used as the initial structure for the MD simulation of the ligand. In addition, the simulation of each protein (trypsin, HIV PR, AChE, and CDK2) was performed in the same manner as that of the complex.

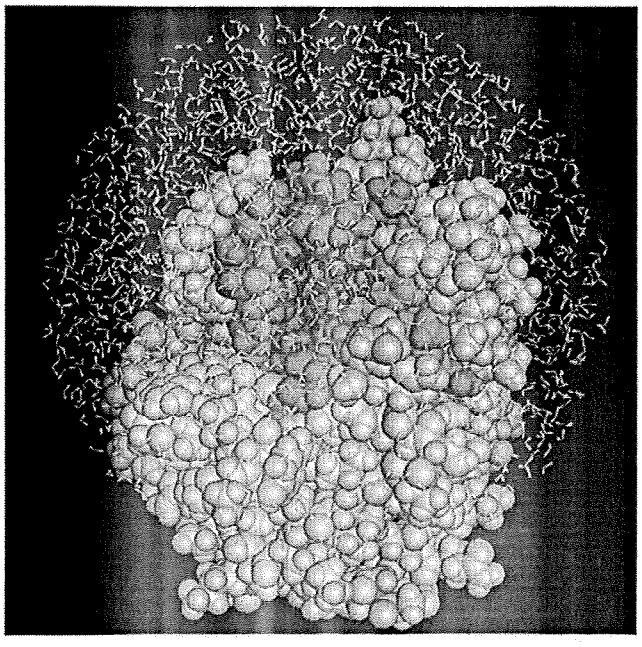

Figure 6. System for MD simulation of trypsin. The protein is shown by the space-filled model, and the ligand is colored blue. The peripheral residues around the active center (red region), a ligand, and water molecules were allowed to move in the MD simulation. The protein residues (grey) were restrained to the X-ray structure by a harmonic energy term. Similar systems were used for the other target proteins.
doi:10.1371/journal.pcbi.1000528.g006

All of the simulations were performed using AMBER 8.0 [63] modified for MDGRAPE-3 [27,28]. The ff03 force field [64] was adopted, and the time step was set at 0.5 fs. To carefully consider the motion of hydrogen atoms in the interactions between the ligands and protein residues, no bond length constraint was applied to solute atoms. The temperature of each system was gradually increased to 300 K during the first 25 ps, and additional MD simulations were performed for 700 ps for equilibration. The temperature was maintained at 300 K by using the method described by Berendsen et al. [65], and the system was coupled to a temperature bath with coupling constants of 0.2 ps. The parameters and charges for the ligands were determined using the *antechamber* module version 1.27 of AMBER 8.0 [63] by utilizing the general atom force field (GAFF) [66] and the AM1-BCC charge method [67,68]. Although the computational cost of the AM1-BCC charge method is low, a some difference between the charge and that of ff03 was noticeable. Since the original GAFF parameters were insufficient to cope with the parameters of all the ligands, we filled the missing parameters on the basis of the information on regarding atom types, bonds, valences, angles, and dihedrals by using an in-house program (see Text S1). (Note: these parameters for proteins and small organic molecules are very important to calculate the binding free energies between proteins and ligands)

Our MDGRAPE-3 system is a cluster of personal computers, each equipped with two MDGRAPE-3 boards. Each board contains 12 MDGRAPE-3 chips and has a peak speed of approximately 2 Tflops. The computations of non-bonded forces and energies for MD simulations were accelerated by MDGRAPE-3, and the other calculations were performed by the host central processing unit (CPU). In this study, we used 50 host computers equipped with 100 MDGRAPE-3 boards. The calculations for an MD simulation and the estimation of the binding free energies by the MM/PB-SA method were performed simultaneously. The average computational time for a single protein-ligand complex was 2.5 h, and the computations for approximately 6,000 protein-ligand conformations obtained by docking for each protein were completed in a week. The total simulation time for each protein was 4 μs, which corresponded to an 8-μs MD simulation with a standard time step of 1 fs. A single MD simulation for the system (Figure 6), without using MDGRAPE-3, requires more than 10 times the abovementioned computational time. Thus, in the current state, it would be quite difficult to use our screening approach without the MDGRAPE-3 system in a practically appropriate time for lead discovery. Therefore, our study can provide important information for MD-based screening.

Calculation of Binding Free Energy by the MM/PB-SA Method

The production MD trajectory was collected for the last period of 210 ps. In the calculation of the binding free energies by the

MM-PB/SA method, the water molecules were replaced with implicit solvation models. The binding free energy was calculated by the following equations.

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand}) \quad (2)$$

$$G = < E_{MM} > + < G_{PB} > + < G_{SA} > - TS \quad (3)$$

$$E_{MM} = E_{\text{int}} + E_{ele} + E_{vdW} \quad (4)$$

$$G_{SA} = \gamma A + b \quad (5)$$

In the above equations, $< >$ denotes the average for a set of 30 conformations along an MD trajectory. E_{int} includes the bond, angle, and torsional angle energies; E_{clc} and E_{vdW} represent the intermolecular electrostatic and van der Waals energies, respectively. G_{PB} was calculated by solving the PB equation with the DelPhi program [69,70], using the PARSE radii [71,72] and AMBER charges. The grid spacing used was 0.5 Å. The dielectric constants inside and outside the molecule were 1.0 and 80.0, respectively. In equation 5, which calculates the nonpolar solvation contribution, A is the solvent-accessible surface area that was calculated using the Michael Sanner's Molecular Surface (MSMS) program [73], and γ and b are 0.00542 kcal/mol-Å^2 and 0.92 kcal/mol, respectively. The probe radius was 1.4 Å. The conformational entropy term of the solute, TS, was approximated by a combination of a classical statistics expression and PCA [74], using the PTRAJ module of AMBER 8.0 [63]. In the PCA calculation, the last 210 ps (3,000 conformations) of each production trajectory were used.

The analysis of the binding free energy involved the calculation of the energies for conformations obtained from the MM (namely, energy-minimized) coordinates or MD trajectories. When the MM calculations or MD simulations of a complex, protein, and ligand were performed, we could obtain various types of binding free energies by combining the respective coordinate sets. The enthalpy contributions of G_{protein} and G_{ligand} in equation 2 were calculated in the following 2 ways: (1) by using the coordinate sets of a protein (or ligand) obtained from the MD simulations (or MM calculations) of the protein (or ligand) and (2) by using the coordinate sets extracted from the MD simulation of a complex. Similar to the enthalpy contribution, the entropy contribution was calculated by using the MD trajectories. When the entropy contributions of G_{complex}, G_{protein}, and G_{ligand} were calculated by using the MD trajectory of only the complex, we considered the entropy contribution of ΔG_{bind} to be zero because the energy components were almost cancelled. In this study, in order to thoroughly investigate which MM/PB-SA energies were suitable for compound screening, we adopted 12 binding free energies, G01–G12, to manage the entropy contributions independently of the enthalpy contributions (see Table 1). It should be noted that the coordinate sets for calculating the entropy contributions were not always consistent with those for calculating enthalpy contributions. Table 1 shows the enthalpy and entropy terms for computing of G_{complex}, G_{protein}, and G_{ligand} in equation 2. We classified the 12 binding free energies into four categories. Category 1 contained the energies obtained by the MM calculations, and categories 2, 3, and 4 contained those obtained by MD calculations. These categories were classified according to the combination of coordinate sets used for enthalpy calculations: G01–G03, G04–G06, G07–G09, and G10–G12 belonged to categories 1, 2, 3, and 4, respectively. Each binding free energy of a ligand adopts the minimum energies from among the energies of multiple poses. Thus, by gathering and arranging their energies, we were able to assess the enrichment performance of the screening approach.

Supporting Information

Figure S1 Active compounds of trypsin. The structural formulae and PDB ids of active compounds used in the seeded compound library are shown in the following figures. The asterisks represent the active compounds in top-scoring 1,000.
Found at: doi:10.1371/journal.pcbi.1000528.s001 (0.08 MB DOC)

Figure S2 Active compounds of HIV PR. The structural formulae and PDB ids of active compounds used in the seeded compound library are shown in the following figures. The asterisks represent the active compounds in top-scoring 1,000.
Found at: doi:10.1371/journal.pcbi.1000528.s002 (0.53 MB DOC)

Figure S3 Active compounds of AChE. The structural formulae and PDB ids of active compounds used in the seeded compound library are shown in the following figures. The asterisks represent the active compounds in top-scoring 1,000.
Found at: doi:10.1371/journal.pcbi.1000528.s003 (0.06 MB DOC)

Figure S4 Active compounds of CDK2. The structural formulae and PDB ids of active compounds used in the seeded compound library are shown in the following figures. The asterisks represent the active compounds in top-scoring 1,000. Compounds 1 and 19 were selected by referencing literatures.
Found at: doi:10.1371/journal.pcbi.1000528.s004 (0.11 MB DOC)

Figure S5 Number of correctly docked conformations in top-scored active compounds. These indicate the number of correctly docked conformations in the top-scoring poses for active compounds obtained from molecular docking, G01, and G06. The red and blue bars indicate the number of poses within the root mean square deviations (RMSDs) of 2.5 and 3.5 Å from those of the experimental structure, respectively. The active compounds in the top 1,000 were investigated. In G06, the final MD structure was used.
Found at: doi:10.1371/journal.pcbi.1000528.s005 (0.10 MB DOC)

Figure S6 Minimal RMSD values of computed poses from experimental poses for active compounds. The horizontal axis indicates the index number of active compounds in the top 1,000 shown in Figures S1, S2, S3, S4 and the vertical axis indicates the minimal RMSD among all the poses. For each protein, the poses obtained from molecular docking, G01, and G06 were investigated. In G06, the final MD structure was used. The red bars indicate the pose within the top-three scoring. For trypsin, HIV PR, and AChE, it was found that MD simulations could improve the binding modes and predict better binding free energies. For CDK2, however, it is suggested that MD simulations lead to structural uncertainties and an inaccurate estimation of the binding free energy.
Found at: doi:10.1371/journal.pcbi.1000528.s006 (0.24 MB DOC)

Figure S7 ROC curves using molecular weight as classifier. This graph shows the sensitivity versus 1-specificity. This indicates

ROC curves when the active compounds in the top 1,000 compounds are considered as total the true positives. ROC curves for trypsin, HIV PR, AChE, and CDK2 were drawn in blue, red, yellow, and orange, respectively. These ROC values for trypsin, HIV PR, AChE, and CDK2 are 0.454, 0.674, 0.462, and 0.430. From statistical analysis, it is obvious that the differences in the ROC values between G06 and molecular weight were statistically significant for trypsin, HIV PR, AChE. The differences in the ROC values between molecular docking and molecular weight were not statistically significant for all proteins.
Found at: doi:10.1371/journal.pcbi.1000528.s007 (0.08 MB DOC)

Text S1 Assignment of missing force field parameters. We filled in the following missing parameters on the basis of the information on regarding atom types, bonds, valences, angles, and dihedrals by using an in-house program.
Found at: doi:10.1371/journal.pcbi.1000528.s008 (0.06 MB DOC)

Author Contributions

Conceived and designed the experiments: NO AS MT. Performed the experiments: NO NF HF. Analyzed the data: NO NF. Contributed reagents/materials/analysis tools: NO NF HF GM RY YO TN. Wrote the paper: NO AS MT.

References

1. Young K, Lin S, Sun L, Lee E, Modi M, et al. (1998) Identification of a calcium channel modulator using a high throughput yeast two-hybrid screen. Nat Biotechnol 16: 946–950.
2. Hamasaki K, Rando RR (1998) A high-throughput fluorescence screen to monitor the specific binding of antagonists to RNA targets. Anal Biochem 261: 183–190.
3. Moore KJ, Turconi S, Miles-Williams A, Djaballah H, Hurskainen P, et al. (1999) A Homogenous 384-Well High Throughput Screen for Novel Tumor Necrosis Factor Receptor: Ligand Interactions Using Time Resolved Energy Transfer. J Biomol Screen 4: 205–214.
4. Dunn D, Orlowski M, McCoy P, Gastgeb F, Appell K, et al. (2000) Ultra-high throughput screen of two-million-member combinatorial compound collection in a miniaturized, 1536-well assay format. J Biomol Screen 5: 177–188.
5. Doman TN, McGovern SL, Witherbee BJ, Kasten TP, Kurumbail R, et al. (2002) Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. J Med Chem 45: 2213–2221.
6. Chen J, Zhang Z, Stebbins JL, Zhang X, Hoffman R, et al. (2007) A fragment-based approach for the discovery of isoform-specific p38alpha inhibitors. ACS Chem Biol 2: 329–336.
7. Carr RA, Congreve M, Murray CW, Rees DC (2005) Fragment-based lead discovery: leads by design. Drug Discov Today 10: 987–992.
8. Hann MM, Leach AR, Harper G (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. J Chem Inf Comput Sci 41: 856–864.
9. Jones G, Willett P, Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 245: 43–53.
10. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267: 727–748.
11. Ewing TJ, Makino S, Skillman AG, Kuntz ID (2001) DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. J Comput Aided Mol Des 15: 411–428.
12. Goodsell DS, Morris GM, Olson AJ (1996) Automated docking of flexible ligands: applications of AutoDock. J Mol Recognit 9: 1–5.
13. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, et al. (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 47: 1750–1759.
14. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, et al. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47: 1739–1749.
15. Rarey M, Kramer B, Lengauer T, Klebe G (1996) A fast flexible docking method using an incremental construction algorithm. J Mol Biol 261: 470–489.
16. Bursulaya BD, Totrov M, Abagyan R, Brooks CL (2003) Comparative study of several algorithms for flexible ligand docking. J Comput Aided Mol Des 17: 755–763.
17. Stahl M, Rarey M (2001) Detailed analysis of scoring functions for virtual screening. J Med Chem 44: 1035–1042.
18. Wyss PC, Gerber P, Hartman PG, Hubschwerlen C, Locher H, et al. (2003) Novel dihydrofolate reductase inhibitors. Structure-based versus diversity-based library design and high-throughput synthesis and screening. J Med Chem 46: 2304–2312.
19. Pearlman DA, Charifson PS (2001) Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system. J Med Chem 44: 3417–3423.
20. Kollman P (1993) Free-Energy Calculations - Applications to Chemical and Biochemical Phenomena. Chemical Reviews 93: 2395–2417.
21. Aqvist J, Luzhkov VB, Brandsdal BO (2002) Ligand binding affinities from MD simulations. Acc Chem Res 35: 358–365.
22. Kollman PA, Massova I, Reyes C, Kuhn B, Huo SH, et al. (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. Accounts of Chemical Research 33: 889–897.
23. Huo S, Wang J, Cieplak P, Kollman PA, Kuntz ID (2002) Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: insight into structure-based ligand design. J Med Chem 45: 1412–1419.
24. Masukawa KM, Kollman PA, Kuntz ID (2003) Investigation of neuraminidase-substrate recognition using molecular dynamics and free energy calculations. J Med Chem 46: 5628–5637.
25. Kuhn B, Gerber P, Schulz-Gasch T, Stahl M (2005) Validation and use of the MM-PBSA approach for drug discovery. J Med Chem 48: 4040–4048.
26. Ferrara P, Curioni A, Vangrevelinghe E, Meyer T, Mordasini T, et al. (2006) New scoring functions for virtual screening from molecular dynamics simulations with a quantum-refined force-field (QRFF-MD). Application to cyclin-dependent kinase 2. J Chem Inf Model 46: 254–263.
27. Narumi T, Ohno Y, Okimoto N, Koishi T, Suenaga A, et al. (2006) A 185 Tflops simulation of amyloid-forming peptides from Yeast Prion Sup35 with the special-purpose computer System MD-GRAPE3. Proc Supercomputing 2006, in CD-ROM.
28. Taiji M (2004) MDGRAPE-3 chip: a 165 Gflops application specific LSI for molecular dynamics simulations.; 2004. IEEE Computer Society. pp. in CD-ROM.
29. Thomas MP, McInnes C, Fischer PM (2006) Protein structures in virtual screening: A case study with CDK2. J Med Chem 49: 92–104.
30. Kontoyianni M, McClellan LM, Sokol GS (2004) Evaluation of docking performance: comparative data on docking algorithms. J Med Chem 47: 558–565.
31. Wang R, Lu Y, Wang S (2003) Comparative evaluation of 11 scoring functions for molecular docking. J Med Chem 46: 2287–2303.
32. Cho AE, Wendel JA, Vaidehi N, Kekenes-Huskey PM, Floriano WB, et al. (2005) The MPSim-Dock hierarchical docking algorithm: application to the eight trypsin inhibitor cocrystals. J Comput Chem 26: 48–71.
33. Erickson JA, Jalaie M, Robertson DH, Lewis RA, Vieth M (2004) Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy. J Med Chem 47: 45–55.
34. Kua J, Zhang Y, McCammon JA (2002) Studying enzyme binding specificity in acetylcholinesterase using a combined molecular dynamics and multiple docking approach. J Am Chem Soc 124: 8260–8267.
35. Witten IH, Frank E (1999) Data mining: practical machine learning tools and techniques with java implementations. New York: Morgan Kaufmann.
36. Habe H, Morii K, Fushinobu S, Nam JW, Ayabe Y, et al. (2003) Crystal structure of a histidine-tagged serine hydrolase involved in the carbazole degradation (CarC enzyme). Biochem Biophys Res Commun 303: 631–639.
37. Dorfman DD, Berbaum KS, Metz CE (1992) Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method. Invest Radiol 27: 723–731.
38. Hillis SL, Berbaum KS (2005) Monte Carlo validation of the Dorfman-Berbaum-Metz method using normalized pseudovalues and less data-based model simplification. Acad Radiol 12: 1534–1541.
39. Hillis SL, Obuchowski NA, Schartz KM, Berbaum KS (2005) A comparison of the Dorfman-Berbaum-Metz and Obuchowski-Rockette methods for receiver operating characteristic (ROC) data. Stat Med 24: 1579–1607.
40. Roe CA, Metz CE (1997) Variance-component modeling in the analysis of receiver operating characteristic index estimates. Acad Radiol 4: 587–600.
41. Roe CA, Metz CE (1997) Dorfman-Berbaum-Metz method for statistical analysis of multireader, multimodality receiver operating characteristic data: validation with computer simulation. Acad Radiol 4: 298–303.
42. Gohlke H, Case DA (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. Journal of Computational Chemistry 25: 238–250.
43. Numata J, Wan M, Knapp EW (2007) Conformational entropy of biomolecules: beyond the quasi-harmonic approximation. Genome Inform 18: 192–205.
44. Chang CE, Chen W, Gilson MK (2005) Evaluating the accuracy of the quasiharmonic approximation. Journal of Chemical Theory and Computation 1: 1017–1028.
45. Davies TG, Bentley J, Arris CE, Boyle FT, Curtin NJ, et al. (2002) Structure-based design of a potent purine-based cyclin-dependent kinase inhibitor. Nat Struct Biol 9: 745–749.
46. Davis ST, Benson BG, Bramson HN, Chapman DE, Dickerson SH, et al. (2001) Prevention of chemotherapy-induced alopecia in rats by CDK inhibitors. Science 291: 134–137.

47. Lamb ML, Jorgensen WL (1997) Computational approaches to molecular recognition. Curr Opin Chem Biol 1: 449–457.
48. Zhou Z, Madura JD (2004) Relative free energy of binding and binding mode calculations of HIV-1 RT inhibitors based on dock-MM-PB/GS. Proteins 57: 493–503.
49. Huang N, Shoichet BK, Irwin JJ (2006) Benchmarking sets for molecular docking. J Med Chem 49: 6789–6801.
50. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, et al. (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33: 889–897.
51. Katz BA, Mackman R, Luong C, Radika K, Martelli A, et al. (2000) Structural basis for selectivity of a small molecule, S1-binding, submicromolar inhibitor of urokinase-type plasminogen activator. Chem Biol 7: 299–312.
52. Ala PJ, DeLoskey RJ, Huston EE, Jadhav PK, Lam PY, et al. (1998) Molecular recognition of cyclic urea HIV-1 protease inhibitors. J Biol Chem 273: 12325–12331.
53. Dvir H, Wong DM, Harel M, Barril X, Orozco M, et al. (2002) 3D structure of Torpedo californica acetylcholinesterase complexed with huprine X at 2.1 A resolution: kinetic and molecular dynamic correlates. Biochemistry 41: 2970–2981.
54. MOE. Montreal: Chemical Computing Group Inc.
55. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46: 3–26.
56. Wang R, Fang X, Lu Y, Yang CY, Wang S (2005) The PDBbind database: methodologies and updates. J Med Chem 48: 4111–4119.
57. Wang R, Fang X, Lu Y, Wang S (2004) The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. J Med Chem 47: 2977–2980.
58. Gray NS, Wodicka L, Thunnissen AMWH, Norman TC, Kwon SJ, et al. (1998) Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. Science 281: 533–538.
59. Schrödinger Inc. PortlandOR).
60. Sørenson T (1948) A method of establishing groups of equal amplitude in a plant based on similarity of species content and its applications to analysis of vegetation on Danish commons. Biologiske Skrifter 5: 1–34.
61. Hoffmann D, Kramer B, Washio T, Steinmetzer T, Rarey M, et al. (1999) Two-stage method for protein-ligand docking. J Med Chem 42: 4422–4433.
62. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. Journal of Chemical Physics 79: 926–935.
63. Case DA, Darden TA, Cheatham TEr, Simmerling CL, Wang J, et al. (2004) AMBER 8 University of California San Francisco.
64. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, et al. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24: 1999–2012.
65. Berendsen HJC, Postma JMP, van Gunsteren WF, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Comput Phys 81: 3684–3690.
66. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25: 1157–1174.
67. Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. J Comput Chem 23: 1623–1641.
68. Jakalian A, Bush BL, Jack DB, Bayly CI (2000) Fast, efficient generation of high-quality atomic Charges. AM1-BCC model: I. Method. Journal of Computational Chemistry 21: 132–146.
69. Rocchia W, Alexov E, Honig B (2001) Extending the applicability of the nonlinear poisson-boltzmann equation: multiple dielectric constants and multivalent ions. J Phys Chem B 105: 6507–6514.
70. Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, et al. (2002) Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. J Comput Chem 23: 128–137.
71. Sitkoff D, Sharp KA, Honig B (1994) Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 98: 1978–1988.
72. Swanson JMJ, Adcock SA, McCammon JA (2005) Optimized radii for Poisson-Boltzmann calculations with the AMBER force field. Journal of Chemical Theory and Computation 1: 484–493.
73. Sanner MF, Olson AJ, Spehner JC (1996) Reduced surface: an efficient way to compute molecular surfaces. Biopolymers 38: 305–320.
74. Levy RM, Karplus M, Kushick J, Perahia D (1984) Evaluation of the Configurational Entropy for Proteins - Application to Molecular-Dynamics Simulations of an Alpha-Helix. Macromolecules 17: 1370–1374.