Theor Chem Acc (2012) 131:1096
DOI 10.1007/s00214-012-1096-z

REGULAR ARTICLE

Use of quantum mechanics/molecular mechanics-based FEP method for calculating relative binding affinities of FBPase inhibitors for type-2 diabetes

R. S. Rathore · R. Nageswara Reddy · A. K. Kondapi · P. Reddanna · M. Rami Reddy

Received: 4 June 2011 / Accepted: 11 July 2011 / Published online: 9 February 2012

Abstract A quantum mechanics (QM)/molecular mechanics (MM)-based free energy perturbation (FEP) method, developed recently, provides most accurate estimation of binding affinities. The validity of the method was evaluated for a large set of diverse inhibitors for fructose 1,6-bisphosphatase (FBPase), a target enzyme for type-II *diabetes mellitus*. The validation set comprises of 22 important structurally different mutations. The calculated relative binding free energies using the QM/MM-based FEP method reproduce the experimental values with exceptional precision of less than ±0.5 kcal/mol. The CPU requirements for QM/MM-based FEP are about fivefold greater than conventional FEP methods, but it is superior in accuracy of predictions. In addition, the QM/MM-based FEP method eliminates the need for time-consuming development of MM force field parameters, which are frequently required for novel inhibitors described by MM. Future automation of the method and parallelization of the code for 128/256/512 cluster computers is expected to enhance the speed and increase its use for drug design and lead optimization. The present application of QM/MM-based FEP method for structurally diverse set of analogs serves to enhance the scope of FEP method and demonstrate the utility of QM/MM-based FEP method for its potential in drug discovery.

Keywords Quantum mechanics · Molecular mechanics · Free energy perturbation · Fructose 1,6-bisphosphatase (FBPase) · QM/MM-based FEP method

1 Introduction

Advances in structure elucidation procedures over the past two decades have produced an unprecedented number of high-resolution protein structures. Efforts to use this structural information to shorten the drug discovery timeline have led to a variety of computational methods for estimating ligand-binding affinities [1, 2]. Most of these methods use algorithms that rely on a multitude of approximations through a simplified treatment of entropy, solvent interactions, and molecular mechanics force fields that are intended for high speed of calculations and compound throughput. Virtual screening of large compound libraries using these automated methods [3–6] is a good starting point for subsequent compound evaluation and synthetic designs. While such methods are quite appealing and convenient to apply for effective guidance to medicinal chemistry, all too often these semi-quantitative calculations lead to erroneous conclusions giving rise to failed clinical compounds.

Dedicated to Professor Eluvathingal Jemmis and published as part of the special collection of articles celebrating his 60th birthday.

R. S. Rathore · A. K. Kondapi
Bioinformatics Infrastructure Facility, School of Life Sciences, University of Hyderabad, Hyderabad 500046, India

R. N. Reddy
Rational Labs Pvt Ltd, Plot # 177, IDA Mallapur, Hyderabad 500076, India

P. Reddanna
Department of Animal Sciences, School of Life Sciences, University of Hyderabad, Hyderabad 500046, India

M. R. Reddy (✉)
RR Labs, Inc., 8013 Los Sabalos Street, San Diego, CA 92126, USA
e-mail: DRMRREDDY@yahoo.com

M. R. Reddy
Centre for Modeling, Simulations and Design, University of Hyderabad, Hyderabad 500046, India

Over the years, methods of free energy calculations have emerged as promising strategy. In contrast to the prevalent methods described earlier, binding affinity calculations using free energy perturbation (FEP)-based methods are even though associated with low throughput, but calculated results are closer to the experimental values than other methods [7–9]. Several calculations on relative differences in lipophilicity, ionization, covalent hydration and solvation have demonstrated the power of the FEP method for correct predictions [10–12]. Further, the approach has also shown good accuracy in predicting relative binding free energies for inhibitors of numerous enzymes [13–16], including those considered as potential drug targets [17]. Despite these impressive results, FEP methods are rarely employed in the pharmaceutical industry. FEP methods suffer from several limitations. First, FEP calculations are CPU demanding and require the availability of validated molecular mechanics force field parameters to achieve high accuracy. Since most drug candidates contain substructures not fully described by existing parameters, the user must develop and input parameters. The process is time-consuming and often limited by the absence of relevant experimental data. Moreover, the process is difficult to automate and requires considerable user expertise and judgment. Second, real-life drug design problems involve the calculation of relative binding affinities for inhibitors with a greater degree of structural dissimilarity. Accordingly, there is a need for methods that enable rapid assessment of large number of structurally unrelated molecules in a reasonably accurate manner.

To address these limitations, we have recently developed a novel method for calculation of binding affinity. The method uses the molecular mechanics (MM) force field for treating the solvent and protein, and quantum mechanics (QM) for treating the ligand. We have in the past reported applications of this QM/MM-based FEP method for calculation of relative solvation free energies of small molecules as well as binding free energy differences of inhibitors [18–22]. To examine the validity of the method for lead optimization and mutations comprises of large structural and conformational changes, we decided to apply this QM/MM-based FEP method for a large set of diverse inhibitors that were designed under an integrated structure-guided drug discovery project for a type 2 diabetes *mellitus* target, fructose 1,6-bisphosphatase (FBPase; EC 3.1.3.11).

FBPase is a rate-controlling enzyme in the gluconeogenesis pathway, catalyzing the hydrolysis of fructose 1,6-bisphosphate to fructose 6-phosphate and inorganic phosphate. Flux through the gluconeogenesis pathway is abnormally high in type 2 *diabetes mellitus*, which is largely responsible for the excessive endogenous glucose production in these patients [23, 24]. Glucose production contributes significantly to the elevated blood glucose levels associated with diabetes, and correspondingly to the disease-related morbidity and mortality, FBPase has long been considered a potential target for treating type 2 *diabetes mellitus*. Discovery of potent FBPase inhibitors, however, proved elusive despite efforts over the past 30 years that involved the screening of large compound libraries [25] as well as synthesis of both substrate and AMP analogs [26–33]. More recently, we reported the use of a structure-guided drug design strategy that led to the first potent inhibitors of FBPase [26] and to the demonstration that these inhibitors result in robust glucose lowering activity in animals of type 2 diabetes [27–33].

2 Methods

2.1 Free energy perturbation method

The statistical perturbation theory is due to the classical work of Zwanzig [7–9], and its detailed implementation in a molecular dynamics program to compute the free energy is available in literature [7–9, 34–38]. For completeness, a very brief description of the FEP method is given here. The total Hamiltonian (H) of a system may be written as the sum of the Hamiltonian (H_0) of the unperturbed and the perturbation (H_1) systems:

$$H = H_0 + H_1 \tag{1}$$

The free energy contribution due to the perturbation is given by

$$G_1 = -\frac{1}{\beta}\langle \exp(-\beta H_1)\rangle_0 \tag{2}$$

where $\beta = 1/kT$ and the mean of exp $(-\beta H_1)$ is computed over the unperturbed ensemble of the system, k is the Boltzmann constant and T is the absolute temperature. To compute ΔG, the difference in free energy between the two solute states, the Hamiltonian for states A and B can be linked by the coupling parameter λ in a linear or nonlinear manner such that, $H(\lambda)$ represents hypothetical coupled Hamiltonian of the system at a given λ $(0 \leq \lambda \leq 1)$. For a simple linear coupling, $H(\lambda)$ can be given as,

$$H(\lambda) = \lambda H_A + (1-\lambda)H_B \tag{3}$$

In the above equation, H_A is the Hamiltonian for the system at state A, and H_B is that for state B. For both simple linear and nonlinear coupling of two solute states, when $\lambda = 0$, $H_\lambda = H_B$, that is, the system is purely in state B and when $\lambda = 1$, $H_\lambda = H_A$ at which point the system is purely in state A. During intermediate value of λ, the solute is a mixture of A and B. This type of coupling ensures a very smooth conversion of two solutes A and B, allowing

the system to read just its configuration smoothly as a function of the state. If we divide the range of λ into N windows, at each window λ_i, the solute state is perturbed between λ_{i+1} and λ_{i-1} states by taking the reference state as $H(\lambda_i)$. The free energy difference between the two solute states A and B is a simple summation over all windows of $[G_1(\lambda_i)]$ as given by,

$$\Delta G = \sum_{i=1}^{N} G_1(\lambda_i) \tag{4}$$

The evaluation of G_1 at λ_{i+1} and λ_{i-1} is a check for possible hysteresis in the calculation and is a measure of the statistical error for the free energy change.

2.2 Free energy decoupling

Even though the free energy difference is a path independent quantity, it is observed that certain sampling difficulties arise when a polar solute is transferred to a non-polar solute accompanied by a large change in molecular volume. Under this circumstance, if one attempts to mutate both the partial charges and the non-bonded parameters simultaneously, the solute–solvent energy increases enormously as a consequence of very close approach of some of the solvent molecules. This artificial limitation can be avoided by decoupling the total free energy as a sum of electrostatic and van der Waals contributions. If we denote the energy due to bond, angle, and dihedral as E_{bad}, the electrostatic energy as E_{ele} and the van der Waals interaction energy as E_{vdw}, the total energy of the system can be written as:

$$E_{\text{tot}} = E_{\text{bad}} + E_{\text{ele}} + E_{\text{vdw}} \tag{5}$$

The conversion of state A to B is achieved trough an intermediate stage A′ such that A → A′ → B and the corresponding net free energy change is:

$$\Delta G_{\text{AB}} = \Delta G_{\text{AA}\prime} + \Delta G_{\text{A}\prime\text{B}} \tag{6}$$

Here, the state A′ has the charge distribution of B, but maintains the molecular geometry and van der Waals parameters of A. Hence, $\Delta G_{\text{AA}'}$ corresponds to the electrostatic contribution to the free energy difference. The conversion of A to A′ is achieved smoothly since the van der Waals parameters are the same for these two states. $\Delta G_{\text{A}'\text{B}}$ is the van der Waals contribution to the free energy, which includes the bond, angle, and dihedral contributions as well as the non-bonded interactions. The conventional FEP calculation entails coupling the MM parameters according to λ and calculating the corresponding MM energies and forces at each λ. In contrast, the QM-based FEP method uses QM methods to calculate the energies and forces for the solute (ligands) in the system and MM to describe the solvent and macromolecule as well as interactions between solvent/macromolecule and solute. To calculate the QM energy and forces, we implemented a procedure [18–22] that separated the threaded molecule into two molecules (A and B) at each dynamic step. The quantum mechanical energies and forces were then computed and the combined energies and forces recomputed using the λ-coupling method. The interaction energy between the perturbing system and the surroundings at any value of λ is calculated by coupling the charges, radius and well depth of the perturbing system as follows:

The partial charge on atom i at any value of λ is given by

$$q_{i\lambda} = \lambda q_i^{\text{A}} + (1-\lambda) q_i^{\text{B}} \tag{7}$$

The radius and well depth of atom $\boldsymbol{i}$ at a given value of λ is given by

$$r_{i\lambda} = \lambda r_i^{\text{A}} + (1-\lambda) r_i^{\text{B}} \tag{8}$$

$$\varepsilon_{i\lambda} = \lambda \varepsilon_i^{\text{A}} + (1-\lambda) \varepsilon_i^{\text{B}} \tag{9}$$

On the other hand, the QM forces and energies are calculated for each molecule, A and B, separately and then coupled based on λ using Eqs. 10 and 11

$$f_\lambda^i = \lambda f_{\text{A}}^i + (1-\lambda) f_{\text{B}}^i \tag{10}$$

$$E_\lambda^{\text{QM}} = \lambda E_{\text{A}}^{\text{QM}} + (1-\lambda) E_{\text{B}}^{\text{QM}} \tag{11}$$

Accordingly, the QM energy and force will correspond to molecule B when $\lambda = 0$ and to molecule A when $\lambda = 1$. It should be noted that the energies and forces are calculated at a given λ and thus enters into the Eqs. 10 and 11 as constant. The total energy for the system is determined using Eq. 12, where the term $E_{\text{QM/MM}}$ represents the interaction energy involving the MM and QM part of the system. The free energy change Eq. 13 is decomposed into the free energy contribution from the subsystem treated by QM and the free energy contribution from the surroundings, that is the subsystem not treated by QM (non-QM or NQM).

$$E_{\text{tot}} = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QM/MM}} \tag{12}$$

$$\Delta G_{\text{tot}} = \Delta G_{\text{QM}} + \Delta G_{\text{NQM}}. \tag{13}$$

2.3 Thermodynamic cycle-perturbation approach

The thermodynamic cycle-perturbation (TCP) approach [7–9, 34–38] may be described as a method for computing the relative changes of free energy for a binding process by the construction of non-physical paths connecting the desired initial and terminal states. This approach enables the calculation of relative change in solvation free energy ($\Delta\Delta G_{\text{sol}}$) and binding free energy ($\Delta\Delta G_{\text{bind}}$) differences between two related compounds, by computationally simulating the "mutation" of one to another. The relative solvation free energy change for two substrates is computed using the solvation cycle shown in Fig. 1, represented in the following equation:

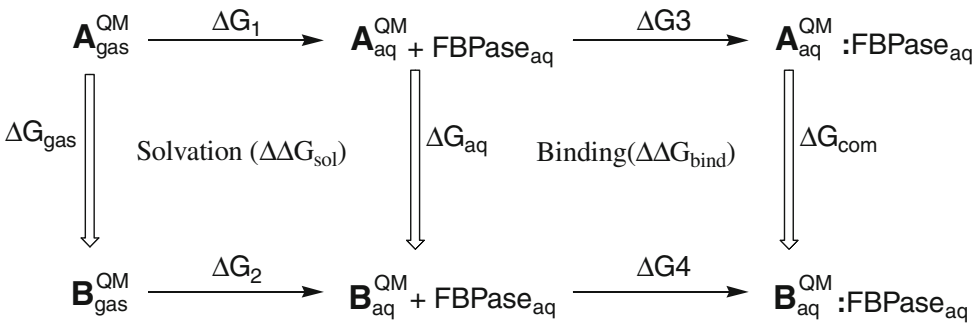

Fig. 1 TCP used for calculating relative solvation and binding free energies of FBPase inhibitors using QM/MM-based FEP method

$$\Delta\Delta G_{sol} = \Delta G_{aq} - \Delta G_{gas} = \Delta G_2 - \Delta G_1 \quad (14)$$

The relative free energy of binding is the difference in two affinities, that is the affinity of the ligand for the protein and the affinity of the ligand for water, which is computed using the binding cycle shown in Fig. 1, represented by the following equation:

$$\Delta\Delta G_{bind} = \Delta G_{com} - \Delta G_{aq} = \Delta G_4 - \Delta G_3 = -kT\ln\left(\frac{k_2}{k_1}\right) \quad (15)$$

where the experimentally measured binding constants k_1 and k_2 refer to the reactions involving A and B inhibitors, respectively, and their corresponding free energy differences are ΔG_3 and ΔG_4, The free energy change for converting A into B is computed by perturbing the Hamiltonian of reactant (initial) state A into that of the product (final) state B. This transformation is accomplished through a parameterization of terms comprising the interaction potentials of the system with a change of state variable that maps onto reactant and product states when that variable is 0 and 1, respectively. The total free energy change for the transformation from the initial to the final state is computed by summing 'incremental' free energy changes over all the windows visited by the state variable, changing from 0 to 1.

3 Computation details for relative solvation and binding free energies of FBPase inhibitors

3.1 Solvent simulations

Solvation free energies were calculated by immersing the inhibitor in a 19 Å box of equilibrated SPC/E water [39, 40] in a rectangular box and using periodic boundary conditions in all directions. Newton's equations of motion for all the atoms were solved using the Verlet algorithm [41] with a 2-fs time step, and SHAKE [42] for constraining all bond lengths. Non-bonded interaction energies were calculated using a 15.0 Å residue-based cutoff.

3.2 Complex simulations

Protein complex simulations began by immersing the FBPase-inhibitor complex [43, 44] in a 30.0 Å radius sphere of solvent centered on the inhibitor. The water sphere was subjected to a half-harmonic restraint near the boundary to prevent evaporation. During the simulation, all atoms of the protein were fixed beyond 30.0 Å. All non-bonded interactions involving the inhibitors and the charged residues of the protein were computed with infinite cutoff. A 15.0 Å non-bonded residue-based cutoff was used for other residues of the system. The algorithm for the complex simulation was identical to the solvent simulation, except for the absence of periodic boundary conditions in the latter.

Free energy simulations, for both solvation and complex phases, began after energy minimization of the system using 500 steps of steepest descent followed by 2,000 steps of conjugate gradient and 20 ps of equilibration using molecular dynamics (MD) simulations. Relative free energy differences were obtained using a two-stage procedure as earlier described [18–22]. Each stage of the simulation was performed using twenty-one windows with each window comprising 10 ps of equilibration and 25 ps of data collection to compute the free energy difference (i.e. a total of 1,470 ps for complete mutation). Each mutation followed the doublewide sampling procedure, and the results reported are based on the averages from the backward and forward simulations of the mutation. Errors were estimated for each window by dividing the window statistics into 5 groups (in both forward and backward directions) and computing the standard deviation for the indicated free energy change. The standard deviation is the root-mean-square of these window errors. All the calculations were carried out using Galaxy program [18–22, 45].

Calculations using the conventional and QM/MM-based FEP methods were performed using procedures previously described [18–22]. In both cases, the λ-coupling method was used in the transformation of inhibitor A to inhibitor B [34]. The thread or double topology method [18–22, 45] (Fig. 2), which is used frequently in conventional FEP calculations for mapping structurally dissimilar molecules,

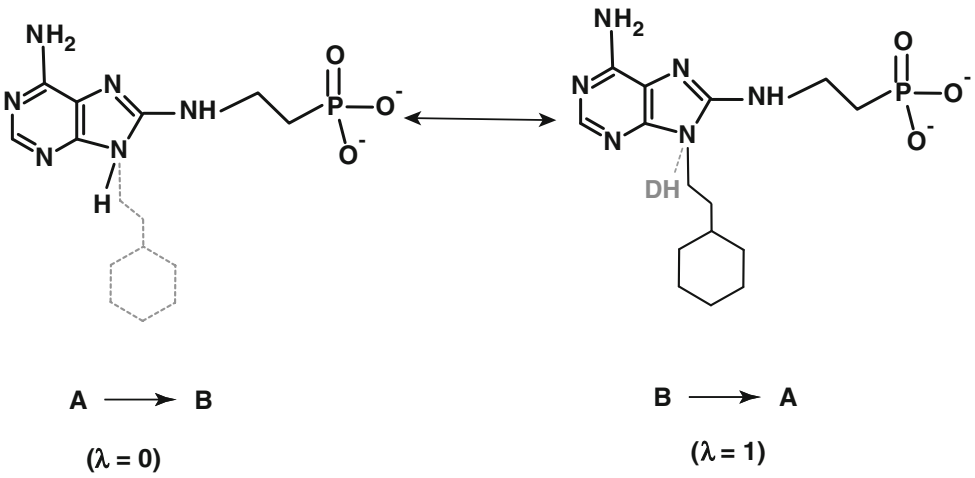

Fig. 2 Dual topology definition between the molecules **4** ($\lambda = 0$) to **6** ($\lambda = 1$). Common atoms are represented by a single topology (amino acid backbone). Prefix 'D' indicates "dummy atoms" as represented by the *dashed* substructure. Hydrogens and dummy hydrogens (DH) associated with R^9-*cyclo*-hexylethyl are omitted for clarity

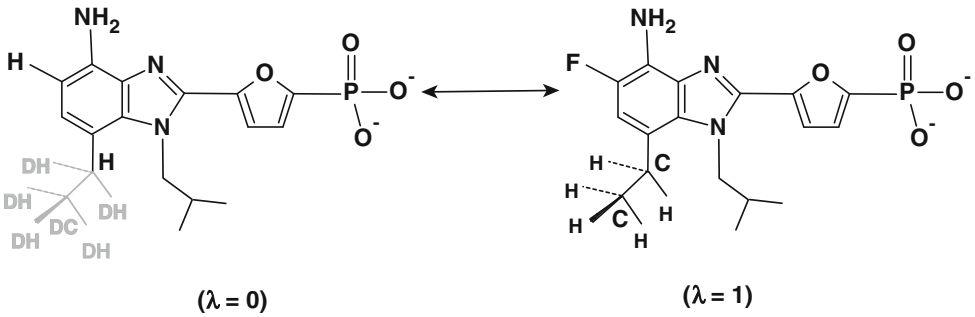

Fig. 3 Single topology definition between the molecules **13** ($\lambda = 0$) and **14** ($\lambda = 1$). Prefix "D" indicates dummy atoms

was used in all transformations involving large structural changes, and single topology [18–22, 45] (Fig. 3) method was used for all the mutations, which require small structural changes. In the conventional method, MM parameters are scaled according to λ and used to calculate MM energies and forces. In the QM/MM-based FEP method, the semi-empirical QM method, AM1, is used to calculate the energies and forces for the ligand, while MM is used to calculate the energies and forces of the solvent and protein. The total energy for the system was determined using Eq. 12 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. 13) is decomposed into the free energy contribution from the subsystem treated by QM and the free energy contribution from the surroundings, that is the subsystem not treated by QM (non-QM or NQM).

4 Results and discussion

Several crystal structures of FBPase complexes with AMP and its analogs have been determined [46]. The three-dimensional structure of the human FBPase-**14** [31, 32] complex solved to 2.0 Å resolution was used for calculating relative binding free energies between all the inhibitors considered in this study, using both conventional as well as QM/MM-based FEP methods. Molecular surface representation of the FBPase:**14** complex is shown in Fig. 4. The stereo diagram showing the interactions between binding site of FBPase and **14** is shown in Fig. 5.

Over the past years, we have developed a new QM/MM-based FEP method that provides very accurate estimation of binding energies (see details in Methods section). The method was successfully applied for calculations of solvation and binding free energies [18–22]. In this article, the calculation of the relative binding free energy differences was carried out for a large and diverse set of FBPase inhibitors shown in Fig. 6. The validation set has been chosen to test the methodology for structurally different inhibitors of FBPase for estimating the relative binding free energy differences between these inhibitors. The validation set comprises 22 important mutations, which involve large structural changes as well as polar and non-polar substitutions with high conformational degree of freedom. Two methods were used for calculating relative binding free energies: (a) QM/MM-based FEP uses AM1 for the gradients and energies (at each MD step) of inhibitors and the HF/6-31G*/ESP for partial atomic charges of the inhibitors in the beginning of the simulation, and solvent and protein

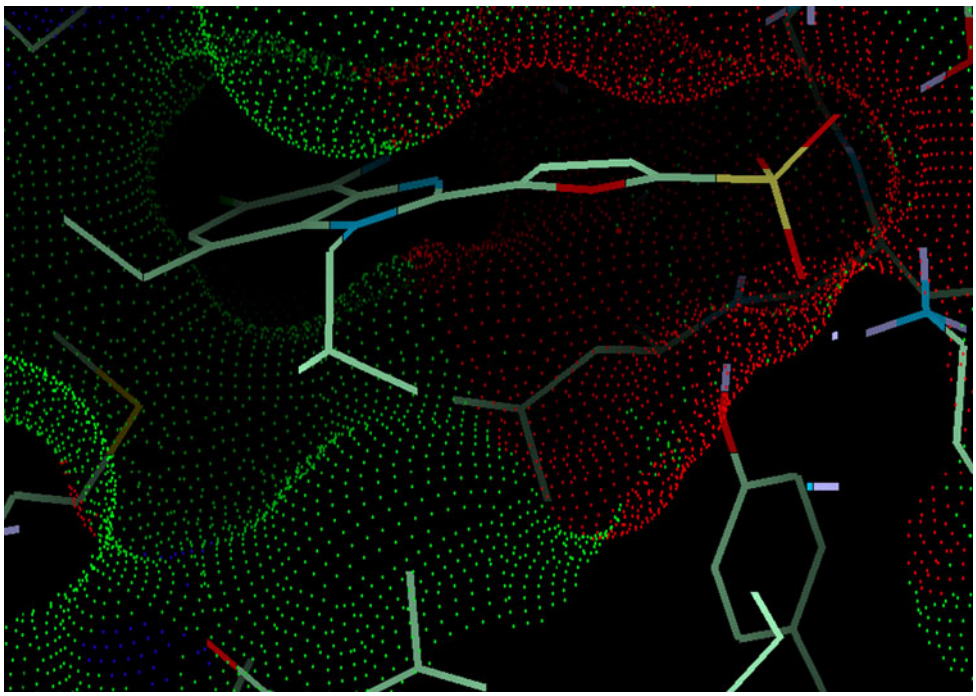

Fig. 4 Connolly dot surface generated using the structure of the human FBPase:**14** complex. The surface was generated using a van der Waals radius of 1.4 Å. The surface is colored based on electrostatic potential with surface points colored *red* indicating positive charge, *green* neutral and *blue* negative charge

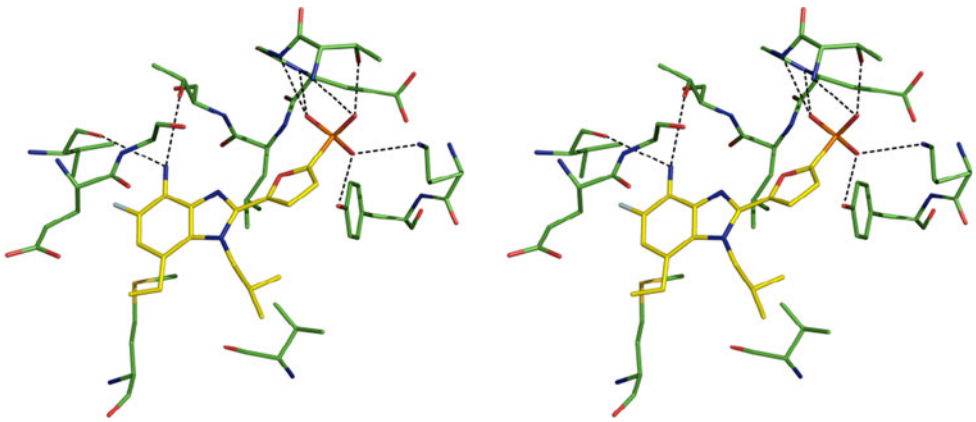

Fig. 5 Stereoview of X-ray structure of the human FBPase:**14** complex showing the interactions between the binding site of FBPase and **14**. Putative hydrogen bonds (considered when distance between the acceptor and donor is less than 3.3 Å) are shown as *dotted lines*

atoms were treated using MM force field; and (b) classical-MM uses conventional FEP method with HF/6-31G*/ESP partial atomic charges for all the inhibitors. We have divided the calculated relative binding free energies between inhibitors into five different classes, based on the size and location of the mutations giving rise to 22 important mutations: (1) mutations (**1** to **3**) consist of substitutions only in the R^9-group and keeping rest of the molecule same, (2) mutations (**4** to **12**), only in C8-spacer group and keeping rest of the molecule same, (3) mutations (**13** to **15**) only in the six-member pyrimidine ring (X and X′) and keeping the rest of molecule same, (4) mutations (**16** to **20**) in both the R^9-group and C8-spacer simultaneously by keeping other parts of the molecule same, and (5) mutations (**21** to **22**) in the R-group, Spacer and Six-member ring simultaneously.

4.1 Substitution at R^9-position

As shown in Figs. 4 and 5, a hydrophobic group is preferred at R^9-position that contributes to the binding affinity by forming favorable van der Waals interactions with the hydrophobic region of the R^9 binding pocket (Fig. 4) within the AMP-binding site. The following three mutations were tested: (1) **Ribose → H**, (2) **H → *cyclo*-hexylethyl**, and (3) **Phenethyl → *cyclo*-hexylethyl**, by keeping all other groups same. The perturbations **1** and **2** are very large (Fig. 2), which are difficult to converge. The

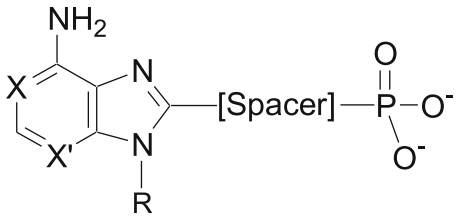

Fig. 6 FBPase inhibitors considered for FEP calculations and their IC_{50} values

Compound #	C8-Spacer	Pyrimidine		R^9	IC_{50} (μM)
		X	X'		
1	$N(H)CH_2$	N	N	ribose	>1000[a]
2	$N(H)CH_2CH_2$	N	N	ribose	118 ± 9
3	$N(H)CH_2CH_2CH_2$	N	N	ribose	>1000[b]
4	$N(H)CH_2CH_2$	N	N	H	500 ± 10
5	$N(H)CH_2$	N	N	*c*hexylethyl	~1000
6	$N(H)CH_2CH_2$	N	N	*c*hexylethyl	97 ± 20
7	$CH_2CH_2CH_2$	N	N	*c*hexylethyl	26 ± 1
8	CH_2OCH_2	N	N	*c*hexylethyl	15 ± 3
9	2,5-furanyl	N	N	phenethyl	5 ± 0
10	2,5-furanyl	N	N	*c*hexylethyl	1.2 ± 0.9
11	2,5-furanyl	N	N	*i*Bu	2.2 ± 0.2
12	2,5-furanyl	N	N	*neo*Pentyl	1.1 ± 0.2
13	2,5-furanyl	CH	CH	*i*Bu	1.6 ± 0.1
14	2,5-furanyl	CF	C-Et	*i*Bu	0.09 ± 0.1

[a]27% inhibition at 1 mM; [b]33% inhibition at 1 mM. Abbreviations: *c*hexylethyl, *cyclo*-hexylethyl, *i*Bu, *iso*-Butyl, C-Et, C-Ethyl, phenethyl, phenylethyl.

Table 1 Relative binding free energies of FBPase inhibitors for mutation **1–3**

No.	Transformation (S1 → S2)		ΔΔ*G* (QM/AM1/MM)[a]	ΔΔ*G* (MM)[b]	$\Delta\Delta G_{bind}$ (exp)[c]
1	**2** → **4**[d]	Ribose → H	1.3 ± 0.9	1.4 ± 0.9	0.9
2	**4** → **6**[d]	H → *cyclo*-hexylethyl	−1.4 ± 0.9	−1.6 ± 0.9	−1.0
3	**9** → **10**[d]	Phenethyl → *cyclo*-hexylethyl	−1.1 ± 0.8	−1.4 ± 0.8	−0.8

[a] Calculated using QM/MM-based FEP method

[b] Calculated using a conventional FEP method

[c] Values obtained from experimental data

[d] Double topology

suitability of FEP methods in such cases is very important for successful application of computer-aided drug design [CADD] methods in drug discovery. The calculated relative binding free energies for these analogs (Table 1) using both conventional as well as QM/MM-based FEP methods suggest that R^9-group substitution such as *cyclo*-hexylethyl group leads to favorable binding affinity due to gain both in desolvation penalty as well as favorable hydrophobic interactions of *cyclo*-hexylethyl with the hydrophobic residues surrounding the R^9–position (Figs. 4, 5) as compared to other two substitutions, and these results are consistent with the experimentally measured values (Table 1).

4.2 Substitution at C8-spacer (mutations 4–12)

The relative binding free energies were performed for the identification of optimal spacer (2–4 atoms) using different combination of substitutions at C8 position. We have used the mutations **4–12** (Table 2) for this study and calculated relative binding free energies for these mutations. Both

Table 2 Relative binding free energies of FBPase inhibitors for the mutations **4–12**

No.	Transformation (S1 → S2)		ΔΔ*G* (QM/AM1/MM)[a]	ΔΔ*G* (MM)[b]	$\Delta\Delta G_{bind}$ (exp)[c]
4	**1** → **2**[d]	$N(H)CH_2$ → $N(H)CH_2CH_2$	−1.7 ± 0.6	−2.0 ± 0.7	<−1.3
5	**2** → **3**[d]	$N(H)CH_2CH_2$ → $N(H)CH_2CH_2CH_2$	1.6 ± 0.6	1.8 ± 0.7	>1.3
6	**5** → **6**[d]	$N(H)CH_2$ → $N(H)CH_2CH_2$	−1.5 ± 0.6	−1.7 ± 0.6	−1.4
7	**5** → **7**[d]	$N(H)CH_2$ → $CH_2CH_2CH_2$	−2.5 ± 0.6	−2.6 ± 0.6	−2.2
8	**5** → **8**[d]	$N(H)CH_2$ → CH_2OCH_2	−2.7 ± 0.6	−3.0 ± 0.6	−2.5
9	**5** → **10**[d]	$N(H)CH_2$ → 2,5-furanyl	−4.3 ± 0.7	−4.7 ± 0.7	−4.0
10	**6** → **7**[e]	$N(H)CH_2CH_2$ → $CH_2CH_2CH_2$	−1.1 ± 0.5	−1.3 ± 0.5	−0.80
11	**7** → **8**[e]	$CH_2CH_2CH_2$ → CH_2OCH_2	−0.7 ± 0.5	−0.8 ± 0.6	−0.40
12	**8** → **10**[d]	CH_2OCH_2 → 2,5-furanyl	−1.8 ± 0.7	−1.9 ± 0.7	−1.5

[a] Calculated using QM/MM-based FEP method

[b] Calculated using a conventional FEP method

[c] Values obtained from experimental data

[d] Double topology

[e] Single topology

Table 3 Relative binding free energies of FBPase inhibitors for mutations **13–15**

No	Transformation (S1 → S2)		ΔΔ*G* (QM/AM1/MM)[a]	ΔΔ*G* (MM)[b]	$\Delta\Delta G_{bind}$ (exp)[c]
13	**11** → **13**[e]	X,X′ = N → CH	−0.5 ± 0.4	−0.7 ± 0.5	−0.2
14	**11** → **14**[d]	X = N → C–F; X′ = N → C-Et	−2.2 ± 0.6	−2.4 ± 0.6	−1.9
15	**13** → **14**[d]	X = CH → C–F; X′ = CH → C-Et	−2.0 ± 0.6	−2.1 ± 0.6	−1.7

[a] Calculated using QM/MM-based FEP method

[b] Calculated using a conventional FEP method

[c] Values obtained from experimental data

[d] Double topology

[e] Single topology

conventional and QM/MM-based FEP methods suggest that 3-atom linker is the best spacer, which were consistent with the experimental results (Table 2). Among various combination of 3-atom spacer examined using QM/MM- and conventional FEP methods, the order of preference is as follows: $N(H)CH_2 < N(H)CH_2CH_2 < N(H)CH_2CH_2 < CH_2CH_2CH_2 < CH_2OCH_2 <$ 2,5-furanyl. Entropic changes associated with the ring spacer as well as favorable electrostatic interactions in the binding site (Fig. 4) give rise to 2,5-furanyl spacer as the most suitable linker.

4.3 Substitutions on pyrimidine ring

Since in pyrimidine ring, N1 and N3 substitutions do not have any hydrogen bonds with the binding site protein residues, it is expected to minimize desolvation penalty as well as gain favorable hydrophobic interactions with protein residues (Fig. 4), when N1 and N3 are replaced by non-polar substitutions shown in the mutations **13–15**. Accordingly, three mutations (Fig. 6, Table 3): (1) X, X′ **= N → CH**, (2) X = **N → C–F**; X′ = **N → C-Et** and (3) X = **CH → C–F**; X′ = **CH → C-Et** were performed using both conventional as well as QM/MM-based FEP methods. The substitution **N → CH** in the mutation **13** produces favorable binding free energy by (−0.5 kcal/mol), and the bulky hydrophobic substitutions shown in the mutations **14** and **15** give even better results with the net gain of approximately −2.0 kcal/mol, in agreement with the experimental data (Table 3). As expected, the non-polar substitutions reduced desolvation penalty and gained hydrophobic interactions with protein residues (Figs. 4, 5) leading to favorable binding energy.

4.4 Simultaneous substitutions on R^9- position and C8-spacer

We examined the mutations **16** to **20** to understand the changes in the binding free energies for simultaneous double substitutions and to evaluate the applicability of the

Table 4 Relative binding free energies of FBPase inhibitors for mutations **16–20**

No.	Transformation (S1 → S2)		$\Delta\Delta G$ (QM/AM1/MM)[a]	$\Delta\Delta G$ (MM)[b]	$\Delta\Delta G_{bind}$ (exp)[c]
16	**1** → **7**[d]	$N(H)CH_2$ → $CH_2CH_2CH_2$; Ribose → *cyclo*-hexylethyl	−2.6 ± 0.9	−2.8 ± 0.9	<−2.2
17	**4** → **10**[d]	$N(H)CH_2CH_2$ → 2,5-furanyl; H → *cyclo*-hexylethyl	−4.0 ± 0.9	−4.3 ± 0.9	−3.6
18	**4** → **12**[d]	$N(H)CH_2CH_2$ → 2,5-furanyl; H → *neo*Pentyl	−4.1 ± 0.9	−4.4 ± 0.9	−3.7
19	**5** → **9**[d]	$N(H)CH_2$ → 2,5-furanyl; *cyclo*-hexylethyl → Phenethyl	−3.6 ± 0.9	−3.8 ± 0.9	−3.2
20	**5** → **12**[d]	$N(H)CH_2$ → 2,5-furanyl; *cyclo*-hexylethyl → *neo*Pentyl	−4.4 ± 0.9	−4.8 ± 0.9	−4.0

[a] Calculated using QM/MM-based FEP method

[b] Calculated using a conventional FEP method

[c] Values obtained from experimental data

[d] Double topology

Table 5 Relative binding free energies of FBPase inhibitors for mutation **21–22**

No.	Transformation (S1 → S2)		$\Delta\Delta G$ (QM/AM1/MM)[a]	$\Delta\Delta G$ (MM)[b]	$\Delta\Delta G_{bind}$ (exp)[c]
21	**5** → **13**[d]	$N(H)CH_2$ → 2,5-furanyl; X = N → CH; X′ = N → CH; *cyclo*-hexylethyl → *i*Bu	−4.2 ± 0.9	−4.6 ± 0.9	−3.8
22	**5** → **14**[d]	$N(H)CH_2$ → 2,5-furanyl; X = N → C–F; X′ = N → C-Et; *cyclo*-hexylethyl → *i*Bu	−5.1 ± 0.9	−4.7 ± 0.9	−5.5

[a] Calculated using QM/MM-based FEP method

[b] Calculated using a conventional FEP method

[c] Values obtained from experimental data

[d] Double topology

method for multiple substitutions in mutations. Relative binding free energy calculations for simultaneous mutations at C8-spacer and R^9-positions were carried out for mutations **16** to **20**. The calculated relative binding free energies using these two FEP methods are consistent with the experimental results (Table 4). These calculations supported the previous observation that analog containing 2,5-furnyl spacer is better as compared to other spacers. And also, the combined substitutions of 2,5-furnyl and *neo*-pentyl substitutions give rise to much better analog with $\Delta\Delta G_{bind}$ of −4.4 kcal/mol (exp. −4.0 kcal/mol).

4.5 Triple substitutions

Finally, simultaneous substitution at all the three places; pyrimidine ring, R^9-position and C8-spacer were performed for the mutations **21** and **22** (Table 5) using both conventional and QM/MM-based FEP methods, and the calculated results are consistent with the experimental data (Table 5). The calculated results for mutation **22** (Spacer → 2,5-furanyl; Pyrimidine **X = N → C–F**; **X′ = N → C-Et**; **R^9 → *cyclo*-hexylethyl**) suggest that compound **14** is the best lead compound, in agreement with our experimental data (Table 5).

5 Conclusion

A QM/MM-based FEP method was validated for predicting binding affinities of a large set of diverse FBPase inhibitors. A method that could be used for structurally diverse compounds enhances the scope for in silico virtual screening. FEP methods are generally employed for structurally similar compounds, and therefore, relative differences in binding energies are typically very small. Hence, successful application of FEP method requires its high accuracy to draw a meaningful conclusion. As demonstrated in the present validation test, comparison with experimental data indicates that the QM/MM-based FEP method is highly accurate that is capable of reproducing the experimental binding affinities within ±0.5 kcal/mol, which is more precise than all other approximate methods including conventional FEP methods.

While the QM/MM-based FEP method requires excessive computational time relative to conventional FEP and other less rigorous methods of CADD, the benefit is increased accuracy, which in turn is expected to provide additional important insights and better guidance for drug design. It is pertinent to note that computational power is expected to improve significantly in the future with

continual advancements in computer hardware along with parallelization of the code to enable simultaneous use of multiple processors. Accurate prediction of inhibitor affinity is expected to shorten the time required to find suitable candidates for drug development by focusing chemistry efforts on more promising compound series, thereby eliminating the time spent on synthesis and characterization of unwanted and incorrectly predicted compounds. The QM/MM-based FEP method is also expected to avoid the need for time-consuming generation of MM force field parameters for novel inhibitor scaffolds, and the inaccuracies originating from parameters derived in the absence of experimental data. Elimination of this step simplifies the process and should enable future automation of these calculations. The results using the present QM/MM method are expected to facilitate and promote the use of FEP calculations in the pharmaceutical industry and among scientists focusing on the identification and optimization of lead compounds for drug discovery.

References

1. Åqvist J, Marelius J (2001) The linear interaction energy method for computation of ligand binding affinities. In: Reddy MR, Erion MD (eds) Free energy calculations in rational drug design. Kluwer/Plenum Press, New York, pp 171–194
2. Reddy MR, Viswanandhan VN, Erion MD (1998) Rapid estimation of relative binding affinities of enzyme inhibitors. In: 3D QSAR in drug design, vol 2. Kluwer publishers, New York, pp 85–96
3. Sundriyal S, Viswanad B, Ramarao P, Chakraborti AK, Bharatam PV (2008) Bioorg Med Chem Lett 15:4959–4962
4. Shaikh S, Jain T, Sandhu G, Latha N, Jayaram B (2007) Curr Pharm Des 13:3454–3470
5. Jain T, Jayaram B (2005) FEBS Lett 579:6659–6666
6. Alvarej J, Shoichet B (2005) Virtual screening in drug discovery. CRC press, Taylor & Francis, Boca Raton, FL
7. Zwanzig RJ (1954) J Chem Phys 22:1420–1426
8. Tembe BL, McCammon JA (1984) Comp Chem 8:281–283
9. Kollman PA (1993) Chem Rev 93:2395–2417
10. Jorgensen WL, Briggs JM, Contreras ML (1990) J Phys Chem 94:1683–1686
11. Jorgensen WL, Briggs JM, Gao J (1987) J Am Chem Soc 109:6857–6858
12. Erion MD, Reddy MR (1998) J Am Chem Soc 120:3295–3304
13. Bash P, Singh UC, Brown FK, Langridge R, Kollman PA (1987) Science 235:574–575
14. Brooks CL III, Fleischman SH (1990) J Am Chem Soc 112:3307–3312
15. Rao BG, Tilton RF, Singh UC (1992) J Am Chem Soc 114:4447–4452
16. Reddy MR, Erion MD, Agarwal A (2000) Free energy calculations: use and limitations in predicting ligand binding affinities. In: Lipkowitz KB, Boyd DB (eds) Reviews in computational chemistry, vol 16. Wiley-VCH Inc., New York, pp 217–304
17. Reddy MR, Erion MD (eds) (2001) Free energy calculations in rational drug design. Kluwer/Plenum Press, New York
18. Reddy MR, Singh UC, Erion MD (2004) J Am Chem Soc 126:6224–6225
19. Reddy MR, Singh UC, Erion MD (2007) J Comp Chem 28:491–494
20. Reddy MR, Erion MD (2007) J Am Chem Soc 129:9296–9297
21. Rathore RS, Aparoy P, Reddanna P, Kondapi AK, Reddy MR (2011) J Comp Chem 32:2097–2103
22. Reddy MR, Singh UC, Erion MD (2011) J Am Chem Soc 133:8059–8061
23. Magnusson I, Rothman DL, Katz LD, Shulman RG, Shulman GI (1992) J Clin Invest 90:1323–1327
24. Villeret V, Huang S, Zhang Y, Lipscomb WN (1995) Biochemistry 34:4307–4315
25. Choe JY, Nelson SW, Arienti KL, Axe FU, Collins TL, Jones TK, Kimmich RD, Newman MJ, Norvell K, Ripka WC, Romano SJ, Short KM, Slee DH, Fromm HJ, Honzatko RB (2003) J Biol Chem 278:51176–51183
26. Erion MD, van Poelje PD, Dang Q, Kasibhatla SR, Potter SC, Reddy MR, Reddy KR, Jiang T, Lipscomb WN (2005) Proc Natl Acad Sci USA 102:7970–7975
27. van Poelje PD, Potter SC, Chandramouli VC, Landau BR, Dang Q, Erion MD (2006) Diabetes 55:1747–1754
28. van Poelje PD, Dang Q, Erion MD (2007) Curr Opin Drug Discov Devel 10:430–437
29. van Poelje PD, Dang Q, Erion MD (2007) Drug discovery today. Ther Strategies 4:103–109
30. Bharatam PV, Patel DS, Adane L, Mittal A, Sundriyal S (2007) Curr Pham Des 13:3518–3530
31. Dang Q, Reddy MR, Brown BS, Liu Y, Rydzewski RM, Robinson ED, van Poelje PD, Erion MD (2009) J Med Chem 52:2880–2898
32. Erion MD, Dang Q, Reddy MR, Kasibhatla SR, Huang J, Lipscomb WN, van Poelje PD (2007) J Am Chem Soc 129:15480–15490
33. Dang Q, Kasibhatla SR, Reddy KR, Jiang T, Reddy MR, Potter SC, Fujitaki JM, van Poelje PD, Huang J, Lipscomb WN, Erion MD (2007) J Am Chem Soc 129:15491–15502
34. Raschke TM, Tsai J, Levitt M (2001) Proc Natl Acad Sci USA 98:5965–5969
35. Pearlman DA (2001) Free energy calculations: methods for estimating ligand binding affinities. In: Reddy MR, Erion MD (eds) Free energy calculations in rational drug design. Kluwer Academic/Plenum Publishers, New York, pp 9–35
36. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagoha G, Profeta S Jr, Weiner PK (1984) J Am Chem Soc 106:765–784
37. Singh UC, Weiner PK, Caldwell JK, Kollman PA (1986) AMBER (Version 3.0), University of California at San Francisco, San Francisco
38. Reddy MR, Erion MD (2001) J Am Chem Soc 12:6246–6252
39. Berendsen HJC, Grigera JR, Straatsma TP (1987) J Phys Chem 91:6269–6271
40. Reddy MR, Berkowitz M (1989) Chem Phys Lett 155:173–176
41. Verlet L (1967) Phys Rev 159:98–103
42. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) J Comp Phys 23:327–341
43. Ke HM, Zhang YP, Lipscomb WN (1990) Proc Natl Acad Sci USA 87:5243–5247
44. Gidh-Jain M, Zhang Y, van Poelje PD, Liang JY, Huang S, Kim J, Elliott J, Erion MD, Pilkis SJ, Raafat el-Maghrabi M, Lipscomb WN (1994) J Biol Chem 269:27732–27738
45. Galaxy Molecular Modeling Software and AM2000. Macromolecular Simulation Package; AM Technologies: San Antonio, Texas, 1995
46. Protein Data Bank. http://www.rcsb.org