Molecular Dynamics Simulations of Protein-Tyrosine Phosphatase 1B. II. Substrate-Enzyme Interactions and Dynamics

Günther H. Peters,* Thomas M. Frimurer,† Jannik N. Andersen,‡ and Ole H. Olsen†
†MedChem Research IV, Novo Nordisk A/S, DK-2760 Måløv; *Department of Chemistry, MEMPHYS, Technical University of Denmark, DK-2800 Lyngby; and ‡Target Cell Biology, Novo Nordisk A/S, DK-2800 Bagsvaard, Denmark

ABSTRACT Molecular dynamics simulations of protein tyrosine phosphatase 1B (PTP1B) complexed with the phosphorylated peptide substrate DADEpYL and the free substrate have been conducted to investigate 1) the physical forces involved in substrate-protein interactions, 2) the importance of enzyme and substrate flexibility for binding, 3) the electrostatic properties of the enzyme, and 4) the contribution from solvation. The simulations were performed for 1 ns, using explicit water molecules. The last 700 ps of the trajectories was used for analysis determining enthalpic and entropic contributions to substrate binding. Based on essential dynamics analysis of the PTP1B/DADEpYL trajectory, it is shown that internal motions in the binding pocket occur in a subspace of only a few degrees of freedom. In particular, relatively large flexibilities are observed along several eigenvectors in the segments: Arg^{24}-Ser^{28}, Pro^{38}-Arg^{47}, and Glu^{115}-Gly^{117}. These motions are correlated to the C- and N-terminal motions of the substrate. Relatively small fluctuations are observed in the region of the consensus active site motif (H/V)CX_5R(S/T) and in the region of the WPD loop, which contains the general acid for catalysis. Analysis of the individual enzyme-substrate interaction energies revealed that mainly electrostatic forces contribute to binding. Indeed, calculation of the electrostatic field of the enzyme reveals that only the field surrounding the binding pocket is positive, while the remaining protein surface is characterized by a predominantly negative electrostatic field. This positive electrostatic field attracts negatively charged substrates and could explain the experimentally observed preference of PTP1B for negatively charged substrates like the DADEpYL peptide.

INTRODUCTION

Protein phosphatases (PPs), including protein-tyrosine phosphatases (PTPs), are key participants in kinase-dependent signal transduction pathways (Stone and Dixon, 1994; Walton and Dixon, 1993). The level of protein-tyrosine phosphorylation and dephosphorylation is critical for normal cell proliferation, differentiation, and metabolism (Fisher et al., 1991; Tonks et al., 1993; Walton and Dixon, 1993) age (Ruiz et al., 1992) and diabetes (Møller et al., 1995; Ide et al., 1994; Boylan et al., 1992; McGuire et al., 1991). There is increasing evidence that aberrant levels of PPs may underlie diseases such as cancer (Zanke et al. 1994; Wiener et al., 1996; Cool and Fischer, 1993; Zheng et al., 1992; Lammers et al., 1993), insulin-resistant states (Posner et al., 1994; Møller et al., 1995), and defects in platelet aggregation response (Frangioni et al., 1993). Thus the ability to inhibit specific protein phosphatases could have clinical importance and may provide alternative treatments of diseases. Furthermore, specific phosphatase inhibitors are invaluable tools for elucidating the function of individual PTPs within cells and for studying phosphatase catalytic mechanisms (Montsera et al., 1996).

Received for publication 23 September 1999 and in final form 22 January 2000.

Address reprint requests to Günther H. Peters, Department of Chemistry, MEMPHYS, The Technical University of Denmark, DK-2800 Lyngby, Denmark. Tel.: +45-4525-2385; E-mail: GHP@KEMI.DTU.DK. Or, Dr. Ole H. Olsen, Novo Nordisk A/S, MedChem Research IV, Novo Nordisk Park, DK Malov, Denmark. Tel.: +45-4443-4511; Fax: +45-4443-4547; E-mail: oho@novo.dk.

0006-3495/00/05/2191/10 $2.00

The development of novel drug candidates to specifically inhibit phosphatases requires not only a detailed structural understanding of the regulation of these enzymes, but also a knowledge of how these enzymes distinguish between the different phosphorylated substrates that they encounter in the cell. The current estimate is that humans have as many as several hundred phosphatase genes (Hunter, 1995). A common approach is the use of inhibitors to probe substrate recognition, to elucidate cellular function of individual PPs, and to study the catalytic mechanisms (Montsera et al., 1996). Moreover, x-ray crystallographic structures of several phosphatases show remarkable structural similarity around the active site, despite very low sequence homology (Fauman and Saper, 1996; Fauman et al., 1996; Barford, 1995). The PTP family is characterized by the consensus active site sequence (H/V)CX_5R(S/T), where X denotes any amino acid residue. The active-site motif includes a nucleophilic cysteinyl residue essential for catalysis and a conserved arginine residue (Pot and Dixon, 1992; Streuli et al., 1990). The structure of the consensus sequence defines the binding site for the tyrosyl phosphate substrate. The catalytic reaction (i.e., hydrolysis of the phosphoester bond) is facilitated by the cysteine thiolate and the closing of the WPD loop. At physiological pH, the cysteine is negatively charged (Denu and Dixon, 1995; Zhang et al., 1994b; Peters et al., 1998) and attacks the electrophilic phosphorus of the phosphotyrosyl residue in the substrate, causing the phosphoester bond to break. At the same time the WPD loop, which upon binding of the substrate moves toward the phosphate moiety of the substrate, essentially causes a tight binding of the tyrosyl phosphate group. This displacement

brings a catalytically active aspartate into position and corresponds to the "activation" of the enzyme (Zhang and Wu, 1997; Eckstein et al., 1996; Schubert et al., 1995; Barford, 1995).

Substrate recognition and the formation of an enzyme-substrate intermediate are complex events that are primarily mediated by residues spanning the binding pocket. The binding event, which is driven by a decrease in the total Gibbs free energy, is generally dictated by a delicate balance of mechanisms of opposing effects involving enthalpic and/or entropic contributions (Salemme et al., 1997; Ajay and Murcko, 1995). Various factors have been suggested as the physical reasons for high binding affinities (Vajda et al., 1994; Krystek et al., 1993; Jencks, 1975; Koshland, 1958). An increase in binding affinity is generally accomplished by favorable interactions between substrate and protein and/or by a net increase in the entropy of the system (i.e., protein and water) (Peters et al., 1997a; Goddette et al., 1993). Typically, entropy is gained when water molecules, which in the unliganded enzyme form a well-defined network in cavities and/or on the protein surface, are displaced upon substrate binding. The net gain is determined by the balance of entropy reduction due to the loss in translational and rotational entropies in the enzyme and/or substrate and entropy increase, e.g., due to the displacement of water. Changes in steric interaction energy on binding or change in conformational energy of the protein and substrate upon binding (induced fit) can make negative or positive contributions to the binding affinity, depending on the architecture of the binding pocket and the molecular structure of the substrate (Koshland, 1958). Intermolecular forces are predominantly determined by van der Waals and electrostatic interactions. The latter is thought often to play an important role in determining the biological function of enzymes (Peters et al., 1997b; Honig and Nicholls, 1995). The physical complexity underlying the free energy change that accompanies the binding of a substrate to an enzyme in aqueous solution makes it difficult to estimate the magnitude of the individual contributions (Ajay and Murcko, 1995). It has been observed that different substrates can bind through either enthalpically or entropically driven processes. For instance, biotin binding to streptavidin is enthalpy driven, whereas azobenzene substrate binding to streptavidin is entropy driven (Weber et al., 1992; Ku et al., 1993).

The prediction of binding energies is a demanding task, and a range of methodologies for this have been suggested. Several theoretical approaches assume fixed substrate and protein structures. These assumptions may be valid for relatively rigid systems, but for macromolecules undergoing conformational changes upon substrate binding and activation, substrate and enzyme flexibilities are important factors. In this category also belong protein tyrosine phosphatases. To elucidate the importance of flexibility and interactions in a complexed structure of a phosphatase, we have performed molecular dynamics simulation of PTP1B in complex with the peptide-based substrate DADEpYL. Furthermore, a molecular dynamics simulation of the free substrate has been conducted to estimate the importance of electrostatic and van der Waals energies to binding. The sequence of this synthetic phosphotyrosine-containing peptide is derived from the autophosphorylation site of the epidermal growth factor receptor ($EGFR_{988-998}$; DADEpYLIPQQG, where pY stands for the phosphorylated tyrosine). This peptide sequence has been widely used to probe the substrate specificity of phosphatases, including PTP1B. In these studies, the amino acids in the peptide were sequentially mutated to Ala (referred to as Ala scan in the literature), and the specific contributions to the binding and catalysis were determined by kinetic measurements (Zhang et al., 1993, 1994a). The data indicated that substitution of acidic residues N-terminal to pY by Ala results in substantial loss in substrate specificity, while substitution of residues on the C-terminal side of pY by Ala only slightly effects the binding affinity and catalytic efficiency. In particular, the substitution Glu $\rightarrow$ Ala at the -1 position changes the kinetic parameters significantly (nomenclature is adapted from Zhang et al. (1993); pY is designated to be at the zero position in the peptide sequence, and the adjacent amino acids are numbered positively (C-terminally to pY) or negatively (N-terminally to pY). The substitution Glu $\rightarrow$ Ala in the peptide sequence causes a 4.7-fold increase in K_M, while k_{cat} and the ratio k_{cat}/K_M are reduced by 1.4- and 6.5-fold, respectively. These experimental studies suggest that a minimum phosphopeptide for optimal binding and catalysis is the hexapeptide DADEpYL. Amidation of the free carboxyl group of Leu increases the catalytic efficiency by 5.8-fold, whereas acetylation of the N-terminal amino group has only a slight effect and increases the k_{cat}/K_M value by 1.2-fold.

In the present study, we have investigated the dynamic and energetic properties of the PTP1B complex in detail. This paper is complementary to paper I (Peters et al., 1999), where we have investigated and identified the concerted motions in PTP1B (open conformation) and PTP1B complexed with DADEpYL (closed conformation). Here we focus on 1) investigating the influence of substrate flexibility on the mobility of the binding pocket, 2) examining the interactions between the substrate and residues in the binding pocket, 3) determining the electrostatic properties of PTP1B, and 4) the contribution from solvation.

MATERIALS AND METHODS

The x-ray crystallographic structure of a Cys^{215} (active cysteine)/serine mutant of the PTP1B structure complexed with the DADEpYL-NH_2 peptide, solved to 2.6-Å resolution (Jia et al., 1995; Barford et al., 1994), was used as model for the closed (active) structure. The structure was obtained from the Protein Data Bank at Brookhaven (Bernstein et al., 1977). The entry code is 1ptu. Before we performed the simulation, the serine (215) was replaced with an ionized cysteine. The rationale for modeling the

cysteine side chain as a thiolate is based on the experimentally observed and theoretically found low pK_a value (Zhang et al., 1993; Peters et al., 1998). We have modified the peptide in the crystal structure from DADEpYL-NH_2 to DADEpYL by adding a negatively charged C-terminus (COO^-).

Molecular dynamics simulation of the PTP1B-peptide complex with explicit SPC water (7222 molecules) was performed for 1 ns, using the charged (C) version of the GROMOS force field (Van Gunsteren and Berendsen, 1987) and applying periodic boundary conditions. The dimension of the octahedral simulation cell was 81.1 Å. Details of the setup and parameters used in the molecular dynamics simulation have been described in part I (Peters et al., 1999). Furthermore, a 1-ns simulation of the free substrate was performed with explicit SPC water (1703 molecules) and ions (six sodium ions), using the same MD simulation protocol as applied for the complexed PTP1B structure.

Analyses of the trajectories and examinations of the molecular structures were carried out using the WHAT IF modeling program (Vriend, 1990). Evaluation of several geometrical properties indicated that the PTP1B-DADEpYL complex is equilibrated after 150 ps. To ease the comparison of the present results to the findings in part I (Peters et al., 1999), we have used the last 700 ps of the trajectory for an essential dynamics analysis (Amadei et al., 1993; Ichiye and Karplus, 1991), evaluation of enzyme-substrate interaction energies, and investigation of the electrostatic properties of the enzyme. The latter refers to the calculation of pK_as of titratable amino acids in PTP1B (Antosiewicz et al. 1994; Gilson, 1993). Similarly, the last 700 ps of the trajectory for the free substrate was analyzed to evaluate the interaction energies between substrate and solvent.

A brief description of the methodologies used to extract information about the concerted motion in the PTP1B complex and to compute the pK_as has appeared in part I (Peters et al., 1999). The electrostatic calculations and estimations of the pK_as were performed using the UHBD program (Madura et al., 1995; Davis et al., 1991) and the "hybrid procedure" developed by Gilson (1993). The parameters used in these calculations have been presented elsewhere (Peters et al., 1998).

RESULTS AND DISCUSSION

One of the striking experimental observations is that the negatively charged peptide DADEpYL shows high binding affinity for PTP1B. The complex structure is displayed in Fig. 1. The high binding affinity is particularly surprising, because the total charge of titratable residues in PTP1B is approximately −6 (Fig. 2).

To elucidate the origin of this observation, we have applied molecular dynamics simulations and macroscopic electrostatic calculations to obtain detailed structural information about the forces involved in the binding process. The binding event can be divided into at least two processes involving diffusion and binding of the substrate. First, the substrate has to diffuse onto the protein surface and into the binding pocket, where binding and hydrolysis eventually take place. To determine the mechanism involved in the diffusion process, we have computed the protonation state of all of the titratable amino acids in the enzyme (Antosiewicz et al., 1994; Gilson, 1993) and used these charges (pH 7) to compute the electrostatic field. Fig. 2 displays contour plots of the field surrounding the active site. Clearly, only the surface above the active site is positively charged, while the remaining part of the surface is predominantly negatively charged. This positively charged patch attracts the negatively charged substrate, and the substrate diffuses toward the active site.

To obtain further details on the atomic level, we have performed molecular dynamics simulations of the complex structure (Fig. 1) and the free substrate to determine the enthalpic contributions to the binding process. The free substrate simulation was included as a reference point for the energy components and to ease comparison between the free and bound states.

Simulations were carried out for 1 ns, and as indicated by several geometrical properties, the protein equilibrated within 150 ps. As mentioned above, the first 300 ps was discarded, and the remaining 700 ps was used in the analysis. Root mean square displacement (rmsd) with respect to

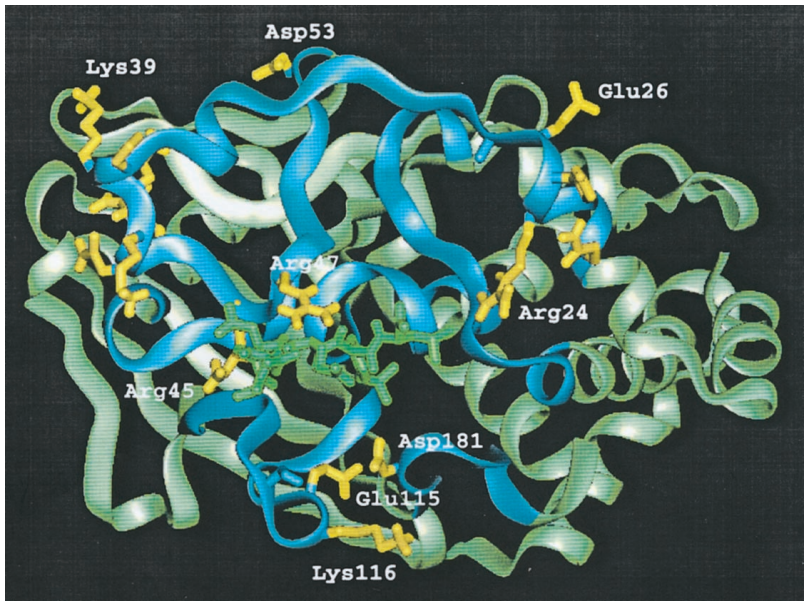

FIGURE 1 Structure of PTP1B complexed with the peptide DADEpYL, displayed in sticks and colored green. The structural parts colored blue and yellow were included in the essential dynamics analysis. The side chains colored yellow show relatively high flexibility and correspond to the lettering in Fig. 6.

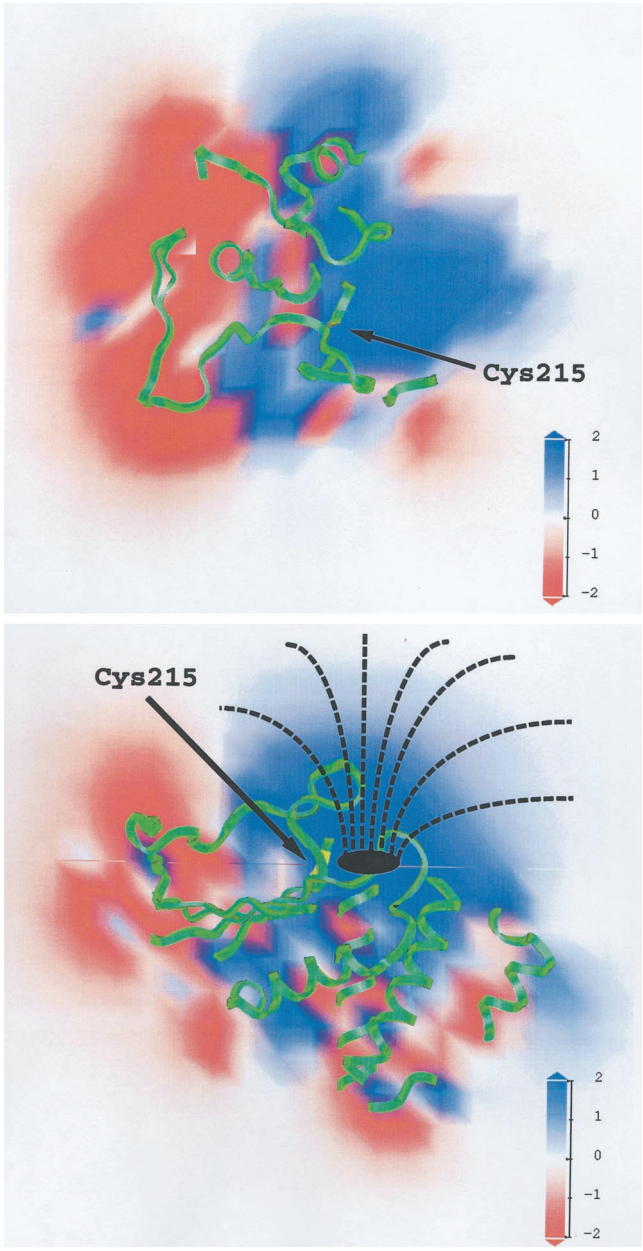

FIGURE 2 Contour plots of the electrostatic field around the active Cys[215]. The levels are indicated by the spectrum (*lower right*). The secondary structure of PTP1B is shown in green. Cys[215] is depicted in yellow. (*A*) The plane of the contour lines is parallel to the central helix and goes through Cys[215]. (*B*) The plane of the contour lines is at a right angle to the central helix and goes through Cys[215].

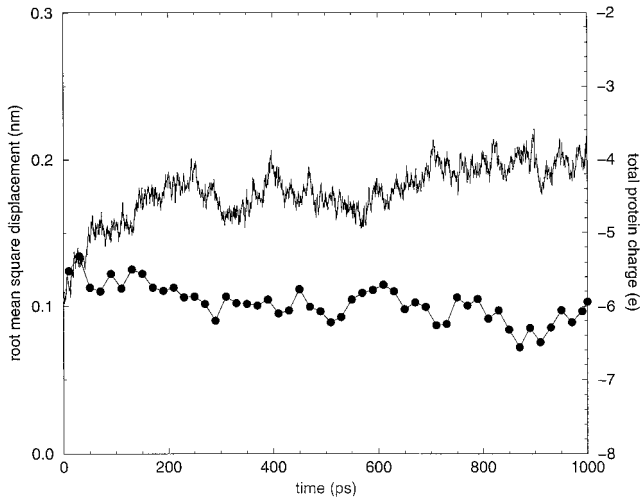

FIGURE 3 Root mean square displacement (*left axis, solid line*) and total protein charge at pH 7.0 (*right axis, connected dots*) as a function of simulation time for the PTP1B-peptide complex.

the starting structure, number of hydrogen bonds, radius of gyration, and accessible surface area (ASA) was calculated along the trajectory. Although the number of hydrogen bonds and the radius of gyration were constant (Peters et al., 1999), relatively large fluctuations were observed in the rmsd (Fig. 3) and ASA (Peters et al., 1999), indicating the inherent flexibility of the protein complexed with the peptide. These fluctuations have only minor influence on the overall protein charge, as shown in Fig. 3. This is a further indication that the protein has adapted to the environment and that a stable trajectory has been obtained (Wlodek et al., 1997).

To further study the effect of a peptide substrate on the dynamics of PTP1B, we have performed an essential dynamics analysis of the PTP1B-substrate complex trajectory. In contrast to our previous study, we have focused on fluctuations in the binding pocket and how these motions are correlated with the motions of the phosphorylated peptide. The analysis was performed by including the substrate and the amino acids of the protein within a distance of ~15 Å from the phosphate moiety. This is also illustrated in Fig. 1, where the structural part of the enzyme considered in the essential dynamics analysis is colored blue. The covariance matrix was constructed from the extracted atom coordinates and subsequently diagonalized. This technique allows us to study the correlated motions within the configurational subspaces spanned by the individual eigenvectors. The eigenvectors describe the directions in the subspace, whereas the eigenvalues reflect the magnitude of these motions. In our previous investigation, we have shown that the first 50 eigenvectors describe ~85% of the motions in the PTP1B complexed with DADEpYL (Peters et al., 1999). Fluctuations along higher eigenvectors are essentially Gaussian (thermal) motions. Thus the correlation coefficient between ideal Gaussian distributions and the distributions calculated from the eigenvector is larger than 0.95 after the first eight eigenvectors (Peters et al., 1999).

Fluctuations within the subspace can be studied by projecting the trajectory onto the individual eigenvector, which provides information about the time dependence of the conformational changes. Two important quantities can be extracted from this dot product. The average value of pro-

jection reflects structural changes, whereas the mean square fluctuation in the projection refers to differences in dynamics in the individual subspaces. For eigenvectors of index larger than 4, the average value of projection was less than 0.01 (Peters et al., 1999), which indicates that no significant structural changes occurred during the simulations and that structures were sampled along an equilibrium molecular dynamics trajectory. The mean square fluctuation of these projections decreases rapidly with increasing eigenvector index and is less than 0.2 nm^2 for eigenvector indices greater than 10 (see Fig. 4 *A*). Further inspection of these projections shows that for eigenvector indices geater than 2 the values fluctuate periodically around zero. Only eigenvector 1 shows a nonharmonic behavior (as shown in Fig. 4 *B*), indicating that the period of these motions exceeds the simulation period. The origin of these fluctuations will be discussed below. With increasing eigenvector indices, motions along the different eigenvectors become increasingly more harmonic (Fig. 4 *B*). There are two ways of illustrating the motions within the subspaces. These are 1) the superpositions of sequential projections of the atom motion onto selected eigenvectors and 2) the absolute value of the components of the eigenvectors as a function of coordinate number. For clarity, we have chosen the latter, and the absolute values of the components of the first five eigenvectors as a function of atom number are shown in Fig. 5. The lower curve corresponds to eigenvector 1, and the subsequent curves are shifted by 0.1 in the *y* direction. The atom numbers 1 to 659 correspond to atoms in the PTP1B, and 660 to 714 refer to atoms in the substrate. As shown in Fig. 5, the correlated motions along different eigenvectors are complex and involve several regions in the protein and substrate. Several of these fluctuations are observed along different eigenvectors. These regions are indicated by the letters *a–j* and involve the following residues: (*a*) Gln^{21} [13–21], (*b*) Arg^{24}-Ser^{28} [38–78], (*c*) Pro^{38}-Asp^{43} [151–202], (*d*) Arg^{47} [234–244], (*e*) Asp^{53} [284–291], (*f*) Glu^{115}-Gly^{117} [391–418], (*g*) Asp^{181} [475–482], (*h*) Asp^{-4} [660–667] (substrate residue; numbering adapted from Zhang et al. (1993); residues placed C-terminally (N-terminally) to the phosphor tyrosine have negative (positive) numbers), (*i*) Glu^{-1} [681–689] (substrate residue), and (*j*) Leu^{1} [694–

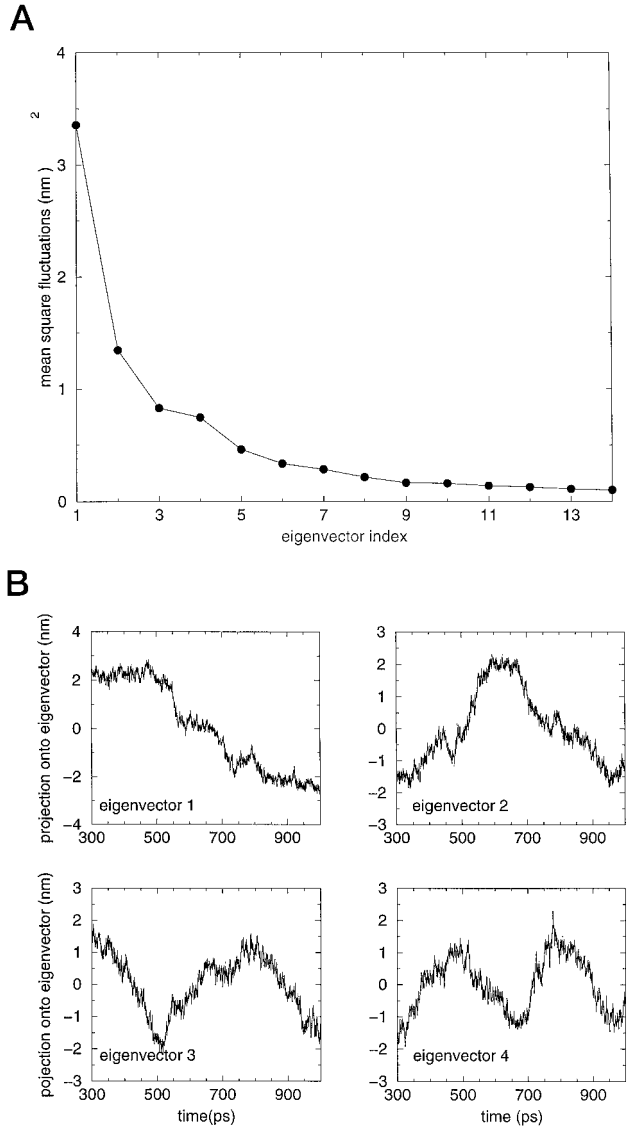

FIGURE 4 (*A*) Mean square fluctuations of projections of the trajectory onto the eigenvectors extracted from the full atom coordinates covariance matrix as a function of the eigenvector index. (*B*) Projection of the full atom trajectory onto the first four eigenvectors as a function of time.

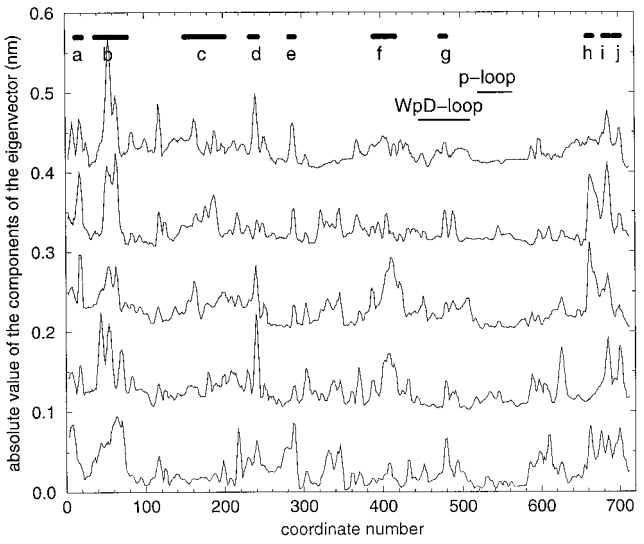

FIGURE 5 Absolute value of the first five eigenvectors (*from the bottom to the top*; individual curves are shifted by 0.1 in the *y* direction) obtained from full atom coordinate covariance matrix as a function of coordinate number. Graphs are sequentially shifted by 0.1. Letters correspond to residues that show relatively high flexibility: (*a*) Gln^{21}, (*b*) Arg^{24}-Ser^{28}, (*c*) Pro^{38}-Asp^{43}, (*d*) Arg^{47}, (*e*) Asp^{53}, (*f*) Glu^{115}-Gly^{117}, (*g*) Asp^{181}, (*h*) Asp^{-4} (substrate), (*i*) Glu^{-1} (substrate), and (*j*) Leu^{1} (substrate).

702] (substrate residue). The numbers in square brackets refer to the coordinate number in Fig. 5. With respect to the substrate, the Glu^{-1} and N- and C-terminals show high flexibility, and, in particular, fluctuations of Asp^{-4} are observed in several subspaces. In all curves, it is noticeable that the P-loop (His^{214} [512–521]-Gly^{219} [552–554]) is rigid, whereas weak fluctuations of the WPD-loop (Thr^{178} [447–453]-Pro^{185} [505–511]) are observed along several eigenvectors. Visualization of the all-atom motions indicates that the fluctuations along eigenvector 1 are triggered by a rotation of the tyrosine benzyl ring (substrate-Ptyr0), providing sufficient space for the phenyl group of the Tyr^{46} residue to move toward the substrate moiety.

Note that the results of the essential dynamics analysis are correlated motions (Amadei et al., 1993). As schematically shown in Fig. 1, the distances between some of the residues are relatively large, and these residues are still influenced by the substrate motion. This may suggest that a tight water network is formed around the active site. To estimate the effect of the solvent, we have calculated the individual energy contributions, which are summarized in Table 1. The relatively large standard deviations found for the different energy distributions are caused by structural fluctuations and the competition between substrate-solvent and substrate-protein interactions. This competitive feature is also observed in Fig. 6, where interaction energies of the complexed structure and the free substrate are compared

TABLE 1 Intermolecular energy components calculated during the molecular dynamics simulation of the substrate-enzyme complex

System	van der Waals energy (kJ/mol)	Electrostatic energy (kJ/mol)
Protein-protein	−8740 ± 109	−21,607 ± 703
Ligand-ligand	−21 ± 11	606 ± 131
Protein-water	−1159 ± 140	−31,055 ± 636
Ligand-protein	−176 ± 19	−2176 ± 254
Ligand-water	10 ± 26	−1607 ± 172
Ligand (Asp^{-4})-protein	−18 ± 8	−430 ± 119
Ligand (Ala^{-3})-protein	−4 ± 2	−1 ± 15
Ligand (Asp^{-2})-protein	−16 ± 9	−518 ± 129
Ligand (Glu^{-1})-protein	−13 ± 4	−331 ± 110
Ligand (Pty^{0})-protein	−15 ± 3	−3 ± 11
Ligand (Leu^{1})-protein	−30 ± 12	−620 ± 94
Ligand (Asp^{-4})-water	13 ± 15	−372 ± 86
Ligand (Ala^{-3})-water	−19 ± 5	−41 ± 17
Ligand (Asp^{-2})-water	14 ± 13	−332 ± 76
Ligand (Glu^{-1})-water	11 ± 14	−519 ± 64
Ligand (Pty^{0})-water	−4 ± 6	−13 ± 14
Ligand (Leu^{1})-water	1 ± 13	−268 ± 64

Errors are standard deviations calculated from the trajectory between 300 and 1000 ps. The numbering of the residues in the ligand is adopted from Zhang et al. (1993) (see text for details).

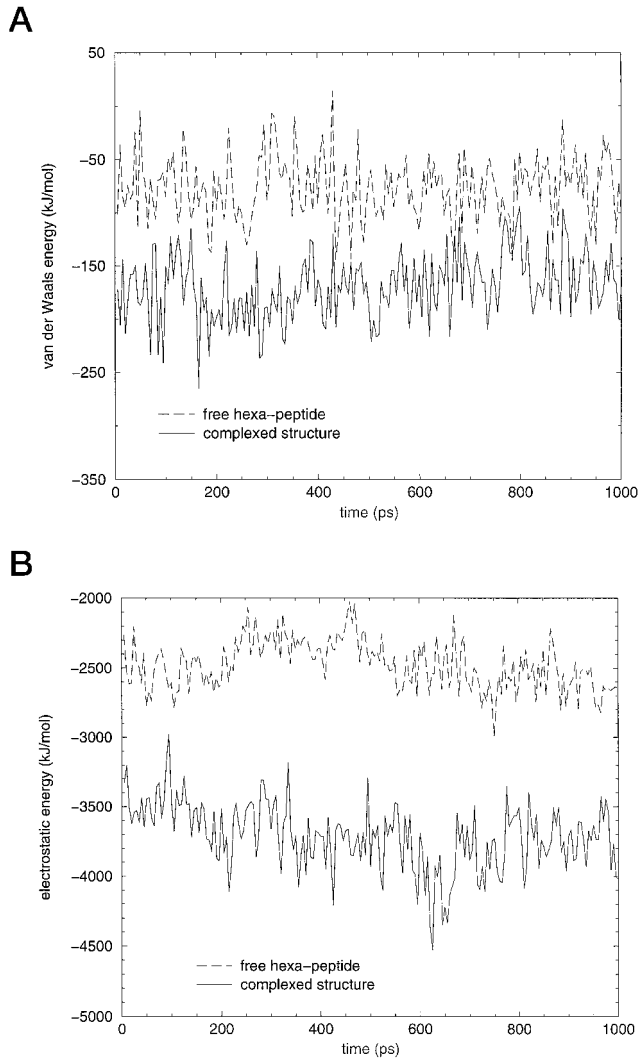

FIGURE 6 (*A*) Van der Waals and (*B*) electrostatic interaction energies between substrate-protein/solvent in the complexed structure and substrate-solvent for the free substrate are shown as a function of simulation time.

along the trajectories. The van der Waals interaction energies between the substrate and its surroundings (Fig. 6 *A*) for the substrate-enzyme complex are consistently lower than for the free substrate. The same tendency is observed for the electrostatic interactions (Fig. 6 *B*). Substrate-protein electrostatic interactions are generally stronger. This large gain in enthalpy will be counteracted by the entropic penalty involved in binding of the substrate, which, of course, also involves contributions from water molecules. However, to evaluate these contributions a detailed calculation of the free energy change during binding has to be performed, which is beyond the scope of the present investigation. In Fig. 7, van der Waals (Fig. 7 *A*) and electrostatic (Fig. 7 *B*) interaction energies are further differentiated for the individual substrate residues (i.e., interaction energy between atoms belonging to one specific substrate residue and pro-

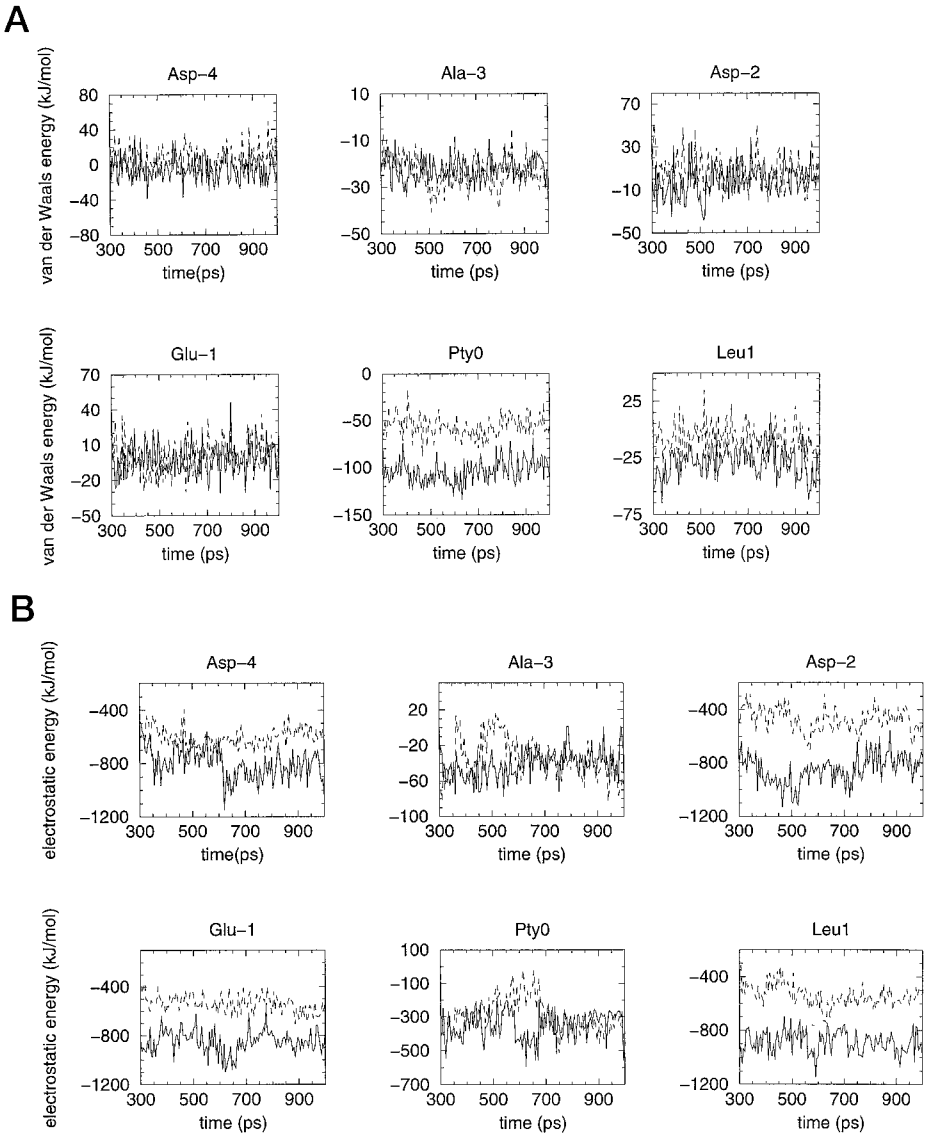

FIGURE 7 (*A*) Van der Waals and (*B*) electrostatic interaction energies are further discriminated for the individual substrate residues. The solid lines present data for substrate residue-protein/solvent interaction energies in the complexed structure. The dashed lines correspond to interaction energies between substrate residues and solvent molecules in the free substrate simulation.

tein/solvent), where the solid and dotted lines refer to the substrate-protein simulation and free substrate simulation, respectively. Significant differences are seen for the different residues in the substrate. Generally, charged residues make a noticeable contribution to the electrostatic enthalpic binding energies. Surprisingly, besides the phosphorylated tyrosine, the van der Waals energies are similar for the substrate residues in the bound or free state. Presumably, the phosphorylated tyrosine (Pty0) is outstanding because it is deeply buried in the enzyme (see Fig. 7 *A*).

For the substrate residues Asp^{-2}, Glu^{-1}, and Leu^{1}, i.e., the substrate positions flanking the phosphor-tyrosine, electrostatic interaction energies for the substrate-protein simulation are (on average) stronger than for the free substrate. The net electrostatic energies (substrate-protein minus substrate-water) for Leu^{1} and Asp^{-2} are -352 and -186 kJ/mol, respectively (see Table 1). This suggests that Leu^{1} is more tightly bound to PTP1B than is Asp^{-2}. Surprisingly, the essential dynamics analysis (Table 2 and Fig. 5) indicates that motions involving Asp^{-2}, Glu^{-1}, and Leu^{1} are pronounced and are observed along several eigenvectors. This is clearly an indication that enthalpic and entropic contributions are important for binding. Recent experimental data (Zhang et al., 1993) have shown that for DADEpY(L→A)IPQQG and DA(D→A)EpYLIPQQG peptides, the K_M values were increased by 1.5-fold and 2.1-fold, respectively. The substitution $Asp^{-2} \rightarrow$ Ala results in a lower binding affinity (higher K_M value), which can be explained in terms of changes in interaction energies. The removal of a charge ($Asp^{-2} \rightarrow$ Ala) decreases the interaction energy between the side chain and protein backbone. However, an explanation of the result for the peptide DADEpY(L→A)IPQQG may involve either entropic or enthalpic contributions. Again, a detailed computation of

TABLE 2 Atoms that show relatively high fluctuations in the subspaces spanned by the first five eigenvectors

Residue	ev 1	ev 2	ev 3	ev 4	ev 5
Tyr^{20}	cd1,ce1,cz,oh	—	—	—	oh
Gln^{21}	—	—	cd,oe1,ne2	cg,cd,oe1,ne2	ne2
Arg^{24}	—	cd,ne,cz,nh2,nh1	—	—	—
His^{25}	o	cb,cg,cd2,ce1,ne2	nd1,ce1	ca,cb,cg,cd2,ne2,c,o	cb,cg,nd1,cd2,ce1,ne2
Glu^{26}	cg,cd,oe1,oe2	—	cd,oe2	cb,cg,cd,oe1,oe2	cd,oe1,oe2
Ala^{27}	cb,o	cb	—	—	—
Arg^{33}	—	—	—	nh2	cz,nh2
Lys^{39}	—	—	—	—	nz
Asn^{42}	—	—	—	—	nd2
Arg^{43}	—	nh1	nh1	—	—
Arg^{45}	cz,nh1,nh2	—	—	—	—
Arg^{47}	nh2	ne,cz,nh1,nh2	cz,nh2	—	ne,cz,nh2
Asp^{53}	cg,od1,od2	—	—	—	od1
Gln^{85}	—	ne2	—	—	—
Met^{114}	—	—	ce	—	—
Lys^{116}	—	o	nz	—	—
Gly^{117}	—	—	ca,c,o	—	—
Thr^{178}	—	—	—	—	—
Asp^{181}	od1,od2	od2	—	—	—
Met^{258}	—	—	—	—	sd
Leu^{260}	cd2	—	—	—	—
Gln^{262}	—	cg,ne2	—	—	—
Asp^{4}	od1,od2	—	cb,od1,cg,od2	od1,cg,od2,cb,ca,n,o	—
Asp^{2}	od2	—	—	—	—
Glu^{1}	oe1	cg,cd,oe1	oe1	cb,cg,cd,oe1,oe2	oe2,oe1
Leu^{1}	cd2	cg,cd1,cd2	—	—	cb,cg,cd2,cd1

Data are extracted from Fig. 6. The listed atoms correspond to an absolute value of the eigenvector (ev) greater than 0.085. The last three entries in the table correspond to residues in the ligand (numbering adopted from Zhang et al. (1993); see text for details).

the change in free energy has to be performed to clarify this. Asp^{-4} and Glu^{-1} show different behavior. The electrostatic interactions of Glu^{-1} with solvent molecules are stronger than those with the protein (see Table 1). For Asp^{-4}, interaction energies with the protein and solvent are the same on average (see Table 1), which explains the high mobility of Asp^{-4}, which was observed along several eigenvectors (Table 2 and Fig. 5). These results suggest that the binding affinity of the peptide could be increased by suitable substitution of Glu^{-1} and Asp^{-4} to increase the interaction energies with the protein. It has been observed experimentally that the substitutions $Glu^{-1} \rightarrow$ Ala and $Asp^{-4} \rightarrow$ Ala in DADEpYLIPQQG decrease the catalytic efficacy by 6.5- and 1.9-fold, respectively (Zhang et al., 1993). The K_M values are increased 4.7-fold for DAD(E→A)pYLIPQQG and 1.9-fold for (D→A)ADEpYLIPQQG, when compared to the parent peptide DADEpYLIPQQG. Smaller differences are found for the turnover number k_{cat}. k_{cat} is decreased 1.4-fold for DAD(E→A)pYLIPQQG, whereas similar rate constants were determined for (D→A)ADEpYLIPQQG and the parent peptide. These experimental results indicate that Glu^{-1} and Asp^{-4} are important for high catalytic efficiency. From the crystal structure it is evident that Arg^{47} located close to the substrate residues Asp^{-4} and Glu^{-1} could form favorable interactions with these residues and hence increase the binding affinity of the substrate. To further analyze the structural environment surrounding these residues, we have calculated distances between atoms in the protein and the substrate. As shown in Fig. 8, substrate residues Ala^{-3} (Fig. 8 *a*), Asp^{-4} (Fig. 8 *b*), and Leu^{1} (Fig. 8 *c*) form hydrogen bonds with atoms in the protein. Favorable interactions (i.e., short distances) are observed for Arg^{47}-(Glu^{-1}), Arg^{47}-(Asp^{-4}), and Gln^{262}-

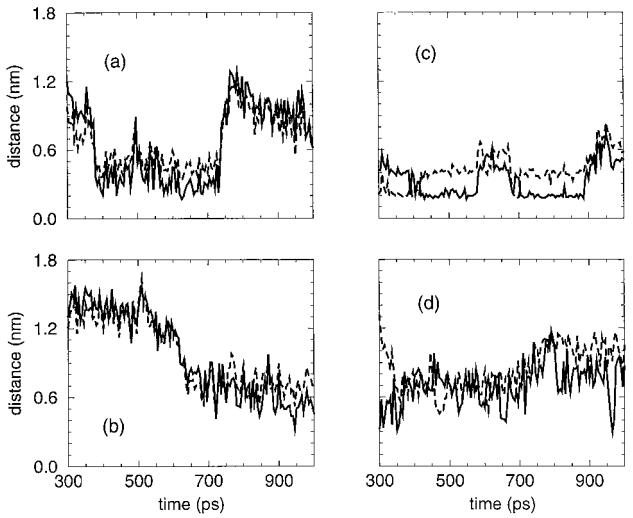

FIGURE 8 Selected distances between protein and substrate atoms. (*a*) Arg^{47} (HH22)-Ala^{-3} (O) (——) and Arg^{47} (HE)-Ala^{-3} (O) (– – –), (*b*) Arg^{45} (HH22)-Asp^{-4} (OD1) (——) and Arg^{45} (HE)-Asp^{-4} (OD1) (– – –), (*c*) Gln^{262} (HE21)-Leu^{1} (OT1) (——) and Gln^{262} (HE21)-Leu^{1} (OT2) (– – –), (*d*) Arg^{47} (HH12)-Glu^{-1} (OE2) (——) and Arg^{47} (HH21)-Asp^{-4} (OD1) (– – –).

(Leu^{1}) throughout the simulation (Fig. 8, *c* and *d*). In particular, the distance between Gln^{262} and (Leu^{1}) is relatively constant, reflecting that Leu accommodates well into the binding pocket. In the course of the simulations Asp^{4} approaches Arg^{45} within ~5 Å (Fig. 8 *c*). These findings are in excellent agreement with the experimentally measured K_M value. K_M is increased 4.7-fold for the DAD(E→A)pYLIPQQG peptide. The mutation from a negatively charged residue (Glu) to an aliphatic amino acid (Ala) results in the elimination of important electrostatic interactions with Arg^{47}. Consequently, the binding affinity of DADApYLIPQQG is reduced.

CONCLUSION

Substrate binding involves a multiplicity of factors, including changes in intermolecular interactions between the substrate, the solvent molecules, and the target protein, as well as changes in polarization, conformation, or flexibility. The analysis of the forces and motions involved in the formation of the enzyme-substrate complex is important for a detailed understanding of the residues/motif imparting substrate affinity and for revealing the structure-function relationship of phosphatases. We have therefore performed molecular dynamics simulations of PTP1B complexed with the high-affinity substrate peptide, DADEpYL, and of the free substrate. Simulations were carried out in periodic boundary conditions, using explicit SPC water, and 700-ps trajectories were used in the analysis.

Several geometrical properties, such as root mean square fluctuations, radius of gyration, and surface accessibility, show relatively large fluctuations, reflecting the inherent flexibility of the protein. These fluctuations are more pronounced than observed for other enzymes; e.g., lipases (Peters et al., 1997b). Both groups of enzymes undergo conformational changes during the enzymatic reaction. However, activation of lipases and the subsequent catalytic reaction occur at the interfacial plane of a lipid surface. PTP1B, on the other hand, belongs to the class of cytosolic enzymes acting in solution. Upon binding of the substrate, a flexible loop closes, bringing an acidic residue into place for the catalytic reaction. It has recently been determined experimentally that an equilibrium exists between the closed and open conformations. Both these experimental observations and the determined geometrical properties may suggest that PTP1B has a relatively high flexibility.

Further analysis of the concerted motions in the protein, including only the substrate and amino acids of the protein within a distance of ~15 Å from the phosphate moiety, indicates that the terminals of the substrate are rather mobile and that the N-terminal is more flexible than the C-terminal. Most fluctuations are observed in the binding pocket, where the N-terminal of the substrate is located and involves polar/charged residues. This observation can be explained in terms of the energy contributions. Binding is predominantly determined by electrostatic interactions and is enhanced by a positively charged patch surrounding the binding pocket, which guides the substrate toward the active site. The contribution of the peptide residues to the binding affinity is significantly different. For Asp^{-4} the electrostatic substrate-protein and substrate-solvent interactions are of approximately the same magnitude, suggesting that the mobility of the substrate is driven by a balance of these two contributions. Another example is Asp^{-2}, whose electrostatic interactions with the protein (on average) are stronger than those with the solvent (Table 2 and Fig. 8 *b*), suggesting that this residue is tightly bound to the protein. This is also supported by the essential dynamics analysis results, where no significant fluctuations were observed in the first 10 eigenvectors. Experimentally determined kinetic data show that the substitution of Glu^{-1} with Ala reduces the K_m value manifold, which is well explained by the interactions between Arg^{47} and Glu^{-1}.

N. P. H. Møller and H. S. Andersen are thanked for inspiring discussions.

GHP was supported by a grant from Novo Nordisk A/S and under grant 97 100 05 from the Danish Cancer Research Foundation.

REFERENCES

Ajay, and M. A. Murcko. 1995. Computational methods to predict binding free energy in ligand-receptor complexes. *J. Med. Chem.* 38:4953–4967.

Amadei, A., A. B. M. Linssen, and H. J. C. Berendsen. 1993. Essential dynamics of proteins. *Proteins Struct. Funct. Genet.* 17:412–425.

Antosiewicz, J., J. A. McCammon, and M. K. Gilson. 1994. Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* 238:415–436.

Barford, D. 1995. Protein phosphatases. *Curr. Opin. Struct. Biol.* 5:728–734.

Barford, D., A. J. Flint, and N. K. Tonks. 1994. Crystal structure of human protein tyrosine phosphatase 1B. *Science.* 263:1397–1404.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rogers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The Protein Data Bank: a computer based archival file for macromolecular structure. *J. Mol. Biol.* 112:535–542.

Boylan, J. M., D. L. Brautigan, J. Madden, T. Raven, L. Ellis, and P. A. Gruppuso. 1992. Differential regulation of multiple hepatic protein tyrosine phosphatases in alloxan diabetic rats. *J. Clin. Invest.* 90: 174–179.

Cool, D. E., and E. H. Fischer. 1993. Protein tyrosine phosphatases in cell transformation. *Semin. Cell Biol.* 4:443–453.

Davis, M. E., J. D. Madura, B. A. Luty, and J. A. McCammon. 1991. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program. *Comp. Phys. Comm.* 62:187–197.

Denu, J. M., and J. E. Dixon. 1995. A catalytic mechanism for the dual-specific phosphatases. *Proc. Natl. Acad. Sci. USA.* 92:5910–5914.

Eckstein, J. W., P. Beer-Romero, and I. Berdo. 1996. Identification of an essential acidic residue in Cdc25 protein phosphatase and a general three-dimensional model for the core region in protein phosphatases. *Protein Sci.* 5:5–12.

Fauman, E. B., and M. A. Saper. 1996. Structure an function of the protein tyrosine phosphatases. *Trends Biochem. Sci.* 21:413–417.

Fauman, E. B., C. Yuvaniyama, H. L. Schubert, J. A. Stuckey, and M. A. Saper. 1996. The x-ray crystal structure of *Yersinia* tyrosine phosphatase with bound tungstate and nitrate. *J. Biol. Chem.* 271:18780–18788.

Fisher, E. H., H. Charbonneau, and N. K. Tonks. 1991. Protein tyrosine phosphatases. *Science.* 253:401–406.

Frangioni, J. V., M. Smith, E. W. Salzman, A. Oda, and B. G. Neel. 1993. Calpain-catalyzed cleavage and subcellular relocation of protein phosphotyrosine phosphatase 1B (PTP-1B) in human platelets. *EMBO J.* 12:4843–4849.

Gilson, M. K. 1993. Multiple-site titration and molecular modeling: two rapid methods for computing energies and forces for ionizable groups in proteins. *Proteins Struct. Funct. Genet.* 15:266–282.

Goddette, D. W., T. Christianson, B. F. Ladin, M. Lau, J. R. Mielenz, C. Paech, R. B. Reynolds, S. S. Yang, and C. R. Wilson. 1993. Strategy and implementation of a system for protein engineering. *J. Biotechnol.* 28:41–54.

Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science.* 268:1144–1149.

Hunter, T. 1995. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell.* 80:225–235.

Ichiye, T., and M. Karplus. 1991. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Protein Struct. Funct. Genet.* 11:205–217.

Ide, R., H. Maegawa, R. Kikkawa, Y. Shigeta, and A. Kashiwagi. 1994. High glucose condition activates protein tyrosine phosphatases and deactivates insulin receptor function in insulin-sensitive rat 1 fibroblasts. *Biochem. Biophys. Res. Commun.* 201:71–77.

Jencks, W. P. 1975. Binding energy, specificity, and enzymic catalysis: the Circe effect. *Adv. Enzymol. Relat. Areas Mol. Biol.* 219–410.

Jia, Z., D. Barford, A. J. Flint, and N. K. Tonks. 1995. Structural basis for phosphotyrosine peptide recognition by protein tyrosine phosphatase 1B. *Science.* 268:1754–1758.

Koshland, D. E., Jr. 1958. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. USA.* 44:98–106.

Krystek, S., T. Stouch, and J. Novotny. 1993. Affinity and specificity of serine endopeptidase-protein inhibitor interactions. *J. Mol. Biol.* 234: 661–679.

Ku, T. W., F. E. Ali, and L. S. Barton. 1993. Direct design of a potent non-peptide fibrinogen receptor antagonist based on the structure and conformation of a highly constrained cyclic rigid peptide. *J. Am. Chem. Soc.* 115:8861–8862.

Lammers, R., B. Bossenmaier, D. E. Cool, N. K. Tonks, J. Schlessinger, E. H. Fischer, and A. Ullrich. 1993. Differential activities of protein tyrosine phosphatases in intact cells. *J. Biol. Chem.* 268:22456–22462.

Madura, J. D., J. M. Briggs, R. C. Wade, M. E. Davis, B. A. Luty, A. Ilin, J. Antosiewicz, M. K. Gilson, B. Bagheri, L. R. Scott, and J. A. McCammon. 1995. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian dynamics program. *Comp. Phys. Comm.* 91:57–67.

McGuire, M., R. M. Fields, B. L. Nyomba, I. Raz, C. Bogardus, N. K. Tonks, and J. Sommercon. 1991. Abnormal regulation of protein tyrosine phosphatase activities in skeletal muscle of insulin-resistant humans. *Diabetes.* 40:939–942.

Møller, N. P. H., K. B. Møller, R. Lammers, A. Kharitonenkov, E. Hoppe, F. C. Wiberg, I. Sures, and A. Ullrich. 1995. Selective down-regulation of the insulin receptor signal by protein-tyrosine phosphatases α and $\epsilon^{\star}$. *J. Biol. Chem.* 270:23126–23131.

Montsera, J., L. Chen, D. S. Lawrence, and Z.-Y. Zhang. 1996. Potent low molecular weight substrates for protein-tyrosine phosphatase. *J. Biol. Chem.* 271:7868–7872.

Peters, G. H., T. M. Frimurer, J. N. Andersen, and O. H. Olsen. 1999. Molecular dynamics simulations of protein-tyrosine phosphatase 1B. I. Ligand-induced changes in protein motions. *Biophys. J.* 77:505–515.

Peters, G. H., T. M. Frimurer, and O. H. Olsen. 1998. Electrostatic evaluation of the signature motif (H/V)CX_5R(S/T) in protein-tyrosine phosphatases. *Biochemistry.* 37:5383–5393.

Peters, G. H., S. Toxvaerd, O. H. Olsen, and A. Svendsen. 1997a. Computational studies of activation of lipases and the effect of a hydrophobic environment. *Protein Eng.* 10:137–147.

Peters, G. H., D. M. F. van Aalten, A. Svendsen, and R. Bywater. 1997b. Essential motions in enzyme binding sites and bound inhibitors: the effect of inhibitors of different chain length. *Protein Eng.* 10:148–156.

Posner, B. I., R. Fauer, J. Burgess, A. P. Bevan, D. Lachance, G. Zhang-Sun, G. Fantus, A. D. Hal, B. S. Lum, and A. Shaver. 1994. Peroxovanadium compounds: a new class of potent phosphotyrosine phosphatase inhibitors which are insulin mimetic. *J. Biol. Chem.* 269:4596–4604.

Pot, D. A., and J. E. Dixon. 1992. A thousand and two protein tyrosine phosphatases. *Biochim. Biophys. Acta.* 1136:35–43.

Ruiz, P., J. A. Pulido, C. Martinez, J. M. Carrascosa, J. Satrustegi, and A. Andres. 1992. Effect of aging on the kinetic characteristics of the insulin receptor autophosphorylation in rat adipocytes. *Arch. Biochem. Biophys.* 296:231–238.

Salemme, F. R., J. Spurlino, and R. Bone. 1997. Serendipity meets precision: the integration of structure-based drug design and combinatorial chemistry for efficient drug discovery. *Structure.* 5:319–324.

Schubert, H. L., E. B. Fauman, J. A. Stuckey, J. E. Dixon, and M. A. Saper. 1995. A substrate-induced conformational change in the *Yersinia* protein tyrosine phosphatase. *Protein Sci.* 4:1904–1913.

Stone, R. L., and J. E. Dixon. 1994. Protein-tyrosine phosphatases. *J. Biol. Chem.* 269:31323–31326.

Streuli, K., N. X. Krueger, T. Thai, M. Tang, and H. Saito. 1990. Distinct functional roles of the two intracellular phosphatase like domains of the receptor-linked protein tyrosine phosphatases LCA and LAR. *EMBO J.* 9:2399–2407.

Tonks, N. K., A. J. Flint, M. F. B. G. Gebbink, H. Sun, and Q. Yang. 1993. Transduction and protein tyrosine dephosphorylation. *Second Messengers Phosphoproteins.* 28:203–210.

Vajda, S., Z. Weng, R. Rosenfeld, and C. DeLisi. 1994. Effect of conformational flexibility and solvation on receptor-substrate binding free energies. *Biochemistry.* 33:13977–13988.

Van Gunsteren, W. F., and H. J. C. Berendsen. 1987. GROMOS: Groningen Molecular Simulation Computer Program Package. University of Groningen, Groningen, the Netherlands.

Vriend, G. 1990. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* 8:52–56.

Walton, K. M., and J. E. Dixon. 1993. Protein tyrosine phosphatases. *Annu. Rev. Biochem.* 62:101–120.

Weber, P. C., J. J. Wendoloski, M. W. Pantoliano, and F. R. Salemme. 1992. Crystallographic and thermodynamic comparison of natural and synthetic substrates bound to streptavidin. *J. Am. Chem. Soc.* 114: 3197–3200.

Wiener, J. R., S. K. Kassim, Y. Yu, G. B. Mills, and R. C. Bast. 1996. Transfection of human ovarian cancer with the HER-2/neu receptor tyrosine kinase induces a selective increase in PTPH1, PTP1B, and PTPα expression. *Gynecol. Oncol.* 61:233–240.

Wlodek, S. T., J. Antosiewicz, J. A. McCammon. 1997. Prediction of titration properties of structures of a protein derived from molecular dynamics trajectories. *Protein Sci.* 6:373–382.

Zanke, B., J. Squire, H. Griesser, M. Henry, H. Suzuki, B. Patterson, M. Minden, and T. W. Mak. 1994. A hematopoietic protein tyrosine phosphatase (HePTP) gene that is amplified and overexpressed in myeloid malignancies maps to chromosome 1q32.1. *Leukemia.* 8:236–244.

Zhang, Z.-Y., D. Maclean, D. J. McNamara, T. K. Sawyer, and J. E. Dixon. 1994a. Protein tyrosine phosphatase substrate specificity: size and phosphotyrosine positioning requirements in peptide substrates. *Biochemistry.* 33:2285–2290.

Zhang, Z.-Y., A. M. Thieme-Sefler, D. Maclean, D. J. Mcnamara, E. M. Dobrusin, T. K. Sawyer, and J. E. Dixon. 1993. Substrate specificity of the protein tyrosine phosphatases. *Proc. Natl. Acad. Sci. USA.* 90: 4446–4450.

Zhang, Z.-Y., Y. Wang, and J. E. Dixon. 1994b. Dissecting the catalytic mechanism of protein-tyrosine phosphatases. *Proc. Natl. Acad. Sci. USA.* 91:1624–1627.

Zhang, Z.-Y., and L. Wu. 1997. The single sulfur to oxygen substitution in the active site nucleophile of the *Yersinia* protein-tyrosine phosphatase leads to substantial structural and functional perturbations. *Biochemistry.* 36:1362–1369.

Zheng, X. M., Y. Wang, and J. C. Pallen. 1992. Cell transformation and activation of pp60c-src by overexpression of a protein tyrosine phosphatase. *Nature.* 359:336–339.