Theor Chem Acc (2011) 129:747–756
DOI 10.1007/s00214-011-0931-y

REGULAR ARTICLE

An atomistic model for simulations of nilotinib and nilotinib/kinase binding

Najl V. Valeyev · Alexey Aleksandrov

Received: 31 January 2011 / Accepted: 22 March 2011 / Published online: 21 April 2011

Abstract Nilotinib is a novel anticancer drug, which specifically binds to the Abl kinase and blocks its signaling activity. In order to model the nilotinib/protein interactions, we have developed a molecular mechanics force field for nilotinib, consistent with the CHARMM force field for proteins and nucleic acids. Atomic charges were derived by utilizing a supermolecule ab initio approach. We considered the ab initio energies and geometries of a probe water molecule that interacts with nilotinib fragments at six different positions. We investigated both neutral and protonated states of nilotinib. The final rms deviation between the ab initio and the force field energies, averaged over both forms, was equal 0.2 kcal/mol. The model reproduces the ab initio geometry and flexibility of nilotinib. To apply the force field to nilotinib/Abl simulations, it is also necessary to determine the most likely protein and nilotinib protonation state when it binds to Abl. This task was carried out using molecular dynamics free energy simulations. The simulations indicate that nilotinib can interact with Abl in protonated and deprotonated forms, with the protonated form more favoured for the interaction. In the course of our calculations, we established that the His361, a titratable amino acid residue that mediates the interaction, prefers to be neutral. These insights and models should be of interest for drug design.

Keywords Molecular recognition · Computer simulation · AMN107 · Drug design · CHARMM program

1 Introduction

Unbalanced Abl kinase activity can lead to excessive cell division and is a cause of several forms of cancer including chronic myeloid leukemia [1]. The development of specific inhibitors to a kinase can be a difficult exercise due to the sequence similarity of kinase active sites [2, 3]. A recently developed compound, Imatinib, specifically inhibits the tyrosine kinase activity of Abl by stabilizing its inactive conformation [4–7]. However, the imatinib-resistant mutations of Abl have been frequently observed. This problem has stimulated the development of new kinase inhibitors [8, 9]. One important novel inhibitor is the anticancer drug, nilotinib (AMN107), which specifically binds to Abl [8]. Nilotinib is not only more potent than imatinib against the wild-type Abl, but has a significantly higher inhibitory activity for most of imatinib-resistant mutants [9, 10].

Given the unique nilotinib-binding properties to Abl and homologous kinases, it is essential to investigate the underlying mechanisms of such interactions at the atomic level. Recently solved crystal structure of the Abl/nilotinib complex has provided valuable biochemical insights [11]. In this study, we employ the experimentally solved structure and undertake a complementary approach by developing

Electronic supplementary material The online version of this article (doi:10.1007/s00214-011-0931-y) contains supplementary material, which is available to authorized users.

N. V. Valeyev
Centre for Molecular Processing, University of Kent, Canterbury, Kent CT2 7NJ, UK

A. Aleksandrov (✉)
Department of Biology, Laboratoire de Biochimie (CNRS UMR7654), Ecole Polytechnique, 91128 Palaiseau, France
e-mail: alexey.aleksandrov@polytechnique.fr

computer simulation models. The developed models allow us to investigate the structure, dynamic, and thermodynamic properties of the nilotinib/Abl kinase complex.

The major requirement for computer simulations is the development of a molecular mechanics force field model. This can be a nontrivial exercise for a molecule as large and complex as nilotinib. In this work, we develop a force field model for nilotinib under different protonation states. The force field is developed in the manner such that it is compatible with the CHARMM27 force field for proteins and nucleic acids [12, 13], as well as with the TIP3P water model [14].

The second major task for computer simulations is the determination of the most favorable nilotinib protonation state required for the interaction with the Abl kinase. The protonation state can have a profound effect on binding [4, 15–17]. Therefore, the protonation states need to be evaluated for both ligands and proteins participating in the interaction. Indeed, 4-methylimidazole group of nilotinib has two protonation states in solution, which may change when nilotinib binds to Abl. Similarly, several protein groups in the nilotinib-binding pocket, in particular a nearby His361, have several possible protonation states, which can also change when nilotinib binds. In this work, two different free–energy methods, that use either an explicit treatment of aqueous solvent [16, 18, 19] or a continuum dielectric treatment [20], have been employed to establish the nilotinib protonation state when it is bound to Abl. Our simulations predict that nilotinib can interact with Abl in protonated and deprotonated forms, with the protonated form more favoured for the interaction. The residue His361 appears to be in the neutral charge state, protonated on the Nϵ atom.

The model is suitable to investigate the interactions of nilotinib with a wide range of protein targets. It also offers a starting point to parameterize potential nilotinib analogues. It should aid efforts both to understand nilotinib/kinase binding and to develop improved kinase inhibitors.

2 Computational methods

2.1 Force field determination for nilotinib

Nilotinib is schematically represented in Fig. 1. The force field is developed in order to be compatible with the CHARMM27 force field for proteins and nucleic acids [12, 13], as well as with the TIP3P water model [14]. The same procedure was recently applied to parameterize a large set of tetracycline variants [21, 22] and imatinib [23] variants. In this study, both neutral and protonated forms of nilotinib were parameterized. Plain nilotinib structures were obtained from the protein data bank.

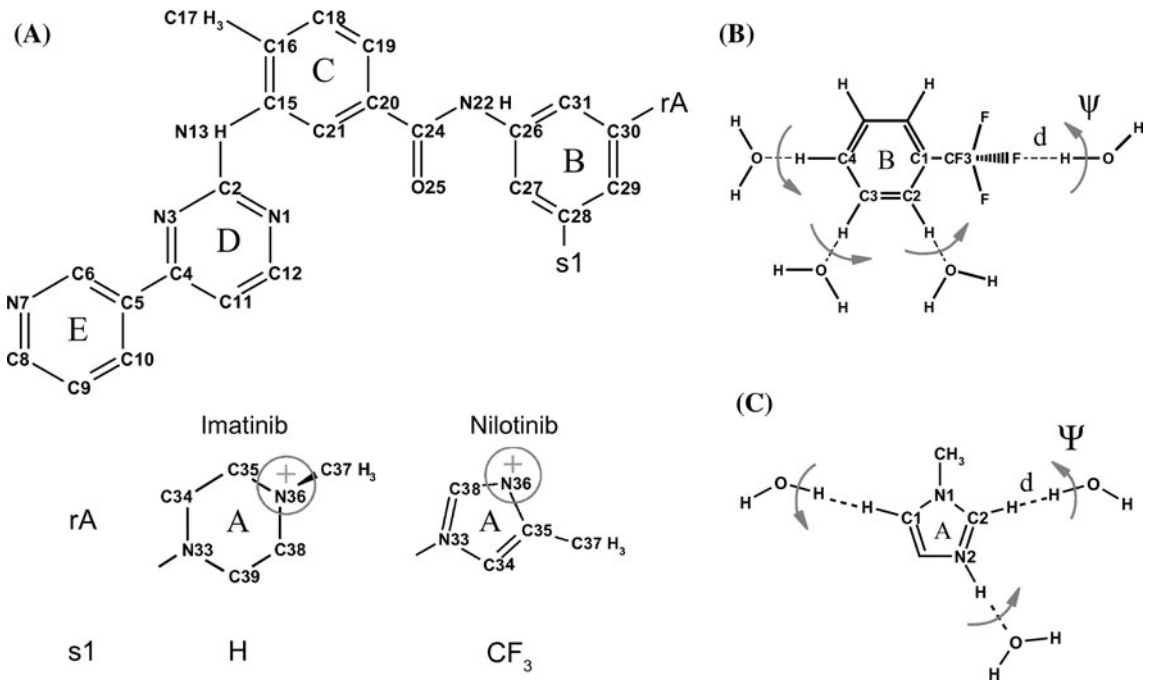

Fig. 1 2D view of nilotinib, imatinib, and fragments used for supermolecule calculations. **a** 2D view of nilotinib and imatinib. Position N36, which can be protonated or deprotonated, is *circled*. **b** A fragment used for supermolecule calculations, corresponding to ring B, with four possible interacting water positions. Each water position is defined by a single distance, *d*, and a single angle, ψ. **c** A fragment used for supermolecule calculations, corresponding to ring A

2.1.1 Optimization of the intermolecular force field parameters

We adopt a force field of the CHARMM27 form [12]. The intermolecular energy terms are Lennard–Jones and Coulomb terms. Lennard–Jones parameters were taken from the CHARMM27 force field, by assigning to each nilotinib atom an existing atomic type within that force field [12, 13, 24]. Consistent with the development of the CHARMM force field, atomic charges were derived from a supermolecule approach. We considered the ab initio energies and geometries of a probe water molecule that interacts with small model compounds, representative of different fragments of nilotinib/trifluoromethylbenzene, and 1-methylimidazole. Partial charges for other groups of nilotinib were adopted from either the CHARMM27 force field or the imatinib force field derived previously [23].

Selected probe positions around each fragment were considered in order to achieve a good balance between solute–water and water–water interactions in the force field [12, 13]. The water geometry required for force field calculations was obtained from the TIP3P model [14]. The geometry of each model compound was first optimized at the HF/6-31G(d) level. Each supermolecule structure was then optimized (at the HF/6-31G(d) level) by varying the interaction distance, to find the local energy minimum for the water position. A single angle ψ defining the orientation of the water molecule was held fixed (see Fig. 1b). For each water position, ψ was varied over the full 180° range in 30° or 60° steps, allowing either six or three distinct supermolecule calculations for each water position. For each optimal distance d, the interaction energy was calculated. No correction for basis set superposition error was made. The ab initio interaction energies were scaled by a factor of 1.16, and the ab initio interaction distances were reduced by 0.2 Å, in order to compensate for overestimated interaction distances with the Hartree–Fock model (due to neglected electron correlation) [12].

The force field charges were then adjusted to reproduce these "corrected" ab initio interaction energies and water positions. The initial partial charges were obtained from a Mulliken population analysis of the HF/6-31G(d) wavefunction. The (corrected) ab initio data were then fitted by manually optimizing the model compound charges. This approach involves the reoptimization of the nilotinib/water distance after the parameter changes. The adjustment of the charges following the model fragments linkage to form larger entities was performed by adding the charge of the deleted hydrogen atom to the heavy atom. This approach maintains the total charge from original compounds. At the same time, the charges of the linking groups were reoptimized for those cases when 2-aminopyrimidine and 4-methylaniline (corresponding to rings D and C) were linked, and when two benzene rings (corresponding to rings B and C) were linked through an amide group to form phenyl-benzamide (Fig. 1).

2.1.2 Optimization of the intramolecular force field parameters

Under the force field consideration, the intramolecular geometry is mainly determined by the minimum energy values of the bond length and bond angle terms, as well as by the phase and multiplicity of the dihedrals. The initial values for the minimum energy bond lengths, bond angles, and torsion angles were taken directly from the ab initio structures, optimized at the HF/6-31G(d) level. The initial parameters for the bond, angle, and dihedral force constants as well as for phases were obtained from the CHARMM27 force field [12] via comparison with small molecular fragments. The geometrical and dihedral parameters were then optimized by fitting to the structures from ab initio optimizations. The structure was minimized with the force field model using a Powell conjugate gradient algorithm, targeting the rms energy gradient equal or less than 10^{-6}kcal/mol/Å. The quality of the molecular structure calculated with the refined parameters was measured by rms coordinate deviation from the ab initio structure and by comparing the force field/ab initio energies. The bond, angle geometries, and the dihedral geometries were then updated manually, and a new round of optimization was carried out. This procedure was repeated until a satisfactory agreement was achieved (the coordinate rms deviation of 0.20 Å or less).

2.2 Molecular dynamics simulations

The crystal structure of the human Abl tyrosine kinase was taken from the protein data bank, entry 3CS9 (with bound nilotinib) [11]. Abl was modeled in the inactive state, competent for nilotinib binding [11]. Protonation states of histidines were assigned by visual inspection: His295 was set neutral, whereas four other histidines were set as double-protonated.

The simulations included protein residues within a 26 Å sphere, centered on the nilotinib-binding site. In addition to crystal waters, a 26 Å sphere of water was overlaid and waters that overlapped protein, crystal waters, or nilotinib were removed. Throughout the MD simulations, protein atoms between 20 and 26 Å from the sphere's center were harmonically restrained to their experimental positions. Simulations were done with the SSBP solvent model [25, 26], which treats the region outside the 26 Å sphere as a uniform dielectric continuum, with a dielectric constant of 80. Newtonian dynamics was used for the innermost region, within 20 Å of the sphere's center; Langevin dynamics was

used for the outer part of the sphere, with a 292 K bath. The CHARMM27 force field was used for the protein [12] and the TIP3P model for water [14]. Electrostatic interactions were computed without any cutoff, using a multipole approximation for distant groups [27]. Calculations were done with the CHARMM program [28].

2.3 Nilotinib protonation state

It has been found that imatinib, a close analogue of nilotinib, spends about two-thirds of the time in the neutral state at a physiological pH of 7.4 [29]. The remaining time imatinib is protonated usually on the methylated nitrogen of the piperazine ring. In nilotinib, N-methyl-piperazine moiety is substituted by the 4-methylimidazole ring. The pK_a of 4-methylimidazole and 1-methylimidazole in solution are found to be 7.52 and 6.95, respectively [30, 31], which are close to the physiological pH. Therefore, nilotinib can exist in both protonation states of the 4-methylimidazole group under physiological pH.

To determine the nilotinib protonation state in the protein complex, we undertook a molecular dynamics free energy simulation study [18, 19]. The nilotinib is gradually protonated both in solution and in complex with Abl, through a series of MD simulations (Fig. 2, with Nil^+ representing the protonated form). This technique has no adjustable parameters and gives good, though not perfect accuracy for acid/base reactions [19].

2.4 Alchemical MD free energy simulations

To compare the neutral nilotinib and its protonated form binding to Abl, we use the thermodynamic cycle shown in Fig. 2. The alchemical MD free energy simulations (MDFE) method follows the horizontal legs of the cycle. Protonated nilotinib is reversibly transformed into its deprotonated form during a series of MD simulations. The corresponding work is derived from a thermodynamic integration formula [32]. All simulations were performed with the spherical boundary conditions described elsewhere (see also [26, 33]). For the lower leg of the thermodynamic

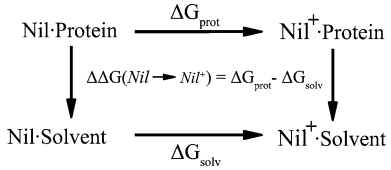

Fig. 2 Thermodynamic cycle to study nilotinib protonation and its effect on binding to a protein (Abl in this case). *Horizontal legs* represent the protonation of nilotinib on its N36 nitrogen, either in complex with Abl (*above*) or alone (*below*). *Vertical legs* represent nilotinib/Abl binding. The MD free energy simulations follow the *horizontal legs*

cycle, we simulate nilotinib in solution. For the upper leg, we simulate a portion of the nilotinib/Abl complex, solvated by the same 26 Å sphere. In each simulation system, the energy function can be expressed as a linear combination of terms associated with nilotinib (Nil) and its protonated form (Nil^+):

$$U(\lambda) = U_0 + (1-\lambda)U(\mathrm{Nil}^+) + \lambda U(\mathrm{Nil}) \quad (1)$$

where λ is a "coupling parameter", and U_0 represents interactions between parts of the system other than nilotinib. The free energy derivative with respect to λ has the form:

$$\frac{\partial G}{\partial \lambda}(\lambda) = \langle U(\mathrm{Nil}) - U(\mathrm{Nil}^+)\rangle_\lambda \quad (2)$$

where the brackets indicate an average over the MD trajectory with the energy function $U(\lambda)$ [32, 34]. We gradually mutated Nil^+ into Nil by changing λ from zero to one. The successive values of λ were as follows: 0.001, 0.01, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 0.9, 0.95, 0.99, 0.999. The free energy derivatives were computed at each λ value from a 150 ps MD simulation, or "window"; the last 120 ps of each window were used for averaging. A complete mutation run thus corresponded to 12 windows and 1.8 ns of simulation. Five runs were performed in each direction (Nil^+ into Nil and the reverse).

The estimation of uncertainty with MDFE is difficult and expensive [35–37]. A widely employed approach which addresses the problem is to perform multiple runs followed by measurements of the dispersion between runs. While such techniques are reasonable, it has a significant downside. Specifically, it does not encounter certain types of systematic error and can lead to the overestimation of uncertainty. It is established that runs performed in opposite directions ("forward" and "backward" transformations; λ increasing or decreasing) exhibit systematic hysteresis effects [38–40]. Thus, we use an error estimator that consider pairs of runs, one in each direction, forming a "forward/backward pair" [33, 41, 42]. In our previous work, the forward/backward averages were much more reproducible than the individual values. The uncertainty is, therefore, calculated as standard deviation of the individual runs obtained from five forward/backward pairs, totaling 18.0 ns. For a more detailed discussion of convergence and some examples, see supplementary material in reference [42].

2.4.1 Poisson–Boltzmann linear response approximation

The Poisson–Boltzmann linear response approximation (PBLRA) is a relatively new method that combines MD simulations with continuum electrostatics [18, 20, 43]. This method is well suited to treat rearrangements of atomic charges, as in the nilotinib or His361 protonation. The free

energy change is approximated by the continuum electrostatic free energy, averaged over the equilibrium states before and after the rearrangement. We applied the method using MD structures of the protonated and neutral nilotinib states. It was also used to compare protonation states of His361 in the nilotinib-binding pocket systematically.

The equilibrium states were simulated for 3 ns for each charge state. Calculations were performed at zero ionic strength by solving the Poisson equation. Nevertheless, we refer to the method as PBLRA. It has been found in similar systems that PBLRA results are weakly dependent on the ionic strength [33]. The protein/solvent dielectric boundary was defined by the molecular surface of the protein. The boundary was constructed using the atomic radii from the CHARMM27 force field, with the exception for the hydrogen radii, which were set to 1.0 Å [44]. The probe sphere for the surface construction was set to 2.0 Å. It has been shown that with such probe radius proteins have no internal cavities [20]. The partial charges on the atoms were taken from either the CHARMM27 molecular mechanics force field [12] or the nilotinib force field developed in the present work. The finite difference Poisson equation was solved in two steps by employing a finite difference method implemented in the CHARMM program.

The first step utilized a cubic-grid spacing of 0.8 Å, whereas the second used the spacing step of 0.4 Å. Energies were averaged over 250 structures for each charge state. The solute dielectric constant was set to 1 or 2, and the solvent dielectric was set to 80, the experimentally determined value for bulk water.

3 Results

3.1 Force field development: nilotinib–water interactions

We first consider the supermolecule calculations for nilotinib fragments interacting with individual water molecules. Detailed results are presented in Tables 1 and 2 for two representative fragments that correspond to the ring B and the protonated ring A, respectively (Fig. 1b, c). Overall, very good agreement was obtained between the ab initio and force field data, similar to our previous result [23].

The rms deviation for the energies averaged over two fragments, 6 water positions considered during optimization, and all water orientations (33 in total) was 0.1 kcal/mol. The rms deviation for the energies averaged over all orientations including the positions that were not considered during optimization is just slightly higher: 0.2 kcal/mol. The rms deviation for the fragment–water distances

Table 1 Interactions between a probe water and trifluoromethylbenzene

Probe site	Ab initio/force field results		
	Energy (kcal/mol)	Distance d (Å)	Angle ψ (°)
Cf3F1	−1.18/−1.21	2.01/2.00	0.0
Cf3F1	−1.10/−1.10	2.01/2.01	30.0
Cf3F1	−1.10/−1.08	2.01/2.01	60.0
Cf3F1	−1.23/−1.17	2.01/2.00	90.0
Cf3F1	−1.36/−1.33	2.01/1.99	120.0
Cf3F1	−1.54/−1.48	2.01/1.98	150.0
Cf3F1	−1.58/−1.58	2.01/1.97	180.0
Cf3F1	−1.58/−1.63	2.01/1.97	210.0
Cf3F1	−1.54/−1.61	2.01/1.97	240.0
Cf3F1	−1.50/−1.55	2.01/1.97	270.0
Cf3F1	−1.41/−1.47	2.01/1.98	300.0
Cf3F1	−1.32/−1.35	2.01/1.99	330.0
C2H	−2.07/−1.96	2.45/2.58	0.0
C2H	−1.85/−1.90	2.49/2.59	60.0
C2H	−1.85/−1.90	2.49/2.59	120.0
C3H	−2.07/−2.03	2.58/2.61	0.0
C3H	−2.25/−2.11	2.54/2.60	60.0
C3H	−2.25/−2.11	2.54/2.60	120.0
C4H	−2.12/−1.96	2.58/2.61	0.0
C4H	−2.29/−2.05	2.54/2.61	60.0
C4H	−2.29/−2.05	2.54/2.61	120.0

Results for trifluoromethyl-benzene (4 sites). The distance d and angle ψ are defined in Fig. 1b. The rms deviation for the energies is 0.10 kcal/mol

was achieved at 0.10 Å via very small adjustments in the atomic charges. The final atomic charges for nilotinib include small adjustments due to the linking of the rings (described in more details in the Sect.2). The final charges can be compared to several groups in the CHARMM27 force field for proteins and nucleic acids [12, 13, 24, 45].

Compatibility with the existing CHARMM force field was ensured by restricting optimization to parameters that did not already exist in the CHARMM force field. For example, the charges of the carbon atoms on the phenyl ring were maintained to the existing values of −0.115. The C_{17} methyl in the C ring has a standard CHARMM methyl charge (−0.27 on the carbon; atomic units) and hydrogens of all methyl groups have standard CHARMM charges of 0.09. The ring nitrogens N_1, N_3, N_7 have charges of −0.74, −0.74, and −0.63, very similar to adenine N_1 (−0.74) and cytosine N_3 (−0.66), and similar to comparable nitrogens in histidine and guanine. The aromatic CH groups are somewhat more polar and more variable than the existing CHARMM groups. Here, the most polar pair is C_{19}–H, with charges of −0.26 and +0.20; the pair with the largest net charge is C_{12}−H, with charges of +0.16 and +0.16. In

Table 2 Interactions between a probe water and 1-methylimidazolium

Probe site	Ab initio/force field results		
	Energy (kcal/mol)	Distance d (Å)	Angle ψ (°)
C1H	−2.32/−2.34	2.47/2.54	0.0
C1H	−2.14/−2.27	2.49/2.55	30.0
C1H	−2.09/−2.24	2.49/2.55	60.0
C1H	−2.27/−2.30	2.47/2.55	90.0
C1H	−2.41/−2.36	2.45/2.54	120.0
C1H	−2.45/−2.38	2.45/2.54	150.0
C2H	−2.49/−2.37	2.41/2.22	0.0
C2H	−2.23/−2.23	2.43/2.23	30.0
C2H	−2.05/−2.12	2.45/2.23	60.0
C2H	−2.18/−2.15	2.44/2.23	90.0
C2H	−2.41/−2.27	2.42/2.23	120.0
C2H	−2.54/−2.39	2.40/2.22	150.0
N2H	−15.23/−15.96	1.76/1.75	0.0
N2H	−15.82/−16.20	1.74/1.75	60.0
N2H	−15.82/−16.19	1.74/1.75	120.0
N2	−7.58/−7.11	1.91/1.91	0.0
N2	−7.78/−7.25	1.90/1.91	60.0
N2	−7.56/−7.21	1.90/1.91	120.0
N2	−7.16/−7.04	1.92/1.91	180.0
N2	−7.56/−7.21	1.90/1.91	240.0
N2	−7.78/−7.25	1.90/1.91	300.0

Results for 1-methylimidazolium (above line; three sites) and 1-methylimidazole (below line; one sites). The distance d and angle ψ are defined in Fig. 1c. The rms deviation for the energies is 0.31 kcal/mol

CHARMM, the Phe and Tyr rings have CH charges of −0.12 and +0.12; uracil has a pair with charges of −0.15 and 0.10; the most positive pair is in (neutral) histidine, with charges of +0.25 and +0.13. The main difference with standard CHARMM27 charges concerns two C-N groups that bridge two pairs of aromatic rings. The C_{24}−N_{22} pair is part of an acetamide group that bridges the B and C rings. The hydrogen and oxygen (H_{22},O_{25}) have charges very close to the CHARMM acetamide groups: 0.37 and −0.55. The C_{24}−N_{22} pair itself has a greater charge separation than in CHARMM: +0.86 and −0.80, compared with +0.47 and −0.55 in CHARMM. The other pair is C_2−N_{13}, bridging the C and D rings, with a similar, large charge separation.

Charges for the protonated A ring were adopted from the CHARMM force field for 4-methylimidazolium. For the neutral nilotinib, charges were adopted from the CHARMM force field for 4-methylimidazole. To adjust the charge of N_{33} due to the linking to the ring B, we used the 1-methylimidazolium model compound (Fig. 1c). To verify the charges adopted from the CHARMM force field, we also computed water interaction energies with 1-methylimidazolium and 1-methylimidazole at position N_2 after the adjustment of the N_1 charge was done (see Fig. 1c and Table 2). N_{33} after adjustment has the charge of −0.07, very similar to adenine N_9 (−0.02), guanine N_9 (−0.05), and cytosine N_1 (−0.13).

Finally, we can compare charges of trifluoromethyl-benzene and CHARMM charges of trifluoroethane. The charge of C_{281} is −0.03, which is lower comparing with the charge of trifluoroethane carbon (−0.38 atomic units). The fluorine charge is −0.07, which is again lower than the CHARMM charge of fluorine in trifluoroethane. Thus, in trifluoromethyl-benzene, the trifluoromethyl group appears to be less polar, which also explains small absolute interaction energies between this group and a water, given in Table 1.

3.2 Force field development: nilotinib structure and flexibility

The force field accurately reproduces the ab initio minimized structure of nilotinib. The superposition of the whole molecule results in the rms deviation between the ab initio and force field structures equal to 0.20 and 0.23 Å for protonated and deprotonated nilotinib, respectively.

To characterize nilotinib flexibility, we have examined two types of fluctuations: small harmonic fluctuations that involve the stiff ring structures, and larger lower-frequency fluctuations, which relate to the soft dihedral angles linking the rings. Figure 3 shows the normal mode frequencies and rms positional fluctuations for a fragment consisting of A and B rings. We found a very good agreement between the ab initio and force field results for the described characteristics.

Figure 4 illustrates the softer, inter-ring, dihedral fluctuations. Energy profiles are shown for two dihedral angles, which link two pairs of nilotinib rings: rings A–B and rings B–C, respectively. Figure 4 highlights agreement between the force field and the ab initio profiles for both the energy wells as well as for the barriers between wells. We found that the softer dihedral degrees of freedom are very important for accurate description of the conformational free energy surface of nilotinib and play an essential role for the induced fit in nilotinib/protein binding.

The force field parameters are given in "supplementary material".

3.3 Nilotinib and His361 protonation

In order to study nilotinib binding to Abl and other kinases, it is essential to establish the most favorable nilotinib protonation state for interactions. Imatinib, a close analogue of nilotinib, reveals that its piperazine ring exists primary in the protonated form while bound to the Abl

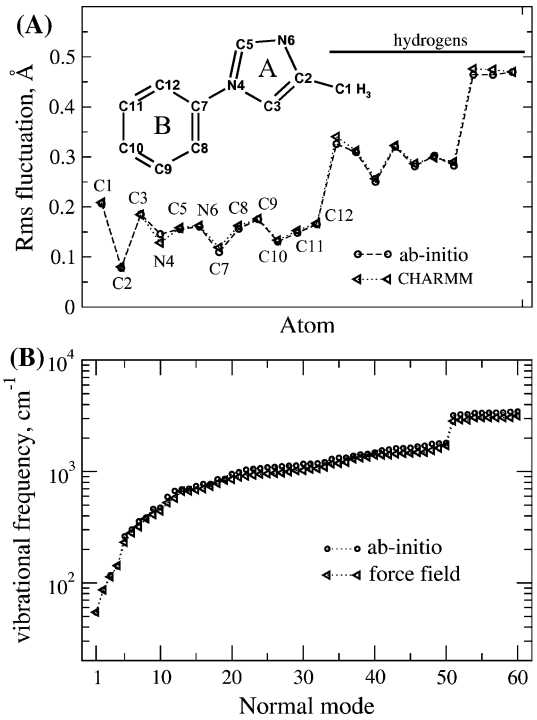

Fig. 3 Rms positional fluctuations (**a**) and normal mode frequencies (**b**) for a fragment representing rings A and B (see *inset*)

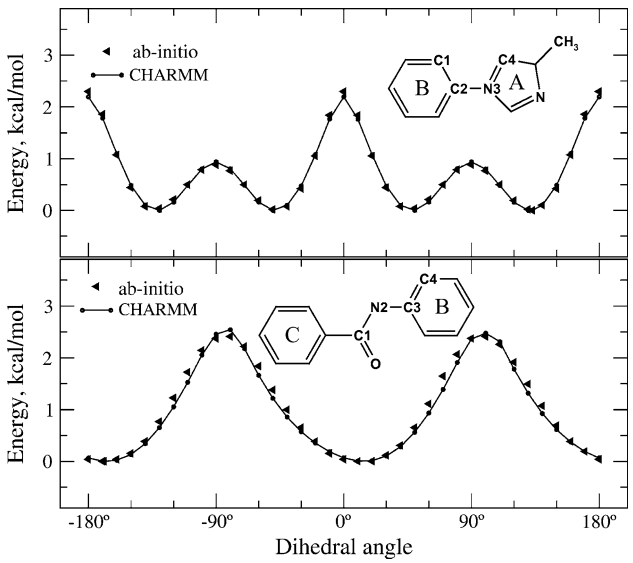

Fig. 4 The comparison of ab initio and force field energies for two dihedral angles linking the A–B (*upper panel*) and B–C (*lower panel*) rings. The dihedral angles are defined by the atoms labeled in each *inset* (C1-C2-N3-C4 and C1-N2-C3-C4). The *solid lines* correspond to the force field energies

kinase, while it can adopt both protonated and deprotonated forms in solution at a physiological pH.

The piperazine moiety in nilotinib is substituted by the 4-methylimidazole ring (see Fig. 1a). In solution, pK_as of 4-methylimidazole and 1-methylimidazole were experimentally found to be 7.52 and 6.95, respectively [30, 31], which are sufficiently close to the physiological pH. This observation suggests that both protonation states of the 4-methylimidazole group of nilotinib can exist under physiological pH: with N36 protonated (in the A ring; see Fig. 1), giving a net positive charge, or with N36 deprotonated. With the developed accurate force field in hand, we investigated the relative free energies of these two forms when nilotinib was bound to Abl, using a rigorous MDFE technique [18, 19]. Unlike widely used continuum electrostatic methods, MDFE has no adjustable parameters [46]. We also employed the approximate PBLRA method to study protonation of nilotinib and His361 systematically. The results are summarized in Table 3. We found that the protonation free energy in the protein is more favorable than in solution with the estimated difference of $\Delta\Delta G_{bind}(\mathrm{Nil}^{+} \rightarrow \mathrm{Nil}) = 1.2 \pm 0.7$ kcal/mol. The positive $\Delta\Delta G_{bind}$ implies that nilotinib prefers to bind in the protonated form, with a net positive charge. However, the relatively small $\Delta\Delta G_{bind}$ suggests that the deprotonated form can also be present in the Abl complex. The PBLRA results for protein dielectric values of one (3.1 kcal mol^{-1}) and two (1.6 kcal mol^{-1}) bracket the MDFE result and show that a protein dielectric of two is more realistic. A dielectric of two can be interpreted as an effect of electronic polarizability of the protein. The preference for the protonated nilotinib can be partially explained by favorable electrostatic interactions between the charged 4-methylimidazole ring of nilotinib (ring A) and the carboxyl group of Asp381.

For the His361 residue, we considered the Abl/nilotinib complexes with two His361 protonation states: His^{ϵ} and the doubly protonated form His^{+}. Complexes with His361 deprotonated on Nϵ were not considered due to the fact that His361 creates a strong hydrogen bond with Ala380 by the protonated Nϵ in the crystal structure. This hydrogen bond was maintained throughout our simulations. Free energy results are summarized in Table 4 and in Fig. 5. The combination of the experimental value of histidine in solution [47], $pK_{a,model}$ = 6.5, with the PBLRA result for the protein dielectric of two yields $\Delta\Delta G(His^{\epsilon} \rightarrow His^{+}) =$ 3.4 kcal/mol. For a protein dielectric of one, PBLRA

Table 3 Nilotinib protonation free energy simulations

Method	ΔG_{prot}	ΔG_{sol}	$\Delta\Delta G$
MDFE	50.5 (0.5)	49.3 (0.5)	1.2 (0.7)
PBLRA ($\epsilon = 1/\epsilon = 2$)	−4.7/−2.3	−7.8/−3.9	3.1/1.6

Protonation free energy in solvent (ΔG_{solv}) or in complex with the protein (ΔG_{prot}) is shown in kcal/mol. $\Delta\Delta G_{\mathrm{MDFE}}$ indicates preferential binding in the protonated form. The uncertainty is shown in parentheses. For PBLRA, results with a solute dielectric of one and two are given before and after the slash

Table 4 PBLRA free energy simulations comparing different protonation states of His361 and nilotinib

Final state	Dielectric constant ($\epsilon = 1/\epsilon = 2$)		
	ΔG_{prot}	ΔG_{sol}	$\Delta\Delta G$
Nil^+, $His361^+$	204.9/102.0	197.8/99.3	7.1/2.7
Nil, $His361^+$	199.9/98.2	192.0/95.3	8.0/2.9

The model with the protonated nilotinib and neutral His361(ϵ) corresponds to the initial state. Protonation free energy in solvent (ΔG_{solv}) or in complex with the protein (ΔG_{prot}) is shown in kcal/mol. The $\Delta\Delta G$ indicates preferential binding of the initial state in both cases. Results with a solute dielectric of one and two are given before and after the slash, respectively

calculations yield $\Delta\Delta G$ equal 7.8 kcal/mol. In the crystal structure, His361 interacts with the backbone amino group of Asp363 by the Nδ nitrogen, with the distance between $N\delta_{His361}$ and N_{Asp363} of 2.9 Å. This interaction was maintained in the simulations with the deprotonated His361, but was disrupted in the simulations with the diprotonated His361. The protonated Nδ does not create a strong hydrogen bond and appears in a hydrophobic pocket formed by residues Leu364 and Met351, which considerably increases the free energy by 2.7 kcal/mol (Table 4; Fig. 5). Thus, the state with the neutral His361 protonated at Nϵ appears to be strongly favored.

3.4 Protein/ligand interactions

We next simulated the protein/ligand complex for 5 ns with protonated nilotinib and His361 singly protonated on Nϵ and compared our calculations with the Xray structure. We computed rms deviations from the Xray structure, averaging over the last 1 ns of the 5 ns. The superposition of the nilotinib on the crystal structure revealed the intramolecular nilotinib deformations of less than 0.8 Å. The superposition of the nilotinib backbone on the crystal structure revealed rms deviations of 1.1 Å, which partly reflects an overall shift of the nilotinib molecule. The Abl kinase backbone superposition revealed the rms deviations of 1.3 Å. The interactions between nilotinib and the protein are also well reproduced (Table 5). Figure 6 shows the Abl/nilotinib complex. Our calculations show that the MD model is comparable with other high-quality simulations.

In the simulations, the ring A of nilotinib preferred to interact with solvent and did not create any strong interaction with Abl. Indeed, experimentally it was found that the superior binding potency of nilotinib over imatinib is mainly explained by the hydrophobic substitution at position C28 [48]. For example, a chloride or methyl group at this position both increase the binding affinity of imatinib to Abl substantially [23, 48, 49]. In agreement with this observation, the hydrophobic trifluoromethyl group binds

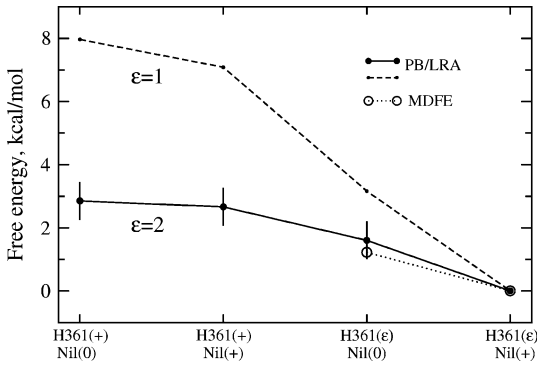

Fig. 5 Relative free energies of all four states. PBLRA is performed with a solute dielectric ϵ of one or two, as indicated. States connected by MDFE are *circled*: His361(ϵ):Nil^+ and His361(ϵ):Nil^0. PBLRA *error bars* are shown for the $\epsilon = 2$ data

Table 5 Selected distances (Å) between atoms of nilotinib and Abl

Atom pair	MD simulation	X-ray structure[a]
$N_{M318}N7_{Nil}$	2.9	3.0
$O\gamma_{T315}N13_{Nil}$	2.9	2.9
$O\epsilon1_{E286}N22_{Nil}$	3.2	3.0
$N_{D381}O25_{Nil}$	3.0	3.1

Abl atoms (left) are labeled by the amino acid to which they belong

[a] PDB entry 3CS9

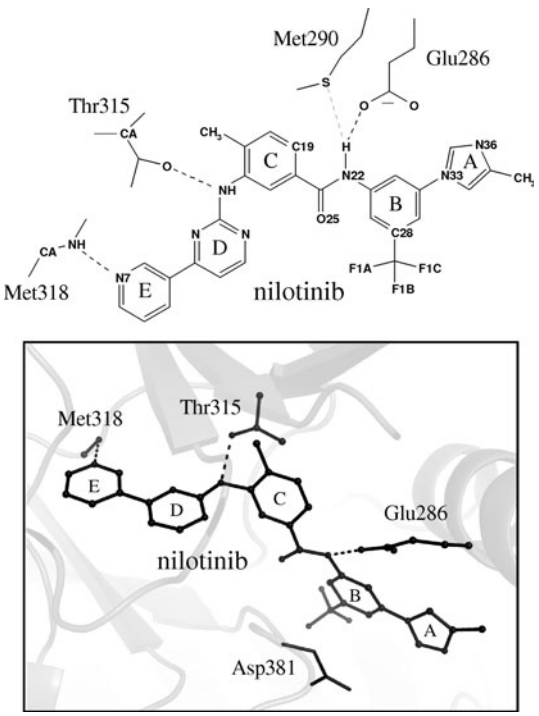

Fig. 6 Binding site residues in Abl; 2D and 3D views

in a hydrophobic pocket formed by residues Ile293, Leu298, Val299, Phe359, Leu354, and His361. It was previously established that this protein environment has a

sufficient plasticity to accommodate large hydrophobic substitutions [23].

4 Conclusions

The major challenge for computer simulations of protein/ligand recognition is the availability of a force field for molecules such as nilotinib, which are nonstandard biopolymers and therefore are not part of standard force fields. In this work, we developed a molecular mechanics force field for nilotinib and several analogues compatible with the CHARMM27 force field for proteins and nucleic acids, as well as with the TIP3P water model. The force field was optimized and tested through careful comparisons of structural, dynamic, and thermodynamic data, from ab initio calculations and experiments. The same procedure was used earlier to parameterize a set of 16 tetracyclines and was shown to give chemical accuracy for relative protein/tetracycline and ribosome/tetracycline binding free energies [21, 42].

The key challenge for the nilotinib/kinase simulations is the determination of the protonation states for both nilotinib and Abl kinase. This problem has been addressed by the employment of two substantially different free energy methods. Our results show that nilotinib can bind to Abl in its protonated or deprotonated forms, with the protonated state somewhat more favoured. Our calculations indicate that the His361 strongly prefers to be neutral, protonated on Nϵ.

The presented approach and the results have implications that extend beyond the example described in this manuscript. Indeed, the proposed methodology is applicable to the investigation of other kinases and provides a starting point for parameterisation of additional nilotinib variants.

Acknowledgments We thank Thomas Simonson for helpful discussions.

References

1. Gambacorti-Passerini CB, Gunby RH, Piazza R, Galietta A, Rostagno R, Scapozza L (2003) Lancet Oncol 4:75–85
2. Levitski A (1996) Curr Opin Cell Biol 8:239–244
3. Chen J, Zhang X, Fernandez A (2007) Curr Drug Targets 7:1443–1454
4. Aleksandrov A, Simonson T (2010) J Biol Chem 285:13807–13815
5. Nagar B, Hantschel O, Young M, Scheffzek K, Veach D, Bornmann W, Clarkson B, Superti-Furga G, Kuriyan J (2003) Cell 112:859–871
6. Deininger M, Buchdunger E, Druker BJ (2005) Blood 105:2640–2653
7. Vajpai N, Strauss A, Fendrich G, Cowan-Jacob S, Manley P, Grzesiek S, Jahnke W (2008) J Biol Chem 283:18292–18302
8. Weisberg E, Manley P, Mestan J, Cowan-Jacob S, Ray A, Griffin JD (2006) Br J Cancer 94:1765–1769
9. DeRemer DL, Ustun C, Natarajan K (2008) Clin Ther 30:1956–1975
10. Giles F, Rosti G, Beris P, Clark R, le Coutre P, Mahon F, Steegmann J, Valent P, Saglio G (2010) Exp Rev Hematol 3:665–673
11. Weisberg E et al (2005) Cancer Cell 7:129–141
12. Mackerell A et al (1998) J Phys Chem B 102:3586–3616
13. Mackerell A, Wiorkiewicz-Kuczera J, Karplus M (1995) J Am Chem Soc 117:11946–11975
14. Jorgensen W, Chandrasekar J, Madura J, Impey R, Klein M (1983) J Chem Phys 79:926–935
15. Trylska J, Antosiewicz J, Geller M, Hodge C, Klabe R, Head M, Gilson M (1999) Prot Sci 8:180–195
16. Aleksandrov A, Proft J, Hinrichs W, Simonson T (2007) ChemBioChem 8:675–685
17. Donnini S, Villa A, Groenhof G, Mark AE, Wierenga RK, Juffer A (2009) Proteins 76:138–150
18. Sham Y, Chu Z, Warshel A (1997) J Phys Chem B 101:4458–4472
19. Simonson T, Carlsson J, Case DA (2004) J Am Chem Soc 126:4167–4180
20. Archontis G, Simonson T (2005) Biophys J 88:3888–3904
21. Aleksandrov A, Simonson T (2006) J Comp Chem 27:1517–1533
22. Aleksandrov A, Simonson T (2009) J Comp Chem 30:243–255
23. Aleksandrov A, Simonson T (2010) J Comp Chem 31:1550–1560
24. Foloppe N, MacKerell A (2000) J Comp Chem 21:86–104
25. Beglov D, Roux B (1994) J Chem Phys 100:9050–9063
26. Simonson T (2000) J Phys Chem B 104:6509–6513
27. Stote R, States D, Karplus M (1991) J Chem Phys 88:2419–2433
28. Brooks B, Bruccoleri R, Olafson B, States D, Swaminathan S, Karplus M (1983) J Comp Chem 4:187–217
29. Szakacs Z, Beni S, Varga Z, Orfi L, Keri G, Noszal B (2005) J Med Chem 48:249–255
30. Vaughan JD, Vaughan VL, Daly SS, Smith WA (1980) J Org Chem 45:3108–3111
31. Bruice T, Schmir G (1958) J Am Chem Soc 80:148–156
32. Simonson T (2001) Free energy calculations. In: Becker O, Mackerell A Jr, Roux B, Watanabe M (eds) Computational biochemistry & biophysics, Chap. 9. Marcel Dekker, New York
33. Thompson D, Plateau P, Simonson T (2006) ChemBioChem 7:337–344
34. Simonson T, Archontis G, Karplus M (2002) Acc Chem Res 35:430–437
35. Hodel A, Simonson T, Fox RO, Brünger AT (1993) J Phys Chem 97:3409–3417
36. Reinhardt W, Miller M, Amon L (2001) Acc Chem Res 34:607–614
37. Shirts MR, Pitera JW, Swope WC, Pande VS (2003) J Chem Phys 119:5740–5761
38. Hermans J (1991) J Phys Chem 95:9029–9032
39. Wood RH (1991) J Phys Chem 95:4838–4842
40. Hummer G (2001) J Chem Phys 114:7330–7337
41. Thompson D, Simonson T (2006) J Biol Chem 281:23792–23803
42. Aleksandrov A, Simonson T (2008) Biochemistry 47:13594–13603
43. Eberini I, Baptista A, Gianazza E, Fraternali F, Beringhelli T (2004) Proteins 54:744–758

44. Simonson T, Brünger AT (1994) J Phys Chem 98:4683–4694
45. Mackerell A Jr (2001) In: Becker O, Mackerell A Jr, Roux B, Watanabe M (eds) Computational biochemistry & biophysics, Chap. 1. Marcel Dekker, New York
46. Simonson T (2001) Curr Opin Struct Biol 11:243–252
47. Richarz R, Wuthrich K (1975) Biopolymers 17:2133–2141
48. Asaki T, Sugiyama Y, Hamamoto T, Higashioka M, Umehara M, Naito H, Niwa T (2006) Bioorg Med Chem Lett 16:1421–1425
49. Puttini M et al (2008) Haematologica 93:653–661