Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry
ELSEVIER
Bioorganic & Medicinal Chemistry 12 (2004) 1867–1880

Inhibition studies with rationally designed inhibitors of the human low molecular weight protein tyrosine phosphatase

Adam P. R. Zabell,[a,b] Steven Corden,[b] Paul Helquist,[b] Cynthia V. Stauffacher[a] and Olaf Wiest[b,*]

[a]*Department of Biological Sciences and the Purdue Cancer Center, Purdue University, West Lafayette, Indiana 47907-1392, USA*
[b]*Department of Chemistry and Biochemistry and the Walther Cancer Research Center, University of Notre Dame, Notre Dame, Indiana 46556-5670, USA*

Received 27 January 2003; revised 19 January 2004; accepted 27 January 2004

Abstract—The human low molecular weight protein tyrosine phosphatase (HCPTP) is ubiquitously expressed as two isoforms in a wide range of human cells and may be involved in regulating the metastatic nature of epithelial tumors. A homology model is presented for the HCPTP-B isoform based on known X-ray crystal structures of other low molecular weight PTPs. A comparison of the two isoform structures indicates the possibility of developing isoform-specific inhibitors of HCPTP. Molecular dynamics simulations with CHARMM have been used to study the binding modes of the known adenine effector and phosphate in the active site of both isoforms. This analysis led to the design of the initial lead compound, based on an azaindole ring moiety, which was then also evaluated computationally. A comparison of these simulations indicates the need for a phosphonate group on the indole and provides insight into inhibitor binding modes. Compounds with varying degrees of structural similarity to the azaindole have been synthesized and tested for inhibition with each isoform. These molecular systems were examined with the program AutoDock, and comparisons made with the kinetics and the explicit simulations to validate AutoDock as a screening tool for potential inhibitors. Two compounds were experimentally found to have sub-millimolar inhibition, but the greater solubility of one reinforces the need for experimental testing alongside computational analysis.

1. Introduction

Tyrosine phosphorylation and dephosphorylation are crucial elements in eukaryotic signal transduction.[1–3] Aberrant cell behavior is correlated with persistent elevated levels of tyrosine phosphorylation in a variety of proteins, which is considered one of the primary indicators for oncogenic transformation.[4,5] Several studies[6–8] have indicated phosphatases are directly involved in cellular transformation, where the action of a phosphatase may increase or decrease the enzymatic activity of an oncogene. This increased understanding of the role dephosphorylation plays in regulation of cell behavior has fostered an interest in using the family of protein tyrosine phosphatases (PTPs) as drug targets.[9] Although many PTPs have been studied with regard to their signaling pathway, only PTP1B and the dual specificity phosphatase Cdc25 have been studied as drug targets.[10–15] That work proposed a set of compounds with remarkable diversity, many of which achieve specificity through bidentate binding.[16]

The Eph family of receptor kinases are involved in a variety of epithelial intracellular contact systems, principally during embryonic development.[17–19] Elevated concentrations of Eph receptors have been found in a broad range of metastatic tumors, but at a reduced phosphorylation state relative to nontransformed cells.[20,21] Regulation of EphB1[22] and EphA2[6] has been shown to be associated with the binding of the low molecular weight protein tyrosine phosphatase (LMW-PTP) to these signaling kinases. In the latter study, Kikawa et al. demonstrated that LMW-PTP overexpression is sufficient to transform non-transformed epithelial cells. This in turn suggests that inhibition of LMW-PTP could provide a means to pharmaceutically

Keywords: EphA2; Human LMW-PTP; AutoDock; Inhibitor design.
* Corresponding author. Tel.: +1-574-631-5876; fax: +1-574-631-6652; e-mail: owiest@nd.edu

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2004.01.042

regulate EphA2 phosphorylation levels and thus control oncogenic potential.

The LMW-PTPs have been characterized in a wide range of organisms[23–36] as 18-kDa enzymes with the conserved CX_5R active site motif found in all PTPs.[37] The enzyme exists as two distinct isoforms in humans, HCPTP-A and HCPTP-B,[38] that are identical in sequence outside of the variable spliced region, which extends from residue 40 to residue 73. Crystal structures have been solved for bovine,[39] yeast,[40] and the human A isoform[41] and show an average root mean squared difference of less than 0.5 Å in backbone coordinates. The variable region in the human isoform makes up one edge of the active site and extends down the side of the presumed specificity cleft seen in all known LMW-PTP structures[41] (Fig. 1a). Variable loop residues of particular interest in the A isoform are Glu50 and Asn53, which are changed to Asn50 and Arg53 in the B isoform. These residues are close to or at the lip of the active site and appear to be well suited to provide substrate or inhibitor specificity.

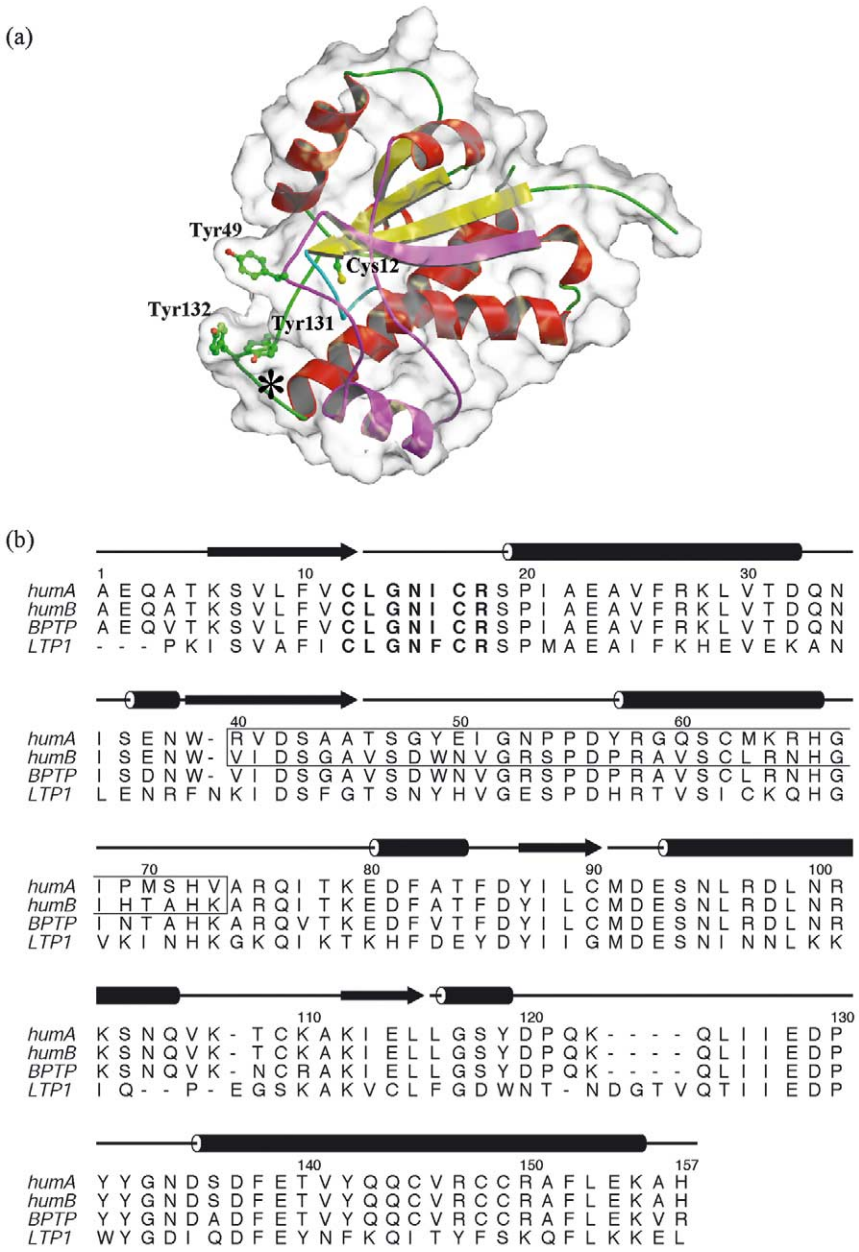
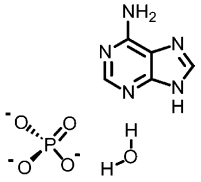


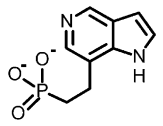
Figure 1. (a) Ribbon drawing of HCPTP-A with a superimposed transparent molecular surface. The Active site loop is colored cyan and the variable region is colored magenta. Active site residue Cys12 and the aromatic residues Tyr49, Tyr131, and Tyr132 at the mouth of the active site are explicitly labeled. An asterisk indicates the location of the Y131 dish. The specificity cleft extends down along the right-hand side of the Y131 dish. This figure and subsequent ribbon figures were drawn with GRASP,[45] Molscript[46] and Raster3-D.[47] (b) Sequence alignment of the two human isoforms of the LMW-PTP (humA and humB) with bovine (BPTP) and yeast (LTP1) as performed by Modeller.[59] The active site loop residues are shown in bold and the variable region for the human isoforms is boxed. Sequence numbering and secondary structure elements come from the HCPTP-A structure. Figure drawn with Alscript.[48]

Previous studies have shown that adenine activates HCPTP-B and the yeast LMW-PTP, but inhibits HCPTP-A.[38,42] Several authors have suggested reasons for the differences in this modulation,[38,42,43] but no complete test of these hypotheses via mutant studies or structural experiments has been reported. The yeast LMW-PTP cocrystalized with adenine[42] provides the only structural evidence to help explain this modulating effect. Bound in the complex with a configuration shown schematically in **1**, the adenine is at the position of the substrate leaving group with a water molecule between the adenine and the phosphate cradled in the active site loop. This water is positioned to act as a nucleophile to release the phosphoenzyme intermediate, hydrolysis of which has been shown to be the rate determining step.[44] While this helps to explain the activation for the yeast and HCPTP-B forms, inhibition of HCPTP-A upon formation of this complex remains unclear.

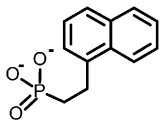


1

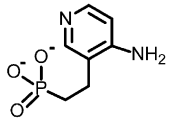
The emerging importance of PTPs in a variety of disease states and the variety of kinetic and structural data for LMW-PTPs has led us to pursue inhibitor design for the human LMW-PTPs with the eventual goal of designing, synthesizing, and evaluating isoform-specific inhibitors. As a guide to our experimental efforts to synthesize HCPTP inhibitors, we use in silico screening of potential ligands. Molecular modeling to design specific inhibitors provides a fast and inexpensive alternative to traditional trial and error drug design approaches.[49–53] Although highly dependent upon the system under investigation, success rates when attempting to reproduce known complexes are reported in the range of 70–80% if a suitable level of approximation is chosen.[53] To validate our rapid screening methodology, we compared results generated from the rigorously developed energy function in CHARMM[54,55] to those from the faster and empirically based energy function in AutoDock.[56] CHARMM can provide detailed information about the structure and dynamics of an enzyme–ligand complex, while AutoDock can quickly generate multiple plausible ligand binding orientations. Molecular dynamics in CHARMM using an explicit water box was performed on both HCPTP isoforms with the bound adenine, inorganic phosphate, and several well-defined water molecules in their initial orientation from the crystal structure for the yeast LMW-PTP. The same calculations were then performed with a phosphonated 5-azaindole (compound **2**), which was designed based on the adenine crystal structure and was expected to mimic those interactions. These systems were then examined using AutoDock, as were several molecules structurally similar to **2** but more synthetically accessible (compounds **3**–**6**). The experimental inhibition of HCPTP by these compounds indicates that the relative ranking predicted by AutoDock is valid provided there are at least three functional groups on the molecule.



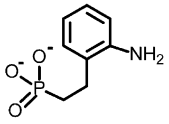
2



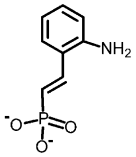
3



4



5



6

2. Computational methods

2.1. Molecular dynamics with compound 2

In order to investigate the interactions of adenine and the proposed inhibitors with both human isoforms, it was necessary to create a homology model for HCPTP-B because there was no experimentally determined structure of this enzyme available. The human A isoform,[41] along with the bovine (BPTP),[57] and yeast (LTP1)[42] enzymes were obtained (PDB[58] codes 5PNT, 1DG9, and 1D2A respectively) and aligned according to their amino acid sequence (Fig. 1b) by the program Modeller 4.[59] The homology model for HCPTP-B as created by Modeller was based on this sequence alignment and visually examined to ensure appropriate placement of the catalytic residues.

2.2. Initial structures

The relative alignment of adenine, inorganic phosphate, and four structurally relevant water molecules in the active site was taken from the crystal structure of adenine bound to LTP1[42] and reproduced with these molecules in the same relative position within the HCPTP-A and HCPTP-B isoforms. Superposition of the enzyme structures was performed with CHARMM28[54,55] using residues 4-39 (HCPTP numbering). This sequence contains the first strand of the β-sheet, the conserved active site loop, and the first α-helix.

The geometry of the phosphonated 5-azaindole was optimized in the neutral charge state with the program Gaussian98[60] using the HF/6-31G(d) method. Hydrogen atoms were removed to model the phosphonate as a dianion, and the charges on the molecule were recalculated in a single-point calculation. The alignment of **2** in the active site was performed in CHARMM from a minimum RMSD overlay of the adenine ring coordinates and corresponding atoms from the azaindole. Energy minimization was then performed in the absence of protein with restraints on the azaindole atoms to

maintain their position. Distance restraints were applied to each atom in the $-PO_3$ group to force them into the corresponding phosphate atom locations.

2.3. Molecular dynamics simulations

Molecular dynamics (MD) simulations were performed in CHARMM using the all-atom parameter set.[55] Protonation states for the catalytic Cys12, the conserved Asp129, and all histidine residues were based on model and pK_a studies.[61–63] The catalytic Cys is the nucleophile that attacks the $-PO_3$ group and becomes the modified residue in the phospho-enzyme intermediate. The conserved Asp donates a hydrogen to the substrate phenolic oxygen, protonating the tyrosyl leaving group. The initial orientation of the Asp129 side-chain was adjusted to permit a hydrogen bond from N9–H9 of adenine to the side-chain carbonyl of the aspartic acid. The partial charges for **2**, the inorganic phosphate, and the anionic Cys12 were constructed based on the values calculated from Gaussian98 and on empirical adjustment to match the values already present in the CHARMM parameter set. Optimized coordinates and partial charges will be made available upon request. Bond, angle, and dihedral forces were similarly determined where necessary. Hydrogen atoms were built using CHARMM onto molecules whose coordinates were determined by X-ray crystallography and the resulting complex minimized while keeping all non-hydrogen atoms at fixed positions.

A 1 ns MD simulation was performed for each isoform with either the adenine, water, and phosphate from the yeast crystal structure (structure **1**) or compound **2** bound in the active site. A truncated octahedron water box with a 9 Å minimal distance between protein and box edge was built for each complex and minimized prior to MD. Dynamics were performed using Vertlet's algorithm[64] with a 2 fs timestep at 298 K. A constant pressure of 1 atm was maintained for the simulation and box volume was monitored to ensure reasonable consistency. For the first 20 ps of MD, a 5 kcal/mol restraint was applied to all backbone non-hydrogen atoms. These initial time-points were not included in the analysis of the simulation.

2.4. Docking simulations

Docking was performed with the program AutoDock3.0.5,[56] applying the Lamarckian genetic algorithm protocol.[65] The simulation space was defined as a box 18.0×20.8×20.8 Å, which includes the active site and all residues of the variable region that line the proposed specificity cleft. The interaction energy between ligand and enzyme was calculated on a 0.2 Å grid using atom affinity potentials, greatly increasing the calculation speed relative to CHARMM. For simulations with inorganic phosphate, the phosphate ion was retained at the base of the active site in the same relative location as seen in the yeast crystal structure and treated as a part of the inflexible protein for the purposes of calculating the grid potential. Polar hydrogen atoms were retained, ring carbons were defined as aromatic, and partial atomic charges were determined by the Gasteiger method[66] with modification to ensure unit charge on each residue. Charge modifications were empirically applied to backbone nitrogens. These changes were rarely more than ± 0.002 charge units except for proline residues, which regularly deviated from unit charge by -0.005. Ligands were allowed full internal flexibility around rotatable (non-ring system) bonds as well as motion within the simulation space. A total of 50 independent simulations with a population size of 50 members were run for each ligand using AutoDock with default parameters. Simulations were ranked according to the docked energy between the protein and the ligand, a summation of internal ligand energy and intermolecular energy terms.

3. Results and discussion

3.1. Molecular dynamics of HCPTP with adenine

The goal of this research is to develop inhibitors that can distinguish between the two HCPTP isoforms. For our initial models, we chose the crystal structure of HCPTP-A which was solved with the buffer molecule 2-(*N*-morpholino)-ethanesulfonic acid in the active site.[41] The buffer molecule roughly approximates the active site position of the phosphate and adenine from the LTP1-adenine structure and allows an initial build of **1** and **2** with the aromatic amino acid side chains of the active site positioned to interact with the effector. Although there is no X-ray or NMR structure of the HCPTP-B isoform currently available, the high structural similarity between the known structures with an average backbone RMSD of 0.5 Å across the secondary structure elements and the extensive sequence identity of 87% between the human isoforms (Fig. 1b) indicated that a homology model of this second isoform should adequately represent the structure. The bovine LMW-PTP, with a sequence identity of 94% to the HCPTP-B isoform, was used to validate plausible side chain orientations in the homology model with satisfactory results. No additional adjustments of the model were performed after checking the structure, and the homology model of HCPTP-B was considered to be a valid structure for all ensuing calculations.

After aligning the active site residues of the yeast complex with those in each human isoform, the adenine, phosphate, and bridging water were placed in the corresponding human enzymes, as shown in Figure 2a. In addition, the three water molecules near the mouth of the active site that form hydrogen bonds between the yeast LMW-PTP and adenine were placed in each human isoform. After hydrating this 'instantly docked' complex and performing energy minimization to alleviate possible bad atom contacts, a 1 ns MD simulation was performed. The total potential energy of the system (Fig. 3a) and the interaction energy between the adenine, phosphate, and bridging water (APW) and the enzyme (Fig. 3b) were examined to determine stability of the system and the complex, respectively. Further measures of complex stability were made by following

characteristic distances between enzyme and APW atoms, or movement of portions of APW during the simulation (Fig. 3c and d). The potential energy of both isoenzyme-inhibitor systems is stable, although the interaction energy between the enzymes and APW is not, and instead shows an increase in energy of 50–75 kcal/mol over the course of the simulation. The distance between the phosphate and the catalytic cysteine remains constant at either 4.0 or 5.5 Å, but the displacement of adenine in both isoforms increases significantly during the simulation to over 20 Å for HCPTP-A and to about 5 Å for HCPTP-B. None of the explicitly placed water molecules remained in their intended hydrogen bonding position (data not shown). This is consistent with a disruption of the interaction between the adenine and the active site residues. A snapshot of the complex taken at 0, 100, 500, and 800 ns during the simulation (Fig. 2b) confirms that both the adenine and the bridging water leave the active site. It is important to note that the phosphate stays in the original location, suggesting that a suitable inhibitor would require covalent linkage between phosphate and adenine mimic.

3.2. Molecular dynamics with inhibitor 2

The results described in the last section, together with the well-ordered position of the APW within the LTP1-adenine complex, suggest the design principle for the lead compound of a new class of inhibitors of HCPTP. The distance between the phosphorus atom and N3 of adenine is 4.0 Å, roughly equal to that of an ethyl group. The use of such a linker would also replace the water molecule thought to be essential for adenine modulation of the enzyme.[44] Retaining the two adenine nitrogens directly involved in hydrogen bonding to the yeast enzyme—the hydrogen bond acceptor atom N1 and the hydrogen bond donor atom N9—and

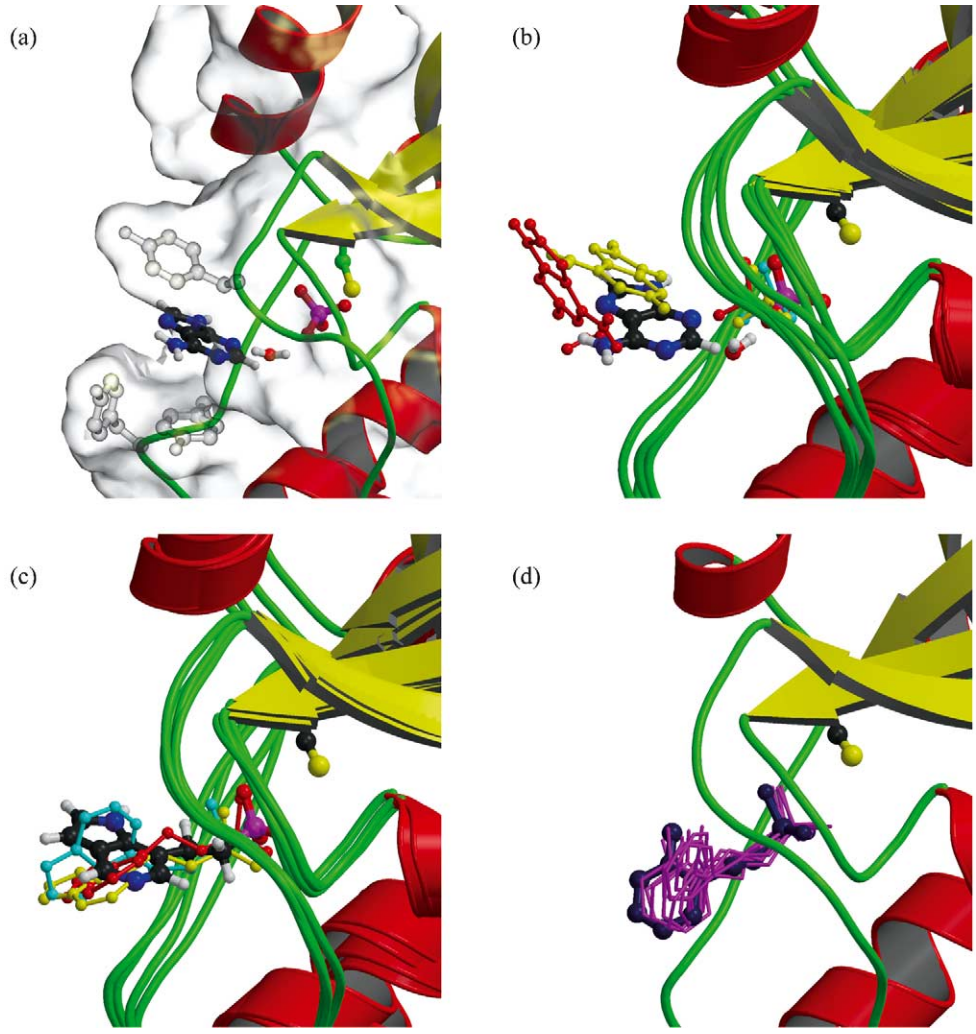


Figure 2. (a) Ribbon drawing of HCPTP-A expanded around the active site with a transparent molecular surface. Residues Cys12, Tyr49, Tyr131, and Tyr132 are explicitly shown. The adenine, phosphate, and bridging water molecules are in the same relative positions in this structure as in the LTP1-adenine crystal structure.[42] The hydrophobic packing between the adenine and the aromatic residues can be easily recognized. (b) Snapshots of the MD with HCPTP-A and the APW complex at 0 (standard atom colors), 100 (red), 500 (yellow), and 800 (cyan) ps. The catalytic Cys12 is shown as a reference for the phosphate motion. Adenine at 800 ps and the water at 500 and 800 ps have left the vicinity of the active site and are not visible from this perspective. (c) Snapshots of the MD with HCPTP-A and **2** at 0 (standard atom colors), 100 (red), 500 (yellow), and 800 (cyan) ps. (d) Top 10 AutoDock structures of **4** docked with HCPTP-A. The structure with the best docked energy is shown as a blue ball and stick model, while the remaining nine structures are drawn in magenta lines.

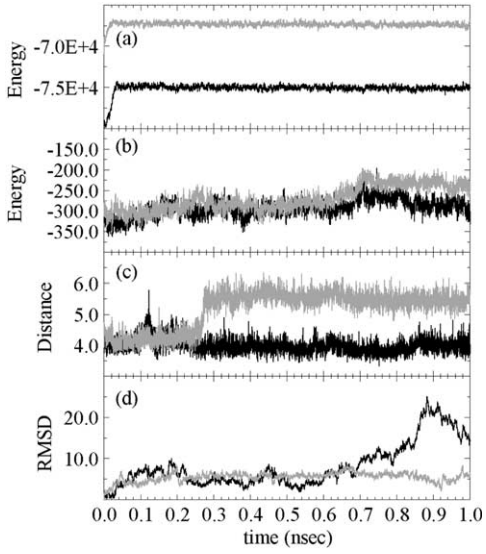


Figure 3. Monitored values for the CHARMM dynamics simulations of HCPTP-A (black) and HCPTP-B (grey) with the adenine, phosphate, and bridging water (APW). (a) Total potential energy of the system, in kcal/mol. (b) Interaction energy between enzyme and the APW, in kcal/mol. (c) Distance in Angstroms between the phosphorus atom of the phosphate and Sγ of Cys12. (d) Distance RMSD between the adenine position at each timepoint and the adenine position at the start of the MD, in Angstroms.

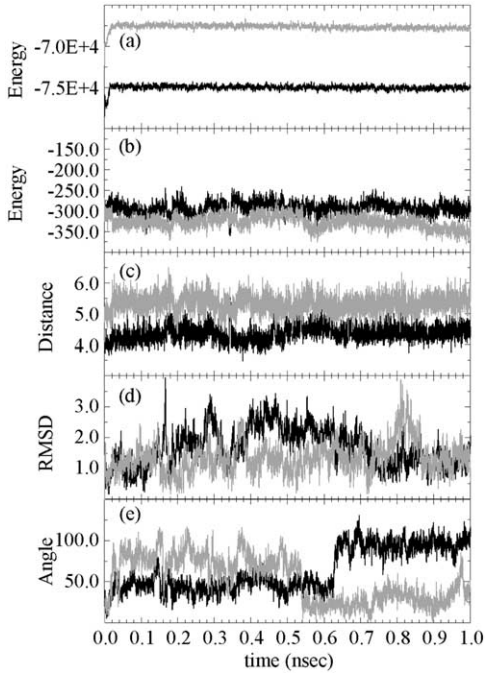


Figure 4. Monitored values for the CHARMM dynamics simulations of HCPTP-A (black) and HCPTP-B (grey) with compound **2**. (a) Total potential energy of the system, in kcal/mol. (b) Interaction energy between enzyme and **2**, in kcal/mol. (c) Distance in Angstroms between the phosphorus atom of the phosphonate and Sγ of Cys12. (d) Distance RMSD between the azaindole position at each timepoint and the azaindole position at the start of the MD, in Angstroms. (e) Relative rotation of the azaindole ring as part of the RMSD calculation of (d).

modifying the phosphate to a phosphonate to improve hydrolytic stability and reduce biological degradation, we proposed that the 5-azaindole **2** would function as an effective inhibitor for LMW-PTPs.

The validity of this design concept was tested by two MD simulations, which differ only in that the APW is replaced with **2**. The total potential energy and interaction energy between the compound and each of the two isoenzymes were monitored as before, with slightly different distance measurements being taken to correspond with the use of a single molecule in the active site. The potential energy of the system is again stable (Fig. 4a), but it is gratifying to note that in contrast to the APW system, the interaction energy is constant in the complex with HCPTP-A and slightly improves with HCPTP-B (Fig. 4b), indicating that **2** is a reasonable inhibitor. The distance measurements lend additional support to the hypothesis of effective inhibition, with small fluctuations between the phosphorus in the phosphonate group and Cys12Sγ of approximately 0.5 Å (Fig. 4c). This is the same degree of motion seen with the phosphate in the simulations with **1**. It is interesting to note that the azaindole ring is not involved in close hydrogen bonding either as an acceptor to residue 50 or a donor to residue 129 (data not shown). Without specific atomic interactions between **2** and HCPTP, the azaindole undergoes a 1–3 Å displacement over the course of the simulation (Fig. 4d). At the same time, the azaindole is limited to a rotation between 30° and 90° (Fig. 4e). This indicates that the shape of **2** is more important than the distribution of hydrogen bond donors and acceptors on the aromatic ring. Consequently, a further simplification of the system for the purpose of an easy experimental proof-of-principle of the design concept should be possible.

Analysis of the geometry in the simulations also provides insight into the nature of inhibitor binding. Figure 2c shows snapshots of **2** in the active site at the same time-points used in Figure 2b. As expected from the trajectories shown in Figure 4, **2** remains tightly bound in the active site. While the phosphonate and the ethyl linker show very similar positioning, the aromatic azaindole ring exhibits torsional flexibility. Along with the distance and energy graphs, these snapshots highlight the three conclusions from our computer-based inhibitor design for HCPTP. Most importantly, the covalent linkage of a phosphate or phosphate mimic with an aromatic system should ensure stronger binding than adenine. Second, hydrogen bond donors or acceptors appeared to make no difference in binding affinity for the inhibitor atoms at the top of the active site. Third, the inhibitor moiety at the top of the active site should have a planar shape and preferably be aromatic. Additional support for these points is the aromatic nature of the natural substrate as well as the roughly planar or aromatic nature of many inhibitors, among them buffer molecules such as 2-(*N*-morpholino)-ethanesulfonic acid (MES) or *N*-2-Hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES), and many of the inhibitors designed for the similarly shaped active site of PTP1B.[10,16]

3.3. Docking of adenine and azaindole to each isoform

The results of the MD simulations show that **2** should effectively bind to either isoform, but a variety of similar

inhibitor structures are possible. The computational requirements of the full 1 ns MD simulation are prohibitive for use as a screening tool. If it could be validated by comparison with the results from the explicit MD simulation, a simplified energy function could be used to increase the speed of analysis for a larger series of compounds. We selected AutoDock for its ability to determine ligand structures with internal flexibility and the improved speed of calculation afforded by pre-calculating a grid of interaction between each atom type and the target molecule. Several recent publications[73–76] have used this program to investigate problems of similar size and scope to ours, with good success.

Potential energy grids were calculated for each HCPTP isoform to represent the enzyme with or without phosphate in the active site. Although phosphate is a weak inhibitor of LMW-PTPs, the yeast crystal structure with adenine clearly indicated the presence of a phosphate coordinated at the base of the active site, so phosphate may play a role in modulating the orientation or activity of adenine. We proposed that AutoDock calculations with and without phosphate should indicate the degree to which it affects binding of adenine. Two major binding sites for adenine were found, one in the active site and the other in a small depression near the active site on the opposite side of aromatic residues Tyr131 and Tyr132 (asterisk, Fig. 1a). This depression, termed the Y131 dish, is formed by aromatic residues and contains a negative charge from a partially exposed Asp56 in the HCPTP-A isoform and both a positive and a negative charge from Asp56 and Arg58 in the HCPTP-B isoform. The Y131 dish can be considered a wide extension of the specificity cleft that extends from the binding site and may act as an initial binding location for adenine from bulk solvent. The relative ranking of each cluster by docked energy, position of binding for that cluster, number of members, and estimated K_i is given in Table 1. It is noteworthy that the lowest energy structure for each isoform when phosphate is present was with adenine bound in the Y131 dish. Of the structures that show adenine bound in the active site, the presence of phosphate promoted a binding in the same relative position as that seen in the LTP1-adenine structure, although the adenine would often be rotated around either the N1-C2 or the C5-C6 bond. The absence of phosphate permitted the adenine to bind deeper in the active site with either the N9 or the amine group positioned in the location where the phosphate was found. One other site with poor interaction energy was infrequently found, where the adenine binds in the specificity cleft, well away from the active site. Taken together, these results suggest that activity modulation of the HCPTP isoforms by adenine may involve multiple binding sites.

The calculated inhibition constants for the adenine complexes are in the μM range and generally 1 to 2 orders of magnitude weaker when phosphate is present. The calculations also predict that adenine is a somewhat better inhibitor for HCPTP-A in the presence of phosphate, but is a better inhibitor for HCPTP-B when phosphate is absent. However, these calculations presume that when the phosphate is present, it is tightly bound to the enzyme. The K_i for phosphate has been experimentally determined to be in the low mM range[57,77] and indicates the phosphate is neither tightly nor constantly bound. Furthermore, experimental data has shown that while adenine inhibits the HCPTP-A isoform, it acts as an activator for the HCPTP-B isoform. Presumably this activation involves a mechanism other than simple binding in the active site of the apoenzyme.[42] Assuming that the calculated inhibition constant, K_i, is equivalent to the dissociation constant, K_D, the results of the AutoDock calculations consistently suggest a stronger binding of adenine to either

Table 1. Relative ranking of AutoDock simulations with adenine and HCPTP-A or HCPTP-B

Rank [a]	# in cluster	Binding site[b]	Docked energy, kcal/mol [c]	K_i, μM [d]
HCPTP-A and Adenine, with PO_4				
1	5	Y131 dish	−6.55	15.9
2	16	Active site	−5.36	119.
3	2	Y131 dish	−5.01	212.
4	1	Active site	−4.94	238.
5	12	Active site	−4.81	299.
6	7	Active site	−4.71	352.
7	2	Active site	−4.71	356.
8	5	binding cleft	−4.09	997.
HCPTP-A and Adenine, without PO_4				
1	19	Active site	−7.24	4.95
2	6	Active site	−7.06	6.65
3	1	Active site	−6.96	7.96
4	1	Active site	−6.88	9.03
5	15	Active site	−6.83	9.87
6	3	Active site	−6.79	10.5
7	1	Active site	−6.67	12.8
8	3	Active site	−6.64	13.7
9	1	Y131 dish	−5.89	48.3
HCPTP-B and Adenine, with PO_4				
1	2	Y131 dish	−5.39	113.0
2	8	Active site	−5.26	139.0
3	3	Y131 dish	−5.25	142.0
4	5	Y131 dish	−5.24	143.0
5	4	Active site	−4.98	223.0
6	2	Y131 dish	−4.92	246.0
7	21	Active site	−4.92	249.0
8	1	Y131 dish	−4.85	277.0
9	1	Y131 dish	−4.74	338.0
10	1	Y131 dish	−4.61	414.0
11	1	Active site	−4.36	642.0
12	1	binding cleft	−3.84	N.B.[e]
HCPTP-B and Adenine, without PO_4				
1	27	Active site	−7.84	1.79
2	2	Active site	−7.68	2.36
3	2	Active site	−7.43	3.56
4	16	Active site	−7.39	3.80
5	1	Active site	−6.99	7.50
6	1	Active site	−6.96	7.97
7	1	Active site	−6.74	11.4

[a] Rank determined by final docked energy.
[b] Placement of cluster in the active site, the Y131 dish, or along the binding cleft.
[c] Docked potential energy of the ligand, from the best complex in the cluster.
[d] Estimated inhibition constant at 298.15 K.
[e] Not bound. The K_I was calculated to be 0.0.

enzyme by 2 or 3 orders of magnitude compared to experimental data. The lowest energy docked structure of adenine in the absence of phosphate was calculated to be 4.95 μM and 1.79 μM, respectively, for HCPTP-A and HCPTP-B. The reported dissociation constant is 2.1 mM and 0.3 mM, respectively.[38] More important than the absolute binding constants is the finding that the relative order of dissociation remains consistent, with adenine acting as a weaker effector of HCPTP-A. We are therefore confident that this methodology can be used as a rapid screening tool, indicating binding constants relative to adenine, and is able to distinguish any preference for a given isoform.

The calculations with **2** were performed only in the absence of phosphate, and the flexibility of the ethylphosphonate chain led to a dramatic increase in the number of independent structures. This apparent diversity in structures, however, comes from a variety of nearly identical structures where the indole is rotated around the N1-C2 bond in a fashion similar to that seen with the adenine complexes. The best 20–30 structures of **2** in the HCPTP isoforms have a docking energy within 2 kcal/mol of each other. The 10 structures with the lowest energy showed each docked complex bound with the phosphonate superimposed on the phosphate position in the LTP1-adenine complex, and the azaindole located almost exactly in the position of the adenine. Consistent with the results from the explicit MD simulations, the predicted hydrogen bonding with the acceptor atom N2 or donor atom N9 was obtained in a limited number of structures. In the HCPTP-A isoform, for example, only one structure has N9 positioned in rough alignment to permit hydrogen bonding to Asp129 while the remaining structures rotate to position N9 near either Tyr131 or Tyr49. As we might predict, the azaindole positioned to form the hydrogen bond to Asp129 does exhibit the best docking energy.

Table 2. Relative ranking of AutoDock simulations with **2** HCPTP-A and HCPTP-B [a]

Rank	Docked Energy, kcal/mol	Binding Energy, kcal/mol [b]	K_i, μM
	HCPTP-A		
1	−10.49	−9.3	0.153
2	−10.45	−9.4	0.128
3	−10.41	−9.29	0.154
4	−10.35	−9.16	0.194
5	−10.30	−9.16	0.193
6	−10.04	−8.87	0.313
7	−10.00	−8.91	0.294
8	−10.00	−8.80	0.357
9	−9.99	−9.03	0.242
10	−9.94	−8.74	0.393
	HCPTP-B		
1	−11.87	−10.69	0.0147
2	−11.78	−10.69	0.0146
3	−11.72	−10.55	0.0185
4	−11.60	−10.40	0.0238
5	−11.57	−10.54	0.0187
6	−11.52	−10.57	0.0178
7	−11.45	−10.37	0.0252
8	−11.36	−10.19	0.0337
9	−11.34	−10.21	0.0329
10	−11.21	−10.18	0.0346

[a] Terms are the same as in Table 1.
[b] Free energy of ligand binding, which incorporates ligand flexibility.

These 10 structures of **2** demonstrated a much stronger binding to the HCPTP isoforms when compared with adenine (Table 2), with a predicted binding energy of approximately −10 kcal/mol. The estimated inhibition constant for the best inhibitor complex is also significantly lower, with a K_i of 0.15 μM for HCPTP-A and 0.015 μM for HCPTP-B. It should be recalled that the estimated K_i for adenine is some 2 or 3 orders of magnitude stronger than experimentally determined, so these values may also be exaggerated.

3.4. Docking of molecular probes for binding interactions

Our docking calculations indicate that **2** can be further simplified while maintaining the key elements of our lead concept. In order to partition out the role of the binding interactions for individual optimization, it would be of interest to study the interactions one at a time. Although the smaller number of interactions in the probes will of course lead to lower inhibition constants, the information provided by this study is very valuable for the longer-term development of more active and selective HCPTP inhibitors. This has for example been demonstrated by the popular 'SAR by NMR' technique, where binding of small molecular probes is used to elucidate binding interactions in the active site, which are then exploited by more complex molecules incorporating these interactions. Four potential inhibitors **3**–**6** were designed to probe the individual importance of those characteristics the simulations predicted to be important for a potent inhibitor: size of the aromatic group, necessity for secondary hydrogen bonding interactions, and flexibility of the phosphonate linker. For example, the heterocycle of **2** is replaced by a naphthalene group in **3**, retaining the size and aromaticity of the initial target while removing the hydrogen bonding donor and acceptor. The relative importance of the hydrogen bond donor and acceptor atoms is measured by reducing the size of the heterocycle to a pyridine in **4** and a benzene in **5**, while an amino group in the ortho position to the phosphonate retains the proposed

Table 3. Calculated and Experimental results with compounds **3**–**6** [a]

Compd	Docked Energy, kcal/mol	Binding Energy, kcal/mol	Calculated K_i, μM	Experimental K_i, mM
	HCPTP-A			
3	−9.88	−8.81	0.346	14.0
4	−10.03	−8.51	0.581	0.30
5	−9.86	−8.43	0.662	0.25
6	−9.28	−8.13	1.10	5.30
	HCPTP-B			
3	−11.60	−10.48	0.0209	10.0
4	−11.45	−9.86	0.0594	2.21
5	−11.46	−9.91	0.0541	1.40
6	−10.61	−9.55	0.100	13.0

[a] Terms are the same as in Table 2.

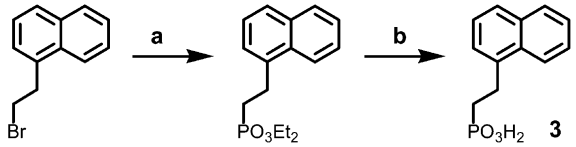
hydrogen bonding to Asp129. In **6**, the ethyl group is modified to an ethylene in order to examine the importance of flexibility in the phosphonated linker. Computational and experimental studies of **3–6** will therefore provide insights into the effect of the different elements of the lead concept.

The estimated docking energy, binding energy, and K_i for the best complex of **3–6** with each HCPTP isoform are shown in Table 3. All four compounds are predicted to have lower binding energies and to be poorer inhibitors than **2**. Compound **3** should be a better inhibitor than the systems with a single aromatic ring, with a predicted K_i of 0.346 μM and 0.021 μM for HCPTP-A and HCPTP-B, respectively, reemphasizing the importance of shape. The acceptor nitrogen of **4** is predicted to have a negligible effect on inhibition, with K_i values essentially equal to those **5** of 0.581 μM and 0.059 μM for either isoform. A further two-fold loss of inhibition is predicted with the loss of flexibility in the phosphonate chain of **6** as the ethyl group is modified to an ethylene. Figure 2d shows the relative orientation of the ten lowest energy structures of **4** bound to HCPTP-A, and is representative of all eight enzyme-inhibitor calculations. In all cases, the ten lowest energy complexes have the phosphonate at the base of the active site and the ring system positioned roughly between the aromatic residues at the mouth of the active site. As was the case with atom N9 of the azaindole, the amino group in **4–6** is typically not aligned to interact with Asp129. Other trends remain the same as those seen in the simulations with **2**, where the predominant difference between docked positions involves the rotation of the ring system relative to that seen in the crystal structure of the yeast form with adenine.

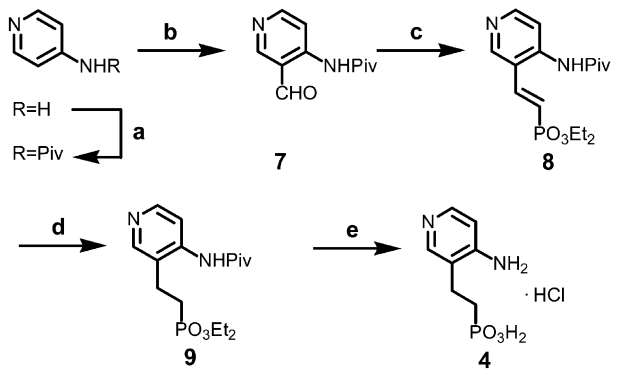
3.5. Synthetic studies

Based on the results for adenine, it can be expected that the real absolute binding constants are 2–3 orders of magnitude lower than the ones predicted by AutoDock. To determine the absolute and relative binding constants for the molecular probes **3–6**, these compounds were synthesized and the inhibition constants for HCPTP-A and HCPTP-B determined. 2-(1-Naphthyl)-ethylphosphonic acid **3** was obtained in two steps as shown in Scheme 1. The Arbusov reaction of commercially available 1-(2-bromoethyl)naphthylene provides the diethyl phosphonate ester, and subsequent deprotection using trimethylsilyl bromide gives the free phosphonic acid **3**.[67]

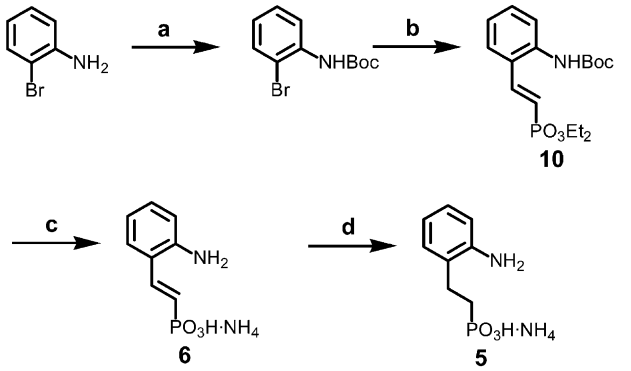
Compound **4** was prepared through regiospecific ortho-metalation of 4-(pivaloylamino)-pyridine and subsequent reaction with dimethylformamide to provide *N*-(3-formyl-4-pyridyl)-2,2-dimethylpropionamide **7** (Scheme 2).[68] The Horner–Wadsworth–Emmons–Wittig reaction of tetraethyl methylenebisphosphonate with this aldehyde carried out in an aqueous two-phase system afforded solely the *E*-isomer **8**.[79] Hydrogenation of the double bond over palladium on carbon provided compound **9**. Simultaneous hydrolysis of the amide and phosphonate esters was achieved by refluxing for 18 h in 6 N hydrochloric acid providing **4**.



Scheme 1. (a) $P(OEt)_3$, 50 °C (90%); (b) TMSBr, CH_3CN, rt (30%).



Scheme 2. (a) Pivaloylchloride, TEA, DCM (63%); (b) BuLi, THF, −78 °C to 0 °C, then DMF at −78 °C, (61%); (c) tetraethyl methylenediphosphonate, 50% NaOH, DCM, (95%); (d) H_2, Pd/C, MeOH (97%); (e) 6N HCl, reflux (54%).



Scheme 3. (a) NaH, THF, $[t\text{-BuOC(O)}]_2O$, reflux (93%); (b) $Pd(PPh_3)Cl_2$, diethyl vinylphosphate, TEA, DMF, 100 °C (41%); (c) TMSBr, CH_3CN, rt (72%); (d) H_2, Pd/C, MeOH (18%).

The anilides **5** and **6** were synthesized from commercially available 2-bromoaniline (Scheme 3). Protection of the aniline with the *t*-butoxycarbonyl (Boc) group,[69] followed by palladium catalyzed olefination of the aryl bromide (Heck reaction)[78] with diethyl vinylphosphate gave the corresponding *E*-2-(2-aminophenyl)vinyl phosphonate **10**. Simultaneous deprotection of the Boc group and hydrolysis of the phosphonate esters was accomplished with trimethylsilyl bromide. Following each hydrolysis, the fully deprotected material was further purified on Dowex 50 ion exchange resin giving **6** as the ammonium salt. Subsequent hydrogenation with palladium on carbon afforded saturated side chain derivative **5**.

3.5.1. Enzyme inhibition of 3–6. The results from the AutoDock calculations with compounds **3–6** were tested by direct comparison with experimentally determined K_i values from a standard phosphatase assay using pNPP

as the substrate. Results from inhibitors **4–6** follow the calculated results relatively well. Compound **5** was found to be a slightly better overall inhibitor than **4**, with a K_i of 0.25 mM and 1.4 mM for HCPTP-A and HCPTP-B, respectively, compared to 0.30 mM and 2.2 mM with the pyridine compound. As predicted, the addition of a hydrogen bonding donor atom in the aromatic ring of **4** does not have a significant effect on enzyme inhibition. The predicted μM to nM inhibition of these compounds would have made them excellent inhibitors, but the observed mM inhibition is more consistent with results for a lead compound. As expected based on the calculated and experimental K_i for adenine, the experimentally determined inhibition constants for **3–6** were orders of magnitude poorer than predicted by AutoDock (Table 3). Compound **4** is much more soluble than **5**, a necessary consideration for effective drug design that AutoDock cannot take into account. One surprise in these experiments involved compound **6**, designed with the expectation that a rigidly positioned distance between the phosphonate and the amine would improve the binding at both the phosphate binding loop and the Asp129 side chain. However, **6** was roughly 10-fold poorer than either **4** or **5**, with a K_i of 5.3 and 13.0 mM for HCPTP-A and HCPTP-B. It is gratifying to note that the relative scale of experimental K_i values between the pyridine and anilide molecules is essentially the same as the prediction, demonstrating that the ethyl group flexibility in **4** and **5** is a critical component of inhibition. It is also interesting to note that even these very simple compounds begin to show selectivity of up to a factor of 7.5 for differential binding to one isoform over the other. A detailed analysis of the structural origin of this selectivity through MD simulation will help in the design of more specific inhibitors of HCPTP.

The unexpected result of the experimental studies was the relatively low activity of **3**, with a K_i of 14.0 mM and 10.0 mM for HCPTP-A and HCPTP-B, respectively, which is in contradiction to the docking results. Although the relatively crude energy function of AutoDock could of course be one reason, it is noteworthy that the secondary plots of $1/V_{max,app}$ and $K_{M,app}/V_{max,app}$ indicate that **3** is a non-competitive inhibitor (not shown). As a non-competitive inhibitor, **3** must be able to bind to the enzyme in the presence and absence of substrate. This implies that a secondary binding site, which was not part of the original design strategy, can bind **3**. It is therefore an instructive example of the breakdown of the model when alternative binding sites interfere with inhibition. This is always a possibility that needs to be evaluated by experimental methods, emphasizing the strengths of the combined computational and experimental approach used here. Finally, the results also establish the importance of hydrogen bonding with the amino function at the mouth of the active site. This can be accomplished by the inclusion of the amino group in compounds **4–6**, which adds a functional group not present in compound **3**, suggesting that a minimum of three interactions are needed for AutoDock to successfully predict the relative levels of inhibition.

4. Conclusions

The CHARMM molecular dynamics and AutoDock docking simulations for **1** and **2** revealed a number of elements necessary for the successful design of inhibitors for the HCPTP isoforms. First, the MD shows that the phosphate ion remains essentially unchanged in its bound location in the active site loop, while adenine is unable to maintain its position in the active site with either isoform of HCPTP. By covalently linking the features of phosphate and adenine in the azaindole compound **2**, the molecule remained bound to the enzyme, supporting this scaffold as a basis for inhibitor design. Second, the MD results with **2** suggested three points of design strategy of a lead compound: a planar shape for the inhibitor at the mouth of the active site, indifference for hydrogen bonds between that moiety and enzyme, and the requirement of the phosphonate to provide the principal binding interaction between inhibitor and enzyme. Third, the correlation of the results from AutoDock and CHARMM suggested that a faster virtual screening of potential compounds with the simple AutoDock scoring function is possible before committing to detailed analysis using extended MD simulations. It is interesting to note that MD showed the azaindole ring of **2** moving away from the initial adenine position while all of the AutoDock solutions show a greater superposition of azaindole on the adenine. This discrepancy is possibly an artifact of the differences between the explicit energy calculation and the implicit, parameterized scoring function. The MD simulation was performed with explicit solvent and permitted free motion of the enzyme, allowing both the indole and the aromatic residues around the active site to move to favorable positions. The docking calculations, on the other hand, maintain a rigid enzyme structure and use an energy function empirically parameterized for favorable binding modes.

Separation of the design elements led to the even further simplified compounds **3–6**, which were studied computationally using AutoDock and experimentally through synthesis and enzyme kinetics. Three of these four compounds show good correlation between computed and experimental binding values with the computed binding energies being consistently overestimated by ~4 kcal/mol, in agreement with the correlation obtained for the well known effector adenine. Of the compounds studied, we found that compound **3** showed poor correlation, most likely because of a different binding mode. This demonstrates that shape complementarity and hydrophobic interactions alone are not sufficient for efficient inhibition and that appropriately positioned functional groups capable of hydrogen bonding are necessary. The initial results for compounds **4** and **5** are also promising in that they start to show a small specificity for HCPTP-A, which will be further explored in future studies.

In summary, we have used computer aided rational design to develop new and structurally simple inhibitors of biologically relevant low molecular weight phosphatases. The relative importance of shape, hydrogen bonding, and charge interactions in the active site were

elucidated and demonstrated separately in a series of simple inhibitors. Although the binding constants of these simplified compounds with only two or three interactions is of course low, the thorough understanding obtained will enable us to combine the individual design elements into more complex, more potent inhibitors. This process is helped by the finding that the binding constants of the AutoDock scoring function correlates reasonably well with the relative binding abilities of the majority of inhibitors. These studies are currently in progress and the results will be reported in due course.

5. Experimental

All reagents were purchased from commercial suppliers and were used without further purification. Melting points were measured on a MEL-TEMP II device without correction. ^{1}H NMR (300 MHz) and ^{13}C NMR (75 MHz) were recorded on a Varian Inova-300 spectrometer. All spectra were recorded in $CDCl_3$, DMSO-d_6 or CD_3OD, and chemical shifts are given relative to TMS. Mass spectrometry was performed with a JEOL JMS-AX505HA double sector mass spectrometer. THF was purified by an Innovative Technologies solvent purification system. CH_2Cl_2 and CH_3CN were both distilled from CaH. Column chromatography was performed with EM Science 60 Å silica gel (230–400 mesh). The DOWEX® 50WX4-400 ion-exchange resin was purchased from Aldrich. The Superdex 75 prep grade gel filtration media and HiPrep 16/10 SP cation exchange column were purchased from Pharmacia.

5.1. Chemical synthesis

5.1.1. 2-(1-Napthyl)ethanephosphonic acid (3). A mixture of 1-(2-bromoethyl)naphthalene (900 mg, 3.8 mmol; Aldrich) and triethyl phosphite (2.6 mL, 15 mmol) was heated to 145–150 °C for 3 h. Excess triethyl phosphite was removed by evaporation under vacuum to give a brown oil (1.0 g, 3.4 mmol, 90%). To this crude oil was added anhydrous acetonitrile (30 mL) and trimethlysilyl bromide (9.0 mL, 68 mmol), and the mixture heated to 60–65 °C for 3 h. All volatiles were removed by evaporation under vacuum, and dichloromethane (5 mL) was added to the resulting oil. The formed precipitate was filtered and recrystallized from ethyl acetate to give white crystals (271 mg, 30%): mp 165–167 °C (lit mp[67] 166–167 °C); ^{1}H NMR (DMSO) δ 1.80–1.94 (2H, m, CH_2P), 3.15–3.25 (2H, m, $ArCH_2$), 7.34–7.59 (4H, m, ArH), 7.73–8.01 (3H, m, ArH); ^{13}C NMR (DMSO) δ 26.92, 28.95 + 30.72 (d), 123.88–131.66 (multiple peaks-undeterminable), 134.18, 138.43 + 138.67 (d); MS (+veFAB) (M + H) 237; HRMS calcd for $C_{12}H_{13}O_3P$ (M + H) 237.0681, found 237.0711.

5.1.2. Diethyl (*E*)-2-{4-[(2,2-dimethylpropanoyl)amino]-pyridin-3-yl}ethenephosphonate (8). To a solution of *N*-(3-formylpyridin-4-yl)-2,2-dimethylpropionamide **7** (485 mg, 2.4 mmol, prepared by the lit.[68] method) and tetraethyl methylenebisphosphonate (0.58 mL, 2.4 mmol) in dichloromethane (3.5 mL) was added sodium hydroxide solution (3.5 mL, 50% in water) and stirred for 1 min, after which a suspension formed. Water (4 mL) was then added, and the organics were extracted with ethyl acetate. The combined organic layers were dried over magnesium sulfate, filtered and the solvent evaporated under vacuum to give the title compound as a yellow oil which slowly solidifies (785 mg, 95%) and was used without further purification: mp: 75–78 °C; ^{1}H NMR (DMSO) δ 1.17–1.25 (15H, m, Piv + $CH_2C\underline{H}_3$), 3.94–4.05 (4H, m, J = 7.4 Hz, $C\underline{H}_2CH_3$), 6.55–6.72 (1H, apparent t, J = 18.0 Hz, PCH_2), 7.21–7.40 (2H, m, Ar-CH + ArH), 8.43–8.48 (1H, d, J = 5.4, ArH), 8.87 (1H, s, ArH), 9.51–9.58 (1H, bs, NH); ^{13}C NMR (MeOH) δ 15.55 + 15.63 (d), 26.53, 39.56, 62.49 + 62.57 (d), 115.34 + 117.85 (d), 120.67, 126.56 + 126.86 (d), 142.16 + 142.25 (d), 144.73, 148.15, 150.29, 178.88; MS (+veFAB) (M + H) 341; HRMS calcd for $C_{16}H_{25}N_2O_4P$ (M + H), 341.1630, found 341.1630.

5.1.3. Diethyl 2-{4-[(2,2-dimethylpropanoyl)amino]pyridin-3-yl}ethanephosphonate (9). A solution of **8** (600 mg, 1.8 mmol) and palladium on carbon (60 mg) in methanol (30 mL) was stirred under an atmosphere of hydrogen for 18 h. The solution was filtered through Celite which was washed further with methanol, and the solvent was evaporated under vacuum to give a grey oil (585mg, 95%) that was used without further purification: ^{1}H NMR ($CDCl_3$) δ 1.15–1.22 (6H, t, J = 7.1, $CH_2C\underline{H}_3$), 1.37 (9H, s, Piv), 2.03–2.15 (2H, dt, J = 17.6, 7.2, CH_2P), 2.85–2.99 (2H, dt, J = 20.1, 7.1, $ArCH_2$), 3.89–4.01 (4H, m, $C\underline{H}_2CH_3$), 7.85–7.89 (1H, d, J = 5.4, ArH), 8.35–8.40 (2H, m, ArH), 8.82 (1H, br s, exchanges with D_2O, NH); ^{13}C NMR ($CDCl_3$) δ 16.41, 21.67 + 25.60 (d), 27.46 + 27.63 (d), 27.80, 40.27, 62.06 + 62.15 (d), 118.23, 126.47 + 126.57 (d), 143.79, 148.85, 151.52, 178.17; MS (+veFAB) (M + H) 343; HRMS calcd for $C_{16}H_{27}N_2O_4P$ (M + H) 343.1787, found 343.1777.

5.1.4. 4-Amino-3-(2-phosphonoethyl)pyridinium chloride (4). A mixture of **9** (300 mg, 0.88 mmol) in hydrochloric acid (4 mL, 6 N solution in water) was refluxed for 18 h after which the water was evaporated under vacuum to give a tan colored solid. The solid was heated in ethanol (1 mL) and the resulting suspension was centrifuged for 10 min. Ethanol was decanted off leaving a fine white powder (113 mg, 54%): ^{1}H NMR (DMSO) δ 1.78–1.90 (2H, dt, J = 18.0, 7.4, P-CH_2), 2.64–2.76 (2H, m, $ArCH_2$), 6.80–6.85 (1H, d, J = 6.9, ArH), 8.02–8.07 (2H, m, ArH); ^{13}C NMR (CD_3OD) δ 21.63, 23.48 + 25.33 (d), 109.09, 121.08, 137.98, 138.47, 159.06; MS (+veFAB) (M + H) 203; HRMS calcd for $C_7H_{11}N_2O_3P$ (M + H) 203.0586, found 203.0587.

5.1.5. Diethyl (*E*)-2-{2-[(*tert*-butoxycarbonyl)amino]phenyl}ethenephosphonate (10). A solution of diethyl vinylphosphate (2.11 mL, 13.5 mmol), 2-bromo-*N*-(*tert*-butoxycarbonyl)aniline (1.84 g, 6.8 mmol, prepared by the lit.[69] method), bis(triphenylphosphine)palladium II chloride (0.487 g, 0.68 mmol), in dimethylformamide (5 mL, degassed) and triethylamine (1.93 mL, 14 mmol) was heated to 90–100 °C for 18–20 h. The solvent was removed by evaporation under vacuum, and the residue was taken up in ethylacetate and water. The precipitated triphenylphosphine was removed by filtration, and the

organic layer was separated from the filtrate solution and extracted further with ethylacetate. The combined organic extracts were dried over magnesium sulfate, and the solvent was removed by evaporation under vacuum. The residue was then subjected to bulb-to-bulb distillation at 90–100 °C at 1mmHg to remove excess diethyl vinylphosphate. The title compound was then further purified by flash chromatography to give a brown solid (992 mg, 41%): mp: 96–98 °C; ^{1}H NMR ($CDCl_3$) δ 1.32–1.39 (6H, t, J = 7.3, $CH_2C\underline{H}_3$), 1.51 (9H, s, *t*Bu), 4.08–4.20 (4H, m, $C\underline{H}_2CH_3$), 6.15–6.29 (1H, apparent t, J = 17.7 Hz, PCH), 6.50–6.64 (1H, br s, NH), 7.08–7.15 (1H, dt, J = 0.6, 7.5, ArH), 7.32–7.39 (1H, dt, J = 1.5, 7.8, ArH), 7.46–7.51 (1H, dd, J = 1.5, 8.1, ArH), 7.58–7.79 (2H, m, ArH + Ar-CH); ^{13}C NMR ($CDCl_3$) δ 16.62 + 16.70 (d), 28.53, 26.17 + 26.24 (d), 81.22, 115.40 + 117.92 (d), 123.43, 124.66, 127.07, 131.03, 135.70, 136.59, 143.83 + 143.92 (d), 153.35; MS (+veFAB) (M + H) 356; HRMS calcd for $C_{17}H_{26}NO_5P$ (M + H) 356.1627, found 356.1626.

5.1.6. 3(*E*)-2 - (2 - Aminophenyl)ethenephosphonic acid monoammonium salt (6). To **10** (500 mg, 0.14 mmol) in acetonitrile (35 mL) was added trimethylsilyl bromide (5.5 mL, 42 mmol) and the mixture was stirred for 15 h. After quenching the reaction with water (5mL), the solvent was removed by evaporation under vacuum. The resulting residue was made slightly acidic (pH 5-6) and purified on ion exchange resin (DOWEX® 50WX4-400), eluting first with water and then ammonium hydroxide solution (5% in water). All fractions that turned positive to ninhydrin solution were combined and water evaporated under vacuum to afford the title compound as a tan colored solid (219 mg, 72%): ^{1}H NMR (DMSO) δ 6.08–6.22 (1H, dd, J = 16.2, 15.3, PCH), 6.45–6.52 (1H, t, J = 7.5, ArH), 6.60–6.65 (1H, d, J = 8.4, ArH), 6.90–6.96 (1H, t, J = 7.7, ArH), 7.03–7.17 (1H, dd, J = 18.8, 20.1, ArCH), 7.17–7.22 (1H, d, J = 7.5, ArH); ^{13}C NMR (DMSO) δ 118.06, 120.44, 121.40 + 123.77 (d), 122.99, 127.23, 130.47, 137.41 (d, ArCH), 143.88; MS (-veFAB) (M-H) HRMS calcd for $C_8H_9NO_3P$ 198.0320, found 198.0316.

5.1.7. 2-(Aminophenyl)ethane phosphonic acid monoammonium salt (5). To **6** (200 mg, 0.93 mmol) in methanol (10 mL) was added palladium on carbon (20 mg, 5% on carbon). The atmosphere was purged with hydrogen, and the solution was stirred overnight. After filtration of the mixture through Celite, the solvent was removed by evaporation under vacuum to give a white solid (205 mg). The solid was taken up in hot methanol and filtered through a sintered glass funnel which was washed further with hot methanol. Solvent was evaporated under vacuum from the filtrate, and the residue was further purified using ion exchange resin (DOWEX® 50WX4-400) eluting with water and then ammonium hydroxide solution (5% sol in water). All ninhydrin positive fractions were combined, and the water was removed by evaporation to give a gray solid. The solid was finally recrystallized from methanol to give a tan solid (36 mg, 18%): ^{1}H NMR (DMSO) δ 1.68–1.89 (2H, m, PCH_2), 2.74–2.72 (2H, m, $ArCH_2$), 6.50–6.62 (1H, d, J = 7.2 Hz, ArH), 6.59–6.66 (1H, d, J = 8.1, ArH), 8.86–7.00 (2H, m, ArH); ^{13}C NMR (DMSO) δ 29.93, 31.71 + 33.47 (d), 120.35, 122.16, 130.92 + 131.13 (d), 132.13, 133.95, 150.86; MS (-veFAB) 200; HRMS calcd for $C_8H_{11}NO_3P$ 200.0477, found 200.0483.

5.2. Protein expression and purification

Cloning, expression, and purification of each isoform was performed as previously described[70] with slight modification. Briefly, the pET-11d vector containing the wild-type sequence for either HCPTP-A or HCPTP-B was transformed into fresh *E. coli* BL21(DE3) cells. Cells were plated, selected, and grown in M9ZB media at 37 °C in the presence of 100 μg/mL ampicillin. When the culture reached an optical density of 0.8–1 at 600 nm, cells were induced for 4 h with the addition of 0.4 mM isopropyl-1-thio-β-D-galactopyranoside. Cells were harvested by centrifugation at 3,000×*g* for 15 min and stored at −80 °C until purification.

The cell pellet was resuspended in the cation exchange loading buffer consisting of 10 mM sodium acetate, pH 4.8, 10 mM sodium phosphate, 50 mM NaCl and 1 mM EDTA. Cells were lysed using a French press and the lysate centrifuged at 27,000×*g* for 30 min. The supernatant was then loaded directly onto a HiPrep 16/10 SP cation exchange column. Elution was done with a buffer of 300 mM sodium phosphate, pH 5.1, and 1 mM EDTA. Fractions with phosphatase activity were concentrated and loaded onto a Superdex 75 gel filtration column using the cation exchange loading buffer.

5.2.1. Enzyme activity assay. All enzymatic assays were performed at 37 °C in 100 mM sodium acetate, pH 5.0, adjusted to an ionic strength of 150 mM with NaCl, as described previously,[71] and the program DNRPEasy[72] was used to perform curve fitting. Activity was monitored with the substrate analogue *p*-nitrophenyl phosphate (pNPP), with production of *p*-nitrophenol measured at 405nm via continuous or discrete assay. In either case, enzyme concentration did not exceed 0.1% of substrate. The discrete assay was quenched with 1N NaOH after a 4 min reaction and used an extinction coefficient of 18,000 $M^{-1}cm^{-1}$, with fitting of the standard Michaelis-Menten equation to provide V_{max} and K_M. The continuous assay is not quenched and is consequently monitored at a pH roughly two units below the pK_a of the product. As such, a standard curve was performed to determine the extinction coefficient of 185 M^{-1} cm^{-1} at pH 5. The Michaelis-Menten parameters were fit to a four variable equation in the continuous assay. Enzyme inhibition was measured at inhibitor concentrations from 0.5 to 5 mM. Compounds **5** and **6** have limited solubility, so the stock solution was prepared as a saturated solution then vortexed to a uniform suspension. The dilutions in the kinetic assay provided homogeneous solutions in all cases and gave reproducible results.

Acknowledgements

The authors gratefully acknowledge the support of the National Cancer Institute (CA82673 to CVS) and col-

laboration with the Walther Cancer Institutes at Purdue University and the University of Notre Dame. We thank Dr. Isamu Katsuyama for preliminary studies on the synthesis of PTP inhibitors. Computing resources were provided by the National Science Foundation (DMR0079647 to OW) and the Office of Information Technology at the University of Notre Dame.

References and notes

1. Stauffacher, C. V.; Charbonneau, H. In *Principles of Molecular Regulation*; Conn, P. M., Means, A. R. (Eds.). Humana Press: Totowa, N.J., 2000, p 323.
2. Tonks, N. K.; Neel, B. G. *Curr. Opin. Cell Biol.* **2001**, *13*, 182.
3. Li, L. W.; Dixon, J. E. *Semin. Immunol.* **2000**, *12*, 75.
4. Hunter, T. *Cell* **2000**, *100*, 113.
5. Li, J.; Yen, C.; Liaw, D.; Podsypanina, K.; Bose, S.; Wang, S. I.; Puc, J.; Miliaresis, C.; Rodgers, L.; McCombie, R.; Bigner, S. H.; Giovanella, B. C.; Ittmann, M.; Tycko, B.; Hibshoosh, H.; Wigler, M. H.; Parsons, R. *Science* **1997**, *275*, 1943.
6. Kikawa, K. D.; Vidale, D. R.; Van Etten, R. L.; Kinch, M. S. *J. Biol. Chem.* **2002**, *277*, 39274.
7. Wang, Q.; Holmes, D. I. R.; Powell, S. M.; Lu, Q. L.; Waxman, J. *Cancer Lett.* **2002**, *175*, 63.
8. Cobb, B. S.; Parsons, J. T. *Oncogene* **1993**, *8*, 2897.
9. Burke, T. R.; Zhang, Z. Y. *Biopolymers* **1998**, *47*, 225.
10. Murthy, V. S.; Kulkarni, V. M. *Bioorg. Med. Chem.* **2002**, *10*, 897.
11. Fu, H.; Park, J.; Pei, D. *Biochemistry* **2002**, *41*, 10700.
12. Shen, K.; Keng, Y. F.; Wu, L.; Guo, X. L.; Lawrence, D. S.; Zhang, Z. Y. *J. Biol. Chem.* **2001**, *276*, 47311.
13. Iversen, L. F.; Andersen, H. S.; Branner, S.; Mortensen, S. B.; Peters, G. H.; Norris, K.; Olsen, O. H.; Jeppesen, C. B.; Lundt, B. F.; Ripka, W.; Moller, K. B.; Moller, N. P. H. *J. Biol. Chem.* **2000**, *275*, 10300.
14. Ducruet, A. P.; Rice, R. L.; Tamura, K.; Yokokawa, F.; Yokokawa, S.; Wipf, P.; Lazo, J. S. *Bioorg. Med. Chem.* **2000**, *8*, 1451.
15. Peng, H. R.; Xie, W. G.; Otterness, D. M.; Cogswell, J. P.; McConnell, R. T.; Carter, H. L.; Powis, G.; Abraham, R. T.; Zalkow, L. H. *J. Med. Chem.* **2001**, *44*, 834.
16. Zhang, Z.-Y. *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 209.
17. Holmberg, J.; Frisén, J. *Trends Neurosci.* **2002**, *25*, 239.
18. Knöll, B.; Drescher, U. *Trends Neurosci.* **2002**, *25*, 145.
19. Dodelet, V. C.; Pasquale, E. B. *Oncogene* **2000**, *19*, 5614.
20. Zantek, N. D.; Azimi, M.; Fedor-Chaiken, M.; Wang, B.; Brackenbury, R.; Kinch, M. S. *Cell Growth & Differ.* **1999**, *10*, 629.
21. Zelinski, D. P.; Zantek, N. D.; Stewart, J. C.; Irizarry, A. R.; Kinch, M. S. *Cancer Res.* **2001**, *61*, 2301.
22. Stein, E.; Lane, A. A.; Cerretti, D. P.; Schoecklmann, H. O.; Schroff, A. D.; Van Etten, R. L.; Daniel, T. O. *Genes Dev.* **1998**, *12*, 667.
23. Heinrikson, R. L. *J. Biol. Chem.* **1969**, *244*, 299.
24. Thomas, C. L.; McKinnon, E.; Granger, B. L.; Harms, E.; Van Etten, R. L. *Biochemistry* **2002**, *41*, 15601.
25. Taga, E. M.; Van Etten, R. L. *Arch. Biochem. Biophys.* **1982**, *214*, 505.
26. Waheed, A.; Laidler, P. M.; Wo, Y. Y. P.; Van Etten, R. L. *Biochemistry* **1988**, *27*, 4265.
27. Manao, G.; Pazzagli, L.; Cirri, P.; Caselli, A.; Camici, G.; Cappugi, G.; Saeed, A.; Ramponi, G. *J. Protein Chem.* **1992**, *11*, 333.
28. Okada, M.; Owada, K.; Nakagawa, H. *Biochem. J.* **1986**, *239*, 155.
29. Chernoff, J.; Schievella, A. R.; Jost, C. A.; Erikson, R. L.; Neel, B. G. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 2735.
30. Lawrence, G. L.; Van Etten, R. L. *Arch. Biochem. Biophys.* **1981**, *206*, 122.
31. Zhang, Z. Y.; Van Etten, R. L. *Arch. Biochem. Biophys.* **1990**, *282*, 39.
32. Modesti, A.; Cirri, P.; Raugei, G.; Carraresi, L.; Magherini, F.; Manao, G.; Camici, G.; Ramponi, G. *FEBS Lett.* **1995**, *375*, 235.
33. Mondesert, O.; Moreno, S.; Russell, P. *J. Biol. Chem.* **1994**, *269*, 27996.
34. Ostanin, K.; Pokalsky, C.; Wang, S.; Van Etten, R. L. *J. Biol. Chem.* **1995**, *270*, 18491.
35. Park, E. K.; Warner, N.; Mood, K.; Pawson, T.; Daar, I. O. *Mol. Cell. Biol.* **2002**, *22*, 3404.
36. Miller, D. T.; Read, R.; Rusconi, J.; Cagan, R. L. *Gene* **2000**, *243*, 1.
37. Raugei, G.; Ramponi, G.; Chiarugi, P. *Cell. Mol. Life Sci.* **2002**, *59*, 941.
38. Dissing, J.; Rangaard, B.; Christensen, U. *Biochim. Biophys. Acta* **1993**, *1162*, 275.
39. Zhang, M.; Van Etten, R. L.; Stauffacher, C. V. *Biochemistry* **1994**, *33*, 11097.
40. Wang, S.; Tabernero, L.; Zhang, M.; Harms, E.; Van Etten, R. L.; Stauffacher, C. V. *Biochemistry* **2000**, *39*, 1903.
41. Zhang, M.; Stauffacher, C. V.; Lin, D.; Van Etten, R. L. *J. Biol. Chem.* **1998**, *273*, 21714.
42. Wang, S.; Stauffacher, C. V.; Van Etten, R. L. *Biochemistry* **2000**, *39*, 1234.
43. Tanizaki, M. M.; Bittencourt, H. M.; Chaimovich, H. *Biochim. Biophys. Acta* **1977**, *485*, 116.
44. Zhang, Z. Y.; Van Etten, R. L. *J. Biol. Chem.* **1991**, *266*, 1516.
45. Nicholls, A.; Sharp, K.; Honig, B. *Proteins* **1991**, *11*, 281.
46. Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946.
47. Merritt, E. A.; Bacon, D. J. In *Macromolecular Crystallography, Pt B*, 1997; Vol. 277, p 505.
48. Barton, G. J. *Protein Eng.* **1993**, *6*, 37.
49. Sippl, W. *Bioorg. Med. Chem.* **2002**, *10*, 3741.
50. Walters, W. P.; Murcko, M. A. *Adv. Drug Deliv. Rev.* **2002**, *54*, 255.
51. Abagyan, R.; Totrov, M. *Curr. Opin. Chem. Biol.* **2001**, *5*, 375.
52. Kuntz, I. D. *Science* **1992**, *257*, 1078.
53. Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput. Aid. Mol. Des* **2002**, *16*, 151.
54. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187.
55. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, 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.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.
56. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comp. Chem.* **1998**, *19*, 1639.
57. Zhang, M.; Zhou, M.; Van Etten, R. L.; Stauffacher, C. V. *Biochemistry* **1997**, *36*, 15.
58. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
59. Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
60. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.;

Montgomery, J., J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A.; A.9 ed.; Gaussian, Inc.: Pittsburgh, PA, 1998.

61. Dillet, V.; Van Etten, R. L.; Bashford, D. *J. Phys. Chem. B* **2000**, *104*, 11321.
62. Zhou, M. M.; Davis, J. P.; Van Etten, R. L. *Biochemistry* **1993**, *32*, 8479.
63. Tishmack, P. A.; Bashford, D.; Harms, E.; Van Etten, R. L. *Biochemistry* **1997**, *36*, 11984.
64. Verlet, L. *Phys. Rev.* **1967**, *159*, 98.
65. Solis, F. J.; Wets, R. J.-B. *Math. Oper. Res.* **1981**, *6*, 19.
66. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
67. Robinson, C. N.; Pettit, W. A.; Walker, T. O.; Shearon, E.; Mokashi, A. M. *J. Heterocyl. Chem.* **1972**, *9*, 735.
68. Janiak, C.; Deblon, S.; Uehlin, L. *Synthesis* **1999**, *6*, 959.
69. Lamba, J. J. S.; Tour, J. M. *J. Am. Chem. Soc.* **1994**, *116*, 11723.
70. Wo, Y.-Y. P.; McCormack, A. L.; Shabanowitz, J.; Hunt, D. F.; Davis, J. P.; Mitchell, G. L.; Van Etten, R. L. *J. Biol. Chem.* **1992**, *267*, 10856.
71. Davis, J. P.; Zhou, M.-M.; Van Etten, R. L. *Biochemistry* **1994**, *33*, 1278.
72. Duggleby, R. G. *Comput. Biol. Med.* **1984**, *14*, 447.
73. Liu, H.; Huang, X. Q.; Shen, J. H.; Luo, X. M.; Li, M. H.; Xiong, B.; Chen, G.; Shen, J. K.; Yang, Y. M.; Jiang, H. L.; Chen, K. X. *J. Med. Chem.* **2002**, *45*, 4816.
74. Rockey, W. M.; Elcock, A. H. *Proteins* **2002**, *48*, 664.
75. Primozic, I.; Hrenar, T.; Tomic, S.; Meic, Z. *J. Phys. Org. Chem.* **2002**, *15*, 608.
76. Buzko, O. V.; Bishop, A. C.; Shokat, K. M. *J. Comput. Aid. Mol. Des* **2002**, *16*, 113.
77. Evans, B. In *Department of Chemistry*; Purdue University: West Lafayette, 1997, p 218.
78. Cristau, H. J.; Pirat, J. L.; Drag, M.; Kafarski, P. *Tetrahedron Lett.* **2000**, *41*, 9781.
79. Bigge, C. F.; Johnson, G.; Ortwine, D. F.; Drummond, J. T.; Retz, D. M.; Brahce, L. J.; Coughenour, L. L.; Marcoux, F. W.; Probert, A. W. *J. Med. Chem.* **1992**, *35*, 1371.