Modeling Studies of the Interactions between the Insulin Receptor Kinase Domain and Protein Tyrosine Phosphatase 1B

Nicholas R. Glover† and Alan S. Tracey*

Contribution from the Department of Chemistry and Institute of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

Received December 17, 1998. Revised Manuscript Received February 19, 1999

Abstract: The sulfotyrosine peptide, IRK1154, is based on the corresponding phosphotyrosine segment of the insulin receptor kinase domain (IRK) and is a known inhibitor of the function of the protein tyrosine phosphatase, PTP1B. Two-dimensional NMR spectroscopy, in the transferred nuclear Overhauser effect (NOE) enhancement experiment, was used to obtain information concerning the bound structure of this peptide. Computer-simulated docking experiments, followed by molecular dynamics simulations in a fully hydrated model, provided information concerning the site-specific interactions influencing the bound peptide. Using the structural and orientational information from the NMR studies as a guide, together with the X-ray coordinates for PTP1B and IRK, a detailed model of the binding of these two proteins was developed. The interface between the two entities is described, and the sites of positive interactions are identified. Potential sources of destabilizing interactions, necessary for dissociation of the two enzymes, were also found.

Introduction

Insulin receptor function, to a large extent, is determined by the interplay between the actions responsible for the phosphorylation state of the receptor. In its active form, the insulin receptor modifies the activity of a variety of enzymes and is important in transport processes that control protein and lipid metabolism.[1] The insulin receptor is a heterotetrameric protein consisting of two α and two β subunits linked by disulfide bonds to form a $\beta-\alpha-\alpha-\beta$ structure.[2] In the presence of insulin, the insulin receptor utilizes ATP in an autophosphorylation reaction to phosphorylate itself on tyrosine residues. The now active insulin receptor has kinase activity and utilizes this capability to phosphorylate a number of substrates and, via this process, initiate various cellular processes.[2] The level of activity of the insulin receptor is tightly controlled in vivo by control of the phosphate levels via function of regulatory enzymes. As a consequence, interest has focused on the potential role of the protein tyrosine phosphatases (PTPases) in insulin receptor function.[3−6] Recent studies have provided credence to this hypothesis and have shown that two PTPases, leucocyte antigen related (LAR) and protein tyrosine phosphatase 1B (PTP1B), undergo strong interactions with the insulin receptor.[4,7−10]

There are at least six β subunit tyrosine phosphorylation sites on the insulin receptor: one in the juxtamembrane domain (Y972), two near the carboxy terminus (Y1328, Y1334), and three in the kinase domain (Y1158, Y1162, and Y1163). Phosphorylation of the insulin receptor kinase domain (IRK) is thought to be critical to biological function.[1,2,9] Furthermore, it appears that phosphorylation of tyrosine residues Y1158, Y1162, and Y1163 is required for association of the kinase domain with PTP1B.[8,9] There is a significant change in structure within the kinase domain upon phosphorylation. In its nonphosphorylated form, IRK is folded in a manner that withdraws the activation loop (A-loop) from the medium and lays it onto the kinase domain surface so that it can make only restricted contact with substrates.[11] Upon phosphorylation, conformational changes occur in the A-loop so that access to the loop tyrosines is unhindered.[12] In particular, from the crystal structure,[13] pY1158 is particularly well isolated from other residues of the IRK domain (Figure 1), and this may indicate an important role for this residue in substrate binding.[13] This has been supported by studies with a peptide model of the A-loop that revealed that pY1158 and pY1162 are preferentially dephosphorylated in the presence of PTP1B.[14]

The eight-amino-acid-residue peptide (Thr1154Arg1155-Asp1156Ile1157sTyr1158Glu1159Thr1160Asp1161) is based on the residues 1154−1161 of the kinase domain A-loop and is designated IRK1154. IRK1154 is a 16 μM inhibitor of the function of PTP1B[15] and will have a residence time in the active site of PTP1B of a few tens of milliseconds. NMR techniques applicable to dynamic systems were used to obtain two-

† Present address: University Medical Discoveries Inc., 100 International Blvd., Toronto, ON M9W 6J6, Canada.

(1) White, M. F.; Kahn, C. R. *J. Biol. Chem.* **1994**, *269*, 1−4.
(2) Lee, J.; Pilch, P. F. *Am. J. Physiol.* **1994**, *266*, C319−C334.
(3) Drake, P. G.; Bevan, A. P.; Burgess, J. W.; Bergeron, J. J. M.; Posner, B. I. *Endocrinology* **1996**, *137*, 4960−4968.
(4) Li, P.-M.; Zhang, W.-R.; Goldstein, B. J. *Cell. Signalling* **1996**, *8*, 467−473.
(5) Kenner, K. A.; Anyanwu, E.; Olefsky, J. M.; Kusari, J. *J. Biol. Chem.* **1996**, *271*, 19810−19816.
(6) Kulas, D. T.; Goldstein, B. J.; Mooney, R. A. *J. Biol. Chem.* **1996**, *271*, 748−754.
(7) Ahmad, F.; Goldstein, B. J. *J. Biol. Chem.* **1997**, *272*, 448−457.
(8) Bandyopadhyay, D.; Kusari, A.; Kenner, K. A.; Liu, F.; Chernoff, J.; Gustafson, T. A.; Kusari, J. *J. Biol. Chem.* **1997**, *272*, 1639−1645.
(9) Seely, B. L.; Staubs, P. A.; Reichart, D. R.; Berhanu, P.; Milarski, K. L.; Saltiel, A. R.; Kusari, J.; Olefsky, J. M. *Diabetes* **1996**, *45*, 1379−1385.
(10) Byon, J. C. H.; Kusari, A. B.; Kusari, J. *Mol. Cell. Biochem.* **1998**, *182*, 101−108.
(11) Hubbard, S. R.; Wei, L.; Ellis, L.; Hendrickson, W. A. *Nature* **1994**, *372*, 746−754.
(12) Hubbard, S. R. *EMBO J.* **1997**, *16*, 5572−5581.
(13) Hubbard, S. R. Personal Communication, 1998.
(14) Ramachandran, C.; Aebersold, R.; Tonks, N. K.; Pot, D. A. *Biochemistry* **1992**, *31*, 4232−4238.
(15) Desmarais, S.; Jia, Z.; Ramachandran, C. *Arch. Biochem. Biophys.* **1998**, *354*, 225−231.

10.1021/ja984354q CCC: $18.00
Published on Web 03/31/1999

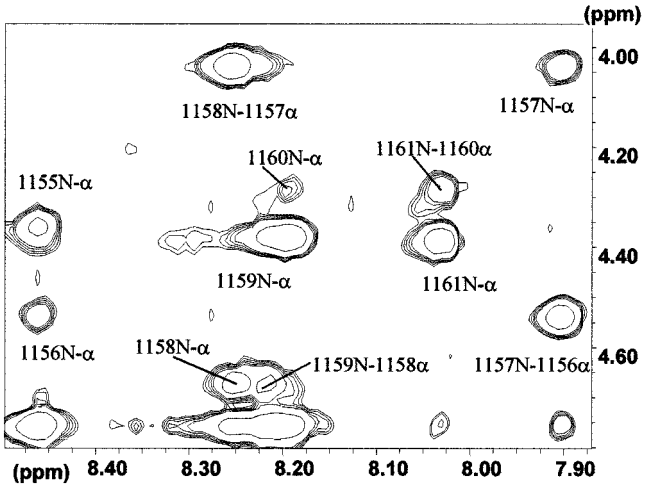

Figure 1. NH−CHα region of the transferred NOE spectrum of IRK1154 in the presence of GST-PTP1B. The concentration of IRK1154 was 2.0 mM, with a GST-PTP1B concentration of 0.2 mM. The spectrum was recorded in H_2O/D_2O (9:1) at a pH of 6.9, a temeperature of 298 K, and a mixing time of 150 ms. Protons are designated by the attached carbon or nitrogen of the numbered residue. Various intraresidue and sequential transferred NOE cross-peaks are indicated.

dimensional transferred NOE spectra of the bound inhibitor. These spectra provided structural information that allowed the bound-state structure of the peptide to be determined. Subsequent molecular and dynamics simulations in a fully hydrated model[16] of the inhibitor/PTP1B complex provided insight into the interactions between the two entities. The structural and orientational information obtained was then utilized in the development of a model of the PTP1B/IR kinase domain complex.

Materials and Methods

Peptide and Protein. GST-PTP1B, and the sulfotyrosine IRK-derived sequence TRDI(sY)ETD, were kindly provided by Merck Frosst Centre for Therapeutic Research, Montreal, Canada. The N-terminal domain of PTP1B as provided (residues 1−298) was expressed as a glutathione *S*-transferase fusion protein, GST-PTP1B, in *Escherichia coli*.[17] Protein concentrations were either assayed according to the method of Bradford[18] or measured by UV spectroscopy at 280 nm.[17,19] The purity of the protein samples was established by SDS−PAGE. The activity of the protein samples was assayed prior to the NMR experiments according to the established procedure.[17,20] The purity of the IRK peptide was checked by ^{1}H NMR spectroscopy. It was used as supplied.

NMR Sample Preparation. All samples for NMR spectroscopy were prepared as aqueous solutions in a buffer containing 2 mM EDA or 2 mM phosphate, pH ∼6.5−7.0, 0.20 M NaCl, 0.1 mM bismercaptoethylsulfone (BMS) or 0.1 mM dithiothreitol (DTT), and 5% D_2O/95% H_2O. Samples of the free peptides were prepared as 2 mM solutions. The pH of each peptide solution was adjusted to an appropriate value by adding a small amount of either HCl or NaOH. The pHs of the samples were measured directly in the NMR tube with a glass electrode (Bradley James, Inc.) and calibrated against freshly prepared standard buffers at pH 4.0 and 7.0. Samples of the peptide in the presence of GST-PTP1B were prepared by titrating an aliquot of the protein into a stock solution of the peptide at the appropriate pH. All samples were filtered into the NMR tube through Celite over a glass wool plug, and the tube was then filled with nitrogen gas and sealed. All chemicals used in the study were reagent grade or better.

NMR Experiments. Nuclear magnetic resonance experiments were carried out on a Bruker AMX600 spectrometer operating at a ^{1}H frequency of 600.13 MHz with the sample nonspinning. All two-dimensional experiments were recorded in phase-sensitive mode using time-proportional phase incrementation (TPPI).[21] The spectral width was equal in both dimensions and varied between 7240 and 8196 Hz, with the carrier frequency set at the water resonance. Except where noted, water suppression was achieved by the WATERGATE gradient-echo method using a 3−9−19 pulse sequence.[22] NOESY experiments[23] were performed on the free peptides at 298 K with typical mixing times of 50, 100, 150, 300, and 600 ms. Transferred NOE experiments of the peptide/protein complexes were carried out with the standard NOESY pulse sequence at 298 K.[23] Transferred NOESY spectra were recorded at a series of mixing times, such as 25, 50, 75, 100, 150, 300, and 600 ms. ROESY and transferred ROESY experiments were performed at 298 K with spin-lock fields of 100−300-ms duration.[24] Water suppression during the ROESY experiments was achieved by continuous low-power presaturation (2 s) during the relaxation delay, with the transmitter frequency set on the H_2O resonance.[25] Residue-specific assignments of peptide resonances were established by analysis of the cross-peak patterns of double-quantum filtered correlation spectroscopy (DQF−COSY) and TOCSY spectra.[24,26] The DQF−COSY experiments,[27,28] and were obtained using presaturation (2 s) during the recycling delay,[25] while the TOCSY spectra were acquired with a spin-locked MLEV-17 sequence (150 ms). Sequence assignments were established from the transferred NOE (NOESY) spectra. The majority of the spectra were collected as 512 t_1 values using 40−128 scans/increment, each with 2K complex data points in t_2 and a relaxation delay of 1−2 s (total experimental time ∼10−22 h/experiment). The data were processed either directly on the AMX600 spectrometer using the standard UXNMR software (Bruker) or off-line on a SGI Indigo2 workstation using XWIN NMR (Bruker) or FELIX software.[29] Data sets were apodized in both directions by either a 30−60° shifted squared sine-bell or a Lorentzian to Gaussian transformation function. Data sets were zero-filled to 1−2 K in the t_1 dimension. Following Fourier transformation and phase correction, a polynomial fifth-order function was used for baseline correction in both dimensions prior to volume integration. All chemical shifts were referenced to an external sample of DSS at 0.0 ppm.

Structure Determination, Molecular Modeling, and Molecular Dynamics. The structural refinement, modeling, and dynamics procedures used for this work have been described in detail elsewhere.[16] They are summarized below.

The NOESY cross-peak intensities were used to apply distance restraints according to very weak (1.8−5.0 Å), weak (1.8−4.0 Å), medium (1.8−3.3 Å), and strong (1.8−2.8 Å) cross-peaks. The ortho proton internuclear distance (2.5 Å) of the sulfotyrosine ring was used

(16) Glover, N. R.; Tracey, A. S. *Biochemistry*, in press.
(17) Huyer, G.; Liu, S.; Kelly, J.; Moffat, J.; Payette, P.; Kennedy, B.; Tsaprailis, G.; Gresser, M. J.; Ramachandran, C. *J. Biol. Chem.* **1997**, *272*, 843−851.
(18) Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248−254.
(19) Mach, H.; Middaugh, C. R.; Lewis, R. V. *Anal. Biochem.* **1992**, *200*, 74−80.
(20) Nxumalo, F. Influence of vanadium complexes on the catalytic activity of the protein tyrosine phosphatases. M.Sc. Thesis, Simon Fraser University, Canada, 1997.
(21) Marion, D.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967−974.
(22) Piotto, M.; Saudek, V.; Sklenár, V. *J. Biomol. NMR* **1992**, *2*, 661−665.
(23) Jeener, J.; Meier, B. H.; Bachman, P.; Earnst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546−4553.
(24) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355−360.
(25) Wider, G.; Macura, S.; Ernst, R. R.; Wüthrich, K. *J. Magn. Reson.* **1984**, *56*, 207−234.
(26) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521−558.
(27) Piatini, U.; Sorensen, O. W.; Ernst, R. R. *J. Am. Chem. Soc.* **1982**, *104*, 4, 6800−6801.
(28) Rance, M.; Sorensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479−485.
(29) Biosym/MSI. *InsightII Molecular Modeling System* (v. 95.0); Biosym/MSI: San Diego, CA, 1995.

for internal calibration of peak intensities. The NOE distance geometry constraints were used in the Molecular Simulations computational package DISCOVER[30] to ascertain possible molecular conformations of the bound peptide.

The consistent valence force field (CVFF) as implemented in the Molecular Simulations software package[29,30] was used for the molecular modeling and dynamics simulations. Starting geometries were minimized in the absence of NMR restraints. One hundred steps of steepest descent followed by 1000 steps of conjugate gradient minimization were initially carried, and out then the NMR constraints were applied using the distance geometry algorithm DGII in Insight_II.[29] Structures were obtained using conjugate gradient minimization of a full-matrix error function. Structural irregularities were then removed using energy minimization to yield the final conformation. Five hundred structures, all satisfying the NMR constraints, were determined.

X-ray coordinates of the catalytically inactive Cys215Ser mutant were modified by replacement of the Ser OH with the SH of the native enzyme. Crystallographically observed water was retained in the model, and the protein was minimized with 2500 steps of deepest descent and then minimized further by the method of conjugate gradients until the maximum energy gradient was less than 0.01 kcal/Å.

Docking of the inhibitor into the active site was carried out with the program DOCK[31] using 500 cycles of rigid body minimization. The 500 calculated structures were docked and ranked according to the minimum of the energies of the geometric and force field parameters.

A three-layer solvent cage was built around the docked peptide and associated PTP1B residues. An inner semisphere of radius 15 Å was built and then surrounded by an adjacent layer of water 10 Å thick. The solvent cage was completed by a final 5-Å-thick outer layer of water. The outer layer of water was minimized and then fixed. This prevented evaporation during subsequent calculations. This restraint on the outer layer of water was not observed to significantly influence the energy or motion of the active site residues of the bound inhibitor.

The energy of the hydrated system was minimized in a three-step process prior to molecular dynamics simulations. The hydrogen atoms of the solvent shell were first minimized to 1 kcal/Å convergence in 1000 steps of steepest descent and to 0.1 kcal/Å convergence in 5000 cycles of conjugate gradient minimization. No cross terms or Morse potentials were included at this stage. With this stage completed, all solvent atoms were minimized by a similar process. Finally, full minimization was carried out in three stages.

With Morse potentials and cross terms being neglected, 10 000 cycles of steepest descent to 0.1 kcal/Å convergence were carried out. The potentials and cross terms were then turned on for an additional 10 000 cycles. Finally, 25 000 iterations of conjugate gradients to a convergence of 0.0001 kcal/Å were carried out. For this final cycle, the side chains of the active site residues, the inhibitor residues, and the water molecules, except for the outer layer, were free to move. The NMR restraints on the peptide structure were maintained during the minimization procedure.

During the molecular dynamics simulations,[30] the active site side chains were not restrained, while the backbone atoms were either fully restrained or subjected to a harmonic restraining function of 500 kcal/Å. A 1-fs time step was employed using the Verlat leapfrog algorithm.[32,33] The system was slowly warmed from 0 to 300 K and then equilibrated for 20 ps. Production dynamics simulations were then carried out for an additional 100 ps. A constant temperature was maintained by coupling the system to a thermal bath. Molecular trajectories were recorded for every picosecond of production dynamics. The ANALYSIS module of Insight_II[29] was used to examine time-dependent fluctuations in energy, distances, and torsion angles during the dynamics trajectory.

In cases where surface areas of specified groups were required, they were calculated according to the method of Lee and Richards,[34] computed using the program NACCESS.[35]

Results and Discussion

NMR Spectroscopy. The one-dimensional NMR spectrum of IRK1154 in the absence of GST-PTP1B is well-resolved, with sharp resonances suitable for two-dimensional NMR experiments. Sequence-specific assignments of the signals were made from considerations of the TOCSY, DQF−COSY, and transferred NOESY spectra.[36] The uncomplexed form of the IRK1154 peptide in aqueous solution gave rise to only a few, low-intensity NOE and ROE cross-peaks. This suggests that the unbound peptide lacks a predominant structural preference in solution. The NOESY cross-peaks observed for the free peptide were estimated to be weak enough to be neglected.

In the presence of GST-PTP1B, the situation changed dramatically, and a number of strong NOE cross-peaks were observed in the 2D NOESY spectrum (Figure 1). A similar transfer of NOE information was not seen in the control experiments performed with GST alone, or with proteins of molecular mass comparable to that of GST-PTP1B (alkaline phosphatase or bovine serum albumin). Similarly, the addition of excess vanadate to a solution containing GST-PTP1B/IRK1154 significantly attenuated the intensity of the observed transferred NOE cross-peaks. Vanadate is a well-known reversible inhibitor of PTPases.[17] These findings suggest that the cross-peaks seen in the presence of GST-PTP1B result from specific binding of the peptide in the active site of the enzyme and are not a consequence of nonspecific binding or viscosity-induced effects. Only cross-peaks between the ligand resonances were observed; no protein−ligand or protein−protein cross-peaks were observed. This can be accounted for by the fact that the very broad signals arising from the efficient NMR relaxation pathways afforded by proteins of this size make observation of such cross-peaks exceedingly difficult.[36] The NOE cross-peaks that were observed were of the same sign as the diagonal peaks and were, therefore, characteristic of enzyme-induced signals.[37]

The pattern of NOE cross-peaks indicated that IRK1154 bound to PTP1B in a predominantly extended β-strand conformation. In general, the cross-peaks observed in the transferred NOESY spectra corresponded to either intraresidue (e.g., NH_i/$CH\alpha_i$ and NH_i/$CH\beta_i$) or sequential (e.g., NH_{i+1}/$CH\alpha_i$ and NH_{i+1}/$CH\beta_i$) transferred NOEs (Figure 1). The fact that the intensity of the intraresidual NH_i/$CH\alpha_i$ transferred NOEs was weaker relative to the intensities of the sequential NH_{i+1}/$CH\alpha_i$ transferred NOEs provided further evidence for an extended structure.[38] For example, the sequential 1157NH−1156CHα, 1158NH−1157CHα, and 1161NH−1160CHα cross-peaks were more intense than the corresponding 1156NH−CHα, 1157NH−CHα, and 1160NH−CHα resonances. An exception to this pattern was observed between sTyr1158 and Glu1159, where the NH−CHα resonances were more intense than the corresponding 1159NH−1158CHα cross-peak. This suggested a different secondary structure at the sTyr1158 position, possible a twist in the strand. Further support for this was provided by the observation of sequential NH_{i+1}/NH_i cross-peaks between Asp1156−Ile1157 and sTyr1158−Ile1157. A weaker NH_{i+1}/

(30) Biosym/MSI. *Discover* (v. 2.9.7./95.0/3.0.0); Biosym/MSI: San Diego, CA, 1995.
(31) Oshiro, C. M.; Kuntz, I. D.; Dixon, J. S. *J. Comput.-Aided. Mol. Des.* **1995**, *9*, 113−130.
(32) Leach, A. R. *Molecular Modelling: Principles and Applications*; Addison-Wesley Longman, Ltd.: Harlow, Essex, UK, 1996.
(33) van Gunsteren, W. F.; Mark, A. E. *Eur. J. Biochem.* **1992**, *204*, 947−961.
(34) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379−400.
(35) Hubbard, S. J.; Thornton, J. M. *NACCESS*; Department of Biochemistry and Molecular Biology, University College: London, 1993.
(36) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.
(37) Ni, F. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 517−606.
(38) Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519−538.

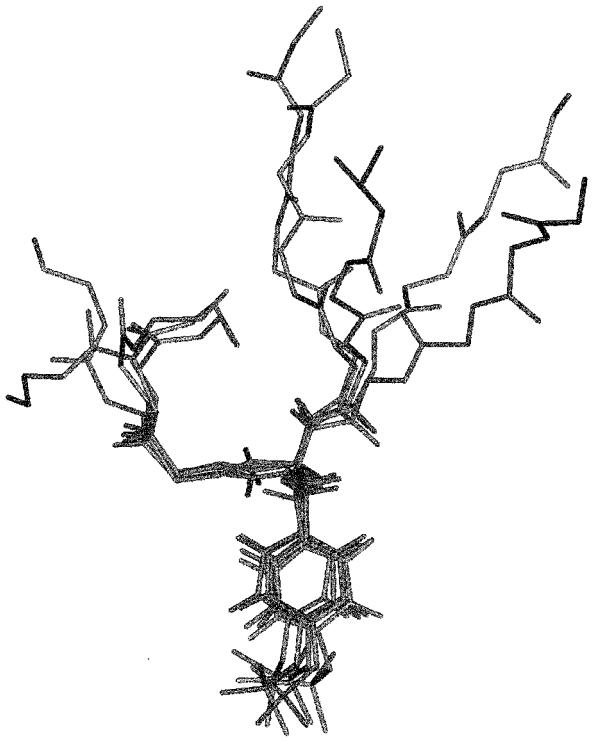

Figure 2. Ensembles of distance geometry-calculated structures for PTP1B-bound IRK1154. The five lowest energy conformers following refinement of the complete ensemble of structures against the active site of PTP1B are shown. For clarity, only the backbone atoms of residues Thr1154 to Asp1161 and the side chain of sTyr1158 are depicted. The distribution reveals excellent structural homology for residues Asp1156 to Glu1159 and poor conformational selectivity outside this region.

NH_i cross-peak was also noted between Thr1160 and Asp1161, possibly indicative of a second twist at the C-terminus. The N-terminal residues (Thr1154 and Arg1155) did not provide much in the way of NOE connectivities, and this suggests that these residues are highly mobile, even when the peptide is bound. In fact, the pattern of 2D cross-peaks suggested that only the residues Asp1156 to Glu1159 are highly restricted by the enzyme, while the remaining residues are relatively unrestrained. The time dependence of the buildup of transferred NOE cross-peaks volumes indicated, up to 150 ms mixing time, that spin diffusion was not significant in these experiments. Analysis of a ROESY spectrum also did not indicate the presence of relayed cross-peaks.

Bound-State Structure. The transferred NOESY spectrum provided a set of 44 distance restraints for use in determining the bound-state structure of the peptide. On the basis of these restraints, 500 structures were generated using the distance geometry procedure. All of the conformers exhibited distance restraint violations less than 0.5 Å, and all had energetically favorable ϕ,ψ angles. Additionally, this ensemble was characterized by segments of amino acid residues that displayed high structural homology and by segments displaying little such homology. Similar behavior has been observed for a peptide-based inhibitor that is, by about 3 orders of magnitude, a better inhibitor of PTP1B than is IRK1154.[16] The 500 structures were sequentially docked into the active site of PTP1B and ranked according to the most energetically favorable fit. An ensemble of the lowest energy conformers determined by the screening process is depicted in Figure 2. As indicated, there is good agreement between the structures for residues Asp1156 to Glu1159, with an RMSD of 1.36 ± 0.4 Å.

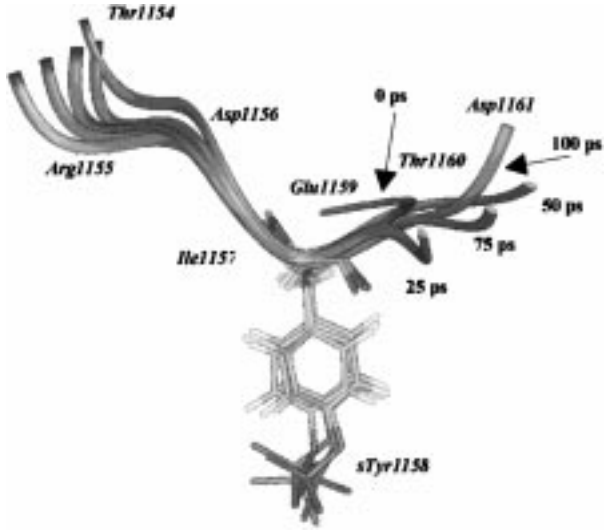

Figure 3. Snapshots of IRK1154 backbone conformations during molecular dynamics. IRK1154 backbone conformers were extracted from the trajectory at time intervals of 0, 25, 50, 75, and 100 ps. Backbone is represented in ribbon form. The relative orientation of the side chain of sTyr1158 is shown for clarity. The approximate positions of the remaining side chains are indicated for illustrative purposes. The backbone atoms of Asp1156, Ile1157, sTyr1158, and Glu1159 are more restrained during the simulation.

On the basis of the set of conformers that had the most favorable interaction score with the enzyme, the peptide binds to PTP1B in an extended β-strand conformation that is twisted around the sTyr1158 amino acid. The backbone twist exposes the acidic side chains of Asp1156 and Glu1159 while defining a sTyr1158 conformation suitable for burial within the active site pocket. The two extreme C-terminal residues adopt a slight twist around Thr1556, which directs Asp1161 away from the other acidic residues. The N-terminal residues were less well-defined due to a paucity of NOE cross-peaks, indicative of a lack of restriction when bound to the enzyme.

The evolution of the production phase dynamics was consistent with a stabilized system with minimal fluctuations in total energy being exhibited as a function of time. There was a spike in energy at about 25 ps of the simulation that arose because of unfavorable contacts between Arg1155 of IRK1154 and surface elements R45 and K41 of PTP1B. Subsequent reorientation alleviated this bad contact, and the remainder of the trajectory was stable, with an acceptable fluctuation in temperature as a function of time.[32,39]

Molecular dynamics "snapshots" of the conformation of the IRK1158−PTP1B complex as a function of time revealed a distribution of peptide conformations virtually indistinguishable from that shown in Figure 2. The time course fluctuation in the backbone conformation of IRK1154 revealed regions of highly restrained conformational selectivity and other regions where little conformational selectivity was imposed by the binding. The fluctuation over the simulation time is shown in Figure 3. The modeling simulation suggests that residues Asp1156 to Glu1159 of IRK1154 are strongly restrained by the enzyme during the course of the dynamics. Conversely, the N- and C-terminal residues undergo considerable time-dependent fluctuation. Some structural changes in the peptide promoted more favorable peptide−protein interaction energies. In particular, the large deviation for Arg1155 was associated with motion of the residue away from positively charged PTP1B amino acids

(39) Daggett, V.; Levitt, M. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 353−380.

Table 1. PTP1B−IRK1154 Interaction Energies (kcal/mol)

	Thr1154	Arg1155	Asp1156	Ile1157	sTyr1158	Glu1159	Thr1160	Asp1161
vdW[a]	−9.1	−9.8	−2.2	−4.7	−31.6	−2.6	−4.9	−0.2
elect.[b]	−8.8	136.8	−183.9	−1.1	−8.3	−18.9	−2.5	12.4
TIE	−17.9	127.0	−186.1	−5.8	−39.9	−21.5	−7.4	12.2

[a] Van der Waals interaction energy. [b] Electrostatic interaction energy. [c] Total interaction energy.

Table 2. Contribution to the Total Interaction Energy between IRK1154 and PTP1B from Various Component Energies

interaction	energy (kcal/mol)
total interaction energy with both solvent and enzyme	−232.3
total interaction energy with solvent	−92.8
total interaction energy with enzyme	−139.5
total van der Waals interaction energy with enzyme	−65.2
total electrostatic interaction energy with enzyme	−74.3

flanking the active site. Similarly, the side chain of Ile1157 was found to collapse onto the surface of the enzyme to promote van der Waals interactions, while sTyr1158 reoriented in the binding pocket to maximize favorable contacts. The active site of PTP1B itself was very stable during MD. The motional characteristics suggested by the dynamics simulations are in full accord with the NMR observations, where regions of restricted motion and regions of highly unrestricted motion were indicated.

Residue−Residue Interactions: IRK1154−PTP1B. Table 1 summarizes the energies of interaction between IRK1154 and the active site components of PTP1B on a residue by residue basis. These energies serve to qualitatively describe the binding characteristics and are not equivalent to free energies of binding, as entropic contributions are not included.[40] These energies are, however, very useful because they help identify contact interactions and give an indication of their strength. Table 2 provides an overview of the contributions that various energetic components make to the overall interaction energy. Interestingly, the van der Waals contribution to stabilization represented almost half the total stabilization energy. This is a much higher proportion than that found for other sulfotyrosine inhibitors we have studied.[41]

Figure 4 depicts IRK1154 bound in the active site of PTP1B. It binds to PTP1B in an extended β-strand conformation along the length of the PTP1B active site. The peptide adopts a pronounced twist about sTyr1158 which, itself, is buried within the binding site pocket. Residues Asp1156 to Glu1159 of IRK1154 make intimate contact with PTP1B amino acids flanking the binding pocket. The terminal amino acids, Thr1154 and Asp1161, extend away from active site and do not significantly interact with the enzyme (Figure 4). The peptide−PTP1B association buries 622 Å^2 of the surface area accessible to the solvent in the free peptide (44%) and 472 Å^2 of the surface area accessible to solvent in unliganded PTP1B. The majority of the buried surface area of the peptide corresponds to residues Asp1156, Ile1157, and, most significantly, sTyr1158.

(a) IRK1154 Backbone Interactions with PTP1B. The only stable hydrogen bond to the backbone of IRK1154 noted during the MD simulations was between the side-chain carboxyl of Asp48 and the amide proton of sTyr1158 (Table 3). Additional, but weaker, water-mediated hydrogen bonds may form intermittently throughout the MD simulations, but they do not persist throughout the trajectory. This suggests that the secondary structure of the bound IRK1154 peptide is determined principally through side-chain−side-chain interactions with PTP1B. This is substantiated to a large extent by the summary of potential hydrogen-bonding interactions suggested in Table 3.

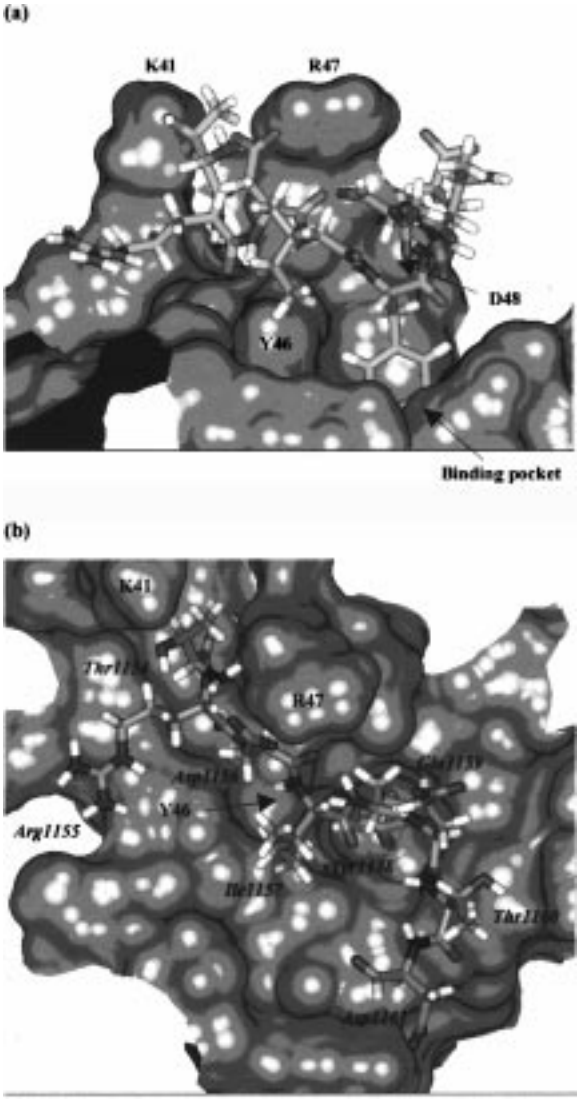

Figure 4. Representation of IRK1154 bound to the active site of PTP1B. IRK1154 is depicted as a stick model, while the surface of PTP1B is depicted as a Connolly surface (probe radius of 1.4 Å).[58,59] PTP1B residues are indicated in bold type; IRK1154 residues are indicated in italics. (a) Side-on view; (b) aerial view (sTyr1158 is buried in the pTyr recognition pocket located on the surface of the enzyme). Asp1156 and Glu1159 make electrostatic contacts to Arg47 of PTP1B. Ile1157 lies across the surface of the enzyme. The N- and C-termini make limited contact with PTP1B.

(b) IRK1154 Side-Chain Interactions with PTP1B. (i) Thr1154 and Arg1155. No NOE cross-peaks were observed for these residues in the transferred NOE spectra. This suggested that there was minimal restriction in the motion of these two residues when the peptide was bound. In fact, the interaction energy (Table 1) calculated for Arg1155 suggests that this residue is actually detrimental to the overall binding of the peptide, and, indeed, truncation of the 1154 and 1155 residues

(40) Kollman, P. A. *Curr. Opin. Struct. Biol.* **1994**, *4*, 240−245.
(41) Glover, N. R.; Tracey, A. S. Structure, modelling and molecular dynamics studies of the inhibition of protein tyrosine phosphatase 1B by sulfotyrosine peptides. Unpublished work, 1998.

Table 3. Principal PTP1B−IRK1154 Hydrogen Bonds[a]

donor atom (D)	hydrogen atom (H)	acceptor atom (A)	distance H−A (Å)	angle D−H−A (deg)
R47:NH1	R47:HH12	D1156:OD2	1.75	134.36
R47:NH2	R47:HH21	E1159:OE1	3.86	135.52
F182:N	F182:HN	sY1158:OH	3.61	141.13
F182:N	F182:HN	sY1158:O2	3.77	172.41
S216:N	S216:HN	sY1158:O1	2.57	123.58
A217:N	A217:HN	sY1158:O1	1.95	168.53
A217:N	A217:HN	sY1158:O3	3.57	128.61
G218:N	G218:HN	sY1158:O1	3.51	138.24
I219:N	I219:HN	sY1158:O3	3.01	126.7
G220:N	G220:HN	sY1158:O2	3.21	100.74
G220:N	G220:HN	sY1158:O3	1.84	153.36
R221:N	R221:HN	sY1158:O2	2.16	169.36
R221:N	R221:HN	sY1158:O3	3.9	130.19
R221:NE	R221:HE	sY1158:O1	3.6	151.52
R221:NE	R221:HE	sY1158:O2	2.54	153.27
R221:NH2	R221:HH22	sY1158:O1	3.43	136.48
R221:NH2	R221:HH22	sY1158:O2	3.72	131.6
Q266:NE2	Q266:HE22	sY1158:O2	3.89	146.94
sY1158:N	sY1158:HN	D48:OD2	3.32	132.31

[a] Hydrogen bonds were determined following energy minimization to convergence. A 4.0-Å cutoff is used. Hydrogen bonds are defined according to a relaxed acceptor−donor distance and D−H−A geometry protocol,[57] consistent with the resolution of the PTP1B X-ray crystal structure (2.8 Å).[43]

results in an enhancement of the ability of the residual IRK1156 peptide to inhibit PTP1B (Table 4). The modeling suggests that Arg1155 makes several unfavorable repulsive electrostatic contacts with positively charged residues flanking the PTP1B binding pocket. Most notably, Arg1155 comes into close contact with K36, K41, R45, and R47 of PTP1B. Interestingly, the interaction energy calculations suggested that Thr1154 at P(−4) has a fairly favorable overall interaction with the enzyme (Table 1). However, no specific interaction was identified that would restrain this residue within the context of the active site.

(ii) Asp1156 and Ile1157. Asp1156 is important for effective binding, and Ile1157 is virtually essential for the binding of IRK1154 to PTP1B. Inhibition data shows that the truncation of Asp1156 yields a 3-fold reduction in binding potency, while the deletion of Ile1157 results in near abolition of binding (Table 4).

The molecular modeling provides a succinct explanation for the results obtained in the inhibition studies. The acidic side chain of Asp1156 makes a strong ion pair with the side chain of R47 (Figure 4). It is possible that, if this aspartate at P(−2) were replaced by a glutamate residue, it could make long contact with K41.[41] The aspartate side chain is too short to reach to the K41 position. However, since there is no acidic residue at the P(−1) position, access to R47 is not restricted. This enables Asp1156 to reach across the surface of PTP1B and form a strong contact with R47. In essence, Asp1156 makes the most favorable interaction that is available to it within the context of the active site.

The aliphatic side chain of Ile1157 makes a number of favorable interactions with PTP1B residues, the most notable of which is a strong van der Waals contact with the backbone of Y46. Favorable van der Waals contacts with R45 and D48 are attenuated by repulsive electrostatic contacts with these residues. In addition, Ile1157 makes weak electrostatic interactions with R47 and K116, as well as with K36 and K41. The net interactions are highly positive and lead to hydrophobic collapse withing the binding pocket. The modeling suggests that, upon collapse, the bulky nonpolar side chain of Ile1157 displaces several loosely bound water molecules clinging to the surface of the enzyme. The collapse of Ile1157 brings its backbone amide proton closer to that of Asp1156, resulting in a twist about Ile1157. This twist accounts for the NH/NH cross-peaks between Ile1157 and Asp1156 and between sTyr1158 and Ile1157 in the transferred NOE spectra.

(iii) sTyr1158. As elucidated by the model, a number of hydrogen bonds are formed between the sTyr1158 oxygens and the amide protons of PTP1B residues lining the base of the active site pocket, most notably those of F182, S216 to R221, and Q266 (Table 3). In addition, R221 forms stabilizing ion pairs with sTyr1158, and Q262 makes a water-mediated hydrogen bond to the sTyr1158 phenolic oxygen. In addition to these interactions, aromatic−aromatic stacking also favors the bound ligand orientation. On the basis of the modeled complex, Y46 and sTyr1158 are oriented in a tilted-T arrangement 5.0 Å apart, while F182 and sTyr1158 stack in a similar tilted-T fashion, with an intercentroid distance of 6.1 Å. Such distances are condusive to strong stabilizing interactions. F182 is part of the "WPD" catalytic loop of PTP1B that swings across to "lock" bound ligands in place during catalysis.[42] Aromatic stacking interactions between F182 and the ligand may help to stabilize the closed-loop conformation.[43]

This work suggests that another source of loop stabilization is mediated via interactions with ligand amino acids located at the P(+1) position. From the binding models, one can surmise that residues at P(+1) lie directly over the plane of the F182 phenyl ring. Van der Waals contacts between F182 and nonpolar residues at P(+1) would contribute to the overall binding affinity and lead to favorable hydrophobic collapse of these residues onto the surface of the active site pocket. This notion is also supported in the literature: a recent study found that nonpolar residues, X, positioned at P(+1) in the general sequence AcDE-(sY)X confer inhibitory potency, with aromatic residues being among the most favored.[15] However, in the IRK1154−PTP1B complex of this study, an acidic glutamate residue is located at the P(+1) position, and it stretches across the binding pocket to interact with R47. Therefore, favorable nonpolar interactions with F182 are not facilitated for IRK1154. As a consequence, F182 was not highly constrained, and this was manifested by substantial motion in the dynamics simulations. This is clearly seen in Figure 5, where all active site residues other than F182 are highly constrained in their motional properties.

(iv) Glu1159. IRK1154 is an interesting peptide for structural studies because it contains an acidic residue at the P(+1) position. The structural and modeling studies suggest that Glu1159 at P(+1) somewhat stabilizes the peptide conformation in the positions C-terminal to sTyr1158. As a consequence, Thr1160 and Asp1161 show a greater involvement in binding than might be expected. The inhibition data presented in Table 4 show that the presence of Glu1159 in the IRK1154 sequence confers a moderate increase in binding potency (2.5 μM), as compared to the peptide in which this residue is truncated (5.3 μM). This suggests a limited overall role for Glu1159 in the binding event, a hypothesis consistent with the relatively modest calculated interaction energy for this amino acid (Table 1). The principal interaction made by Glu1159 is the formation of a long ion pair to the basic side chain of R47 (Figure 4). Conversely, Glu1159 experiences unfavorable charge repulsion with the acidic side chains of D48 and D181, which adversely affect the net interaction energy for this residue.

(v) Thr1160 and Asp1161. As noted above, both Thr1160 and Asp1161 significantly influence peptide binding. As

(42) Zhang, Z.-Y. *Curr. Top. Cell. Regul.* **1997**, *35*, 21−68.
(43) Jia, Z.; Barford, D.; Flint, A. J.; Tonks, N. K. *Science* **1995**, *268*, 1754−1758.

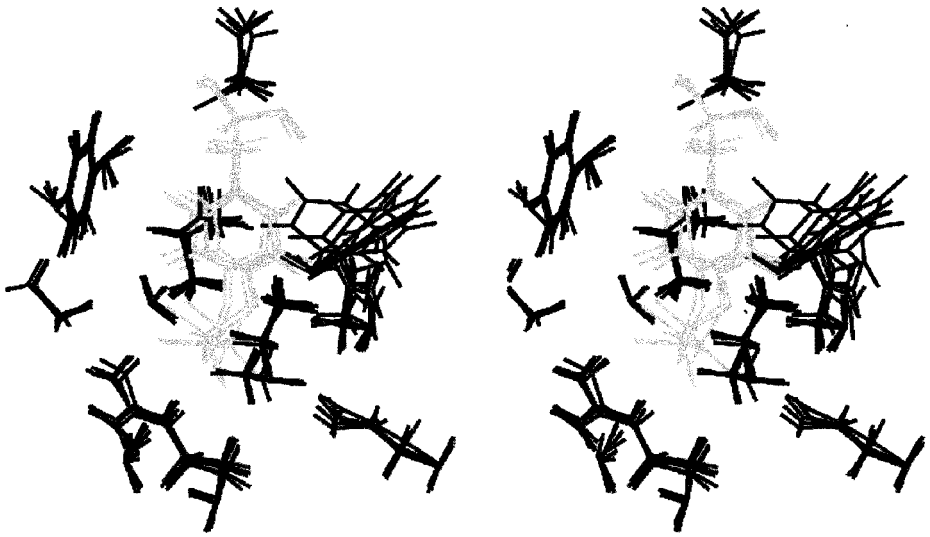

Figure 5. IRK1154−PTP1B binding site interactions. Stereo representation of sTyr1158 in the PTP1B binding pocket. Molecular dynamics snapshots were taken at 0, 25, 50, 75, and 100 ps of production MD. sTyr1158 is depicted in gray, and the PTP1B active site side chains are in black. The extensive rearrangement with simulation time in the position of F182 compared to other active site residues, including Y46 and sTyr1158, is shown.

Table 4. Inhibitory Properties of Insulin Receptor Kinase-Related Peptides

peptide sequence	IC_{50} (μM)[a]
TRDI(sY)ETD	16.1
DI(sY)ETD	5.2
I(sY)ETD	17
(sY)ETD	>200
DI(sY)ETD	1.4
DI(sY)ETD	2.5
DI(sY)ETD	5.3

[a] Inhibition versus GST-PTP1B; data are taken from ref 15.

evidenced by its moderate interaction energy with the enzyme (Table 1), Thr1160 has a stabilizing role in binding. One source of this stabilization is the formation of nonspecific van der Waals contacts between Thr1160 and PTP1B amino acid residue backbone atoms, such as those of R24 and K116. However, the predominant interaction experienced by Thr1160 is a van der Waals contact with the aromatic ring of F182. The upfield chemical shift in the 1D spectrum of the NH proton of Thr1160 in the presence of PTP1B (data not shown) provides support for this close association. Inhibition data indicate that peptides that contain threonine at P(+2) (e.g., DI(sY)ET) are slightly better inhibitors than ones that do not (e.g., DI(sY)E; IC_{50} = 1.4 vs 2.5 μM; Table 4).

Inhibition data show that peptides containing aspartic acid (e.g., DI(sY)ETD) are poorer inhibitors than those without (e.g., DI(sY)ET; IC_{50} = 5.2 vs 2.5 μM).[15] This finding corroborates the finding of a poor interaction energy calculated for Asp1161 in this work. In the modeled complex, Asp1161 experiences unfavorable electrostatic repulsion with the side chains of D48 and D181, interactions that destabilize its contact with PTP1B. These repulsive forces cause Asp1161 to twist away from the surface of the enzyme, which consequently brings Thr1160 and Asp1161 closer together in space. This twist provides a facile explanation for the NH/NH cross-peak observed between these residues in the transferred NOE spectra.

(c) Molecular Modeling of the Potential Interaction between PTP1B and the Insulin Receptor Kinase Domain. With evidence accumulating to support the hypothesis that PTP1B directly associates with and regulates the insulin receptor kinase domain (IRK),[8,10] we have constructed a molecular model in order to probe the potential interactions between PTP1B and IRK. This model has been assessed according to structural criteria presented in the literature for other protein−protein complexes and is a structurally feasible representation of a possible form of association. Corroboration of the proposed mode of interaction awaits further study.

To a first approximation, the putative interaction between PTP1B and IRK was simulated by the protein−protein docking algorithm GRAMM.[44] GRAMM performs a rigorous 3D search for correlations between protein surfaces and ranks the resulting set of possible associations according to the best physical match. The routine was applied at high- (2.5 Å), medium- (4.3 Å), and low-resolution (7 Å) grid searches. A pronounced trend toward a preferred orientation was identified with increasing step size.[45] These potential interaction sites were then correlated with associations obtained from the more rigorous high-resolution data. In this manner, a refined set of potential association sites was determined. The resulting set of protein association "hits" generated by GRAMM suggested an optimal IRK orientation N-terminal to the pTyr active site on PTP1B. Such an arrangement placed pTyr1158 of IRK in the vicinity of the PTP1B binding pocket. Studies have shown that pTyr1158 of the IRK A-loop is preferentially dephosphorylated by PTP1B,[14] suggesting that pTyr1158 interacts within the PTP1B binding pocket.[13] Therefore, the mode of association suggested by GRAMM appears to be reasonable and would account for the aforementioned experimental observations.

The conformation adopted by the IRK1154 residues within the active site cleft of PTP1B is very similar to the structure, determined by X-ray studies, for the corresponding residues of the analogous A-loop region of the insulin receptor kinase domain (Figure 6). This is particularly true for residues Asp1156 to Glu1159. These correspond to critical A-loop residues identified from the crystal structure, and this suggests that, by using the structure of the A-loop from the crystal structure and the orientation of IRK1154 in the active site pocket, a good initial approximation for the orientation of the kinase domain in the PTP1B/IRK complex would be obtained. As previously

(44) Katchalski-Katzir, E.; Shariv, I.; Eisenstein, M.; Friesem, A. A.; Aflalo, C.; Vasker, I. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2195−2199.

(45) Vasker, I. A. *Protein Eng.* **1995**, *8*, 371−377.

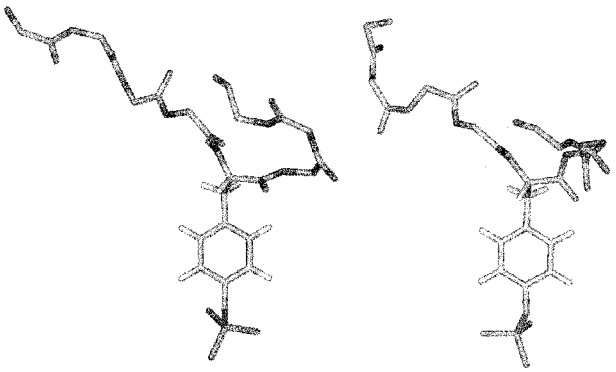

Figure 6. Comparison of the crystallographic A-loop (1154−1161) IRK conformation (a) and the NMR-derived solution structure of IRK1154 (b). Enzyme/peptide are depicted as backbone structures. pTyr/sTyr1154 side chains are shown for clarity. There is a significant correlation between the overall backbone-folds in the crystallographic and solution conformations. The major deviations occur for the residues of IRK1154 that are not strongly restrained by binding interactions with active site residues.

noted, outside of this loop region, the amino acid residues of IRK1154 are more mobile. This is to be expected given the disparate nature of the two types of measurements.

The backbone of the A-loop in the crystallographic structure of IRK was then placed in a position corresponding to that of the docked IRK1154 peptide, and this orientation was used to provide a starting point for the development of a putative PTP1B−IRK complex. Following this initial docking step, bad atom contacts between PTP1B and IRK were alleviated by manual rotation of protein side chains, and an interfacial zone between both enzymes was determined and solvated. The complex was then energy minimized to convergence and subjected to molecular dynamics simulations. The resulting protein−protein complex is depicted in Figure 7. As suggested by the figure, the two proteins form a very complementary association. In the proposed model, the A-loop of IRK is located over the PTP1B binding pocket, so that pY1158 of IRK is buried within the active site. The putative complex is stabilized by interactions between the two enzymes, as discussed below.

(i) Interface Size. The size of the interface between IRK and PTP1B was determined according to the difference in solvent-accessible areas between the bound and free states. In dimeric proteins, the size of the interface area varies widely but typically ranges between ∼350 and 5000 Å^2/subunit. This represents between 7 and 30% of the total accessible surface area of the monomers.[46−48] The solvent-accessible area (calculated according to the procedure of Hubbard and Thornton[35]), which was buried upon formation of the proposed PTP1B−IRK complex, was estimated to be ∼2474 Å^2 (19%) for PTP1B and ∼2642 Å^2 (17%) for IRK, for a total of 5116 Å^2 of total buried accessible surface area (18%). This corresponds to a relatively large interface, as defined by Nussinov's group,[48] and is in excellent agreement with their survey of ∼1700 structures deposited in the Protein Data Bank. This suggests that the interface in the modeled complex has a physically reasonable total surface area. It has been estimated that a minimum of 1200 Å^2 of buried surface area is required in order to stabilize a protein−protein complex,[49] which correlates well with these findings.

(ii) Number of Amino Acids in Interfaces. The need to bury ∼600 Å^2/subunit to form a stable protein−protein complex suggests that a significant number of amino acids are likely to be involved at the interfacial zone between macromolecules. A recent survey of protease inhibitor complexes found that, on average, 25−40 residues were involved in the formation of interfaces on the order of ∼2000 Å^2.[50] This would suggest that ∼80 amino acids should comprise the PTP1B−IRK interface. When interfacial residues were defined according to a 0.5 Å cutoff distance,[50] a total of 136 amino acids were determined to be in the interface. However, a more informative measure of interfacial involvement is provided by the change in accessible surface area that occurs for a residue upon complexation.[47,51] It follows that amino acids involved in the formation of an interface will experience a greater change in accessible surface area than those that are not. With this in mind, analysis of the PTP1B−IRK interface was modified to consider only those residues that had greater than 50% of their accessible surface area buried upon complexation. In this manner, a revised estimate of 81 interfacial residues was obtained, consistent with the value suggested in the literature.[50] Of these, 46 are on PTP1B, and 35 are on the IRK domain (Table 5).

Examination of protein−protein associations suggests a central role for hydrophobic residues in the formation of protein−protein complexes.[48,52] Specifically, it has been suggested that there is a propensity for protein−protein interfaces to be more hydropathic than the remainder of the surfaces of the constituent proteins.[49] Moreover, there is a tendency for interfaces to have complementary hydropathic surfaces, with hydrophobic regions matched with hydrophobic regions, and so forth.[48] In terms of amino acid composition, it has been noted that there is a preference for larger, nonpolar residues at protein−protein interfaces, particularly aromatic amino acids. In addition, there is also a surprising preference for charged residues at interfacial regions.[48,52] The amino acid composition of the PTP1B−IRK complex tabulated in Table 5 shows that there is an overall preference for uncharged amino acids at the interface (56 residues), with charged residues nearly identically matched across the interfacial plane (12 acidic, 13 basic). The total hydropathy of the interface on a residue-by-residue basis is 30% charged, 30% polar, and 40% hydrophobic, which is consistent with estimates in the literature.[47,50,51] The amino acid composition of the proposed PTP1B−IRK interfacial region is, therefore, physically reasonable and is consistent with the view that interfacial regions are typically somewhat hydrophobic in nature.[52] According to a survey of protein−protein association, it would seem that only a few key residues are actually critical to protein−protein complexation, regardless of the total number of amino acids interacting across the interface.[49] In terms of the modeled complex, two patches of interacting hydrophobic residues (M133 to L140 on PTP1B, and V1059 to V1066 on IRK) may represent a key source of association energy. This interaction site is designated "A" in Figure 7c,d. It may be possible to corroborate this hypothesis through mutation studies,

(46) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155−164.
(47) Jones, S.; Thornton, J. M. *Prog. Biophys. Mol. Biol.* **1995**, *63*, 31−65.
(48) Tsai, C.-J.; Lin, S. L.; Wolfson, H. J.; Nussinov, R. *Protein Sci.* **1997**, *6*, 53−64.
(49) Stites, W. E. *Chem. Rev.* **1997**, *97*, 1233−1250.
(50) Janin, J.; Chothia, C. *J. Biol. Chem.* **1990**, *265*, 16027−16030.
(51) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13−20.
(52) Young, L.; Jernigan, R. L.; Covell, D. G. *Protein Sci.* **1994**, *3*, 717−729.

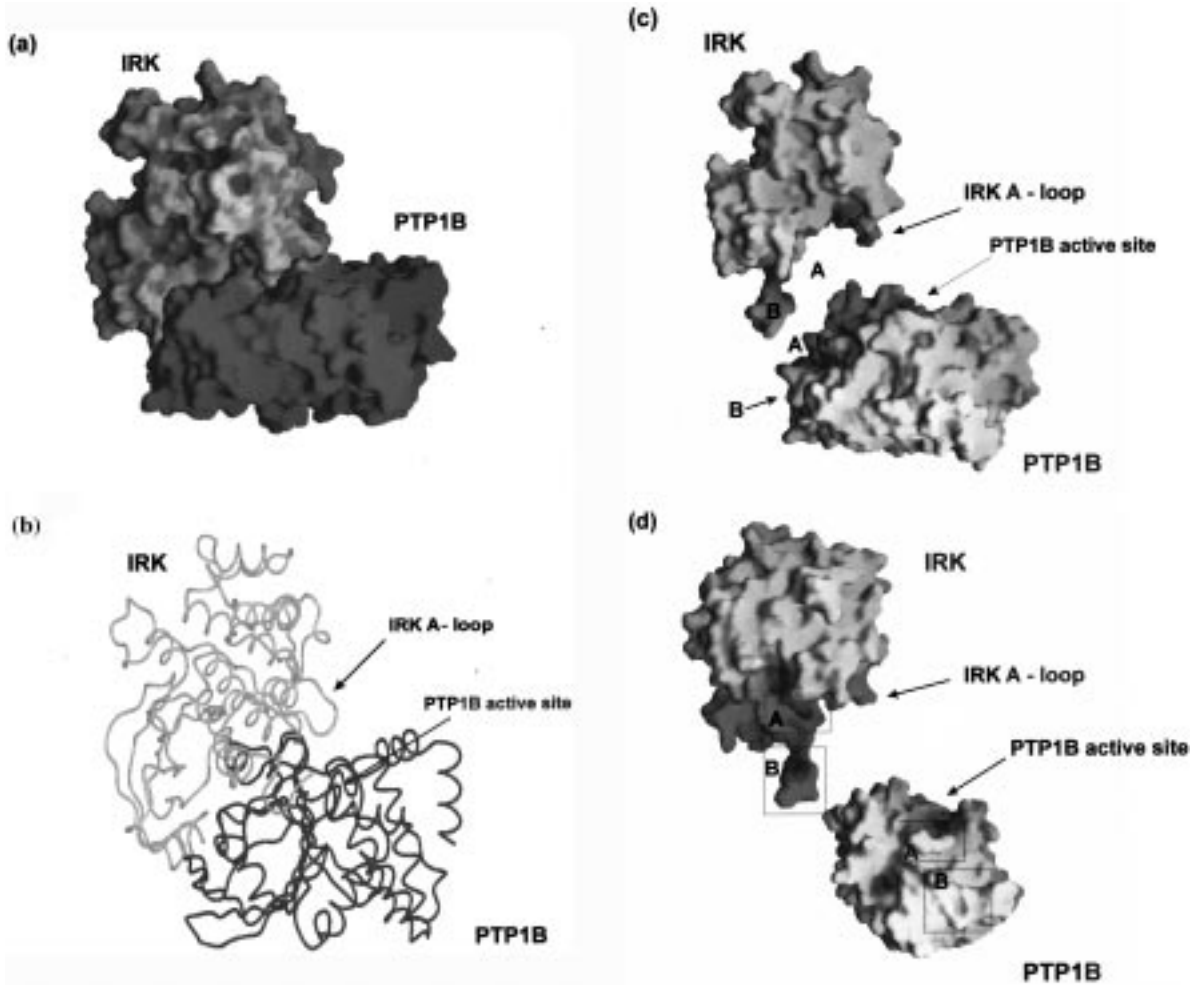

Figure 7. Representations of the putative complex formed between PTP1B and the insulin receptor kinase domain (IRK). (a) The surface of IRK is represented according to surface curvature. For clarity, in (a), PTP1B is depicted in dark gray. The surface complementarity of the two structures is apparent. (b) Ribbon representation of the IRK−PTP1B complex; view as for (a). Backbone of IRK is in light gray; backbone of PTP1B is in dark gray. (c) Representation of the IRK−PTP1B complex. IRK has been "lifted" away from the surface of PTP1B to illustrate the surface complementarity between the molecules. The A-loop of IRK and the active site of PTP1B are indicated. Key: A, overlapping hydrophobic region; B, complementary electrostatic region. (d) View as for (c), rotated by 45°. Boxes A and B are as above. Panel a, c, and d were rendered using the program GRASP.[60]

Table 5. Amino Acid Composition of the PTP1B−IRK Interface

residue type	PTP1B	IRK	total	residue type	PTP1B	IRK	total
GLU	4	4	8	LEU	2	2	4
GLY	5	2	7	PRO	2	1	3
LYS	2	5	7	THR	2	1	3
VAL	2	4	6	TYR	1	2	3
ARG	4	1	5	CYS	2	0	2
ASN	4	1	5	ILE	1	1	2
ASP	3	2	5	HIS	1	0	1
PHE	2	3	5	MET	1	0	1
SER	2	3	5	TRP	0	1	1
ALA	2	2	4				
GLN	4	0	4	total	46	35	81

whereby these residues are substituted by mutation to those of another class of amino acid.[49] Mutation of key residues found in the putative interfacial zone could lend credence to this model of binding.

(iii) Structural Motifs. Apparently, there is no general preference for a specific interaction motif.[48] It does not appear, for instance, that large single sheets or single helices interact in a complementary manner.[47,51] Instead, many distinct secondary structural elements appear to contribute to favorable interface formation. It has been suggested that between 2 and 15 distinct segments typically comprise each face of an interface.[47,51] In the modeled PTP1B−IRK complex, the interface is formed from about 12 secondary structural elements, and this is compatible with a favorable interface.

(iv) Hydrogen Bonding and Electrostatics. In general, most of the electrostatic interactions noted between associating proteins arise from hydrogen-bonding interactions between protein side chains. Ion-pairing interactions are less common.[49] It has been demonstrated that the total number of hydrogen bonds formed between protein subunits is approximately proportional to the area of the subunit interface, with one hydrogen bond found on average for each 200 Å^2 of interface surface.[50] Alternative surveys have suggested that this estimate should be more on the order of 1−1.5 hydrogen bonds/100 Å^2.[47,51] In the modeled PTP1B−IRK complex, a total of 22 well-defined hydrogen bonds are formed between the two complexes, with four ion pairs formed. This value is, therefore, consistent with the literature. The ion pairs noted in the model consist of favorable interactions between R43-D987, K41-E1043, R24-E1159, and R47-D1156. The latter two of these interactions correspond to those observed in the structure of IRK1154 bound in the active site of PTP1B (Figure 5). A putative complementary electrostatic interaction is depicted in Figure 7c,d as region

"B", in which two structurally and electrostatically complementary elements correspond.

(v) Packing, Cavities, and Shape Complementarity. It is widely assumed that there should be both shape and electrostatic complementarity between associating molecules.[49] As noted above, there is favorable electrostatic complementarity between PTP1B and IRK, as evidenced by the presence of several overlapping patches of hydrophobic residues in the interface, as well as by the formation of favorable ion pairs and hydrogen bonds. In terms of shape complementarity, it is generally found that dimeric proteins have the most complementary shapes, with protein−ligand and antibody−protein complexes less so.[49] The points of interface between two proteins define a surface, whose shape (relative length and breadth) can be assessed in order to gain insight into the complementarity of fit between its components. Topologically, the PTP1B−IRK interface is quite irregular. However, the planarity of the interface and the length/breadth ratio (0.68) are consistent with those of other heterocomplexes.[47,51]

The volume in the gaps that exist between molecules after complexation provides a useful means of assessing the geometric favorability of geometric association.[53] The gap volume index is a measure of how complementary the two surfaces are to one another, with an index of 1.0 indicating a perfect match. The PTP1B−IRK complex has a gap volume index of 1.62, which demonstrates a strikingly favorable site of intersection. By comparison, the mean gap volume index derived for a set of 27 diverse heterodimers was 2.48,[51] while antibody−antigen complexes are the least complementary, with a gap volume index of 3.02.[51] As depicted in Figure 7, the association between the two macromolecules studied here involves the fitting together of several complementary structural segments. The degree of fit lends credence to the feasibility of the modeled complex. No significant cavities are found in the interface region of the complex.

(vi) Backbone and Side-Chain Rearrangements. In general, only minor conformational changes from the uncomplexed state are observed in the backbone structure of constituent proteins as the complex forms.[49] Side-chain rearrangements are found when such changes facilitate more favorable packing or hydrogen-bonding interactions.[49] During energy minimization of the PTP1B−IRK model, the backbone structure and side-chain conformations of several PTP1B and IRK residues underwent a degree of reorientation. Given the static nature under which the structures were initially docked, this was not surprising. These changes were generally minimal (∼2.2 Å RMSD for PTP1B; ∼2.6 Å RMSD for IRK) and do not appear to represent a serious deficiency of the model. During molecular dynamics simulations of the complex, the backbone atoms of PTP1B active site residues were fixed to their minimized position. This was found to stabilize the complex over the course of the trajectory.[39,54] Only minor changes in the backbone structure of IRK were noted during the MD simulation. However, there were several interesting side-chain rearrangements noted during the trajectory. For example, the acidic side chain of D1160 reoriented with respect to R47 in order to make a favorable electrostatic contact. This appears to stabilize the integrity of the PTP1B−IRK complex at the PTP1B active site. In addition, there were also some side-chain rearrangements noted at the interfacial site, such as for the complementary hydrophobic patches corresponding to M133 to L140 on PTP1B and V1059 to V1066 on IRK.

(53) Laskowski, R. A. *SURFNET*; Department of Biochemistry and Molecular Biology, University College: London, 1991.

(54) van Gunsteren, W. F. *Curr. Opin. Struct. Biol.* **1993**, *3*, 277−281.

Conclusions

Measurements of transferred signal intensity in two-dimensional transferred NOE NMR spectroscopy have allowed the development of a structural and binding model that explains the inhibitory properties of an insulin receptor kinase domain-based peptide (IRK1154). In the model, IRK1154 binds to PTP1B in an extended β-strand conformation that is twisted about sTyr1158 and Thr1160. Amino acids along the length of the peptide make contact with residues constituting the active site of PTP1B. The sulfotyrosine moiety of IRK1154 is buried within the phosphotyrosine recognition pocket and is held in position by a number of hydrogen bonds and ion pairs formed with PTP1B active site residues. Aromatic−aromatic stacking and nonpolar interactions also stabilize this interaction. The principal mode of peptide recognition is mediated by the formation of a strong ion pair between Asp1156 and R47 of PTP1B, together with a weaker ion pair between Glu1159 and R47. Residues Ile1157 and Thr1160 of IRK1154 also stabilize the bound peptide conformation through van der Waals contacts to the enzyme. Conversely, Arg1155 and Asp1161 experience repulsion from similarly charged PTP1B residues flanking the binding pocket and, therefore, have a destabilizing effect on overall binding. The N-terminal residue, Thr1154, appears to have a limited role in binding. The presence of a polar amino acid at the P(+1) position results in substantial motion of F182 of the WPD loop. A nonpolar amino acid in this position would be expected to undergo favorable hydrophobic constraints with the phenyl ring of F182 and, by doing so, restrict its motion.

The interaction between the IRK1154 peptide and PTP1B is driven by interactions with IRK1154 side chains rather than with the backbone of the peptide. Van der Waals contacts are clearly very important in the overall binding event (Table 2), a hypothesis that is supported by inhibition data (Table 4). In the modeled complex, interactions of Ile1157, Thr1160 with PTP1B clearly confer considerable binding affinity (Table 1). Again, this contrasts with the other peptides, for which electrostatic interactions were predominant. Although the active site of PTP1B is encircled by charged residues, these findings suggest that nonpolar interactions should be taken into account when attempting to design optimal inhibitors of PTP1B. Specifically, introduction of larger, nonpolar moieties at positions P(−1), P(+1), and P(+2) may confer additional binding affinity for PTP1B inhibitors. Indeed, recent evidence suggests that optimization of nonpolar contacts at the P(+1) site can result in improved potency of peptide-based inhibitors.[15]

A putative model of the interaction between PTP1B and the insulin receptor kinase domain has been constructed to help explain recent evidence demonstrating in vitro association between PTP1B and IRK.[9] The proposed model was assessed according to criteria developed from recent surveys of protein−protein complexation.[47,51] The principal findings of this analysis suggest that the modeled PTP1B−IRK complex is reasonable from a structural and physical basis. In particular, the total area of the interfacial zone, the number of interacting amino acids, and the composition of the interface region are all consistent with expectations based on the literature surveys. Hydrophobic interactions appear to be of major importance to the proposed interprotein interaction, while hydrogen bonding and electrostatics also contribute significantly to the association. The surface complementarity between the macromolecules in the model is surprisingly good, considering the irregular general shape of

the two proteins, and no significant cavities were formed upon complexation. Only minor conformational changes from the crystallographic parent structures occur in the model.

The complex between PTP1B and IRK proposed in this study provides critical information that points to the way in which the two proteins associate. While no experimental evidence exists to directly corroborate or refute the proposed model of association, analysis of the complex supports its general plausibility. In particular, the model accounts for the fact that PTP1B specifically dephosphorylates on pTyr1158, pTyr1162, and pTyr1163.[14] While it is likely that some conformational change accompanies the protein−protein interaction,[49] gross structural modifications of one or both of the macromolecules would be necessary to accommodate an alternative binding mode to the one proposed. Therefore, the proposed complex appears to represent a rational mode of interaction based upon the known crystallographic structures of PTP1B and IRK.

The modeled complex between IRK and PTP1B appears to provide a reasonable interpretation of the structural and physical elements that potentially stabilize an association between the two enzymes. Indeed, in the complex, pTyr1158 is optimally positioned for catalysis within the PTP1B active site. However, protein−protein interactions are expected to involve more of a delicate balance between entropic and enthalpic contributions than those of protein−ligand complexes.[55] In particular, in terms of an interaction between PTP1B and IRK, one must also consider how the protein−protein complex is broken apart after completion of catalysis. It is interesting to note that there are several regions of less than favorable interaction between PTP1B and IRK in the modeled complex. For example, there is a region of overlapping negative charge in region "A" of Figure 7c,d. Moreover, it was determined in the modeling studies of IRK1154 that Arg1155 was located in a region of unfavorable charge repulsion in the bound model. R1155 of the A-loop sequence is located in an analogous position in the PTP1B−IRK complex. In addition, Glu1159 of IRK1154 experiences charge repulsion with D55 of PTP1B in the bound peptide model, an interaction that is also noted in PTP1B−IRK model. It is possible that these elements provide a counterbalance to the associative forces and subsequently facilitate dissociation of IRK from the complex.

This model of association between PTP1B and IRK provides a structural basis to account for the known dephosphorylation event that occurs at pY1158 of the A-loop of IRK. Kinetics studies of a model peptide system revealed that PTP1B preferentially dephosphorylates at the pY1158 and pY1162 residues, while it is less effective for pY1163.[14] The model proposed in this paper only accounts for the first dephosphorylation event (pY1158). However, further modeling suggests that structural modifications in the A-loop can readily accommodate dephosphorylation of the phosphorylated tyrosine residues of Y1162 and Y1163. Specifically, a rearrangement of the A-loop backbone, accompanied by rotation of residues T1158 to T1161 of IRK away from the active site of PTP1B, can position residue pY1162 within the region of the active site. Similarly, a further rotation of the A-loop backbone may facilitate the dephosphorylation of pY1163. The latter rearrangement appears to be the most difficult to accommodate structurally. The modeling suggests that, when pY1162 is bound in the PTP1B binding pocket, residue pY1163 of the IRK A-loop is able to reside within the second, low-affinity PTP1B binding pocket, which is located adjacent to the primary site.[56] Conversely, when pY1163 is modeled into the binding site, pY1162 is forced to make several unfavorable contacts to PTP1B residues flanking the binding pocket. Apparently, then, the conformation adopted during the third dephosphorylation event is less structurally favored than for the first two dephosphorylations.

It is hoped that this model will be useful in experimental studies of the interaction between PTP1Band IRK. In particular, the interfacial residues proposed in this paper could be investigated through mutagenesis experiments. Such work may corroborate the involvement of specific amino acid residues in the protein−protein interaction[49] and provide support for the proposed model.

Acknowledgment. This material was abstracted, in part, from the doctoral thesis of N.R.G. Thanks are gratefully extended to Merck Frosst Canada, Inc. for the provision of materials and for the financial support of this work. Thanks are especially extended to Dr. Michael Gresser and Dr. Chidambaram Ramachandran for their support and encouragement.

JA984354Q

(55) Muegge, I.; Schweins, T.; Warshel, A. *Proteins: Struct., Funct. Genet.* **1998**, *30*, 407−423.

(56) Puius, Y. A.; Zhao, Y.; Lawrence, D. S.; Almo, S. A.; Zhang, Z.-Y. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13420−13425.

(57) McDonald, I. K.; Thornton, J. M. *J. Mol. Biol.* **1994**, *238*, 777−793.

(58) Connolly, M. L. *J. Appl. Crystallogr.* **1983**, *16*, 548−558.

(59) Connolly, M. L. *Science* **1983**, *221*, 709−713.

(60) Nicholls, A.; Sharp, K. A.; Honig, B. *Proteins* **1991**, *11*, 281−296.