molecular informatics

DOI: 10.1002/minf.201000007

Free Energy Calculations of Mutations Involving a Tightly Bound Water Molecule and Ligand Substitutions in a Ligand-Protein Complex

Alfonso T. García-Sosa*[a] and Ricardo L. Mancera[b]

Abstract: The accurate calculation of the free energy of interaction of protein-water-ligand systems has an important role in molecular recognition and drug design that is often not fully considered. We report free energy thermodynamic integration calculations used to evaluate the effects of inclusion, neglect, and targeting and removal (i.e., systematic substitution by ligand functional groups) of an important, tightly bound, water molecule in the SH3 domain of Abl tyrosine kinase. The effects of this water molecule on the free energies of interaction of several Abl-SH3 domain-ligand systems reveal that there is an unfavourable free energy change associated with its removal into the bulk solvent. Only three substitutions by an additional functional group (out of methyl, ethyl, hydroxyl, amino, and amide groups) in the phenyl ring of a tyrosine in the peptide ligand resulted in a favourable change in the free energy of binding upon replacement of the ordered water molecule. This computational approach provides a direct route to the systematic and rigorous prediction of the thermodynamic influence of ordered, structural water molecules on ligand modification and optimization in drug design by calculating free energy changes in protein-water-ligand systems.

Keywords: Free energy of binding · Water · Hydration · Drug design · Thermodynamic integration · Protein-ligand complex

1 Introduction

It is well known experimentally that water molecules have important effects on biomolecular binding, recognition and drug design.[1–7] There are different ways of incorporating these effects into computational drug design. Water molecules can be included in docking and scoring strategies,[8–15] and in pharmacophores.[16] The importance of including the full entropic effects of water molecules is also being recognized.[17,18] This is due to the pivotal role and effects that ordered, tightly-bound, crystallographically-observed water molecules can have in protein-ligand binding energies and structure.

Consideration of the thermodynamic effects of these structural, ordered water molecules is critical for the correct analysis of the structure and energetics of many biological complexes.[19,20] Some water molecules are loosely associated with the protein surface, whereas others can be strongly associated with the biomolecule and ligand in a ternary complex.[21,22] Water molecules can be targeted for substitution by ligands, with resulting favourable free energies of binding as in the classical cases of substituted cyclic ureas for HIV-1 aspartic protease (which also benefited from a loss of ligand entropy),[23] and the inhibitors of histone deacetylase.[24] However, there have also been cases reported where either the substitution group on the ligand did not replace the water molecule, and/or it resulted in an unfavourable substitution[25] or only a modest improvement in binding affinity.[26] In addition, the entropy of these water molecules can have a wide range of values, from ice to bulk water,[22] so their displacement will not always bring about a decrease in entropy. Ordered, bridging water molecules can also greatly contribute to the selectivity of ligands.[27] Clearly, an accurate and systematic modelling of the inclusion, neglect, or substitution of water molecules is required in order to predict if the change in free energy of binding will be favourable.

The energetic contributions of structural water molecules have been calculated, for example, through molecular mechanics calculations,[10] molecular dynamics,[28] free energy perturbation,[17,29–33] hydropathic interactions,[34] and statistical mechanical formulas,[35] amongst others.[36,37] We previously determined the energy neglected by ignoring important water molecules bridging between the protein poly(-ADP-ribose)-polymerase (PARP) and inhibitor, as well as dif-

[a] *A. T. García-Sosa*
Institute of Chemistry, University of Tartu
Jakobi 2-319, Tartu 51014, Estonia
tel: + 372 737 5270; fax: + 372 737 5264
**e-mail: alfonsog@ut.ee*
t.alfonso@gmail.com

[b] *R. L. Mancera*
Curtin Health Innovation Research Institute, Western Australian Biomedical Research Institute, School of Biomedical Sciences and School of Pharmacy, Curtin University of Technology
GPO Box U1987, Perth WA 6865, Australia

Supporting Information for this article is available on the WWW under http://dx.doi.org/10.1002/minf.200900007

ferent substitutions by functional groups on ligands using molecular mechanics calculations.[10] Improper treatment of the solvent in the setup of molecular dynamics simulations can lead to large scale deformities or collapses in the protein structure.[28] A related molecular dynamics technique, free energy perturbation, allows full treatment of solvent degrees of freedom and has been used to replace a whole molecule of camphor in the binding site of cytochrome P450cam for water molecules in order to calculate its absolute binding free energy.[29,30] Monte Carlo simulations with replica exchange thermodynamic integration were used to determine the absolute binding free energies of different classes of water molecules observed in a selection of crystal structures, with the observation that conserved water molecules had tighter binding than those displaced by a ligand.[17] Free energy perturbations were carried out on two water molecule systems, in a polar protein landscape environment and in a non-polar one.[31] Olano et al. found that the water in the polar cavity was preferred to an empty polar cavity, and that the addition of water to either cavity made the protein more flexible.[31] Free energy perturbations also revealed the important energetic contribution of water molecules in *wetspots*, regions between the interfaces of proteins, as well as their longer residence times compared to other surface water molecules, and how they allow sequence variability in the binding partner.[33] Their energetic contribution was of the same magnitude as that provided by protein residues, and their mobility was even lower than the latter for the several cases studied.[33] Molecular dynamics with statistical mechanical formulas to combine functionals for the entropy and enthalpy of molecular correlation functions in inhomogeneous infinitely dilute solutions (using spatially dependent correlation functions on the origin-fixed solute instead of averages over the entire fluid as the standard homogeneous correlation functions) and the inclusion of desolvation and entropy of conformational restriction have been used to determine the effect of substituting a bound water molecule by a hydroxyl group on a ligand in concanavalin A.[35] The results were in agreement with experiment, showing the unsubstituted case as the thermodynamically preferred situation due to the stronger interaction of an ion with the water molecule than with the extra hydroxyl group.[35] Hydropathic calculations also showed that four water molecules in the structure of HIV-1 protease were crucial for a proper account of the binding energy.[34] Empirical functions were able to provide a degree of quantification of the free energy contributions provided by ligands evacuating adjacent hydration sites in bound complexes.[37] However, the functions are simple and do not include the electrostatic and van der Waals interaction energies between protein and ligand, nor the ligand solvation free energy (in addition to other terms such as ligand configurational entropy and protein-reorganization free energy).[37]

A water molecule is particularly sensitive to its environment when bridging a biomolecule-ligand interaction or when located on the surface of a biomolecule. This environment, together with the differences in both desolvation energy and the interactions (enthalpic and entropic contributions) that a substitution group may have, will dictate the behaviour and energetics of different biomolecular associations. These in turn, have direct implications for molecular recognition, which is at the heart of biochemical and biological processes and for the rational design of ligands as inhibitors, agonists or antagonists. It is important to take into account interactions with water molecules in order to accurately predict the free energy of interaction in ligand-protein complexes, which is otherwise likely to be incorrectly estimated. Therefore, water molecules that are tightly-bound and have strong interaction energies with the protein-ligand surface should be either retained or carefully substituted in drug design and protein mutation strategies.

The Abelson (Abl) tyrosine kinase (TK) mediates protein–protein interactions and therefore cellular signaling, and is disrupted in chronic myelogenous leukaemia (CML), which can be a deadly disease. The Abl forms a protein-protein complex with the SH3 domain. The cellular Abelson leukemia virus, c-Abl, is made constitutively active by mutation or deletion of the SH3 domain. High-affinity ligands for the SH3 domain can have anti-tumour activity and enhance the effects of Abl TK-binding drugs such as Imatinib (Gleevec).[38] The specificity in SH2 and SH3 domain interactions is of significant interest for the understanding of tyrosine phosphate signal-transduction pathways and the discovery of drugs that can interfere with the cellular functions of SH2 or SH3 domain-mediated processes in disease pathways.[39] Palencia et al. studied the SH3 domain through isothermal titration calorimetry, and established a critical role for the hydrogen bonding network of water molecules in the protein–ligand interface to explain the large favourable enthalpy of binding accompanied by a large unfavourable entropy of binding for a peptide-protein complex dominated otherwise by hydrophobic interactions.[39] They also proposed substituting a tyrosine residue in the ligand for methionine, which improves binding, as well as substituting a serine in the ligand by a short polar residue capable of making strong hydrogen bonds, such as a glutamine.[39]

In the present work, a peptide ligand in complex with the SH3 domain of the Abl tyrosine kinase was studied by adding functional groups to a tyrosine side chain in the peptide ligand in a region of the protein containing a tightly-bound water molecule (W2041) that bridges the protein-ligand interaction. We determined the thermodynamic influence of this particular hydration site, when occupied by a tightly-bound water molecule bridging the protein-ligand interaction (inclusion), when this water molecule is removed (neglect), and when this water molecule is substituted by different substituent groups on the ligand (targeting). This water molecule in the ligand binding site was 'annihilated' (mutated into nothing and then re-introduced into the bulk solvent) to measure the associated free

energy change in the protein interior as well as in protein-ligand complexes, allowing the measurement of its contribution to the total free energy of binding. Water molecules have been mutated to calculate the energy of their displacement by a ligand in the protein-ligand binding site in crystal structures of complexes,[17] or to calculate their energy in the SH3 domain interface.[33] However, the method described here uses free energy thermodynamic integration techniques to calculate the predicted changes in binding affinity of several different functional groups on the same ligand after displacing a tightly-bound water molecule which has not been previously considered. These functional groups are chemical probes that can characterize the binding properties of a variety of molecules: hydroxyl, amide, amine, methyl and ethyl groups. By combining different techniques, this approach provides an accurate and formally rigorous method of determining the likely appropriate modifications of a ligand structure, and may thus prove useful in the optimization of ligands and peptides, since the most suitable functional groups can be chosen to target one or more water molecules in the protein binding site. This approach is superior to the generally practiced neglect of water molecules in molecular docking studies. We present a systematic use of ligand group substitutions as chemical probes using thermodynamic integration free energy perturbation calculations to predict the effects of inclusion, neglect, and substitution of an ordered, tightly-bound water molecule bridging a protein-ligand complex by different, chemically meaningful and representative ligand functional groups.

2 Computational Methods

The structure of the enzyme Abl-SH3 domain tyrosine kinase in complex with a peptide was retrieved from the Protein DataBank (PDB code 1bbz[40]), and chains A and B were extracted. The p41 peptide ligand (APSYSPPPPP) in this structure is a member of a group of peptide ligands designed to bind specifically to the Abl-SH3 domain. It has a dissociation constant, K_d, of 1.5 μM,[40] and the system is sensitive to changes in its peptide environment.[39]

Minimizations and molecular dynamics simulations were carried out using the sander module in Amber7.[41] Free energy perturbations (alchemical mutations)[42] were performed with the Gibbs module in Amber7. Hydrogens and protonation states were assigned using Maestro (Schrödinger, Inc).[43,44] Partial atomic charges for the ligands were derived from HF/6–31g* ab initio geometry optimizations with Gaussian[45] and adapted for the parm99 forcefield[46,47] through the RESP procedure.[48–50]

In the present work, a number of mutations were carried out on Tyr4 (chain B) in the ligand by adding various functional groups to the phenyl ring of the side chain. The neighbouring water molecule (W2041) that interacts with both the ligand and the protein surface was also 'annihilated' by being mutated into nothing and then introduced into bulk water. Tyr4 was modified to include six different functional groups (methyl, ethyl, hydroxyl, amino, and amide) on its aromatic phenyl ring. W2041 is well surrounded by protein and ligand atoms (it is not accessible to the solvent and has a WaterScore = 1, indicating that it is a tightly-bound water molecule with moderate to low B-factor, low solvent accessible surface area and that involves contacts to the protein[9]). Mutations in the complex and the free ligands were carried out in explicit TIP3P water, under periodic boundary conditions, at a constant pressure of 1 atm, and with a minimum distance of 8 Å between the solute and the edges of the box. Sodium counter-ions were added to neutralize the overall charge of the systems, as needed. Gas phase mutations were carried out in the vacuum.

The complexes were slowly heated from 0 to 300 K with the use of 31 gentle equilibration runs of 310 ps in total, followed by constant pressure equilibration with a 1.0 fs timestep for up to 2.0 ns. When the systems were determined to have non diverging RMSDs under 1 Å and bound structures, stable temperatures and energies, they were considered to be ready for the subsequent free energy simulations. The particle-mesh-Ewald[51] method was used for the calculation of long-range electrostatics. A cut-off threshold of 8.0 Å was used to directly calculate electrostatic interactions. Constant pressure was ensured using isotropic scaling and constant temperature was ensured using the Berendsen coupling algorithm.[52]

Thermodynamic integration free energy calculations were carried out with dynamically modified windows that allowed varying the size of the reaction coordinate parameter λ according to the size of the energy change. The SHAKE algorithm (constraining of bonds to hydrogens) was not applied. A total of 500 equilibration steps were performed between each λ sampling window. Three sampling procedures were chosen using 50 000, 75 000, and 100 000 steps of sampling for each λ window, corresponding to 50 ps, 75 ps and 100 ps, respectively. Transformations were performed from the mutated ligand back to tyrosine, i.e., by annihilating the extra atoms (and mutating the last attached atom into an aromatic hydrogen HA atom type, as in tyrosine) instead of creating them. This allowed a smooth simulation, adequate sampling and prevented instabilities. The energy value was then taken as the changed sign energy value for this mutation. The first mutation was the deletion of water molecule W2041, which corresponds to the transition from state **1** (wild type state, see Table 1 and Figure 3) to state **2**. The rest of the ligand transformations involved annihilating the additional functional group to reach state **2**. The previously calculated energy (from state **1** to state **2**) was then added.

The van der Waals contributions to the free energy were separated from the electrostatic contributions in two consecutive runs by electrostatic decoupling.[53] This also helped to create a more gradual transition, which avoids

the problem of having charges on a particle with a very small radius and resulting 'blow-ups' (due to the presence of a very small atom with a high charge for its size that does not repel enough other atoms through nonbonded interactions).[54]

For the annihilation of water molecule W2041, the absolute binding free energy was calculated, whereas relative free energy binding differences were calculated between the different ligands. The water molecule W2041 was annihilated from the protein-ligand complex, and then also mutated from bulk water into the gas state. The water molecule W2041 was constrained as it was unable to leave the space between the protein and the ligand peptide due to steric repulsions with both, in addition to the use of a harmonic positional restraint of 2 kcal mol^{-1} to prevent the water molecule escaping from the region centred on W2041. Importantly, no direct replacement of this water molecule was observed by any other water molecule from the bulk solvent during annihilating. To annihilate W2041, an intermediate step was used where the partial charges on the water molecule were reduced to a state termed "hydrophobic water",[30,55] which comprises the same water molecule with reduced charges from those of the TIP3P three point charge, rigid model of water,[56] of +0.417 on the hydrogen atoms and −0.834 on the oxygen atom, to +0.175 and −0.350, respectively. After this intermediate step, the charges were finally reduced to zero. At this point the atomic radii were gradually brought to zero, completing the process. These series of transformations allowed the mutation to progress smoothly and numerical integration to be accurate. The free energy change obtained upon removing the water molecule from the binding site was corrected by a term to include the standard free energy of a water molecule constrained in the binding site, which is given by $-RT \ln(C^0(2\ \pi RT/k)^{3/2})$, where k is the effective restraining constant and C^0 is the standard concentration of 55.55 mol/L.[53,57]

During the ligand–protein simulations, a harmonic potential restraint of 2 kcal $mol^{-1}Å^{-1}$ was placed on the backbone atoms of the protein, including those in the binding site. Setting the restraint improves convergence and impacts favourably on the smoothness of the free energy changes. Since this restraint is a soft-harmonic potential, it allows enough movement in the protein backbone and captures the structural and dynamic adaptations of the protein to the removal of W2041. At the same time, a replacement water molecule from the bulk solvent cannot re-enter the former hydration site, as otherwise this would imply that the hydration site water molecule was not replaced and/or a different hydration site would have been created. All energetic contributions to the free energy from any constraint whose equilibrium value changed with lambda were calculated and added back to the final energy at the end of the simulation. All calculations and models were also inspected visually. In addition, for comparative purposes a series of runs were calculated that did not include this harmonic restraint on all the protein backbone atoms.

Various precautions were taken to ensure that the transitions between changes were conducted in a smooth manner and to minimise any nonsystematic errors, such as performing the simulations in explicit solvent, using periodic boundary conditions, using dynamically modified windows performing extensive sampling, using electrostatic decoupling, gradually replacing charges on atoms, applying thermodynamic integration with Gaussian quadrature to avoid an abrupt behaviour when annihilating atoms, and placing smooth restraints on the system (and verifying that they did not affect the temperature, energy or structure).

The relative free energy of binding between two ligands and a protein is defined within a double closed thermodynamic cycle, where in each closed cycle the sum of the free energy equals zero, as shown in Figure 1.

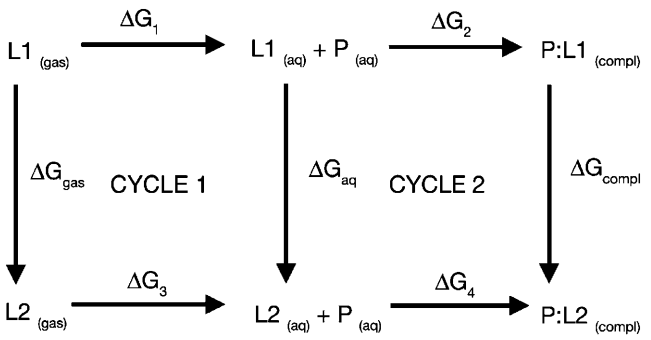

Figure 1. Thermodynamic cycles for the calculation of the relative free energy of binding of two different ligands L1 and L2 that interact with the same protein P. aq = aqueous phase, compl = in the binding complex.

The difference in the free energy of binding in aqueous solution between two ligands can be calculated as[58,59]

$$\Delta\Delta G_{bind} = \Delta G_4 - \Delta G_2 = \Delta G_{compl} - \Delta G_{aq}$$

while the change in the free energy of hydration of the ligands can be calculated as:

$$\Delta\Delta G_{hyd} = \Delta G_3 - \Delta G_1 = \Delta G_{aq} - \Delta G_{gas}$$

In this case, the energy required to annihilate the water molecule must be considered, in order to return to the initial state **1** which includes W2041. ΔG_{water_mut} is the sum of the free energy change associated with annihilating the water molecule from its location in the protein and the free energy change associated with re-introducing this water molecule into the bulk of the solution (the free energy of hydration of water):

$$\Delta G_{water\_mut} = \Delta G_{water\_annihilation} + \Delta G_{water\_hydration}$$

In the present case, we are interested in calculating the relative free energy of binding ΔG_{rel} that arises from alchemical mutations in a ligand in aqueous solution and that result in the displacement of a neighbouring water molecule, and which are measured by $\Delta\Delta G_{bind}$ and ΔG_{water_mut}, respectively. Hence, ΔG_{rel} is calculated as:

$$\Delta G_{rel} = \Delta\Delta G_{bind} + \Delta G_{water\_mut}$$

The free energy of ligand mutation in the ligand-protein complex (ΔG_{compl}) and the free energy of ligand mutation in aqueous solution (ΔG_{aq}) can be calculated by slowly perturbing the states of initial and final atoms through a reaction coordinate lambda (λ), gradually turning off the interactions in the initial state, $\lambda = 0$, and adapting to the final state when $\lambda = 1$.

Thermodynamic integration (TI) techniques for free energy calculations are based on rigorous equations from classical statistical mechanics, and are based on the integral[60]

$$\Delta G = \int_0^1 \left\langle \frac{dV(\lambda)}{d\lambda} \right\rangle_\lambda d\lambda \qquad (1)$$

where λ is the change parameter, V is the potential energy measured in a molecular dynamics simulation, and $\langle\ \rangle_\lambda$ is the ensemble average produced at each λ step. In practice, numerical integration is used to evaluate the integral. This requires that the integrand (ensemble) be evaluated at a series of λ intermediates. At each λ point, an equilibration of the system is performed, followed by data collection to determine the value of the energy of the ensemble for the equilibrated system.

3 Results and Discussion

After inspecting the crystal structure (PDB code 1bbz) and its water molecule network and environment, an ordered water molecule, W2041, was chosen as it is buried within the protein-ligand interface (chains A and B) and is excluded from the solvent (see Figure 2) in a moderately hydrophobic binding site. W2041 has strong interactions with the ligand-protein complex system such as hydrogen bonds to Ser12 OG (chain A), and to the Tyr4 phenolic OH (and in the crystal structure to Ser58 OG, chain E).

The structure of the peptide (p41, chain B) that binds to the ABL-SH3 domain is shown in Figure 3.

Several functional group modifications were made to Tyr4 (chain B) in order to place a functional group in the same location that W2041 has in the binding site. These modifications in the *ortho* position in the phenyl ring of Tyr4 included an extra hydroxyl (modified residue named Tya), methyl (Tyc), amine (Tyd), ethyl (Tye) and formamide (Tyf), as part of states **1** to **7**, represented in Figure 4.

The systems were equilibrated after heating. For the simulation of the original complex, analysis of the trajectories

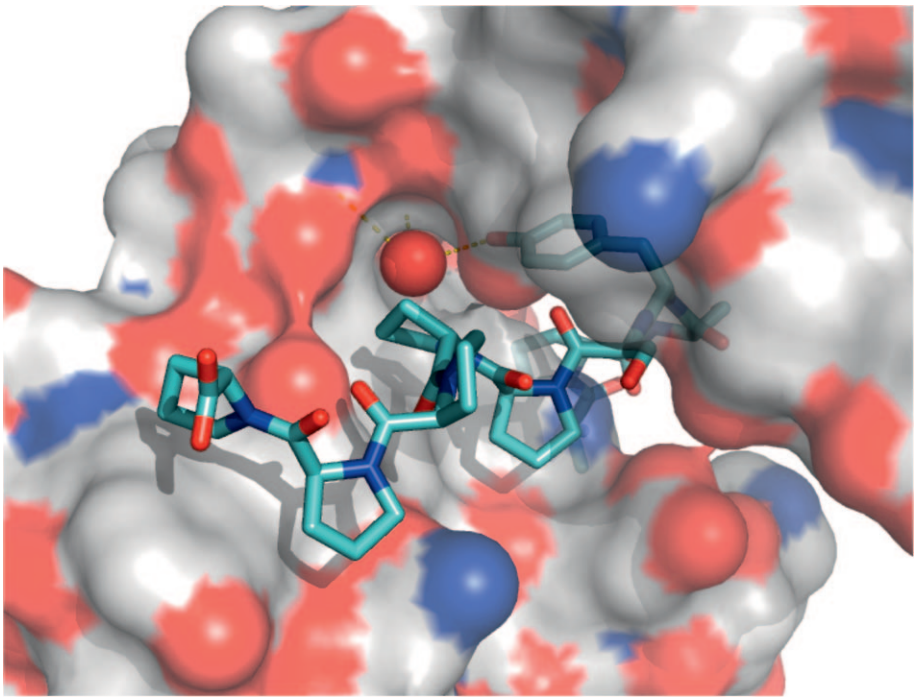

Figure 2. Complex between the ABL-SH3 domain and a ligand peptide (PDB structure 1bbz). The peptide ligand (chain B) is coloured in cyan, water molecule W2041 is shown as a red sphere.

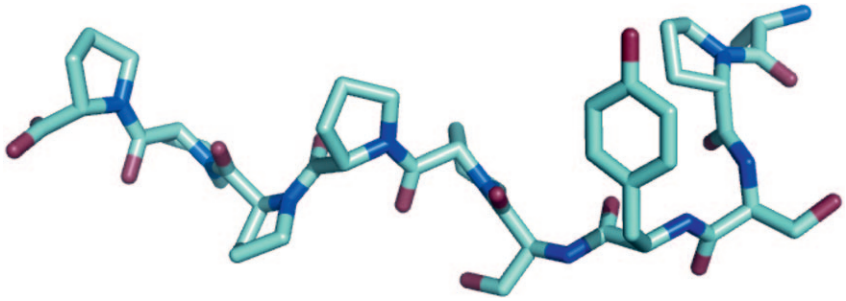

Figure 3. Structure of the peptide APSYSPPPPP (chain B) bound to the ABL-SH3 domain in PDB structure 1bbz.

for hydrogen bonds of W2041 with the ptraj module in Amber showed that the maximal occupancies were 95% during the trajectory snapshots for Ser12 OG (chain A) donating a hydrogen bond to the water, around 80% for the water donating to Ser12 OG, and around 30% for the water donating to Tyr4 OH (chain B).

For the mutation simulations, three independent procedures were carried out for each mutation and the final values were then taken from the mean of the three calculations. Table 1 shows the differences in energy calculated for each complex state **2–7**, relative to the wild-type, water-mediated state **1**. The correction term $-RT\ \ln(C^0(2\ \pi RT/k)^{3/2})$ was equal to -2.9 kcal mol^{-1} (effective $k = 0.7$ kcal mol^{-1}Å^{-2}), which was applied to obtain state **2**. The total simulation times for the transformations ranged from 4.1 ns (Tye) to 17.5 ns (Tyf).

Table 2 shows the partial free energy changes for each mutation.

Hysteresis in the calculations has been suggested to not be reliable in determining error estimates given the different convergence properties of the forward and backward simulations.[61] The error was calculated as in Shirts et al.,[62] resulting in 0.4 kcal mol^{-1} for Tyr → Tya, 0.6 kcal mol^{-1} for Tyr → Tyc, 0.6 kcal mol^{-1} for Tyr → Tyd, 0.8 kcal mol^{-1} for Tyr → Tye, and 0.5 kcal mol^{-1} for Tyr → Tyf. For comparison, hysteresis values were computed resulting in small errors, for example for Tyr ↔ Tyc, of 0.4 kcal mol^{-1}. Another measure of the error in these calculations was implemented by running a full cycle of transformations (Tyr → Tyc → Tye → Tyr), which gave a total result of (5.2–3.1–1.6 = 0.5 kcal mol^{-1}), in agreement with previous estimates. In addition, the series of calculations performed without including the restraining harmonic potential on all protein backbone atoms showed reasonably close values to those calculated with the restraint.

The number of λ windows varied from 36 to 344 in the different systems, depending on the size of the free energy change. Figure S1 in the Supplementary Information shows the free energy changes per window in the reaction coordinate (λ), for each of the three sampling methods. It can be seen from Figure S1 that the variations in free energy changes are small and smooth, which is required for a gradual transformation between systems. In addition, the three different sampling procedures provided comparable results. The qualitative behaviour of the simulations was as expected, since Figures S1(a) to S1(f) show that there were no significant energy or structural drifts, fluctuations, or distortions. The respective number of λ windows for the 50, 75 and 100 ps mutations in the complex were for Tya: 122, 132, and 139; for Tyc: 63, 64, and 75; for Tyd: 251, 54, and 100; for Tye: 50, 37, and 36; and for Tyf: 140, 153, and 169 windows.

The second transformation reported in Table 1 is the calculated $\Delta\Delta G_{hyd}$ of a TIP3P water molecule[63] or, in other words, the solvation energy of a TIP3P water molecule, equivalent to the reverse process of extracting a water molecule from bulk TIP3P water and placing it in the gas phase. This process can be difficult to model accurately due to polarization effects which may underestimate the energy.[64] Our calculated value of 6.6 kcal mol^{-1} is nonetheless in excellent agreement with the experimental value of 6.32 kcal mol^{-1}.[65]

Table 1 reveals that the differences in free energy relative to the initial (wild-type, unperturbed) state **1** suggest that a significant amount of energy is required to remove W2041 from the protein-ligand complex and return it to bulk water. This amount of energy is around 1.8 kcal mol^{-1}. This result is comparable to the reported value of 3.1 ± 0.6 kcal mol^{-1} for the energy of annihilating individual water molecules calculated by Hamelberg et al. in the complex of HIV-1 protease with an inhibitor, as well as their value of 1.9 ± 0.5 kcal mol^{-1} in the case of a trypsin/benzamidine complex using the double-decoupling method.[57]

The positive value for our calculated difference in free energy indicates that it is unfavourable for the system to lose this specific water molecule to the bulk solvent. Therefore, it is possible to conclude that docking or structure-based generation of ligands within this binding site of the Abl-SH3 domain tyrosine kinase should consider W2041 and its site, and if it is to be replaced by a group on the ligand or protein, this energetic contribution cannot be neglected and may need to be compensated for in energetic terms by the incoming group. This is an important finding, since it is commonplace to remove all water molecules from protein binding sites during molecular docking studies.

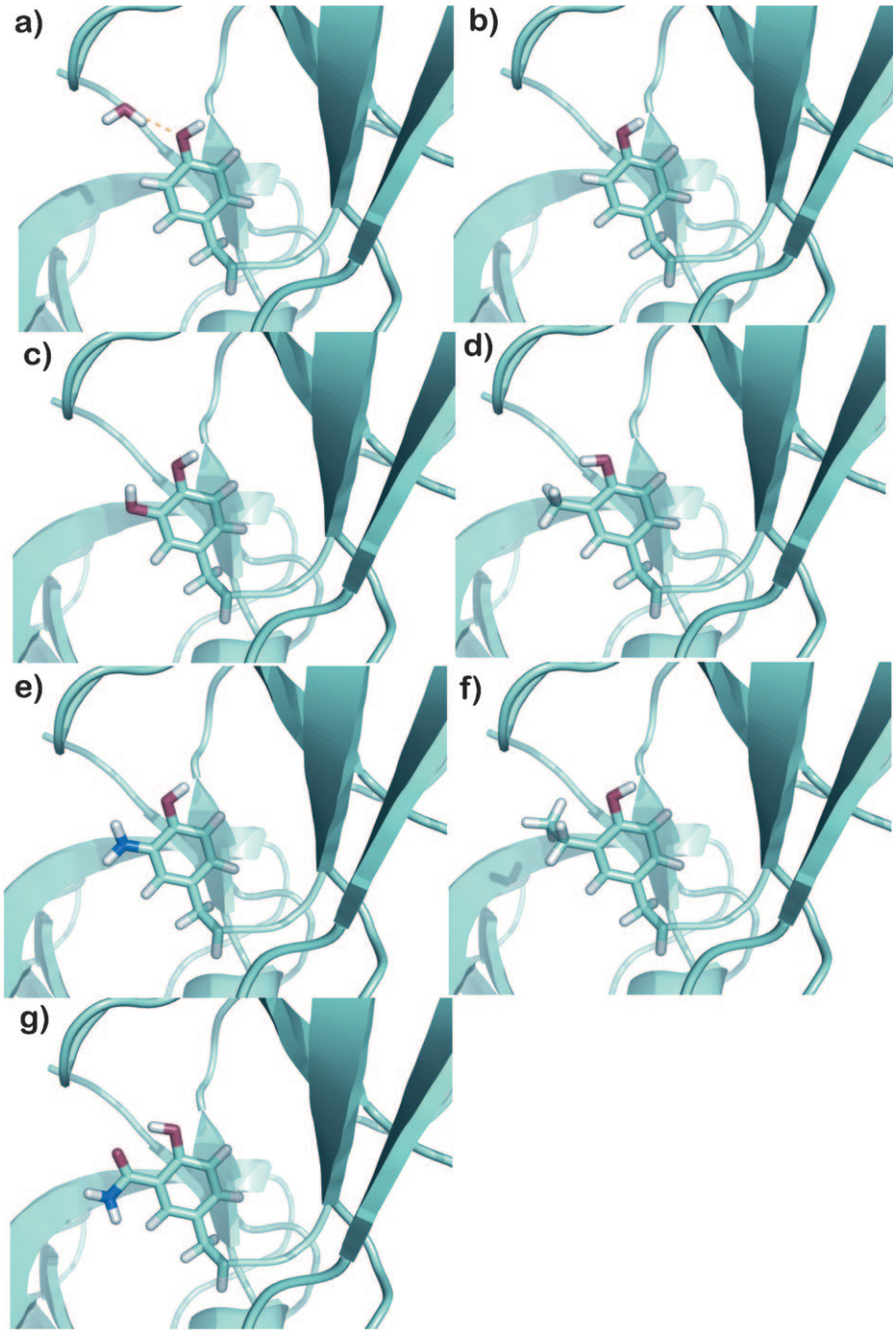

Figure 4. Representation of states **1–7** achieved through alchemical mutations. Residues 1–3 in chain B omitted for clarity. a) **1**, Tyr4 + W2041. b) **2**, Tyr4. c) **3**, Tya4. d) **4**, Tyc4. e) **5**, Tyd4. f) **6**, Tye4. g) **7**, Tyf4.

The effects of targeting the hydration site of this water molecule by six different groups on the ligand were measured in the transformations of states **3** to **7**. Many of the calculations revealed that there was an unfavourable free energy cost associated with the mutation of the ligand from tyrosine to a modified form, including the energy associated with annihilating W2041 and returning it to the bulk solvent. Only three specific mutations in the ligand-protein system were favourable: the addition of a second *ortho*-hydroxyl (*o*-OH) to the phenyl ring of tyrosine to target W2041, i.e. from Tyr to Tya (–OH) ($\Delta G_{rel} = -8.9$ kcal mol^{-1}), state **3**; the addition of an *ortho*-ethyl group to this ring, i.e. from Tyr to Tye ($-CH_2CH_3$) ($\Delta G_{rel} = -3.2$ kcal mol^{-1}), state **6**; or the addition of a formamide, i.e. from Tyr to Tyf ($-CONH_2$) ($\Delta G_{rel} = -0.4$ kcal mol^{-1}), state **7**. The free energy changes for the other mutations varied starting from the

Table 1. Thermodynamic integration free energy changes for the transformations of the protein-water-ligand system. Energies are shown in kcal mol^{-1} ± standard deviation values where available. aq = in aqueous environment, compl = in the binding complex environment.

Transformation	State	$\Delta\Delta G_{bind}$	ΔG_{rel}(Tyr → mut., w.r.t. Tyr + H_2O)
$Tyr_{compl} + W2041_{compl}$	**1**		0
$W2041_{gas} \rightarrow W2041_{aq}$			−6.6
$Tyr_{compl} + W2041_{compl} \rightarrow Tyr_{compl}$			8.4
$Tyr_{compl} + W2041_{compl} \rightarrow Tyr_{compl} + W2041_{aq}$	**2**		1.8
$Tyr_{compl} + W2041_{compl} \rightarrow Tya\_OH_{compl} + W2041_{aq}$	**3**	−10.7	−8.9 ± 0.7
$Tyr_{compl} + W2041_{compl} \rightarrow Tyc\_CH_{3compl} + W2041_{aq}$	**4**	−1.2	0.6 ± 0.7
$Tyr_{compl} + W2041_{compl} \rightarrow Tyd\_NH_{2compl} + W2041_{aq}$	**5**	−0.6	1.2 ± 0.9
$Tyr_{compl} + W2041_{compl} \rightarrow Tye\_CH_2CH_{3compl} + W2041_{aq}$	**6**	−4.9	−3.2 ± 0.9
$Tyr_{compl} + W2041_{compl} \rightarrow Tyf\_CONH_{2compl} + W2041_{aq}$	**7**	−2.2	−0.4 ± 2.8

slightly unfavourable Tyr to Tyc (–CH_3) ($\Delta G_{rel} = +0.6$ kcal mol^{-1}). It can be seen that both small and large hydrophobic and hydrophilic mutations resulted in unfavourable changes to the free energy of binding when substituting W2041. This suggests that this ligand does not easily tolerate changes in order to target W2041. It is worth pointing out that the resulting calculated relative free energies (ΔG_{rel}) varied due to different reasons in each case (see below), and that the energy contribution from the leaving water molecule can be decisive to the free energy change between states.

For comparison, we also calculated the relative free energies of hydration, ΔG_{hyd}, for the ligands calculated with respect to the model compound phenol (−10.5 kcal mol^{-1}) with results in relative agreement with the experimental values[66]: for Tya (OH, −5.35 kcal mol^{-1} cf. −3.6 kcal mol^{-1} – *o*-methylphenol aka catechol), Tyc (CH_3, 2.9 kcal mol^{-1} cf. 0.7 kcal mol^{-1} – *o*-methylphenol), Tyd (NH_2, −0.2 kcal mol^{-1} cf. −1.2 kcal mol^{-1} – *o*-aminophenol), Tye (CH_2CH_3, 4.5 kcal mol^{-1} cf. 1.3 kcal mol^{-1} – *o*-ethylphenol), and Tyf: ($CONH_2$, −2.7 kcal mol^{-1} cf. −4 kcal mol^{-1} – *o*-hydroxybenzamide).

The OH group can best replicate the interactions that W2041 has with the ligand-protein complex system, such as hydrogen bonds to Ser12 OG (chain A) and to the Tyr4 phenolic OH, to provide a $\Delta G_{compl} = -11$ kcal mol^{-1}. It also has acceptable energy characteristics for the transformation in aqueous solution (moderately favourable $\Delta G_{aq} = -0.3$ kcal mol^{-1}) which together with the free energy required to remove W2041, determine an overall favourable relative free energy ($\Delta G_{rel} = -8.9$ kcal mol^{-1}). The ethyl group (state **6**), on the other hand, despite not replacing the hydrogen bonds that W2041 formed in the binding site, increases the van der Waals contacts between ligand and protein; at the same time, the hydrogen bonds that W2041 can make in the bulk solvent compensate for some of those lost in the binding site. State **6** had a moderately unfavourable transformation in solvent with $\Delta G_{aq} = 3.2$ kcal mol^{-1}. Including the moderately favourable changes in the binding site, $\Delta G_{compl} = -1.7$ kcal mol^{-1}, and the free energy to remove W2041, the transformation to Tye resulted in a moderately favourable overall relative free energy, $\Delta G_{rel} = -3.2$ kcal mol^{-1}. For state **8** (formamide), ΔG_{compl} and ΔG_{aq} nearly balance each other, and the final ΔG_{rel} is a very slightly favourable −0.4 kcal mol^{-1}.

On the other hand, for state **5** (amino) the balance of mutation energy in the complex and in aqueous solution is slightly favourable, $\Delta\Delta G_{bind} = -0.6$ kcal mol^{-1}, but this is not enough to overcome the 1.8 kcal mol^{-1} free energy required to replace W2041. For state **4** (methyl) the $\Delta\Delta G_{bind}$ of −1.2 kcal mol^{-1} cannot overcome the free energy required to displace W2041, even if it has a moderately favourable free energy change in the binding site $\Delta G_{compl} = -5.1$ kcal mol^{-1}. For state **7**, the functional group substituting W2041 is large (formamide), which is reflected by steric effects in the energy transformations within the complex as well as in aqueous solution. In addition to the steric effect, it also formed a strong intramolecular hydrogen bond with the *ortho* –OH group, which is reflected in the transformation free energies for this compound.

These results indicate that the environment in the Abl-SH3 domain around W2041 is important for the overall stability of the system. Also, it is not straightforward to predict the free energy changes arising from modifications to the ligand, particularly while substituting a bound water molecule, since neither some hydrophilic groups (such as –NH_2) nor a particular hydrophobic group (–CH_3) were favourable substituents. These results are shown in Figure 5.

It is worth noting that addition of the hydroxyl, formamide and ethyl groups resulted in relative free energies in favour of replacing water molecule W2041, but that the addition of the methyl group did not. This fact may be explained, in the former pair, by the need of the substituting functional group to have a hydrogen bond donor as well as a hydrogen bond acceptor (similar to the binding properties of W2041 seen in the simulations: Ser12 OG donating a hydrogen bond to the water molecule and the water molecule donating a hydrogen bond to Ser12 OG and to Tyr62 OH) in order to best mimic or replace the protein-W2041-ligand interactions; and in the latter, by the need for more extensive van der Waals contacts than the methyl group can provide, as well as in a more propitious change in aqueous solution, as is the case with the ethyl substituent. Another possibility is that W2041 is not displaced completely and is instead nudged into a different hydration site

Table 2. Partial thermodynamic integration free energy changes for the transformations of the protein-water-ligand system. Energies are shown in kcal mol^{-1}. I = 50 ps, II = 75 ps, III = 100 ps. C = electrostatic leg, vdW = non-electrostatic leg.

Transformation	State	In the complex		In aqueous solution		In the gas phase	
		C	vdW	C	vdW	C	vdW
Tya_OH	**3**						
I		−10.2	−0.6	−0.8	1.2	5.1	−0.1
II		−10.7	−0.3	−2.2	1.3	5.1	−0.03
III		−11.4	0.2	−1.8	1.3	5.1	−0.06
average		−11		−0.3		5.05	
hysteresis		−10.9		−0.8		5.4	
without backbone restraint		−10					
$\Delta\Delta G_{bind}$	−10.7						
$\Delta\Delta G_{hyd}$	−5.35						
Tyc_CH_3	**4**						
I		−4.2	−1	−5.8	2.4	−6.7	0.3
II		−4.4	−0.9	−5.8	2.4	−7.1	0.1
III		−5.1	0.2	−6.3	1.5	−7.1	0.2
average		−5.1		−3.9		−6.8	
hysteresis		−4.6		−3.5		−6.2	
without backbone restraint		−4.47					
$\Delta\Delta G_{bind}$	−1.2						
$\Delta\Delta G_{hyd}$	2.9						
Tyd_NH_2	**5**						
I		−23.3	0.3	−24.6	2.5	−24	0.7
II		−23.5	0.2	−24.7	2.5	−22.4	0.8
III		−23.2	0.4	−24.4	1.6	−22.4	0.8
average		−23		−22.4		−22.2	
hysteresis		−21.9		−22		−23.3	
without backbone restraint		−22.65					
$\Delta\Delta G_{bind}$	−0.6						
$\Delta\Delta G_{hyd}$	−0.2						
Tye_CH_2CH_3	**6**						
I		−0.6	−2.3	−1.4	5.7	−1.9	0.8
II		−0.6	−1	−1.5	4.4	−2	0.7
III		−1.2	0.5	−1.4	3.9	−2	0.6
average		−1.7		3.2		−1.3	
hysteresis		−0.8		3.3		−0.6	
without backbone restraint		−2.2					
$\Delta\Delta G_{bind}$	−4.9						
$\Delta\Delta G_{hyd}$	4.5						
Tyf_$CONH_2$	**7**						
I		8.3	6.4	6.3	11.7	12.9	7.2
II		8.3	7.4	4.4	10.9	13.5	7.2
III		8.4	7.9	9.9	10.3	13.5	7.1
average		15.6		17.8		20.5	
hysteresis		15.9		17.2		26.4	
without backbone restraint		14.5					
$\Delta\Delta G_{bind}$	−2.2						
$\Delta\Delta G_{hyd}$	−2.7						

on the protein surface, although this was not observed in the present simulations. It is important that no water molecule occupies the same location as W2041 during the alchemical transformation, as otherwise the results would then be confounded. Indeed, if W2041 were mutated to nothing, but another water molecule occupied its hydration site, one could argue that the water had not disappeared.

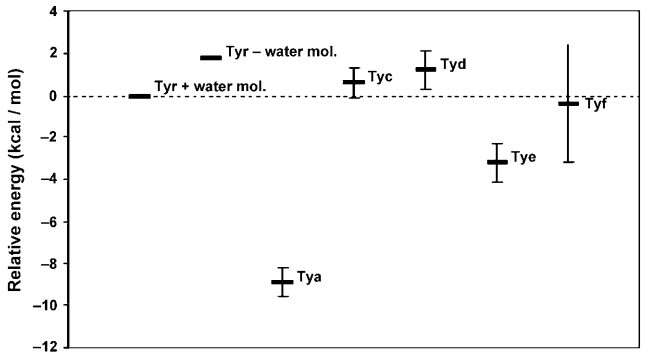

Figure 5. Energy level diagram of the states with energies relative to the initial state. Error bars are calculated as in Table 1.

This set of results suggest that when a substituent group on a ligand displaces water molecule W2041 into the bulk solvent, this will only be energetically favourable depending on the strength of the protein–ligand (or protein–water–ligand) interactions formed, including enthalpic and entropic effects, in addition to desolvation effects and the interactions that such a bridging water molecule has when in the bulk solvent. All of these contributions were determined to be significant in the predicted relative free energy changes for the cases reported here. Systematic functional group substitutions on the ligand shed light on the structural and thermodynamic properties of a protein–water–ligand complex that are observed to be critical to the relative free energies between complexed states.

The consideration of selected tightly-bound water molecules in protein-ligand complexes in molecular recognition studies or structure-based drug design may modify the size, type and number of possible functional groups on a ligand or residue, since the retention of such water molecules may be more energetically favoured. In addition, this type of systematic free energy calculations can also be used to determine if certain tightly-bound water molecules should be retained during docking simulations, which may result in more accurate virtual screening and predictions of ligand binding modes.[67]

4 Conclusions

An ordered water molecule (W2041) bridging the protein-peptide complex in the binding site of the Abl-SH3 tyrosine kinase domain (PDB structure 1bbz) was mutated using free energy thermodynamic integration molecular dynamics simulations. The calculations revealed an unfavourable free energy cost of placing this tightly-bound water molecule into the bulk solvent. A number of mutations with chemical probe fragments (functional group substituents) were carried out on the phenyl ring of a tyrosine side chain (Tyr4) in the peptide ligand to displace this water molecule and demonstrate the efficiency and implications of such an approach for computational drug design to account for the changes in standard state, desolvation energy, enthalpy, and entropy. These calculations showed that a hydroxyl, formamide, and an ethyl group were the only substituent groups capable of favourably displacing the water molecule. All the other mutations (addition of methyl and amine groups) were predicted to be unfavourable with respect to retaining the tightly bound water molecule. This was due mainly to either a large unfavourable free energy change of the substitution in the complex, since some mutations had a favorable desolvation free energy change with respect to the native ligand, or due to an unfavorable desolvation free energy change from the free to the bound complex, or else due to the unfavourable free energy change of removing W2041 from the complex which was not compensated by a large enough favourable free energy change from the ligand.

This work demonstrates the importance of the thermodynamic and structural effects that ordered water molecules bridging a biomolecular complex have, as well as the importance of the manner in which ligand functional group replacements are conducted and evaluated in the complex and in solution. It is important to know when and how, i.e., the best manner in which a bridging water molecule can be effectively used (targeted and displaced, or conversely, conserved) since this has a crucial effect on the optimization of drugs or ligands. In our study, specific functional groups were found to be the best to substitute a bridging water molecule present in a moderately hydrophobic environment that is typical of many binding complexes. However, the affinity of ligands and water molecules to proteins is particularly sensitive to the biomolecular environment of each isolated and bound component. Hence it is possible that there are no general, empirical rules that can easily be applied to predict whether the inclusion or targeting of an ordered water molecule will result in favourable changes to the free energy of binding in a ligand-protein system.[7] The presence of certain water molecules can be considered more reliable than others in their structural determination (due to having lower B-factors, lower solvent exposed surface area and/or higher number of hydrogen bonds), and therefore, may also be more conserved across protein crystal structures.[9] In other cases, especially for water molecules with higher B-factors, higher solvent exposure and/or lower number of hydrogen bonds, i.e., weaker attachment to the protein surface, water molecules may be easily displaced at a low energetic cost. For others, however, there needs to be a more detailed account of all the enthalpic, entropic, solvation and ground state energy contributions in order to make a proper assessment. Rigorous and systematic free energy calculations, such as the one described in this work, may need to be conducted on large sets of ligand substitutions and different hydration environments to attempt to determine if a pattern can be discerned, such as the approximate average binding energy provided per additional specific functional group, similar to the analysis

determined by Andrews et al.[68] where average free energies of binding provided by each typical functional group in drug molecules were rationalized. Once patterns are established, empirical rules may be set which would eliminate the need for expensive MD simulations. This may establish if in a given ligand optimization setting (for example, with certain protein binding groups or a certain binding site shape) a specific functional group such as a hydroxyl, formamide or ethyl can provide enough additional free energy of binding to successfully displace a tightly-bound water molecule or if there is no general substitution pattern and each case must be considered separately.

Acknowledgements

Estonian Science Foundation Grant JD80, Estonian Ministry for Education and Research Grant SF0140031Bs09, the *Institute of Advanced Study*, Berlin, for hosting May–July 2010, and the University of Tartu High Performance Computing Centre for computing resources. The authors declare no financial nor personal conflicts of interest that might bias this work.

References

[1] R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373 – 380.
[2] J. R. H. Tame, S. H. Sleigh, A. J. Wilkinson, J. E. Ladbury, *Nat. Struct. Biol.* **1996**, *3*, 998 – 1001.
[3] E. Chung, D. Henriques, D. Renzoni, M. Zvelebil, J. M. Bradshaw, G. Waksman, C. V. Robinson, J. E. Ladbury, *Structure* **1998**, *6*, 1141 – 1151.
[4] J. E. Ladbury, *Chem. Biol.* **1996**, *3*, 973 – 980.
[5] S. Bergqvist, M. A. Williams, R. O'Brien, J. E. Ladbury, *J. Mol. Biol.* **2004**, *336*, 829 – 842.
[6] J. Böttcher, A. Blum, A. Heine, W. Diederich, G. Klebe, *J. Mol. Biol.* **2008**, *383*, 347 – 357.
[7] R. L. Mancera, *Curr. Opin. Drug Discovery Dev.* **2007**, *10*, 275 – 280.
[8] C. Hetényi, D. van der Spoel, *Prot. Sci.* **2002**, *11*, 1729 – 1737.
[9] A. T. García-Sosa, R. L. Mancera, P. M. Dean, *J. Mol. Model.* **2003**, *9*, 172 – 182.
[10] A. T. García-Sosa, S. Firth-Clark, R. L. Mancera, *J. Chem. Inf. Model.* **2005**, *45*, 624 – 633.
[11] A. T. García-Sosa, R. L. Mancera, *J. Mol. Model.* **2006**, *12*, 422 – 431.
[12] M. Rarey, B. Kramer, T. Lengauer, *Proteins* **1999**, *34*, 17 – 28.
[13] M. L. Verdonk, G. Chessari, J. C. Cole, M. J. Hartshorn, C. W. Murray, J. W. M. Nissink, R. D. Taylor, R. Taylor, *J. Med. Chem.* **2005**, *48*, 6504 – 6515.
[14] B. C. Roberts, R. L. Mancera, *J. Chem. Inf. Model.* **2008**, *48*, 397 – 408.
[15] R. Thilagavathi, R. L. Mancera, *J. Chem. Inf. Model.* **2010**, *50*, 415 – 421.
[16] D. G. Lloyd, A. T. García-Sosa, I. L. Alberts, N. P. Todorov, R. L. Mancera, *J. Comput.-Aided Mol. Des.* **2004**, *18*, 89 – 100.
[17] C. Barillari, J. Taylor, R. Vinter, J. W. Essex, *J. Am. Chem. Soc.* **2007**, *129*, 2577 – 2587.
[18] J. Qvist, M. Davidovic, D. Hamelberg, B. Halle, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6296 – 6301.
[19] I. Haq, *Arch. Biochem. Biophys.* **2002**, *403*, 1 – 15.
[20] H. Yu, S. W. Rick, *J. Am. Chem. Soc.* **2009**, *131*, 6608 – 6613.
[21] M. A. Karplus, C. Faerman, *Curr. Opin. Struct. Biol.* **1994**, *4*, 770 – 776
[22] V. P. Denisov, K. Venu, J. Peters, H. D. Horlein, B. Halle, *J. Phys. Chem. B* **1997**, *101*, 9380 – 9389.
[23] P. Y. S. Lam, P. K. Jadhav, C. J. Eyermann, C. N. Hodge, Y. Ru, L. T. Bacheler, J. L. Meek, M. J. Otto, M. M. Rayner, Y. N. Wong, C. H. Chang, P. C. Weber, D. A. Jackson, T. R. Sharpe, S. Ericksonviitanen, *Science* **1994**, *263*, 380 – 384.
[24] T. Suzuki, A. Matsuura, A. Kouketsu, H. Nakagawa, N. Miyata, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 331 – 335.
[25] V. Mikol, C. Papageorgiou, X. Borer, *J. Med. Chem.* **1995**, *38*, 3361 – 3367.
[26] D.- G. Liu, Y. Gao, J. H. Voigt, K. Lee, M. C. Niklaus, L. Wu, Z.- Y. Zhang, T. R. Burke, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3005 – 3007.
[27] B. A. Katz, C. Luong, J. D. Ho, J. R. Somoza, E. Gjerstad, *J. Mol. Biol.* **2004**, *344*, 527 – 547.
[28] K. L. Meagher, H. A. Carlson, *Proteins* **2005**, *58*, 119 – 125.
[29] V. Helms, R. C. Wade, *Biophys. J.* **1995**, *69*, 810 – 824.
[30] V. Helms, R. C. Wade, *Proteins* **1998**, *32*, 381 – 396.
[31] L. R. Olano, S. W. Rick, *J. Am. Chem. Soc.* **2004**, *126*, 7991 – 8000.
[32] Y. Lu, C.-Y. Tang, S. Wang, *J. Am. Chem. Soc.* **2006**, *128*, 11830 – 11839.
[33] S. Samsonov, J. Teyra, M. T. Pisabarro, *Proteins* **2008**, *73*, 515 – 525.
[34] M. Fornabaio, F. Spyrakis, A. Mozzarelli, P. Cozzini, D. J. Abraham, G. E. Kellogg, *J. Med. Chem.* **2004**, *47*, 4507 – 4516.
[35] Z. Li, T. Lazaridis, *J. Phys. Chem. B* **2005**, *109*, 662 – 670.
[36] S. Geroult, M. Hooda, S. Virdee, G. Waksman, *Chem. Biol. Drug Des.* **2007**, *70*, 87 – 99.
[37] R. Abel, T. Young, R. Farid, B. J. Berne, R. A. Friesner, *J. Am. Chem. Soc.* **2008**, *130*, 2817 – 2831.
[38] B. Nagar, W. G. Bornmann, P. Pellicena, T. Schindler, D. R. Veach, W. T. Miller, B. Clarkson, J. Kuriyan, *Cancer Res.* **2002**, *62*, 4236 – 4243.
[39] A. Palencia, E. S. Cobos, P. L Mateo, J. C. Martínez, I. Luque, *J. Mol. Biol.* **2004**, *336*, 527 – 537.
[40] M. T. Pisabarro, L. Serrano, M. Wilmanns, *J. Mol. Biol.* **1998**, *281*, 513 – 521.
[41] D. A. Case, D. A. Pearlman, J. W. Caldwell, *AMBER*, Version 7.0, University of California, San Francisco, **2002**.
[42] T. Rodinger, R. Pomès, *Curr. Opin. Struct. Biol.* **2005**, *15*, 164 – 170.
[43] M*aestro*, Version 8.0, Schrödinger LLC, New York, **2007**.
[44] J. C. Shelly, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin, M. Uchimaya, M. Epik, *J. Comput.-Aided Mol. Des.* **2007**, *21*, 681 – 691.
[45] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, *Gaussian*, Version 2003, Gaussian, Pittsburgh, PA, **2003**.
[46] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179 – 5197.
[47] J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, *21*, 1049 – 1074.
[48] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269 – 10280.
[49] W. D. Cornell, P. Cieplak, C. I. Bayly, P. A. Kollman, *J. Am. Chem. Soc.* **1993**, *115*, 9620 – 9631.
[50] P. Cieplak, W. D. Cornell, C. Bayly, P. A. Kollman, *J. Comput. Chem.* **1995**, *16*, 1357 – 1377.

[51] T. A. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089 – 10092.
[52] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684 – 3690.
[53] M. K. Gilson, J. A. Given, B. L. Bush, J. A. McCammon, *Biophys. J.* **1997**, *72*, 1047 – 1069.
[54] P. A. Bash, U. C. Singh, R. Langridge, P. A. Kollman, *Science* **1987**, *236*, 564 – 568.
[55] V. Helms, R. C. Wade, *J. Am. Chem. Soc.* **1998**, *120*, 2710 – 2713.
[56] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926 – 935.
[57] D. Hamelberg, J. A. McCammon, *J. Am. Chem. Soc.* **2004**, *126*, 7683 – 7689.
[58] M. D. Erion, M. R. Reddy, *J. Am. Chem. Soc.* **1998**, *120*, 3295 – 3304.
[59] Y. Deng, B. Roux, *J. Phys Chem. B* **2009**, *113*, 2234 – 2246.
[60] J. G. Kirkwood, *J. Chem. Phys.* **1935**, *3*, 300 – 313
[61] C. Chipot, X. Rozanska, S. B. Dixit, *J. Comput.-Aided Mol. Des.* **2005**, *19*, 765 – 770.
[62] M. R. Shirts, J. W. Pitera, W. C. Scope, V. S. Pande, *J. Chem. Phys.* **2003**, *119*, 5740 – 5761.
[63] C. Chipot, C. Millot, B. Maigret, P. A. Kollman, *J. Chem. Phys.* **1994**, *101*, 7953 – 7962.
[64] H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.* **1987**, *91*, 6269 – 6271.
[65] A. Ben-Naim, Y. Marcus, *J. Chem. Phys.* **1984**, *81*, 2016 – 2027.
[66] C. J. Cramer, D. G. Truhlar, *Free Energy Calculation in Rational Drug Design*, (Eds: M. R. Reddy, M. D. Erion), Kluwer Academic/ Plenum, New York, **2001**, pp. 63 – 95.
[67] N. Huang, B. K. Shoichet, *J. Med. Chem.* **2008**, *51*, 4862 – 4865.6057 – 6063.
[68] P. R. Andrews, D. J. Craik, J. L. Martin, *J. Med. Chem.* **1984**, *27*, 1648 – 1657.

Received: January 11, 2010
Accepted: July 29, 2010
Published online: August 30, 2010

