A Molecular Mechanics Model for Imatinib and Imatinib:Kinase Binding

ALEXEY ALEKSANDROV, THOMAS SIMONSON
Department of Biology, Laboratoire de Biochimie (CNRS UMR7654), Ecole Polytechnique, Palaiseau, 91128, France

Received 11 June 2009; Revised 21 August 2009; Accepted 22 September 2009
DOI 10.1002/jcc.21442
Published online 17 December 2009 in Wiley InterScience (www.interscience.wiley.com).

Abstract: Imatinib is an important anticancer drug, which binds specifically to the Abl kinase and blocks its signalling activity. To model imatinib:protein interactions, we have developed a molecular mechanics force field for imatinib and four close analogues, which is consistent with the CHARMM force field for proteins and nucleic acids. Atomic charges and Lennard-Jones parameters were derived from a supermolecule *ab initio* approach. We considered the *ab initio* energies and geometries of a probe water molecule interacting with imatinib fragments at 32 different positions. We considered both a neutral and a protonated imatinib. The final RMS deviation between the *ab initio* and force field energies, averaged over both forms, was 0.2 kcal/mol. The model also reproduces the *ab initio* geometry and flexibility of imatinib. To apply the force field to imatinib:Abl simulations, it is also necessary to determine the most likely imatinib protonation state when it binds to Abl. This was done using molecular dynamics free energy simulations, where imatinib is reversibly protonated during a series of MD simulations, both in solution and in complex with Abl. The simulations indicate that imatinib binds to Abl in its protonated, positively-charged form. To help test the force field and the protonation prediction, we did MD free energy simulations that compare the Abl binding affinities of two imatinib analogs, obtaining good agreement with experiment. Finally, two new imatinib variants were considered, one of which is predicted to have improved Abl binding. This variant could be of interest as a potential drug.

 J Comput Chem 31: 1550–1560, 2010

Key words: molecular recognition; computer simulation; Gleevec; drug design; Charmm program

Introduction

Tyrosine kinases regulate essential aspects of cell growth and differentiation.[1–5] They modify other proteins by transferring a phosphate group from ATP to a tyrosine sidechain; this usually results in a functional change of the target protein and a cellular signal. Unregulated kinase activity can lead to excessive cell division and is a cause of many forms of cancer, such as chronic myeloid leukemia, where the kinase Abl is improperly activated.[6] Obtaining inhibitors that are specific for a particular kinase is difficult, because of the sequence similarity of kinase active sites.[7,8] One important inhibitor is the anticancer drug, imatinib, which binds specifically to Abl and blocks its activity by stabilizing its inactive conformation.[9–11]

It is thus of interest to increase our understanding of imatinib binding to Abl and homologous kinases. Crystallography has provided essential information, with X-ray structures of both Abl and Src bound to imatinib.[9,12] A complementary approach is to develop computer simulation models, which can be used to investigate the structure, dynamics, and thermodynamics of imatinib:protein complexes. The main requirement for computer simulations is to develop a molecular mechanics force field model. For a large, complex molecule like imatinib, this is a difficult task. In this study, we developed a force field model for imatinib and several close analogues. The force field is developed so as to be compatible with the CHARMM27 force field for proteins and nucleic acids,[13,14] and with the TIP3P water model.[15] It should thus be of general interest.

A second requirement for computer simulations is to determine the preferred protonation state of imatinib when it binds to a kinase. Indeed, imatinib has two main protonation states, whose relative populations are known in solution,[16] but not in the protein-bound form. The protonation state can have a profound effect on binding,[17–19] so it must be established for each protein of interest. This is possible using rigorous, molecular dynamics free energy simulations (MDFE).[18,20,21] Below, we report MDFE simulations that

Additional Supporting Information may be found in the online version of this article.

Correspondence to: T. Simonson; e-mail: thomas.simonson@polytechnique.fr

Contract/grant sponsor: French Agence Nationale pour la Recherche (Biology program).

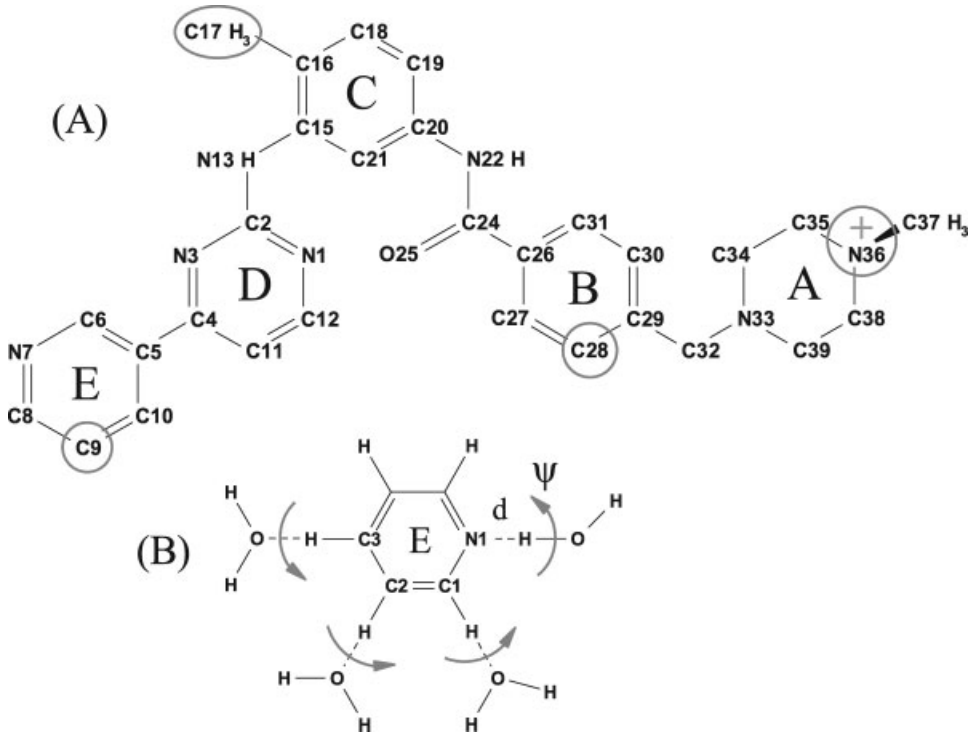

Figure 1. (A) 2D view of imatinib. Positions that will be modified in this study are circled, including N36, which can be protonated or deprotonated. (B) A fragment used for supermolecule calculations, corresponding to ring E, with four possible interacting water positions. Each water position is defined by a single distance, d, and a single angle, ψ.

establish the imatinib protonation state when it is bound to Abl. In solution at a pH of 7.4, imatinib spends 2/3 of its time in a neutral state.[16] The simulations predict that it binds to Abl (and probably to the other, homologous kinases) in a protonated, positively-charged state.

To test the force field and the protonation prediction, we then performed MD simulations of imatinib bound to Abl and Src, and MDFE simulations that compare Abl binding by imatinib and two analogs for which experimental data are available. Agreement with experiment is very good. Finally, to illustrate the usefulness of the model, we used MDFE to explore Abl binding by two new imatinib analogs, one of which is methylated on C8 in the B ring (see below). The simulations predict that this last analog binds more strongly to Abl than plain imatinib, and could thus be of interest as a potential drug.

The model is suitable to investigate the interactions of imatinib and its analogs with a wide range of protein targets. It also provides a starting point to parameterize further imatinib analogues. It should aid efforts both to understand imatinib:kinase binding and to develop improved kinase inhibitors.

Computational Methods

Force Field Determination for Imatinib and Four Analogs

Imatinib is schematized in Figure 1. The force field is developed so as to be compatible with the CHARMM27 force field for proteins and nucleic acids,[13,14] and with the TIP3P water model.[15] The same procedure was used recently to parameterize a large set of tetracycline variants.[22,23] Four imatinib variants were considered (in addition to plain imatinib). The imatinib variations involve three positions, circled in panel A of Figure 1. We refer to them as R2H, X1N, R1CL, and R1M. In R2H, the C17 methyl is replaced by a hydrogen. In X1N, the C9H group in ring E is replaced by a nitrogen atom. In R1CL, chloride replaces the hydrogen in the C28H group. In R1M, there is a new methyl on C28. For each imatinib variant, force field parameters were developed for the most important protonation state at neutral pH: protonated on N36 (with a net positive charge of one). For plain imatinib, both the neutral and the protonated forms were parameterized. Plain imatinib structures were taken from the PDB.

Optimization of the Intermolecular Force Field Parameters

We adopt a force field of the CHARMM27 form.[13] The intermolecular energy terms are Lennard-Jones and Coulomb terms. Lennard-Jones parameters were taken from the CHARMM27 force field, by assigning to each imatinib atom an existing atomic type within that force field.[13,14,24] In accord with the development of the CHARMM force field, atomic charges were derived from a supermolecule, *ab initio* approach. We considered the *ab initio* energies and geometries of a probe water molecule interacting with small model compounds that are representative of different fragments of imatinib and its variants: benzene, chlorbenzene, pyridine,

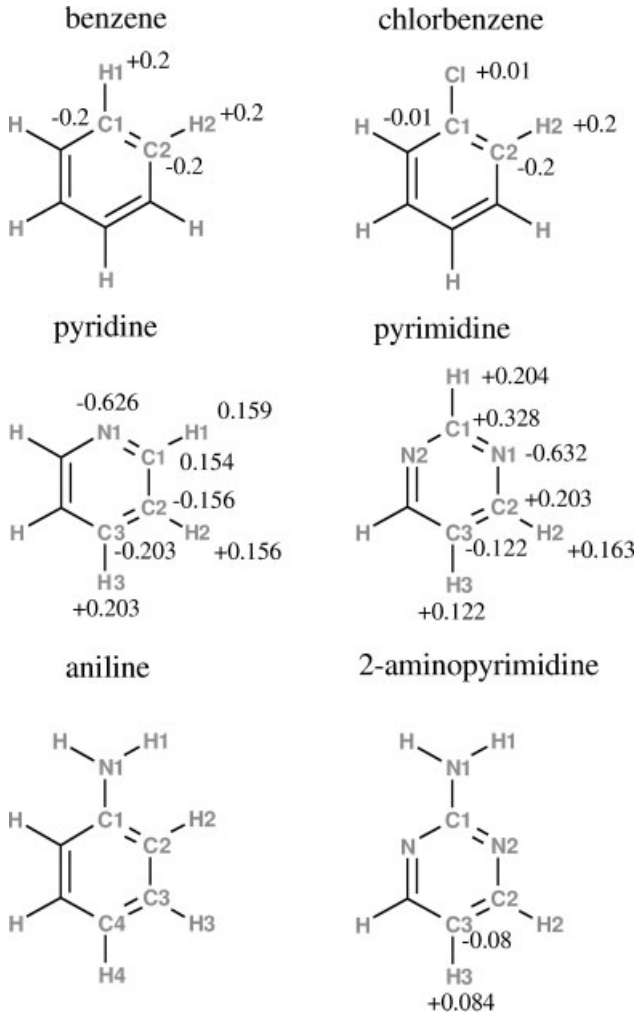

Figure 2. Small molecules used for supermolecule calculations, with atomic charges indicated.

pyrimidine, aniline, and 2-amino pyrimidine (Fig. 2). Selected probe positions around each fragment are considered. This supermolecule approach is known to give a good balance between solute–water and water–water interactions in the force field.[13,14] For the force field calculations, the water geometry was taken from the TIP3P model.[15] 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 that water position. A single angle ψ defining the orientation of the water molecule was held fixed (see Fig. 1). For each water position, ψ was varied over the full 180° range in 30° or 60° steps, giving 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, to compensate for overestimated interaction distances with the Hartree-Fock model (due to neglected electron correlation).[13] The force field charges were then adjusted to reproduce these "corrected" *ab initio* interaction energies and water positions. 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 varying manually the model compound charges. This involved reoptimizing the imatinib-water distance after each parameter change. Adjustment of the charges upon linking the model fragments to form larger entities was performed by adding the charge of the deleted hydrogen atom to the heavy atom to which it was previously attached. This approach maintains the total charge from each, original, model compound. Charges of linking groups were reoptimized 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 (Figs. 1 and 2).

Optimization of the Intramolecular Force Field Parameters

In a force field treatment, the intramolecular geometry is mainly determined by the minimum energy values of the bond length and bond angle terms, and by the phase and multiplicity of the dihedrals. 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. Starting guesses for the bond and angle force constants, and for the dihedral force constants and phases were taken from the CHARMM22 force field[13] for related, small molecular fragments. The geometrical and dihedral parameters were then optimized by fitting to the structures from *ab initio* optimizations. At each parameter optimization step, the structure was minimized with the force field model, using a Powell conjugate gradient algorithm, stopping when the RMS energy gradient reached 10^{-6} kcal/mol/Å. The quality of the parameters was measured by the RMS coordinate deviation from the *ab initio* structure, and the force field/*ab initio* energy difference. The bond and angle geometries, and the dihedral geometries were then updated manually and a new round of optimization was started. This procedure was repeated until a satisfactory agreement was achieved (coordinate RMS deviation of 0.16 Å).

For the imatinib variants, R2H, X1N, R1CL, and R1M, very few additional bonded parameters were needed. For R2H (demethylated C16), parameters were taken from the benzene fragment. For X1N (nitrogen in ring E), parameters were taken from pyrimidine. For R1CL, the chlorinated benzene fragment was used. For R1M (methylated C28), parameters were taken from the fragment used to parameterize ring C.

Molecular Dynamics Simulations

The crystal structure of the human Abl tyrosine kinase was taken from the Protein Data Bank, entry 2HYY (with bound imatinib).[25,26] We use human Abl numbering throughout this article. Abl was modelled in its inactive state, which is the state competent to bind imatinib.[25,26] Protonation states of histidines were assigned by visual inspection: His361 is neutral; five other histidines are doubly-protonated. For Src, a structure of chicken Src with bound imatinib was used (entry 2OIQ)[12]). In this structure, residues 407–421 in the activation loop are disordered. In the Src simulations, they are neglected. To alleviate the effect of the missing residues, residues

406 and 422 are harmonically restrained to stay in their experimental positions.

The simulations included protein residues within a 26 Å sphere, centered on the imatinib binding site. In addition to crystal waters, a 26 Å sphere of water was overlaid and waters that overlapped protein, crystal waters, or imatinib were removed. Throughout the MD simulations, protein atoms between 20 Å and 26 Å from the sphere's center were harmonically restrained to their experimental positions. Since the model is centered on the 20 Å-long imatinib, the shortest distance between the ligand and any harmonically-restrained group is a bit more than 10 Å. Simulations were done with the SSBP solvent model,[27,28] which treats the region outside the 26 Å sphere as a uniform dielectric continuum, with a dielectric constant of 80. This is reasonable, since most of the outer region is water. A more sophisticated alternative would be to include the heterogeneous outer medium through a more complex continuum treatment,[29] however, the simpler method should be sufficient here. Newtonian dynamics were used for the innermost region, within 20 Å of the sphere's center; Langevin dynamics were used for the outer part of the sphere, with a 292 K bath. The CHARMM22 force field was used for the protein[13] and the TIP3P model for water.[15] Electrostatic interactions were computed without any cutoff, using a multipole approximation for distant groups.[30] Calculations were done with the CHARMM program.[31]

Imatinib Protonation State

At a physiological pH of 7.4, imatinib spends about 2/3 of its time in the neutral state[16]; the rest of the time it is protonated—usually on the methylated nitrogen N36 of its piperazine ring (Fig. 1). To determine its protonation state in the protein complex, we undertook a molecular dynamics free energy simulation study.[20,21] The imatinib is gradually protonated both in solution and in complex with Abl, through a series of MD simulations (Fig. 3A, with Ima^+ representing the protonated form). The MD simulation lasted for 3.6 ns for both forward and backward runs (7.2 ns total). This technique has no adjustable parameters and gives good, though not perfect accuracy for acid/base reactions.[21]

Alchemical MD Free Energy Simulations

To compare imatinib and its analogs binding to Abl, we use the thermodynamic cycle in Figure 3B. The MDFE method follows the horizontal legs of the cycle. Imatinib is reversibly transformed into its analog Ima′ 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 previously (see also refs. 28, 33).

For the lower leg of the thermodynamic cycle, we simulate imatinib in solution. For the upper leg, we simulate a portion of the imatinib: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 imatinib (Ima) and its analogue (Ima′):

$$U(\lambda) = U_0 + (1-\lambda)U(\mathrm{Ima}) + \lambda U(\mathrm{Ima}') \qquad (1)$$

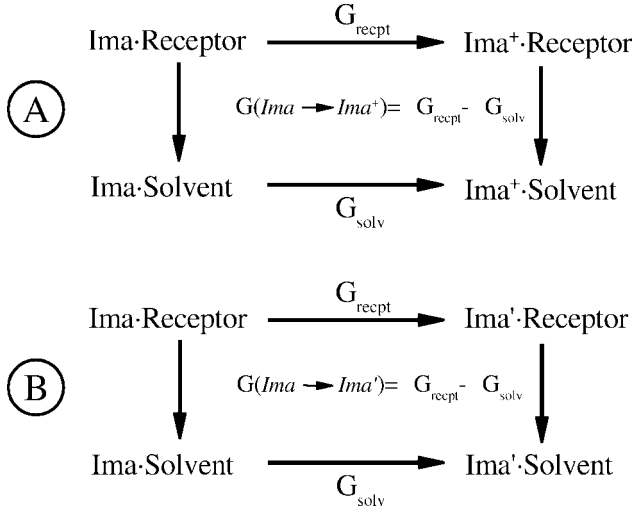

Figure 3. (A) Thermodynamic cycle to study imatinib protonation and its effect on binding to a receptor (Abl in this case). Horizontal legs represent the protonation of imatinib on its N36 nitrogen, either in complex with Abl (above) or alone (below). Vertical legs represent imatinib:Abl binding. The MD free energy simulations follow the horizontal legs. (B) Thermodynamic cycle to study imatinib (Ima) and its analog (Ima′) binding to a receptor. Horizontal legs represent the alchemical transformation of imatinib into its analog, either in complex with a receptor (above) or alone (below). Vertical legs represent ligands:Receptor binding. The MD free energy simulations follow the horizontal legs.

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

$$\frac{\partial G}{\partial \lambda}(\lambda) = \langle U(\mathrm{Ima}') - U(\mathrm{Ima})\rangle_\lambda \qquad (2)$$

where the brackets indicate an average over an MD trajectory with the energy function $U(\lambda)$.[32,34] We gradually mutated Ima into Ima′ by changing λ from 0 to 1. The successive values of λ were: 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. One run was performed in each direction (Ima into Ima′ and the reverse). To obtain the free energy change ΔG from the derivatives, we used trapezoidal integration, except for van der Waals contributions close to the endpoints of the mutation (λ = 0.0–0.001 and 0.999–1.0). For these endpoint contributions, analytical integration was performed.[32,35]

Accurate uncertainty estimation with MDFE is difficult and expensive.[36–38] A widely-used approach is to perform multiple runs and measure the dispersion between runs. Although this seems plausible, it neglects some forms of systematic error, and can actually lead to an overestimated uncertainty. In particular, it is well known that runs performed in opposite directions ("forward" and "backward" transformations; λ increasing or decreasing) exhibit systematic hysteresis effects.[39–41] Thus, in the past, we have used

Table 1. Interactions Between a Probe Water and Benzene or Chlorobenzene.

	Ab initio/Force field results		
Probe site	Energy (kcal/mol)	Distance d (Å)	Angle ψ (°)
C1H1	−1.35/−1.38	2.69/2.64	0.0
C1H1	−1.40/−1.41	2.68/2.64	30.0
C1H1	−1.48/−1.47	2.65/2.63	60.0
C1H1	−1.53/−1.50	2.63/2.63	90.0
C1H1	−1.48/−1.47	2.65/2.63	120.0
C1H1	−1.40/−1.41	2.68/2.64	150.0
C1Cl	−0.03/−0.04	3.49/4.08	0.0
C1Cl	−0.03/−0.05	3.48/4.07	30.0
C1Cl	−0.07/−0.06	3.47/4.06	60.0
C1Cl	−0.07/−0.07	3.46/4.05	90.0
C1Cl	−0.07/−0.06	3.47/4.06	120.0
C1Cl	−0.03/−0.05	3.48/4.07	150.0
C2H2	−1.94/−1.62	2.48/2.62	0.0
C2H2	−1.89/−1.67	2.48/2.61	60.0
C2H2	−1.89/−1.67	2.48/2.61	120.0

Results for benzene (upper 6 lines; one site) and chlorobenzene (bottom 9 lines; two sites).
The distance d and angle ψ are defined in Figure 1B.

Table 2. Interactions Between a Probe Water and N,N-methyl-piperazine.

	Ab initio/Force field results		
Probe site	Energy (kcal/mol)	Distance (Å)	Angle (°)
N1H1	−15.25/−14.82	1.81/1.81	0.0
N1H1	−15.17/−14.85	1.80/1.81	60.0
N1H1	−15.25/−14.82	1.81/1.81	120.0
C1H2	−8.30/−8.08	2.27/2.48	0.0
C1H2	−8.55/−8.15	2.25/2.48	60.0
C1H2	−8.83/−8.35	2.24/2.47	120.0
C2H3	−8.51/−8.72	2.27/2.46	0.0
C2H3	−8.34/−8.68	2.29/2.46	60.0
C2H3	−8.89/−8.89	2.25/2.46	120.0
C2H4	−8.18/−7.50	2.26/2.49	0.0
C2H4	−8.30/−7.37	2.25/2.49	60.0
C2H4	−7.53/−7.23	2.32/2.50	120.0
C3H5	−6.44/−5.66	2.35/2.56	0.0
C3H5	−6.78/−5.78	2.32/2.55	60.0
C3H5	−6.34/−5.54	2.36/2.56	120.0
C3H6	−6.75/−6.83	2.45/2.55	0.0
C3H6	−6.31/−6.71	2.50/2.56	60.0
C3H6	−6.34/−6.61	2.50/2.56	120.0
C4H7	−4.45/−3.99	2.52/2.62	0.0
C4H7	−4.53/−4.01	2.52/2.62	60.0
C4H7	−4.98/−4.23	2.45/2.61	120.0

Results for positively-charged N,N-methyl-piperazine (seven sites).
The distance d and angle ψ are defined in Figure 1B.

error estimators that consider pairs of runs, one in each direction, forming a "forward/backward pair".[33,42] The forward/backward averages were much more reproducible than the individual values. Here, we perform just one forward and one backward run, so this method is not possible. Instead, we use a simpler approach, based on a standard block-averaging of each simulation window.[43,44] Each window is split into two equal blocks, and the derivative values computed. The difference between the two blocks yields an uncertainty estimate for that window. The uncertainty for the overall free energy is then obtained by standard error propagation. We have found in the past that for simple transformations, such as those involving the imatinib variants, this block averaging error agrees with the above method (comparing forward/backward pairs of runs).[45,46]

Results

Force Field Development: Imatinib–Water Interactions

We first consider the supermolecule calculations for imatinib fragments interacting with individual water molecules. Detailed results are given in Tables 1 and 2 for two representative fragments. The first corresponds to both rings B and C; the second corresponds to the protonated piperazine ring A (Fig. 1). Results for the other two rings are similar. Overall, very good agreement was obtained between the *ab initio* and force field data, as shown in Figure 4. The RMS deviation for the energies was 0.2 kcal/mol, averaged over all fragments, all 32 water positions, and all water orientations (six for each water position). The RMS deviation for the fragment–water distances was 0.17 Å. To obtain this agreement, only small adjustments were required for the atomic charges. Thus, for plain, protonated imatinib, the RMS deviation between the initial guess (Mulliken charge) and the final charge was 0.04 (in units of a proton's charge); the largest deviation was 0.11.

The charges for the small fragments are reported in Figure 2. The final atomic charges for imatinib include some small adjustments due to the linking of rings (see Methods); they are reported

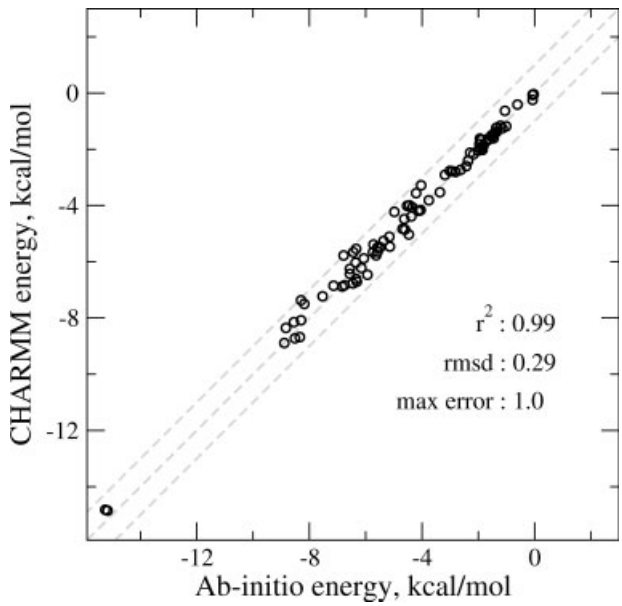

Figure 4. Interaction energies between the probe water molecules and the imatinib fragments from the supermolecule calculations.

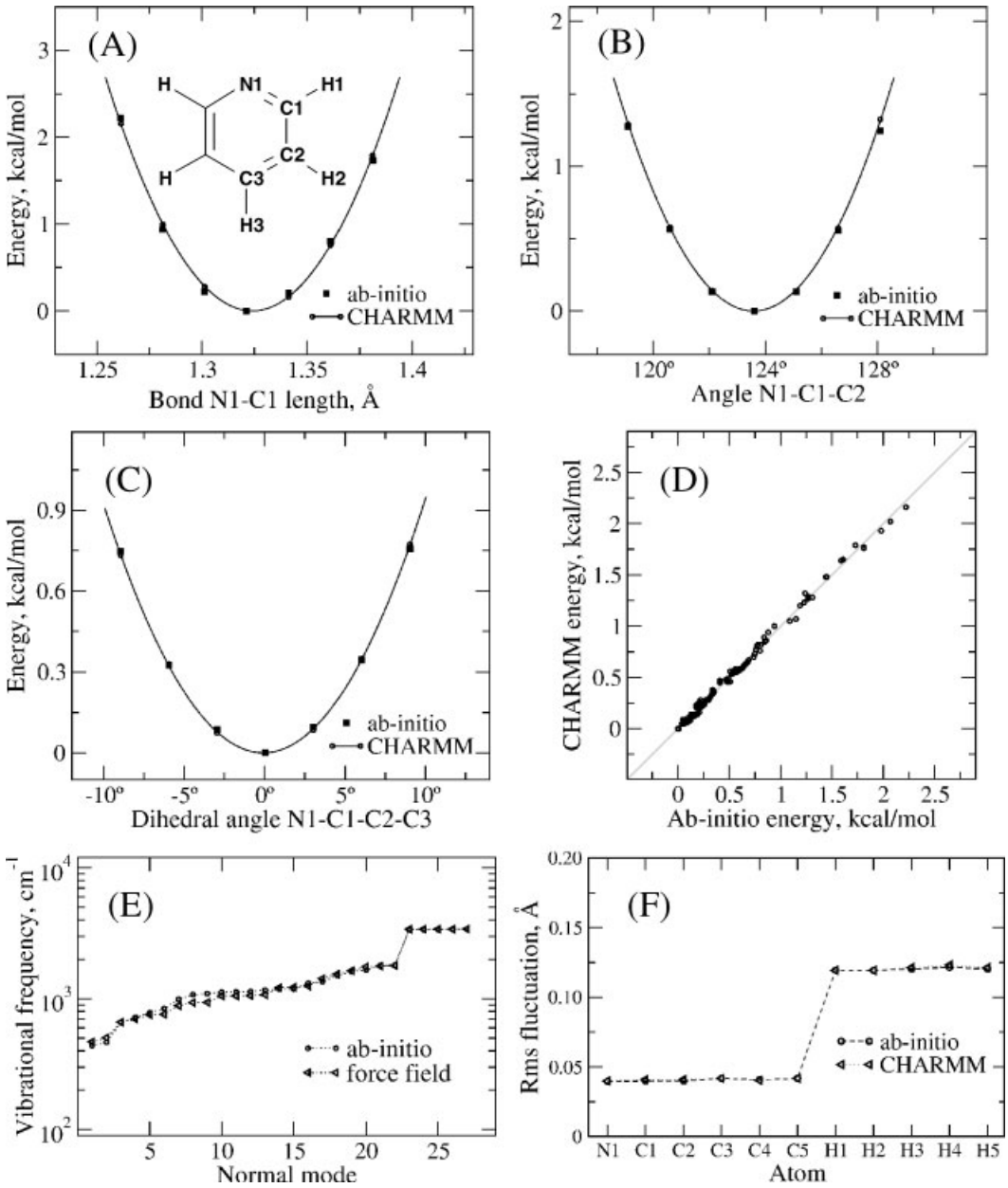

Figure 5. Comparing *ab initio* and force field results for the internal degrees of freedom of ring E. (A) Bond stretching energy for N1-C1. (B) Angle bending energy for N1-C1-C2. (C) Torsional energy for the intraring angle N1-C1-C2-C3. (D) 161 points corresponding to small ring deformations, involving five bonds, eight bond angles, and 10 intraring dihedrals (including the ones shown in the panels A–C). In panels A–D, the solid lines correspond to force field energies. (E) Normal mode frequencies. (F) RMS positional fluctuations from the normal mode calculation.

in Supporting Information. These final charges can be compared to several groups in the Charmm27 force field for proteins and nucleic acids.[13,14,24,47] The C_{17} methyl in the C ring has standard Charmm methyl charges (−0.27 on the carbon; atomic units). 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 to comparable nitrogens in histidine and guanine. The C_{32} and C_{39} methylenes have a typical carbon charge (−0.14, −0.16) but a 0.22 net positive charge; this differs from the most common Charmm methylene, which is neutral, but is identical to the serine sidechain methylene and similar to glycine. 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_{19}, with charges of −0.26 and +0.29; the pair with the largest net charge is C_{12}-H_{12}, with charges of +0.16 and +0.16. In Charmm, the Phe and Tyr rings have CH charges of −0.12 and +0.12; uracil has

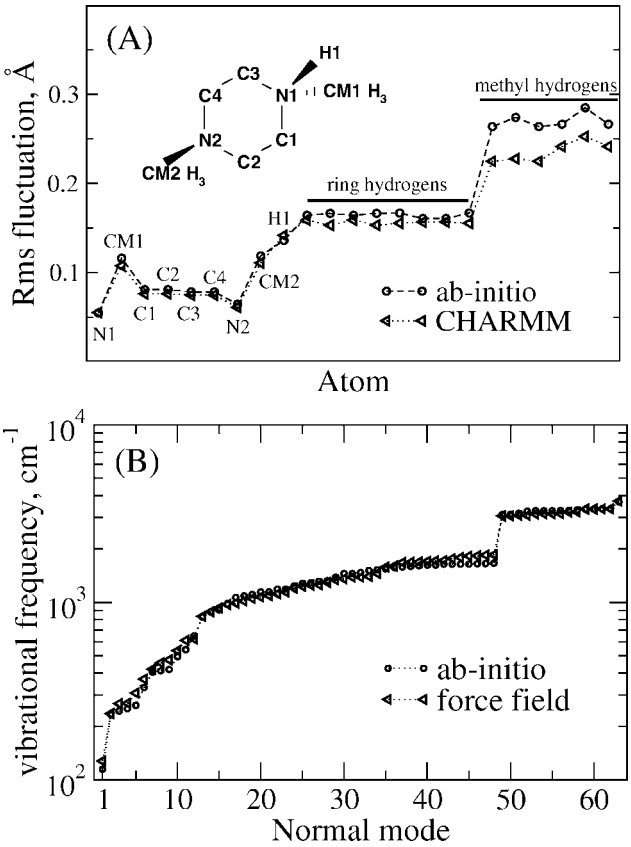

Figure 6. Normal mode frequencies (A) and RMS positional fluctuations (B) for a fragment representing ring A (see inset).

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. Finally, 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 to +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. It may be that with some adjustment of nearby charges, we could obtain an optimized charge set for these atoms that is more similar to standard Charmm; we did not explore this possibility.

Force Field Development: Imatinib Structure and Flexibility

The force field accurately reproduces the *ab initio* minimized structure of imatinib. For protonated imatinib, the RMS deviation between the *ab initio* and force field structures, after superposition of the whole molecule, is 0.16 Å; for deprotonated imatinib, the RMS deviation is also 0.16 Å.

To characterize imatinib flexibility, we have examined two kinds of fluctuations: small, harmonic fluctuations, which involve the stiff ring structures, and larger, lower-frequency fluctuations, which involve the soft dihedral angles linking the rings. Figure 5 illustrates the small, harmonic fluctuations for a single fragment, representing the E ring. Energy profiles are shown for bond stretching, angle bending, and an intra-ring torsion angle. Normal mode frequencies are also shown, along with the corresponding RMS positional fluctuations of the ring atoms. Figure 6 shows the normal mode frequencies and RMS positional fluctuations for a second fragment, representing ring A. For all these properties, agreement between the *ab initio* and force field results is very good.

Figure 7 illustrates the softer, inter-ring, dihedral fluctuations. Energy profiles are shown for two dihedral angles, which link two pairs of imatinib rings: rings A–B and rings D–E, respectively. Again, agreement is very good between the force field and the *ab initio* profiles, not only for the energy wells, but also for the barriers between wells. These softer, dihedral degrees of freedom are very important for accurately describing the conformational free energy surface of imatinib and the role of induced fit in imatinib–protein binding.

The force field parameters are given in Supporting Information.

Imatinib Protonation and Protein:Ligand Interactions

To study imatinib binding to Abl and other kinases, it is essential to first establish which imatinib protonation state is relevant. In solution, two protonation states are predominant:[16] with N36 protonated (in the piperazine A ring; see Fig. 1), giving a net positive charge, or with N36 deprotonated. The protonated form is occupied about 1/3 of the time; the deprotonated form about 2/3 of the time. With an accurate force field in hand, we may investigate the relative free energies of these two forms when imatinib is bound to Abl, using a rigorous, MDFE technique.[20,21] Unlike widely-used continuum electrostatic methods,[48] MDFE has no adjustable parameters.

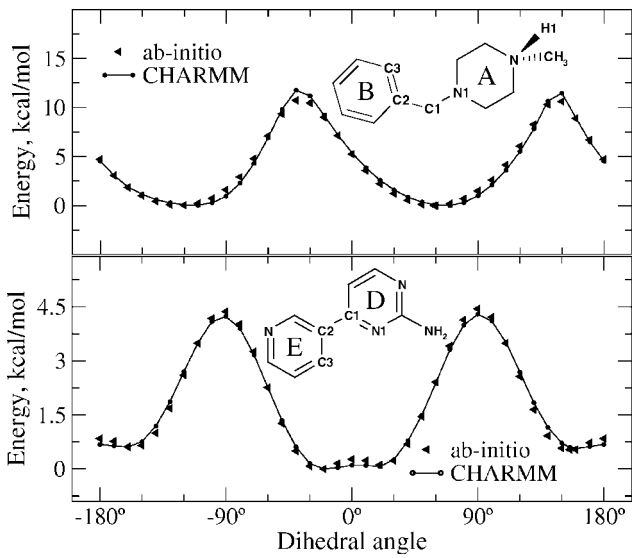

Figure 7. Comparing *ab initio* and force field energies for two dihedral angles that link rings A–B (upper panel) and D–E (lower panel). The dihedral angles are defined by the atoms labelled in each inset (C3-C2-C1-N1 in both cases). In each panel, the solid line corresponds to force field energies.

Table 3. Imatinib Protonation Free Energy Simulations.

	Individual MDFE runs[a]	Average[b]
ΔG_{solv}	33.4, 34.0/30.2, 30.4	32.0 (0.4)
ΔG_{prot}	27.8, 27.2/28.8, 28.1	28.0 (0.7)
$\Delta\Delta G$		**−4.0 (0.8)**

Protonation free energy in solvent (ΔG_{solv}) or in complex with the protein (ΔG_{prot}), in kcal/mol. The $\Delta\Delta G_{\text{MDFE}}$ indicates preferential binding of the protonated form.
[a]Values before the slash are "forward" runs (protonation); values after the slash are "backward" runs.
[b]Uncertainty in parentheses.

Table 4. Selected Distances (Å) Between Atoms of Imatinib and Abl.

Atom pair	MD Simulation	X-ray structure[a]
N_{M318} $N3_{\text{Ima}}$	3.1	3.0
$O\gamma_{T315}$ $N8_{\text{Ima}}$	3.0	3.0
$O\epsilon1_{E286}$ $N21_{\text{Ima}}$	3.1	3.1
N_{D381} $O29_{\text{Ima}}$	3.0	3.0
O_{I360} $N36_{\text{Ima}}$	2.8	2.7
O_{H361} $N36_{\text{Ima}}$	3.5	3.1

Abl atoms (left) are labelled by the amino acid to which they belong.
[a]PDB entry 2HYY.

Results are summarized in Table 3. The protonation free energy in the protein is more favorable than in solution; the difference is estimated to be $\Delta\Delta G_{\text{bind}}(\text{Imat} \rightarrow \text{Imat}^{+}) = -4.0 \pm 0.8$ kcal/mol. The negative $\Delta\Delta G_{\text{bind}}$ implies that imatinib prefers to bind in its protonated form, with a net positive charge. At physiological pH, therefore, the complex with neutral imatinib represents only a small fraction of the overall population (less than 0.1%). Notice that this disagrees with the assumption of a neutral imatinib made in a recent simulation study.[49] The strong preference for the protonated imatinib is partly due to favorable electrostatic interactions between

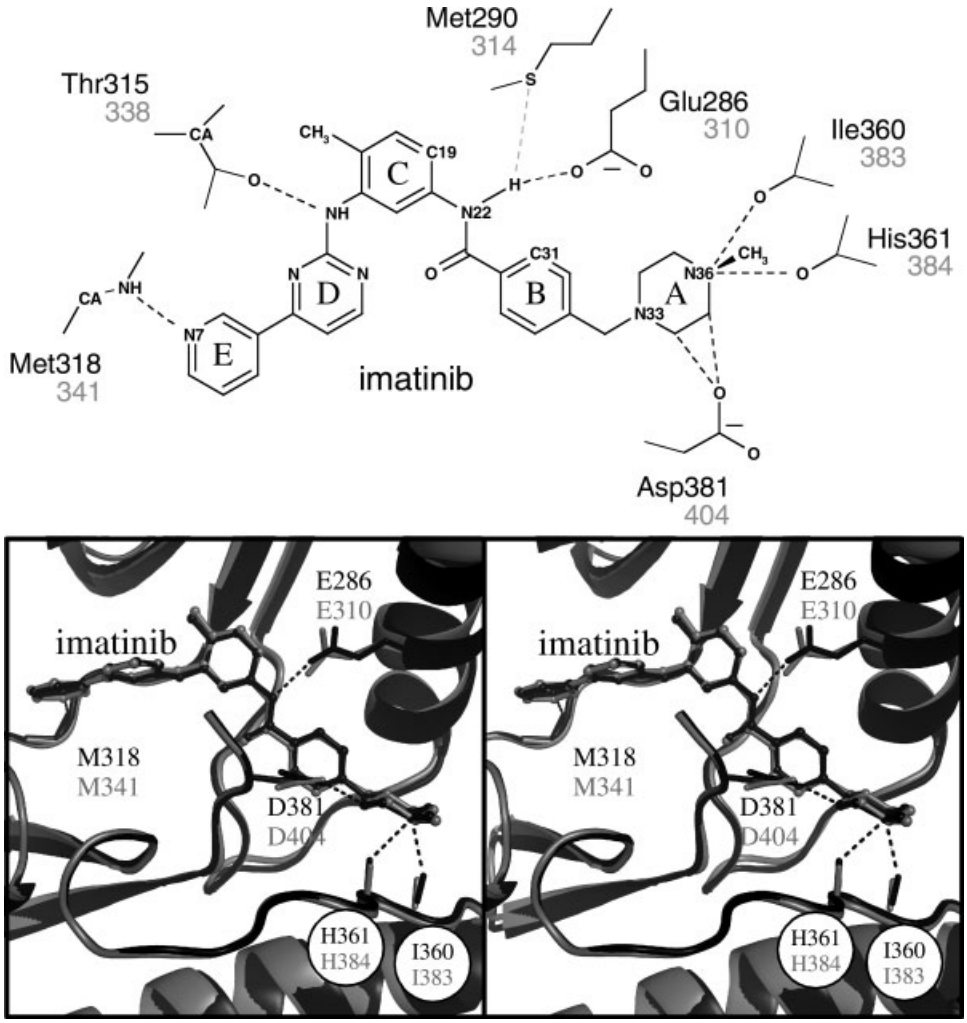

Figure 8. Binding site residues in Abl (black) and Src (gray); 2D and 3D (stereo) views. Both human Abl (black) and chicken Src (gray) numbering are given.

the charged *N*-methyl-piperazine ring of imatinib (ring A) and the carbonyl groups of Ile359 and His360.

For the protonated imatinib, we then simulated the protein:ligand complex for 5 ns, for comparison with the X-ray structure. We computed RMS deviations from the X-ray structure, averaging over the last 50 ps of the 5 ns. By superimposing the imatinib on the crystal structure, we obtain intramolecular imatinib deformations of less than 0.5 Å. Superimposing the protein backbone on the crystal structure, we obtain imatinib RMS deviations of 0.7 Å, which correspond partly to an overall shift of the imatinib molecule. The protein backbone RMS deviations are 1.0 Å. The interactions between imatinib and the protein are also well-reproduced (Table 4).

Figure 8 shows the Abl:imatinib and also the Src:imatinib complex; the latter complex was also simulated for 5 ns. The backbone RMS deviation between the Src MD and X-ray structures is 1.0 Å, identical to the Abl complex. Overall, both MD models are similar to other good quality, current simulations.

Table 5. Imatinib Analogs Binding to the Abl Kinase: MDFE Results.

	MDFE			Experiment
Imatinib modification	ΔG_{prot} [a]	ΔG_{sol} [a]	$\Delta\Delta G$	$\Delta\Delta G_{exp}$
X1N	−5.4	−5.4	0.0	−0.1[50 b]
R1CL	4.6	7.4	−2.8	−2.8, −1.8[50,51 b]
R2H	−4.2	−5.8	1.6	–
R1M	0.5	3.2	−2.7	–

Free energies in kcal/mol.
[a]Columns 2, 3 are the free energy to reversibly transform plain imatinib into one of its analogs, either in the protein (ΔG_{prot}) or in solution (ΔG_{sol}). $\Delta\Delta G = \Delta G_{Abl} - \Delta G_{sol}$ is the relative binding free energy. A negative $\Delta\Delta G$ means the imatinib analog binds more strongly. MDFE error bars are all close to 0.3 kcal/mol.
[b]Derived from minimum inhibitory concentration *in vivo*.

Comparing Imatinib Analogs: MDFE Simulations

In this section, we report MD free energy simulations comparing the binding of imatinib and three analogs to Abl. For two of the analogs, the binding free energy differences with respect to imatinib are known experimentally. Comparing the MDFE results to experiment helps support the accuracy of the imatinib force field model developed here. We also consider two new imatinib analogs. The imatinib variations involve three positions, circled in panel A of Figure 1. Following the protonation calculations above, all the ligands are assumed to be protonated on N36 and positively-charged. The results are summarized in Table 5.

In the first variant ("X1N" in Table 5), the C9H group in the pyridine E ring is replaced by a nitrogen atom. The computed binding free energy difference is 0.0 (0.3) kcal/mol, in close agreement with the experimental value, −0.1 kcal/mol.[50] The C9H group is solvent-exposed in the Abl complex, and does not contribute directly to the binding. The computed solvation free energy change is −5.5 kcal/mol, favoring solvation of the new ligand over imatinib. Thus, while this modification does not change imatinib's affinity for Abl, it changes its solubility, which could be important for drug admission.

In the second variant, "R1Cl," chloride replaces the hydrogen in the C28H group.[50,51] This group binds in a hydrophobic pocket formed by residues Ile293, Leu298, Leu354, His361, and Val379[9]; see Figure 9. The computed binding free energy difference is −2.8± 0.2 kcal/mol, favoring the new ligand. This is in close agreement with the experimental value, −2.8 kcal/mol.[51]

In the third variant, "R2H," the methyl group $C17H_3$ in the C ring is replaced by a hydrogen. The computed binding free energy

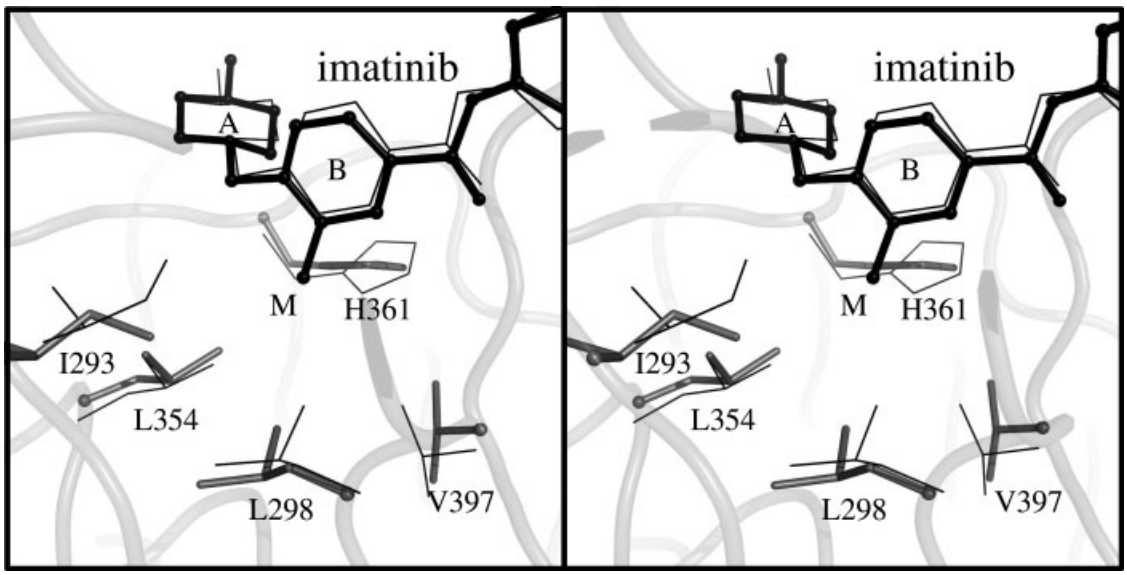

Figure 9. Wall-eyed stereo view of the R1M analogue bound to Abl. The structure in bold is taken from the end of the molecular dynamics free energy simulation. The lighter structure is the initial, crystal structure. Relaxation of the protein in response to the new methyl group (M) can be seen. Residues that form the hydrophobic pocket are shown in gray.

difference is 1.6 (0.3) kcal/mol, which reflects the superior solvation of the demethylated variant.

Finally, we considered a new variant of imatinib, "R1M," with a methyl group at the C28 position in the B ring. The computed binding free energy difference is -2.7 ± 0.2 kcal/mol, strongly favoring the new ligand over imatinib. The new methyl group fits well into the hydrophobic pocket discussed earlier. When the methyl is introduced during the MDFE simulations, imatinib does not shift; rather, the methyl is accomodated by a 1.3 Å shift of the L298 Cδ2 atom and a 1.1 Å shift of the L293 Cδ atom (Fig. 9). This plasticity of the protein environment could not be deduced directly from the available X-ray structures.

Conclusions

Despite its importance as an anti-cancer drug, imatinib has rarely been the object of theoretical studies.[49,52,53] Indeed, a major hurdle for computer simulations of protein:ligand recognition is force field availability for molecules like imatinib that are not standard biopolymers, and are not part of standard force fields. In this study, we developed a molecular mechanics force field for imatinib and several analogs that is compatible with the CHARMM27 force field for proteins and nucleic acids, and with the TIP3P water model. The force field was optimized and tested by comparison to a range of structural, dynamic, and thermodynamic data, from *ab initio* calculations and experiment. 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.[22,45]

A difficulty for imatinib:kinase simulations is that the imatinib protonation state must be determined. For this, we used MDFE simulations, showing that the protonated, positively-charged state is predominant when imatinib is bound to Abl, and probably to other, homologous kinases. This prediction disagrees with the assumption of a neutral imatinib made in a recent simulation study.[49]

The force field and protonation state were tested by MD simulations of the Abl:imatinib and Src:imatinib complexes, and by MDFE simulations that compared two imatinib variants binding to Abl, yielding structures and relative binding free energies in good agreement with experiment.

The present methods and results are of general interest. Indeed, they are applicable to other kinases, and provide a starting point to parameterize and study additional imatinib variants. Finally, the new, methylated imatinib variant proposed here has an enhanced Abl binding affinity and could be of interest as a potential drug.

Acknowledgments

The authors thank Christine Bathelt and Winfried Hinrichs for discussions, Christine Bathelt for preliminary calculations, and John Kuriyan for providing data before publication.

References

1. Hunter, T. Semin Cell Biol 1994, 5, 367.
2. Hubbard, S. R.; Hill, J. H. Ann Rev Biochem 2000, 69, 373.
3. Pawson, T.; Nash, P. Science 2003, 300, 445.
4. Bose, R.; Holbert, M. A.; Pickin, K. A.; Cole, P. A. Curr Opin Struct Biol 2006, 16, 668.
5. Kuriyan, J.; Eisenberg, D. Nature 2007, 450, 983.
6. Gambacorti-Passerini, C. B.; Gunby, R. H.; Piazza, R.; Galietta, A.; Rostagno, R.; Scapozza, L. Lancet Oncol 2003, 4, 75.
7. Levitski, A. Curr Opin Cell Biol 1996, 8, 239.
8. Chen, J.; Zhang, X.; Fernandez, A. Curr Drug Targets 2007, 7, 1443.
9. Nagar, B.; Hantschel, O.; Young, M. A.; Scheffzek, K.; Veach, D.; Bornmann, W.; Clarkson, B.; Superti-Furga, G.; Kuriyan, J. Cell 2003, 112, 859.
10. Deininger, M.; Buchdunger, E.; Druker, B. J. Blood 2005, 105, 2640.
11. Vajpai, N.; Strauss, A.; Fendrich, G.; Cowan-Jacob, S.; Manley, P.; Grzesiek, S.; Jahnke, W. J Biol Chem 2008, 283, 18292.
12. Seeliger, M. A.; Nagar, B.; Frank, F.; Cao, X.; Henderson, M. N.; Kuriyan, J. Structure 2007, 15, 299.
13. Mackerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586.
14. Mackerell, A. D.; Wiorkiewicz-Kuczera, J.; Karplus, M. J Am Chem Soc 1995, 117, 11946.
15. Jorgensen, W.; Chandrasekar, J.; Madura, J.; Impey, R.; Klein, M. J Chem Phys 1983, 79, 926.
16. Szakacs, Z.; Beni, S.; Varga, Z.; Orfi, L.; Keri, G.; Nosza, B. J Med Chem 2005, 48, 249.
17. Trylska, J.; Antosiewicz, J.; Geller, M.; Hodge, C.; Klabe, R.; Head, M.; Gilson, M. Prot Sci 1999, 8, 180.
18. Aleksandrov, A.; Proft, J.; Hinrichs, W.; Simonson, T. Chem Bio Chem 2007, 8, 675.
19. Donnini, S.; Villa, A.; Groenhof, G.; Mark, A. E.; Wierenga, R. K.; Juffer, A. Proteins 2009, 76, 138.
20. Sham, Y.; Chu, Z.; Warshel, A. J Phys Chem B 1997, 101, 4458.
21. Simonson, T.; Carlsson, J.; Case, D. A. J Am Chem Soc 2004, 126, 4167.
22. Aleksandrov, A.; Simonson, T. J Comput Chem 2006, 13, 1517.
23. Aleksandrov, A.; Simonson, T. J Comput Chem 2008, 30, 243.
24. Foloppe, N.; MacKerell, A. D. Jr., J Comput Chem 2000, 21, 86.
25. Levinson, N. M.; Kuchment, O.; Shen, K.; Young, M. A.; Koldobskiy, M.; Karplus, M.; Cole, P. A.; Kuriyan, J. Plos Biol 2006, 4, 753.
26. Cowan-Jacob, S.; Fendrich, G.; Floersheimer, A.; Furet, P.; Liebetanz, J.; Rummel, G.; Rheinberger, P.; Centeleghe, M.; Fabbro, D.; Manley, P. Acta Cryst D 2007, 63, 80.
27. Beglov, D.; Roux, B. J Chem Phys 1994, 100, 9050.
28. Simonson, T. J Phys Chem B 2000, 104, 6509.
29. Im, W.; Bernèche, S.; Roux, B. J Chem Phys 2001, 114, 2924.
30. Stote, R.; States, D.; Karplus, M. J Chim Phys 1991, 88, 2419.
31. Brooks, B.; Brooks, C. L. III.; Mackerell, A. D. Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. J Comput Chem 2009, 30, 1545.
32. Simonson, T. Free Energy Calculations. In Computational Biochemistry & Biophysics, Becker, O.; Mackerell, A. Jr.; Roux, B.; Watanabe, M. Eds.; Marcel Dekker: New york, 2001; chapter 9.
33. Thompson, D.; Plateau, P.; Simonson, T. Chem Bio Chem 2006, 7, 337.
34. Simonson, T.; Archontis, G.; Karplus, M. Acc Chem Res 2002, 35, 430.
35. Simonson, T. Mol Phys 1993, 80, 441.
36. Hodel, A.; Simonson, T.; Fox, R. O.; Brünger, A. T. J Phys Chem 1993, 97, 3409.
37. Reinhardt, W.; Miller, M.; Amon, L. Acc Chem Res 2001, 34, 607.

38. Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. J Chem Phys 2003, 119, 5740.
39. Hermans, J. J Phys Chem 1991, 95, 9029.
40. Wood, R. H. J Phys Chem 1991, 95, 4838.
41. Hummer, G. J Chem Phys 2001, 114, 7330.
42. Thompson, D.; Lazennec, C.; Plateau, P.; Simonson, T. J Biol Chem 2007, 42, 30856.
43. Straatsma, T.; Berendsen, H.; Stam, A. Mol Phys 1986, 57, 89.
44. Allen, M.; Tildesley, D. Computer Simulations of Liquids; Clarendon Press: Oxford, 1991.
45. Aleksandrov, A.; Simonson, T. Biochemistry 2008, 47, 13594.
46. Aleksandrov, A.; Schuldt, L.; Hinrichs, W.; Simonson, T. Biophys J (in press).
47. Mackerell, A. Jr., Atomistic Models and Force Fields. In Computational Biochemistry & Biophysics; Becker, O.; Mackerell, A. Jr.; Roux, B.; Watanabe, M. Eds; Marcel Dekker: New York, 2001; chapter 1.
48. Simonson, T. Curr Opin Struct Biol 2001, 11, 243.
49. Shan, Y.; Seeliger, M. A.; Eastwood, M. P.; Frank, F.; Xu, H.; Jensen, M. O.; Dron, R. O.; Kuriyan, J.; Shaw, D. E. Proc Natl Acad Sci USA 2009, 106, 139.
50. Asaki, T.; Sugiyama, Y.; Hamamoto, T.; Higashioka, M.; Umehara, M.; Naito, H.; Niwa, T. Bioorg Med Chem Lett 2006, 16, 1421.
51. Puttini, M.; Redaelli, S.; Moretti, L.; Brussolo, S.; Gunby, R.; Mologni, L.; Marchesi, E.; Cleris, L.; Donella-Deana, A.; Drueckes, P.; Sala, E.; Lucchini, V.; Kubbutat, M.; Formelli, F.; Zambon, A.; Scapozza, L.; Gambacorti-Passerini, C. Haematologica 2008, 93, 653.
52. Fernandez, A.; Sanguino, A.; Peng, Z.; Crespo, A.; Ozturk, E.; Zhang, X.; Wang, S.; Bornmann, W.; Lopez-Berestein, G. Cancer Res 2007, 67, 4028.
53. Lee, T. S.; Potts, S. J.; Kantarjian, H.; Cortes, J.; Giles, F.; Albitar, M. Cancer 2008, 112, 1744.