Structural Characterization of the Phosphotyrosine Binding Region of a High-Affinity SH2 Domain−Phosphopeptide Complex by Molecular Dynamics Simulation and Chemical Shift Calculations

Ming-Hsiang Feng,[†] Marios Philippopoulos,[‡] Alexander D. MacKerell, Jr.,[§] and Carmay Lim*[,†]

Contribution from the Department of Chemistry, National Tsing-Hua University, and Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan 11529, Republic of China, Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada, and School of Pharmacy, University of Maryland at Baltimore, Baltimore, Maryland 21201

Received May 7, 1996[⊗]

Abstract: Three molecular dynamics simulations of the free, phosphate ion-bound and phosphopeptide-bound C-terminal SH2 domain of phospholipase C−γ1 (PLCC·pY) have been performed to aid in the interpretation of chemical shift data and the elucidation of interatomic interactions at the phosphotyrosine (pTyr) binding region. The simulation of the phosphopeptide complex was carried out with newly developed CHARMM force-field parameters for pTyr, optimized against experimental data and *ab initio* calculations. The lack of NOEs involving phosphate in the binding pocket had necessitated a chemical shift analysis of the pTyr binding region for a more detailed characterization of the hydrogen bonding interactions involving pTyr. Although most of these interactions are not present in the NMR structure used as the simulation starting point, the system converges early in the simulation to a structure more compatible with the chemical shift data. This is supported by *ab initio* determination of the $^1H^{\eta}$ and $^1H^{\epsilon}$ chemical shifts of the three arginines (Arg 18, 37, and 39) in the pTyr binding pocket based on the PLCC·pY MD structure, which are in accord with the experimental values. The simulation structure of the PLCC·pY complex reveals a more complete picture of interatomic interactions in the pTyr binding pocket than is possible with current chemical shift and NOE approaches alone, thereby permitting the identification of the primary pTyr-recognition residues. This pattern of interactions is strikingly similar to those of crystal structures of related SH2 domains. The simulations also suggest several alternative interpretations of the chemical shift data to those suggested in the experimental investigation (Pascal, S. M., et al. *Biochemistry* **1995**, *34*, 11353). This insight is valuable as the observed chemical shifts could result from a number of possible pictures of interactions. The present study demonstrates that the combination of molecular dynamics simulations and *ab initio* chemical shift calculations can enhance the hydrogen-bonding, amino−aromatic, and aliphatic−aromatic information content of NOE- and chemical-shift-based protein structures and serve as a complementary tool for the interpretation of chemical shift data at the atomic level.

Introduction

Many cellular processes such as proliferation, growth, differentiation, and metabolism are triggered by external stimuli that bind receptor molecules at the cell membrane and induce their clustering and autophosphorylation.[1−3] The activated receptors elicit a cascade of intracellular reactions through which the mitogenic signal is transmitted to the nucleus, and the specific function is initiated.[4,5] A first step in this array of events is the sequence-specific binding of Src-homology-2 (SH2) domains of a series of cytoplasmic proteins to the receptor tyrosine phosphorylation sites.[6−9]

$PLC_{\gamma 1}$ is a tyrosine-phosphorylation-activated signal transduction enzyme that catalyzes the hydrolysis of phosphatidylinositol 4,5-bisphosphate, thereby producing second messengers inositol trisphosphate and diacylglycerol; these in turn activate protein kinase C and raise intracellular calcium levels.[10] $PLC_{\gamma 1}$ contains two SH2 domains through which it binds to various growth factor receptors following hormonal stimulation. In particular, $PLC_{\gamma 1}$ is a substrate for the platelet-derived growth factor receptor (PDGFR).[11,12] Tyr 1021 of PDGFR has been identified as a binding site of $PLC_{\gamma 1}$, since its mutation to phenylalanine dramatically reduced $PLC_{\gamma 1}$ binding and catalytic activity.[11] Furthermore, the high-affinity binding of $PLC_{\gamma 1}$ to Tyr-1021 of PDGFR is an important interaction in cell signaling processes as mutation of Tyr 1021 has been shown to result in a reduction in the mitogenic response to PDGF in vivo.[12]

The two SH2 domains of $PLC_{\gamma 1}$ have different sequence specificities:[13] the N- and C-terminal SH2 domains prefer an

* To whom correspondence should be addressed.
† National Tsing-Hua University and Academia Sinica.
‡ University of Toronto.
§ University of Maryland at Baltimore.
⊗ Abstract published in *Advance ACS Abstracts*, November 1, 1996.

(1) Ullrich, A.; Schlessinger, J. *Cell* **1990**, *61*, 203.
(2) Cantley, L. C.; Auger, K. R.; Carpenter, C.; Duckworth, B.; Graziani, A.; Kapeller, R.; Soltoff, S. *Cell* **1991**, *64*, 281−302.
(3) Pawson, T.; Gish, G. *Cell* **1992**, *71*, 359−362.
(4) Kazlauskas, A.; Ellis, C.; Pawson, T.; Cooper, J. A. *Science* **1990**, *247*, 1578.
(5) Mayer, B. J.; Baltimore, D. *Trends Cell Biol.* **1993**, *3*, 8−13.
(6) Anderson, D.; Koch, C. A.; Grey, L.; Ellis, C.; Moran, M. F.; Pawson, T. *Science* **1990**, *250*, 979−982.
(7) Matsuda, M.; Mayer, B. J.; Fukui, Y.; Hanafusa, H. *Science* **1990**, *248*, 1537.
(8) Kashishian, A.; Kazlauskas, A.; Cooper, J. A. *EMBO J.* **1992**, *11*, 1373−1382.
(9) Fantl, W. J.; Escobedo, J. A.; Martin, G. A.; Turck, C. W.; Rosario, M.; Mccormick, F.; Williams, L. T. *Cell* **1992**, *69*, 413−423.
(10) Rhee, S. G.; Suh, P. G.; Ryu, S. H.; Lee, S. Y. *Science* **1989**, *244*, 546−550.
(11) Valius, M.; Bazenet, C.; Kazlauskas, A. *Mol. Cell. Biol.* **1993**, *13*, 133−143.
(12) Valius, M.; Kazlauskas, A. *Cell* **1993**, *73*, 321−334.

S0002-7863(96)01530-2 CCC: $12.00

Scheme 1

pY1021	D	N	D	pY	I	I	P	L	P	D	P	K
Number	−3	−2	−1		+1	+2	+3	+4	+5	+6	+7	+8

Scheme 2

	βA	AA	αA		AB			βB		BC
Src	WYF	GKI	TRRESERLLL		NPE	NP	RG	TFLVRES		ETTKG
Lck	WFF	KNL	SRKDAERQLL		APG	NT	HG	SFLIRES		ESTAG
SypN	WFH	PNI	TGVEAENLLL		TRG	V	DG	SFLARPS		KSNPG
PLCC	WYH	ASL	TRAQAEHMLM		RVP	RD	G	AFLVRKR		NEPN
	11	14	17		27			33		40

	βC	CD			βD	βD′	DE		βE
Src	AYCLSVSD	F	DNAK	GL	NVKHYKI	RKL	DS	G	GFYI
Lck	SFSLSVRD	F	DQNQ	GE	VVKHYKI	RNL	DN	G	GFYI
SypN	DFTLSVRR		N	G	AVTHIKI	QNT		GD	YYDL
PLCC	SYAISFRA		E	G	KIKHCRV	QQE		GQ	TVML
	44				54	61			66

	EF	βF	FB	αB	BG		βG
Src	TSR	TQF	S	SLQQLVAYYSKH	ADGL	CHRLT	NVC
Lck	SPR	ITF	P	GLHDLVRHYTNA	SDGL	CTRLS	RPC
SypN	YGG	EKF	A	TLAELVQYYMEH	HGQLKEKNGD	VIELK	YPL
PLCC	GN	SEF	D	SLVDLISYYEKH	PLYR	KMKLR	YP
	70	72	75	76	88		97

acidic residue and a hydrophobic residue, respectively, at the +2 position after the pTyr. Since the sequence of the Tyr 1021 $PLC_{\gamma 1}$ site of PDGFR, pTyr-Ile-Ile-Pro, matches the sequence specificity of the C-terminal SH2 domain, the interaction of $PLC_{\gamma 1}$ with Tyr-1021 of PDGFR is thought to be due to the C-terminal SH2 domain of the enzyme. Thus, to gain a deeper understanding of the molecular details of this biologically important interaction in signal transduction, the solution structure of the C-terminal SH2 domain of $PLC_{\gamma 1}$ (PLCC) complexed with a 12-residue phosphotyrosyl peptide (pY1021) has been determined by nuclear magnetic resonance (NMR) spectroscopy.[14] Scheme 1 gives the pY1021 peptide sequence, corresponding to the PDGFR Tyr 1021 binding site, while Scheme 2 lists the PLCC sequence aligned with the sequences of the Src, Lck, and Syp N-terminal SH2 domains, whose X-ray structures complexed with high-affinity phosphopeptides have been determined.[15−18]

The NOE- and chemical-shift-based NMR structure of the PLCC SH2 domain[14,19] complexed with a high-affinity pTyr-containing peptide (PLCC·pY) reveals that the binding region of the SH2 domain consists of a pTyr-binding pocket and a hydrophobic groove contacting peptide residues +2 to +6 C-terminal to the pTyr. In particular, the pTyr-binding region is a deep depression on the protein surface lined by four arginine side chains. Binding studies of peptides with randomized sequences at the +1, +2, and +3 positions[13] and of peptides truncated down to three residues (pY-1, pY, pY+1)[20] have shown that the PLCC SH2 domain retains considerable binding affinity under these conditions. The pTyr-binding site therefore appears to account for most of the binding free energy of the phosphopeptide complex.[21] In light of these findings, the detailed characterization of the interatomic interactions at the site is an issue of outstanding interest. On the basis of the NMR structure alone, some or all of the four arginines in the region are expected to be involved in electrostatic interactions with the phosphate group of the pTyr. However, it has not been possible to study the precise nature of these interactions by NOE methods alone for several reasons.[22,23] First, the presence of four arginine side chains in the binding site as well as the degeneracy and overlap of most arginine N^{η} and H^{η} resonances inhibit NOE assignment of these atoms. Second, rapid solvent exchange involving H^{ϵ} and H^{η} protons causes attenuation of their respective NOE's. Third, intermediate exchange dynamical processes and rotation about the $N^{\epsilon}-C^{\zeta}$ bond on the microsecond-to-millisecond time scale cause broadening of most arginine aliphatic and amino side chain resonances, respectively, and consequently reduce their NOE intensities. Finally, the lack of hydrogens on the phosphate group, which is predominantly dianionic at the pH of the NMR structure determination,[23,24] creates difficulties for the generation of protein−peptide NOE distance restraints at the site based on $^1H-^1H$ experiments.

To further elucidate the interactions at the pTyr-binding site, arginine $^{15}N^{\epsilon}$ T_1 and T_2 relaxation times, steady-state $N^{\epsilon}-H^{\epsilon}$ NOEs, $^1H^{\epsilon\bullet}$ exchange rates with water, and chemical shift measurements in different environments have been obtained.[23] The effects of the pTyr aromatic ring and phosphate group on the arginine chemical shifts were isolated by recording $^{15}N-^1H$ HSQC spectra of 0.3 mM PLCC SH2 at pH 6.0 in (a) 100 mM imidazole buffer, (b) 100 mM phosphate ion buffer, and (c) 0.3 mM pY1021 peptide in 100 mM imidazole buffer. (Extrinsic chemical shifts in proteins measure the degree of shielding of the surrounding electron cloud on the nucleus in question, and originate primarily from participation of the nucleus in hydrogen bonds, or proximity to aromatic rings and carbonyl groups[25].) In the presence of imidazole buffer (a above) all $N^{\eta}-H^{\eta}$ and $N^{\epsilon}-H^{\epsilon}$ resonances of the four arginines in the binding pocket (Arg 18, 37, 39, and 59) were degenerate, implying similar chemical environments. However, of all $N^{\eta}-H^{\eta}$ resonances, only those of Arg 37 were affected upon binding of phosphate ion (b), indicating involvement of both Arg 37 NH_2 groups in binding phosphate. Addition of phosphate ion also caused downfield shifts to the $N^{\epsilon}-H^{\epsilon}$ resonances of Arg 18, 39, and 59, suggesting formation of hydrogen bonds with the phosphate, while the $N^{\epsilon}-H^{\epsilon}$ crosspeak of Arg 37 underwent a slight upfield shift, pointing to the possible disruption of a pre-existing hydrogen bond in the free SH2 domain. Upon substitution of phosphate ion buffer (b) with phosphopeptide imidazole buffer (c), upfield shifts of an Arg 18 $N^{\eta}-H^{\eta}$ peak and of the $N^{\epsilon}-H^{\epsilon}$ peaks of Arg 39 and 59 were suggested to indicate an axial orientation of these groups relative to the pTyr ring plane, whereas the downfield shift of an Arg 39 $N^{\eta}-H^{\eta}$ peak was interpreted to arise from a H^{η} proton lying in the plane of the pTyr ring.

Although the chemical shift data, as summarized above, provide information on the contacts between pTyr and SH2 domain, and therefore serve to identify the residues involved

(13) Songyang, Z.; Shoelson, S. E.; Chauduri, M.; Gish, G.; Pawson, T.; Haser, W. G.; King, F.; Roberts, T.; Ratnofsky, S.; Lechleider, R. J.; Neel, B. G.; Birge, R. B.; Fajardo, J. E.; Chou, M. M.; Hanafusa, H.; Schaffhausen, B.; Cantley, L. C. *Cell* **1993**, *72*, 767−778.

(14) Pascal, S. M.; Singer, A. U.; Gish, G.; Yamazaki, T.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E. *Cell* **1994**, *77*, 461−472.

(15) Waksman, G.; Kominos, D.; Robertson, S. C.; Pant, N.; Baltimore, D.; Birge, R. B.; Cowburn, D.; Hanafusa, H.; Mayer, B. J.; Overduin, M.; Resh, M. D.; Rios, C. B.; Silverman, L.; Kuriyan, J. *Nature* **1992**, *358*, 646−653.

(16) Eck, M. J.; Shoelson, S. E.; Harrison, S. C. *Nature* **1993**, *362*, 87−91.

(17) Waksman, G.; Shoelson, S. E.; Pant, N.; Cowburn, D.; Kuriyan, J. *Cell* **1993**, *72*, 779−790.

(18) Lee, C.-H.; Kominos, D.; Jacques, S.; Margolis, B.; Schlessinger, J.; Shoelson, S. E.; Kuriyan, J. *Structure* **1994**, *2*, 423−438.

(19) Pascal, S. M.; Singer, A. U.; Pawson, T.; Kay, L. E.; Forman-Kay, J. D. **1996**, Manuscript in preparation.

(20) Wolf, G.; Lynch, A.; Chaudhuri, M.; Gish, G.; Pawson, T.; Shoelson, S. E. *J. Biol. Chem.* submitted.

(21) Kay, L. E.; Muhandiram, D. R.; Farrow, N. A.; Aubin, Y.; Forman-Kay, J. D. *Biochemistry* **1996**, *35*, 361−368.

(22) Yamazaki, T.; Pascal, S. M.; Singer, A. U.; Forman-Kay, J. D.; Kay, L. E. *J. Am. Chem. Soc.* **1995**, *117*, 3556−3564.

(23) Pascal, S. M.; Yamazaki, T.; Singer, A. U.; Kay, L. E.; Forman-Kay, J. D. *Biochemistry* **1995**, *34*, 11353−11362.

(24) Singer, A. U.; Forman-Kay, J. Personal communication, 1995.

(25) Haigh, C. W. and Mallion, R. B. *Prog. NMR Spectrosc.* **1980**, *13*, 303−344.

in pTyr recognition, several important questions remain to be answered. (i) It is difficult, on the basis of chemical shift differences alone, to identify interactions other than those between the pTyr and the SH2 domain, e.g., contacts stabilizing the arginines in the binding pocket. (ii) Protein−solvent interactions that may be present at the binding site have not been clearly identified. (iii) There is no experimental information on one-to-one interatomic interactions, so that neither the distribution nor the number of hydrogen bonds among the four arginines in the binding pocket and the four pTyr phosphate oxygens is known. To address these questions and thus provide a structural picture of the pTyr binding pocket at a level of detail not attainable by chemical shift and NOE methods alone, a molecular dynamics (MD) simulation of the PLCC·pY complex has been carried out and analyzed. In addition, an *ab initio* evaluation of proton shieldings at the pTyr binding pocket has been performed to verify the interactions at the pTyr-binding site found in the MD complex. Furthermore, to aid in the interpretation of chemical shift changes upon substitution of imidazole with phosphate ion buffer and upon substitution of phosphate ion with peptide-containing imidazole buffer, MD simulations of PLCC bound to isolated phosphate ion (PLCC·P_i) and free in solution (PLCC) have also been performed. The pTyr side chain is modeled by newly developed CHARMM empirical force-field parameters optimized against crystallographic data and *ab initio* calculations; the parameterization procedure is described in the Appendix. The simulation protocol and chemical shift calculations are summarized in the next section. The structural features of the pTyr binding site found in the three simulations are presented in the Results. Interactions found in the pTyr-binding pocket as well as predicted proton chemical shifts based on the MD PLCC·pY structure are discussed against experimental evidence including NMR data and X-ray structures of related SH2 domains in the Discussion. Finally, the key findings of this work are highlighted in the Conclusion.

Methods

Simulation Procedure. Three molecular dynamics simulations, using the program CHARMM,[26] were carried out on the fully solvated C-terminal SH2 domain of phospholipase C-γ1 in different environments: (i) PLCC·pY consisting of 1710 protein atoms and 198 peptide atoms fully solvated with 10086 TIP3P[27] water atoms; (ii) PLCC·P_i soaked with 9546 water atoms; (iii) PLCC free in solution, i.e., with the phosphate peptide removed, solvated with 9558 water atoms. The NMR structure[19] (a minimized average one based on an ensemble of 30 structures compatible with the NOE data, henceforth referred to as the NMR structure) was used as the starting point for the PLCC·pY simulation. Comparison of the crystal structures of the Src SH2 domain in the absence of peptide or bound to the isolated phosphate ion with those complexed with the phosphopeptide[17] as well as a study[28] reporting PLCC backbone amide chemical shift changes upon binding phosphopeptide reveal only localized and relatively small structural changes. Thus, the NMR structure of the PLCC·pY complex with the phosphopeptide removed was directly used as the starting point for the PLCC·P_i and free PLCC simulations. These simulations were carried out at pH 6.4, the pH at which the NMR structure was solved and near the pH of the chemical shift experiments (pH 6.0).[23] At this pH (6.4), His 13, 23, and 57 are neutral, whereas His 87 and 6 are protonated, as deduced from ^{15}N chemical shift titration experiments.[24] Furthermore, the measured pK_a value of the pTyr phosphate group (less than 4.5)[24] indicates that it is almost entirely in its 2− charge state at pH 6.4. The pK_a values of the other ionizable groups in free and bound PLCC have not been measured and are assumed to be normal: all lysines and arginines are positively charged and all aspartates and glutamates are negatively charged. This yields a total charge of 1+ for PLCC·pY, 3+ for PLCC·P_i, and 5+ for free PLCC.

The forces on the atoms and their dynamics were calculated with the program CHARMM[29] (version 23) and an all-hydrogen parameter set.[30] All protein and solvent atoms were propagated according to Newton's equations using a leap-frog integrator and a timestep of 1 fs. The water molecules were subjected to a deformable mean-field potential,[31,32] which accounts for interactions with infinite bulk waters not treated explicitly in the simulation. Electrostatic interactions were truncated by applying a atom-based force-switching function[33] in the region between 11 and 15 Å, whereas van der Waals interactions were truncated at 15 Å by a shifted potential.[29]

The initial structure was subjected to 100 steps of steepest−descent and 200 steps of adopted−basis Newton−Raphson minimization in the presence of strong (10 kcal/Å^2) harmonic constraints on all heavy atoms for PLCC·pY in order to relieve energetically unfavorable contacts in the protein without disrupting its overall conformation. For PLCC·P_i and free PLCC, weaker constraints (2 kcal/Å^2 on the BC, BG, and EF binding loops and 4 kcal/Å^2 on adjacent residues) were employed to allow for small conformational changes resulting from the removal of the phosphopeptide. At this stage, the root mean square deviation from the NMR coordinates was 0.12 Å for protein heavy atoms and 0.09 Å for backbone atoms for PLCC·pY, 0.38 and 0.31 Å for PLCC·P_i and 0.45 and 0.39 Å for free PLCC. The protein was then immersed in a 30.5 Å radius sphere of TIP3P water molecules[27] constructed from an equilibrated configuration of TIP3P waters at a pressure of 1 atm and a density of 0.0334 molecules/Å^3. Overlaid water molecules whose oxygens were within 2.7 Å of protein heavy atoms were deleted. The solvated system was then subjected to 500 steps of steepest descent minimization, during which harmonic restraints on the protein heavy atoms were gradually reduced to zero; the root mean square deviation from the NMR structure at this point was 0.33 Å for protein heavy atoms and 0.29 Å for backbone atoms for PLCC·pY, 0.56 and 0.46 Å for PLCC·P_i and 0.63 and 0.54 Å for free PLCC.

The minimized solvated system was prepared for subsequent dynamics by an initial 6 ps heating phase, during which initial velocities were randomly assigned from a Maxwell−Boltzmann distribution at 25 K and rescaled at temperatures increasing from 25 to 300 K. This was followed by a 24 ps equilibration stage, during which velocities were rescaled to 300 K at 2 ps intervals if the average temperature fell outside a 10 K window (for 14 ps) and a 5 K window (for 10 ps). About 3 ps of production dynamics was completed per day on an IBM RS/6000 590 workstation.

Solvent-Accessible Surface Area (ASA). Solvent-accessible-surface-area (ASA) calculations were performed on the dynamics-averaged MD structures using the Shrake & Rupley algorithm[34] and a probe radius of 1.4 Å. The percentage residue accessibility is defined[35] as the % ratio of the water accessible surface area of the side chain X to that of its accessible surface area in the tripeptide −Gly−X−Gly−.

Chemical Shift Calculations. Chemical shielding calculations were carried out using the program GAUSSIAN 94.[36] A basis set of gauge

(26) Brooks, B. R.; Bruccoleri, R. D.; Olafson, B. O.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(27) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(28) Farrow, N. A.; Muhandiram, R.; Singer, A. U.; Pascal, S. M.; Kay, C. M.; Gish, G.; Shoelson, S. E.; Pawson, T.; Forman-Kay, J. D.; Kay, L. E. *Biochemistry* **1994**, *33*, 5984−6003.
(29) Brooks, C. L., III; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312.
(30) MacKerell, A. D., Jr.; Bashford, D.; Bellot, M.; Dunbrack, R. L.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Wiorkiewicz-Kuczera, J.; Karplus, M. *FASEB J.* **1992**, *6*, A143.
(31) Berkowitz, M.; McCammon, J. A. *Chem. Phys. Lett.* **1982**, *90*, 215−217.
(32) Beglov, D.; Roux, B. *J. Chem. Phys.* **1994**, *100*, 9050−9063.
(33) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667−683.
(34) Shrake, A.; Rupley, J. A. *J. Mol. Biol.* **1973**, *79*, 351−371.
(35) Chothia, C. *J. Mol. Biol.* **1976**, *105*, 1−14.
(36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Petersson, T. Keith G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*; Gaussian Inc.: Pittsburgh, PA, 1995.

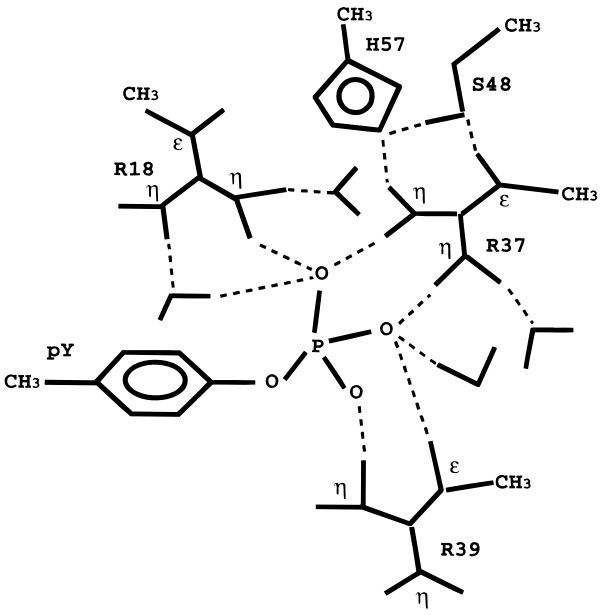

Figure 1. Schematic representation of the hydrogen-bonding pattern in the pTyr-binding pocket. Hydrogen bonds are represented by broken lines. The fragment shown corresponds to model I. For model II, the benzyl ring was replaced by a methyl group.

invariant atomic orbitals (GIAO)[37,38] was employed to give gauge-invariant results of magnetic properties. Since the magnetic shielding tensor at a given nucleus is a local property, models were chosen to contain only the nuclei of interest, fragments hydrogen bonding to these nuclei as well as those having short range contributions through bonds. Model I (Figure 1) was extracted from the "minimized" MD structure of the PLCC·pY complex (see also the Discussion) consisting of a total 91 atoms (including 40 heavy atoms) depicting the hydrogen-bonding interactions in the pTyr binding pocket. Model II is the same as model I except that the benzyl ring of pTyr has been replaced by a methyl group. The contribution from aromatic ring current effects was then estimated by the shielding difference between models I and II. The calculations were carried out at the HF/3-21+G* level and each calculation took about 100 h on an IBM RS/6000 590 workstation.

Results

The backbone root-mean-square deviations (rmsd's) from the initial NMR coordinates as a function of simulation time for the three MD trajectories in Figure 2 show that equilibrium is reached around 100−110 ps. The rmsd of the four arginines (i.e., Arg 18, 37, 39, and 59) forming the pTyr binding pocket also indicate that the conformation of the pTyr binding pocket is equilibrated in all three cases (Figure 2). Therefore, the average MD structures for PLCC·pY, PLCC·P_i, and free PLCC were computed from the 110−240, 110−230, and 100−230 ps portion of the trajectories, respectively. Assuming that the initial NMR structure is close enough to the "true" structure, the current trajectories should be long enough for convergence of the structural properties of the pTyr binding region, as opposed to dynamic properties, for which the simulation length needs to be at least one order of magnitude longer than the relaxation time of the dynamic property of interest for sufficient sampling.

Tables 1−3 list interactions of the four arginines in the pTyr binding pocket of PLCC·pY, PLCC·P_i, and free PLCC. A hydrogen bond is defined by a heavy atom−heavy atom distance ≤3.5 Å, a light donor−heavy acceptor distance ≤2.5 Å, and a deviation of less than ±60° from linearity. Table 4 gives the predicted ^{1}H chemical shifts of arginine guanidino groups based on the MD structure and the corresponding experimental values.[22,23]

(37) Ditchfield, R. *J. Chem. Phys.* **1972**, *56*, 5688−5691.
(38) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.

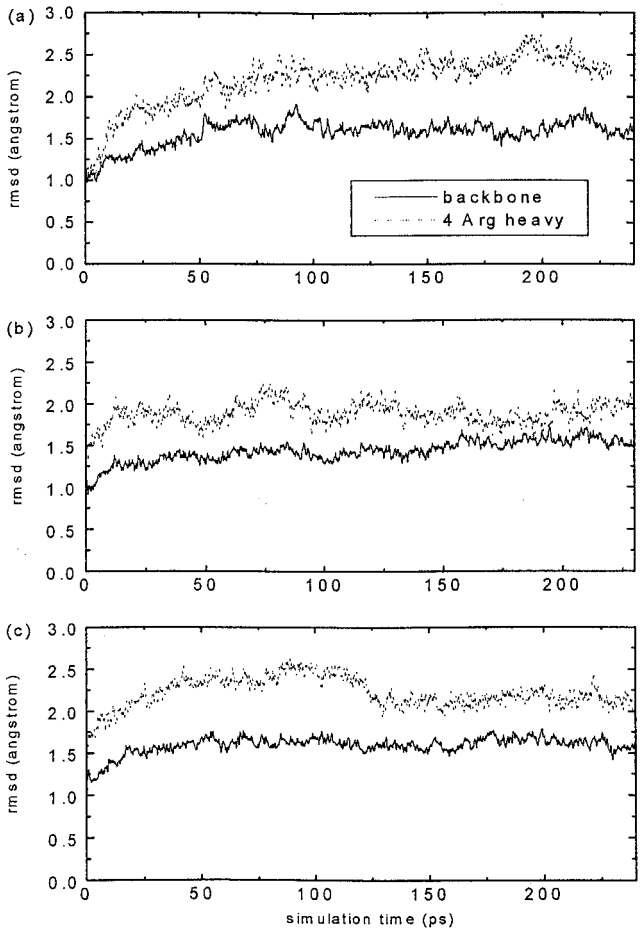

Figure 2. Plots of the root-mean-square deviation of (a) PLCC·pY, (b) PLCC·P_i, and (c) free PLCC structures from the the initial NMR structure as a function of simulation time.

Residue Solvent Accessibility. The residue %ASA values of Arg 18, 37, 39, and 59 and highly conserved Ser 48 and His 57 are listed in Table 5. Arg 37, Ser 48, and His 57 are solvent-protected, and Arg 18, 39, and 59 are solvent-exposed residues with %ASA < ≈40% and > ≈40%, respectively, in all three states of PLCC. A decrease in the %ASA of Arg 18, Arg 39, and His 57 is observed as the pTyr-binding pocket of the PLCC (free) structure is filled with phosphate ion and pTyr. In contrast, the %ASA values of solvent-protected Arg 37 and Ser 48 and the most solvent-exposed residue Arg 59 are overall unaffected by these changes.

Structural Analysis of PLCC·pY Binding Pocket. Of the four arginines located in the pTyr binding pocket, Arg 18, Arg 37, and Arg 39 hydrogen-bond directly to one or two of the four phosphate oxygens (Table 1). These interactions are illustrated in Figure 3: O^H and O^4 and interact with Arg 39 H^ϵ. O^4 forms hydrogen bonds with $H^{\eta 11}$ and $H^{\eta 21}$ from both amino groups of Arg 37. O^2 interacts with 2 arginines (Arg 37 $H^{\eta 11}$ and Arg 18 $H^{\eta 21}$), whereas O^3 binds to Arg 39 $H^{\eta 12}$. The phosphate group is held rigidly throughout the simulation, since none of the four bonds connecting the phosphorus atom to O^H, O^2, O^3, and O^4 changed orientation throughout the production trajectory. Apart from Arg 18, Arg 37, and Arg 39, no other protein residues are found to interact directly with the phosphate oxygens except for three water molecules (one NMR[19] and two overlay waters), each of which hydrogen-bonds to a phosphate oxygen (Figure 3). Bul 981, specifically, forms a hydrogen-bonding bridge between O^2 and Arg 18 $H^{\eta 11}$.

The orientation of an atom with respect to the pTyr or a histidine ring is defined by the distance R_c of the atom to the ring centroid, and its inclination angle θ with respect to the

Table 1. Interactions of the Four Arginines in the Binding Pocket of PLCC·pY MD and NMR Structures with PTyr, Protein, and/or Solvent[a]

donor/acceptor	acceptor/donor	$\langle r \rangle_{MD}$ (Å)[b]	rmsd (Å)	r_{NMR} (Å)
Arg 18(O)	Glu 22(H^N)	2.00	0.19	2.27
Arg 18($H^{\eta 11}$)	Bul 981(O)	1.90	0.16	
Arg 18($H^{\eta 21}$)	pY (O^2)	1.65	0.10	(3.95)
Arg 18($H^{\eta 22}$)	Bul 1855(O)	2.05	0.31	
Arg 37(O)	Ala 46(H^N)	1.92	0.14	2.19
Arg 37(O)	Wat 2(H)	(8.40)	0.67	2.48
Arg 37(H^N)	Ala 46(O)	1.91	0.13	2.08
Arg 37(H^ϵ)	Ser 48($O^{\gamma 1}$)	2.00	0.19	(3.87)
Arg 37(H^ϵ)	Ile 47(O)	(4.86)	0.42	2.19
Arg 37($H^{\eta 11}$)	pY (O^2)	2.00	0.57	(3.30)
	pY (O^4)	2.62	0.45	(4.23)
Arg 37($H^{\eta 12}$)	His 57($N^{\delta 1}$)	1.98	0.15	(2.70)
Arg 37($H^{\eta 12}$)	pY (O^4)	(4.27)	0.44	2.03
Arg 37($H^{\eta 21}$)	pY (O^4)	1.77	0.23	(1.98)
Arg 37($H^{\eta 22}$)	Bul 2011(O)	1.91	0.14	
Arg 39(O)	Glu 41(H^N)	2.00	0.17	(3.55)
Arg 39(H^N)	Ser 44(O)	1.94	0.15	2.18
Arg 39(H^ϵ)	pY (O^4)	2.08	0.63	(3.58)
	pY (O^H)	2.62	0.30	(3.15)
Arg 39($H^{\eta 12}$)	pY (O^3)	1.73	0.16	(3.30)
Arg 59(H^N)	Wat 1(O)	(8.85)	5.22	2.28
Arg 59(H^N)	Bul 546(O)	2.63	0.99	
Arg 59(O)	Bul 135(H2)	2.33	1.08	

[a] Wat is a water determined by NMR;[19] Bul is a bulk−overlaid water. [b] Distance is averaged over the 110−240 ps portion of the trajectory; numbers in parentheses indicate absence of a hydrogen bond according to the criterion stated in the Results.

Table 2. Interactions of the Four Arginines in the Binding Pocket of PLCC·P_i MD Structure with Phosphate Ion, Protein, and/or Solvent[1]

donor/acceptor	acceptor/donor	$\langle r \rangle_{MD}$ (Å)[a]	rmsd (Å)
Arg 18(O)	Glu 22(H^N)	2.12	0.22
Arg 18(H^N)	Wat 2(O)	2.22	0.37
Arg 18(H^ϵ)	P_i (O^2)	1.84	0.23
Arg 18($H^{\eta 11}$)	P_i (O^2)	1.88	0.29
Arg 37(O)	Ala 46(H^N)	1.92	0.15
Arg 37(H^N)	Ala 46(O)	1.92	0.13
Arg 37(H^ϵ)	Ser 48($O^{\gamma 1}$)	2.02[b]	0.27[b]
Arg 37($H^{\eta 11}$)	His 57($N^{\delta 1}$)	2.02	0.16
Arg 37($H^{\eta 12}$)	P_i (O^2)	1.67	0.11
Arg 37($H^{\eta 21}$)	Bul 2021(O)	1.94	0.17
Arg 37($H^{\eta 22}$)	P_i (O^4)	1.68	0.17
Arg 39(O)	Glu 41(H^N)	2.13	0.23
Arg 39(H^N)	Ser 44(O)	1.92	0.15
Arg 39(H^ϵ)	P_i (O^4)	1.69	0.11
Arg 39($H^{\eta 11}$)	P_i (O^3)	1.81	0.24
Arg 39($H^{\eta 21}$)	Asn 40($O^{\delta 1}$)	2.15	0.62
Arg 59(O)	Bul 388(H)	2.60	0.89

[a] Distance is averaged over the 110−230 ps portion of the trajectory. [b] Distance is averaged over the last 60 ps of the trajectory (170−230 ps).

normal to the ring plane. The R_c distances and inclination angles θ of the Arg 18, 37, 39, and 59 side chain atoms are listed in Tables 6 and 7. Geometric analyses of 33 protein structures[39] show that amino−aromatic interactions are characterized by preferred R_c distances of 3.0−6.0 Å with the side chain amino groups preferrentially found above or below the plane of the aromatic ring. Tables 6 and 7 show that, based on these criteria, only the amino side chains of Arg 18 and 59 are involved in amino−aromatic interactions, whereas those of Arg 37 and 39 are either too far away (>6 Å) and/or are situated roughly in the plane of the pTyr ring. Furthermore, the aliphatic side chain of Arg 59 is also found to interact with the pTyr aromatic ring in the MD structure.

In addition to the pTyr−arginine interactions; the MD PLCC·pY structure also permits identification of additional interactions not present in the NMR structure, that may serve in stabilizing the pTyr recognition elements. In particular, the side chain atoms of two highly conserved residues, Ser 48 $O^{\gamma 1}$ and His 57 $H^{\delta 1}$, interact with the Arg 37 H^ϵ and $H^{\eta 12}$ and protons, respectively (see Table 1). The former two residues also hydrogen-bond to each other. Furthermore, several water molecules are found to interact with the 4 arginines in the binding pocket, e.g., Bul 2011 forms a hydrogen-bonding bridge between Arg 37 $H^{\eta 22}$ and the carbonyl oxygen of Lys 38. Several water molecules interact transiently with the H^η atoms from both amino groups of Arg 18 and Arg 39. Finally, the backbone and side chain atoms of Arg 59 do not participate in

Table 3. Interactions of the Four Arginines in the Binding Pocket of Free PLCC MD Structure with Protein and/or Solvent

donor/acceptor	acceptor/donor	$\langle r \rangle_{MD}$ (Å)[a]	rmsd (Å)
Arg 18(O)	Glu 22(H^N)	1.96	0.15
Arg 18(H^N)	Bul 866(O)	2.49	0.91
Arg 18(H^ϵ)	Glu 22($O^{\epsilon 2}$)	2.04	0.90
Arg 37(O)	Ala 46(H^N)	2.03	0.20
Arg 37(H^N)	Ala 46(O)	1.89	0.14
Arg 37(H^ϵ)	Bul 1061(O)	1.89	0.20
Arg 37($H^{\eta 12}$)	Bul 551(O)	1.91	0.27
Arg 37($H^{\eta 21}$)	Leu 16(O)	2.24	0.42
Arg 37($H^{\eta 22}$)	Wat 2(O)	1.94	0.25
Arg 39(O)	Glu 41(H^N)	2.06	0.20
Arg 39(H^N)	Ser 44(O)	2.27	0.27
Arg 59(O)	Bul 498(H)	1.93	0.35

[a] Distance is averaged over the 100−230 ps portion of the trajectory.

Table 4. ^1H Chemical Shifts of Arginine Guanidino Groups in Models I and II Compared to Experiment

residue	nucleus	expt[a]	model I[b]	model II[b]
Arg 18	H^ϵ	7.4	2.7	2.9
	$H^{\eta 11}$	(6.1)	5.9	6.3
	$H^{\eta 12}$		2.9	3.2
	$H^{\eta 21}$	–	10.5	10.7
	$H^{\eta 22}$	–	7.0	7.2
Arg 37	H^ϵ	6.9	6.9	6.9
	$H^{\eta 11}$	9.5	10.3	10.2
	$H^{\eta 12}$	7.5	7.2	7.2
	$H^{\eta 21}$	9.7	9.8	9.8
	$H^{\eta 22}$	5.8	6.3	6.2
Arg 39	H^ϵ	8.1	9.7	9.5
	$H^{\eta 11}$	(7.3)	3.6	3.5
	$H^{\eta 12}$		11.1	10.9
	$H^{\eta 21}$	(6.7)	4.2	4.1
	$H^{\eta 22}$		4.2	4.1

[a] Experimental values taken from Yamazaki *et al.*;[22] ^1H chemical shifts are calibrated relative to TSP. A dash means the shift is not observed due to broadening of the H^η. A number in parentheses is the average shift of the two H^η protons of an amino group. [b] ^1H chemical shifts calculated relative to TMS. All chemical shifts are calculated with the 3-21+G* basis set. TMS and TSP resonate within 0.1 ppm to each other.[50]

Table 5. Percent Solvent Accessible Surface Area of the Residues in the PTyr-Binding Pocket in the Three States

residue	PLCC	PLCC·P_i	PLCC·pY
Arg 18	60	46	39
Arg 37	11	0	10
Arg 39	65	46	47
Arg 59	74	79	72
Ser 48	0	0	0
His 57	33	16	5

(39) Burley, S. K. and Petsko, G. A. *Adv. Prot. Chem.* **1988**, *39*, 125−189.

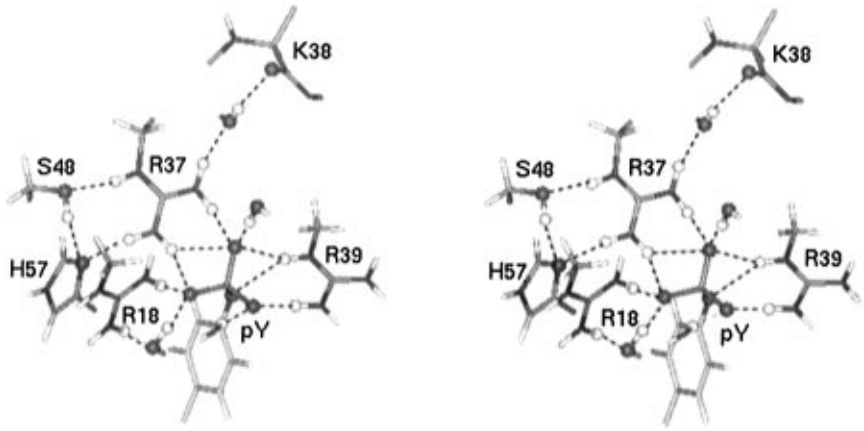

Figure 3. Stereoview of the hydrogen-bonding network around the pTyr-binding pocket in the average PLCC·pY MD structure.

Table 6. Centroid Distances and Inclination Angles of the Amino Side Chain Atoms of Arg 18, 37, and 39 in the PLCC·pY Binding Pocket with Respect to the PTyr or His 57 Rings

		MD(pTyr)[a,b]		MD(His 57)[a,c]		NMR(pTyr)[b]	
residue	atom	R_c (Å)	θ (deg)	R_c (Å)	θ (deg)	R_c (Å)	θ (deg)
Arg 18	N^{ϵ}	8.0	30.8	4.8	43.4	4.8	19.9
	H^{ϵ}	8.4	24.7	4.8	49.6	5.0	23.2
	$N^{\eta 1}$	6.0	21.8	3.6	24.4	2.8	5.7
	$H^{\eta 11}$	5.1	21.1	3.4	13.5	3.1	23.4
	$H^{\eta 12}$	6.6	15.5	3.9	37.7	2.2	12.6
	$N^{\eta 2}$	6.3	42.1	4.3	25.7	4.0	29.2
	$H^{\eta 21}$	5.3	45.0	4.3	18.5	3.6	38.1
	$H^{\eta 22}$	7.0	47.8	4.9	34.6	4.8	19.9
Arg 37	N^{ϵ}	7.5	69.3	4.9	83.0	5.8	69.3
	H^{ϵ}	7.5	65.0	4.2	88.0	5.5	68.8
	$N^{\eta 1}$	5.3	63.3	3.8	60.2	4.7	49.6
	$H^{\eta 11}$	5.5	55.9	2.9	47.1	4.5	43.3
	$H^{\eta 12}$	4.7	62.8	4.1	66.2	4.7	46.8
	$N^{\eta 2}$	6.6	77.9	6.0	61.4	6.1	65.1
	$H^{\eta 21}$	6.0	79.6	6.1	52.6	6.1	61.7
	$H^{\eta 22}$	7.5	81.0	6.8	65.3	6.7	71.9
Arg 39	N^{ϵ}	5.5	66.7	9.0	42.0	5.5	78.5
	H^{ϵ}	5.0	75.9	8.1	40.2	5.4	80.1
	$N^{\eta 1}$	5.4	71.9	9.5	29.6	6.8	95.1
	$H^{\eta 11}$	5.9	70.6	8.6	26.9	6.8	96.3
	$H^{\eta 12}$	4.9	80.7	9.9	27.4	7.6	99.1
	$N^{\eta 2}$	6.7	55.8	11.1	38.4	6.7	85.8
	$H^{\eta 21}$	7.1	55.7	11.7	35.2	6.2	87.8
	$H^{\eta 22}$	7.2	52.3	11.5	42.5	7.4	90.8

[a] Averaged over the 110−240 ps trajectory of the PLCC·pY complex. [b] Parameters were measured with respect to the pTyr ring. [c] Parameters were measured with respect to the His 57 ring.

strong inter-residue charge−charge interactions but weak ones with water molecules, consistent with its high %ASA value (Table 5).

Structural Analysis of PLCC·P_i Binding Pocket. In the PLCC·Pi structure (Figure 4), all three pTyr recognition residues, Arg 18, 37, and 39, are still found to bind isolated phosphate dianion and most of the binding interactions with phosphate found in the PLCC·pY structure are preserved. Differences between phosphate binding in the PLCC·P_i vs PLCC·pY average MD structures lie in O^H interacting with a water molecule in place of Arg 39 H^{ϵ}, O^2 hydrogen-bonding to Arg 18 H^{ϵ} instead of a water molecule (Bul 981), and O^4 forming hydrogen bonds with only one Arg 37 H^{η} proton rather than two protons from each amino group.

Most of the hydrogen-bonding contacts stabilizing the 4 arginines in the binding pocket of the PLCC·pY structure, including those characteristic of elements of regular secondary structure, have been maintained (Table 2). In addition, the hydrogen bond between Arg 37 H^{ϵ} and Ser 48 $O^{\gamma 1}$ found in the PLCC·pY MD structure is also formed during the last 60 ps of the PLCC·P_i trajectory (170−230 ps). There is no significant

Table 7. Centroid Distances and Inclination Angles of the Arg 59 Side Chain Atoms in the PLCC·pY-Binding Pocket with Respect to the PTyr or His 57 Rings

		MD(pTyr)[a,b]		MD(His57)[a,c]		NMR(pTyr)[b]	
residue	atom	R_c (Å)	θ (deg)	R_c (Å)	θ (deg)	R_c (Å)	θ (deg)
Arg 59	N^{ϵ}	5.0	28.9	10.9	61.5	6.8	27.7
	H^{ϵ}	5.6	33.2	11.3	65.5	7.1	30.5
	$N^{\eta 1}$	5.8	48.1	11.6	59.1	8.8	33.5
	$H^{\eta 11}$	6.0	54.4	12.0	63.2	9.0	34.5
	$H^{\eta 12}$	6.3	47.9	11.8	56.5	9.5	35.1
	$N^{\eta 2}$	4.4	40.9	10.7	50.9	7.5	31.8
	$H^{\eta 21}$	4.8	51.1	10.9	48.2	6.8	30.6
	$H^{\eta 22}$	4.1	33.7	10.3	49.0	8.4	34.9
	C^{β}	4.5	33.0	9.2	70.4	4.0	17.8
	$H^{\beta 1}$	3.5	35.3	8.3	66.4	3.0	11.0
	$H^{\beta 2}$	5.2	27.9	9.9	68.3	4.7	11.1
	$C^{\gamma 1}$	4.7	27.2	9.9	69.0	4.7	26.1
	$H^{\gamma 1}$	5.7	28.1	10.9	71.9	5.6	30.6
	$H^{\gamma 2}$	4.3	38.9	9.4	71.5	4.2	37.8
	$C^{\delta 1}$	4.5	14.8	10.3	61.2	5.5	21.0
	$H^{\delta 1}$	3.6	9.7	9.6	57.1	5.1	28.0
	$H^{\delta 2}$	5.2	6.8	11.1	59.8	5.8	11.1

[a] Averaged over the 110−240 ps trajectory of the PLCC·pY complex. [b] Parameters were measured with respect to the pTyr ring. [c] Parameters were measured with respect to the His 57 ring.

overall conformational change of the protein in the PLCC·P_i structure relative to the PLCC·pY structure, as evidenced by an rmsd of 1.10 Å for the backbone atoms of residues 11−98 between the two structures.

Structural Analysis of Free PLCC Binding Pocket. In the absence of phosphopeptide or phosphate ion, the four arginines interact mainly with solvent. As expected, the protein−protein hydrogen-bonding contacts stabilizing the four arginines in the PLCC·P_i and/or PLCC·pY structures are not preserved (Table 3). Arg 18 H^{ϵ} forms an unstable hydrogen bond with the carboxylate oxygens of Glu 22 in the free PLCC structure instead of a phosphate oxygen in the PLCC·P_i structure. Arg 37 H^{ϵ} and H^{η} protons do not interact directly with the side chain atoms of Ser 48 and His 57. Instead, penetration of solvent into the pTyr binding pocket has resulted in a water molecule bridging the Arg 37 H^{ϵ} proton and His 57 imidazole nitrogen. This may be a stable water bridge over time scales longer then the present simulation length, as both Arg 37 and His 57 are relatively solvent-protected in the PLCC (free) structure with %ASA values <40% (Table 5). The backbone oxygen of Leu 16 replaces the side chain nitrogen of His 57 in interacting with an H^{η} proton of Arg 37. These interactions are consistent with the observation that Leu 16, Ser 48 and His 57 exhibit a N−H or C^{α}−H proton chemical shift change of greater than 0.6 ppm upon binding peptide.[28]

All the backbone interactions involving the four arginines found in the PLCC·pY structure are still maintained. The rmsd

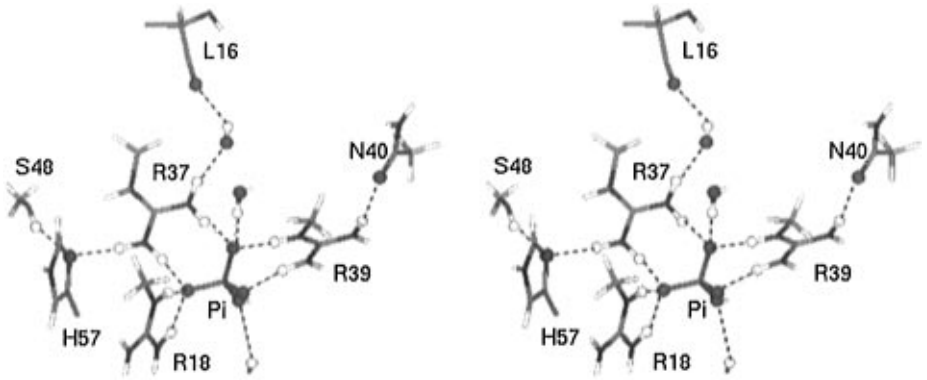

Figure 4. Stereoview of the hydrogen-bonding network around the phosphate binding pocket in the average PLCC·P_i MD structure.

of 1.14 Å for the backbone atoms of residues 11−98 from the PLCC·pY structure indicates that the overall protein conformation has not changed significantly in the absence of phosphopeptide during the course of the simulation. The largest rmsd's occur in the BC and BG loops involved in binding. These changes are slight compared to those found for the BC, EF, and BG loops in the Src free and complex structures,[17] but are consistent with the NMR study[28] of PLCC, in which where no significant changes in the chemical shifts for the BC loop were observed.

Discussion

Validation of pTyr-Binding Interactions in PLCC·pY MD Structure. *Comparison with NOE Data.* Due to technical difficulties in the structural determination of PLCC·pY by NOE methods (outlined in the Introduction) detailed information of the specific electrostatic interactions between the SH2 domain and the pTyr based only on the NMR structure is problematic. In particular, NOEs alone cannot yield a detailed hydrogen-bond distribution among the arginines in the binding pocket and the phosphate oxygens, due to the lack of protons in the phosphate portion of the pTyr residue. Only one hydrogen bond (according to the criteria stated in the Results) involving phosphate oxygens exists in the NMR structure, whereas seven protein−phosphate and three solvent−phosphate hydrogen bonds were found in the MD PLCC·pY structure (Figure 3). During simulated annealing of the NOE-based structure,[14,19] *nonspecific* hydrogen-bond restraints derived from the chemical shift studies[23] were imposed: there are four and two restraints for the two amino groups of Arg 37 and Arg 39 $N^{\epsilon}-H^{\epsilon}$, respectively, and any phosphate oxygen. Thus, the near absence of hydrogen bonds found in the NMR structure seems to be due to the lack of specific restraints between the arginine protons and phosphate oxygens.

A total of 45 NOE restraints involving pTyr were imposed during the restrained MD protocol used in refining the current NMR structure.[19] The majority of these NOE's involve Arg 59 (13), His 57 (7), and Cys 58 (7). All four observed NOE's among Arg 37, Ser 48, and His 57 and all 13 NOE restraints between Arg 59 and pTyr (the majority of them between Arg 59 aliphatic and pTyr aromatic protons) are satisfied in the MD structure. Of the 45 NOE restraints involving pTyr there are only five "apparent" NOE violations, all pertaining to nonspecific hydrogen-bond restraints, where the distance averaged over the PLCC·pY MD trajectory and over all proton pair combinations is outside the range of the NOE restraint (Table 8). However, in each case, the $\langle r \rangle_{MD}$ distances in Table 8 indicate that the non-pTyr hydrogen is closer to one of the two pTyr protons and the corresponding distance is within the range of the NOE restraint, except for the Arg 37 $H^{\epsilon}-$pTyr $H^{\epsilon 2}$ distance, which exceeds the NOE restraint by 0.3 Å. Therefore, the present structural model is on the whole consistent with the available NOE data.

Table 8. "Apparent" NOE Violations

	NOE restraint[b]				
atom pair[a]	R_{ij}	r^{l}_{ij}	r^{u}_{ij}	$\langle r \rangle_{MD}$	$\langle r \rangle_{MD,av}$[c]
pTyr H^{ϵ}*−Arg 37 H^{ϵ}	4.0	2.2	1.0	5.3, 8.4	6.9
pTyr H^{ϵ}*−Arg 39 H^{α}	5.0	3.2	3.0	8.0, 9.1	8.6
pTyr H^{ϵ}*−Ala 46 H^{α}	3.3	1.5	3.0	4.5, 8.2	6.4
pTyr H^{δ}*−His 57 H^{α}	3.3	1.5	3.0	4.6, 8.2	6.4
pTyr H^{δ}*−Val 60 H^{γ}**	5.0	3.2	5.4	*d*	10.9

[a] All violations belong to nonspecific proton pairs; H^{ϵ}* refers to $H^{\epsilon 1}$ and $H^{\epsilon 2}$ protons, H^{δ}* refers to $H^{\delta 1}$ and $H^{\delta 2}$, and Val 60 H^{γ}** refers to $H^{\gamma 11}$, $H^{\gamma 12}$, $H^{\gamma 13}$, $H^{\gamma 21}$, $H^{\gamma 22}$, and $H^{\gamma 23}$. [b] R_{ij} is the calculated NOE distance between protons *i* and *j*; r^{l}_{ij} and r^{u}_{ij} are the values of the upper and lower limits of the target distances, respectively. [c] $\langle r \rangle_{MD,av}$ is the average of the $\langle r \rangle_{MD}$ distances. [d] The corresponding 12 numbers are 10.5, 13.5, 10.3, 13.2, 10.3, 13.5, 8.7, 12.4, 8.2, 11.7, 7.0, and 10.6.

Comparison with Chemical Shift Data. When comparing the computed and experimental chemical shifts three approximations in the calculations need to be taken into consideration. First, the calculated chemical shift is based on a model extracted from a minimized average structure of the PLCC·pY complex, whereas the measured chemical shift is an average over the shieldings for all configurations. It is computationally prohibitive (see Methods) to calculate the shieldings for each snapshot of the trajectory and average them. On the other hand, averaging the MD configurations over the trajectory to obtain a mean structure may yield unreasonable internal coordinates (bond lengths and bond angles) for mobile waters or aliphatic side chains, thus yielding incorrect shielding constants. Therefore, the average structure of the entire complex was minimized first by molecular mechanics and the part of interest was then extracted. It is not necessary to subject the latter to further geometry optimization at the *ab initio* level as the absence of the whole protein architecture will have large boundary effects. Second, the four H^{η} protons of an arginine guanidinium group are equivalent due to rapid rotation of $N^{\epsilon}-C^{\zeta}$ and $C^{\zeta}-N^{\eta}$ bonds in cases where the guanidinium group does not interact with nearby hydrogen-bonding acceptors or the interaction is not strong enough to inhibit free rotation on the time scale of chemical shift measurements. However, the bond rotation can not be accounted for in the chemical shift calculations, which will thus generate four nonequivalent H^{η} for each arginine residue. Third, the present calculations exclude effects of the medium (in particular solvent), resulting in a more upfield shift for protons especially when the proton does not hydrogen bond to phosphate or other hydrogen bonding acceptors.

In the case of Arg 37, the NMR spectrum[22,23] showed unique values for all two N^{η} and four H^{η} nuclei, suggesting that the rotation of $N^{\epsilon}-C^{\zeta}$ and $C^{\zeta}-N^{\eta}$ bonds was restricted. On the other hand, the hydrogen-bonding network observed in the

average MD structure of the PLCC·pY complex (Figure 3) showed two H^{η} protons interacting with phosphate oxygens and the other two interacting with His 57 imidazolium nitrogen and a water, respectively. This is an excellent case where the calculated chemical shifts can be compared with experimentally measured values as the effects of bond rotation and solvent are minimized. As shown in Table 4, the computed Arg 37 chemical shifts are in good agreement with the experimental values, in particular the $^1H^{\epsilon}$ resonance. The two H^{η} protons hydrogen bonding to the phosphate oxygens have the most downfield chemical shifts (10.3 and 9.8 ppm). The $H^{\eta 12}$ proton hydrogen bonding to His 57 $N^{\delta 1}$ has a chemical shift around 7.2 ppm, while the one interacting with water gives the most upfield shift at 6.3 ppm. Note that although a water molecule was found to hydrogen bond to Arg 37 $H^{\eta 22}$ throughout the simulation, this interaction may not always exist on the time scale of the chemical shift measurement, thus accounting for the more upfield shift in the experimental value relative to the computed one (5.8 vs 6.3 ppm of $H^{\eta 22}$, Table 4).

In the case of Arg 39, only two $N^{\eta}-H^{\eta}$ crosspeaks were observed,[23,22] suggesting that the $N^{\epsilon}-C^{\zeta}$ bond has been slowed down considerably but the two $C^{\zeta}-N^{\eta}$ bonds can still undergo rapid rotation on the chemical shift time scale. The $N^{\epsilon}-C^{\zeta}$ bond restriction is consistent with the hydrogen bond found in the MD structure between a phosphate oxygen and Arg 39 $H^{\eta 12}$, whose computed chemical shift (11.1 ppm) is the most downfield relative to other H^{η} protons. The average of the $H^{\eta 12}$ (11.1 ppm) and $H^{\eta 11}$ (3.6 ppm) chemical shifts is in quantitative agreement with the more downfield Arg 39 H^{η} resonance observed at 7.3 ppm. The other two Arg 39 H^{η} protons, which have no hydrogen-bonding partners, exhibit lower chemical shift values (4.2 ppm) and correspond to the second experimentally observed set of $N^{\eta}-H^{\eta}$ crosspeaks at 6.7 ppm. The more upfield calculated shifts relative to the experimental values may be attributed to the effect of solvent deshielding. The calculated $^1H^{\epsilon}$ chemical shift (9.7 ppm) is consistent with the formation of a hydrogen bond to a phosphate oxygen. However, it is shifted more downfield relative to the experimental value (8.1 ppm).

The unobservable (due to intermediate exchange broadening) Arg 18 $N^{\eta}-H^{\eta}$ crosspeak,[22,23] associated with the more downfield shift of $^{15}N^{\eta}$ (73.2 ppm) relative to the other N^{η} (71.0 ppm), probably corresponds to $H^{\eta 21}$ (10.5 ppm) hydrogen-bonded to a phosphate oxygen and $H^{\eta 22}$ (7.0 ppm) in the MD PLCC·pY structure (see also discussion below). The computed chemical shifts for Arg 18 $H^{\eta 11}$ (5.9 ppm) and $H^{\eta 12}$ (2.9 ppm), which are not protected by a nonsolvent hydrogen bond, are upfield relative to the experimental average value (6.1 ppm). This is likely due to the deshielding effect of solvent. The Arg 18 H^{ϵ} proton, which does not interact with any hydrogen-bond acceptors in the MD structure, has an upfield chemical shift (2.7 ppm) relative to the experimental value (7.4 ppm). The experimental result suggests participation of the H^{ϵ} proton in a hydrogen bond; however, the order parameter of the Arg 18 $N^{\epsilon}-H^{\epsilon}$ bond (0.72 $\pm$ 0.01) is significantly lower than that of Arg 37 (0.80 $\pm$ 0.01),[23] indicating higher mobility of the former on the picosecond time scale.

The difference between calculated chemical shifts in models I and II can be attributed to ring current effects from the pTyr ring. Table 4 shows that pTyr ring current effects do not seem significant: the largest effect is only 0.3−0.4 ppm from Arg 18 $H^{\eta 11}$ and $H^{\eta 12}$ protons located above/below the pTyr ring plane. Overall, the calculated results are generally in good agreement with experiment. Hydrogen-bonding effects are the dominant factors causing the chemical shift nonequivalence of protons, whereas ring current effects do not appear to contribute significantly.

Comparison with Crystal Structures of Related SH2 Domains. The structures of Lck, Src, and SypN SH2 domains bound to high-affinity phosphopeptides have already been solved by X-ray crystallography.[16−18] These structures present a detailed picture of the interatomic interactions at the pTyr binding site, which can be compared with the hydrogen-bond network at the pTyr-binding region observed in the MD and NMR[14,19] structures. To facilitate the following discussion, the nomenclature proposed by Eck *et al.*[16] (Scheme 2), based on the common scaffolding of these SH2 domains, will be used.

A key difference between PLCC and the Lck, Src, and SypN SH2 domains lies in the variability of three out of the four PLCC arginines that constitute the pTyr-binding pocket. In Src and Lck there is a serine at position 39 (βB7) and a lysine at position 59 (βD6). In SypN Arg 18 (αA2) is replaced by glycine, whereas positions 39 and 59 are the same as in Src and Lck. Arg 37 (βB5) is, in sequences from 67 SH2 domains,[40] the only one strictly conserved out of the four arginines in the pocket.

A common feature among the three aforementioned SH2-phosphopeptide crystal structures and the PLCC·pY MD structure is the ion-pair interactions between Arg βB5 and the pTyr phosphate oxygens. The Arg βB5 interactions in the MD structure (Figure 3) are analogous to those of the Src and SypN structures:[17,18] all three structures show one Arg βB5 N^{η} nitrogen binding to a terminal oxygen and the other one forming bidentate ion-pair interactions simultaneously with the same terminal oxygen and another phosphate oxygen. This two-pronged interaction may explain why the charge-conserving mutation of Arg βB5 to lysine in the c-Abl SH2 domain led to a loss in binding affinity,[41] and why this residue is totally conserved among 67 sequences.[40] Furthermore, the interactions among Arg βB5 and highly conserved residues Ser βC5 and His βD4 are also remarkably similar in the MD PLCC·pY, Src and SypN structures. Notably, in the MD structure of PLCC free SH2, the interaction between Arg βB5 and His βD4 is maintained via a bridging water, which is likely to play a structural role since these two residues are relatively buried (Table 5). Thus, highly conserved Ser βC5 and His βD4 appear to play an important role in positioning Arg βB5 so as to interact favorably with the phosphate.

Another common feature among the three SH2-phosphopeptide crystal structures and the PLCC·pY MD structure is the interactions between the aliphatic part of the βD6 side chain

(40) Russell, R. B.; Breed, J.; Barton, G. J. *FEBS* **1992**, *304*, 15−20.
(41) Mayer, B. J.; Jackson, P. K.; Van Etten, R. A.; Baltimore, D. *Mol. Cell. Biol.* **1992**, *12*, 609−618.
(42) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink-Peters, T.; Hummelinkand B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr.* **1979**, *B35*, 2331−2339.
(43) Cook, R. L.; DeLucia, F. C.; Helminger, P. *J. Mol. Spectrosc.* **1974**, *53*, 62−76.
(44) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *Gaussian 92, Revision A*; Gaussian Inc.: Pittsburgh, PA, 1992.
(45) MacKerell, A. D., Jr.; Wiorkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.
(46) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(47) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 1276−1284.
(48) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 569−578.
(49) Almlof, J.; Fischer, T. H.; Gassman, P. G.; Ghosh, A.; Haser, M. *J. Phys. Chem.* **1993**, *97*, 10964−10970.
(50) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomolec.NMR* **1995**, *6*, 135−140.

and the pTyr ring. Thus, the substitution of arginine with lysine at the βD6 position in Src, Lck, and SypN seems to have no effect on the aliphatic−aromatic interactions with pTyr. In addition, both the Src and MD complexes show the βD6 terminal *amino* group interacting with the pTyr ring. Furthermore, the βB7 side chain hydrogen-bonds to one or more terminal phosphate oxygens in the Src, SypN, and MD PLCC·pY structures. However, in the Src and SypN structures, the short, neutral side chain of Ser βB7 does not interact with the pTyr ring while in the MD structure, Arg βB7 $N^{\epsilon}-H^{\epsilon}$ and one of the amino groups are located in proximity and equatorially with respect to the pTyr ring.

Some discrepancy is found for the contacts of the Arg αA2 guanidinium protons with pTyr among the MD PLCC·pY and the crystal structures. In Src and Lck, the N^{ϵ} and one of the N^{η} interact with the same (terminal) phosphate oxygen. In addition, both N^{η} nitrogens serve as hydrogen-bond donors to the backbone carbonyl oxygen of the peptide residue immediately N-terminal to pTyr (pY−1), resulting in three simultaneous interactions for one of the N^{η} groups with the pY−1 carbonyl oxygen, a phosphate oxygen, and the aromatic ring. Two of these interactions are preserved in the MD PLCC·pY complex: an $N^{\eta}-H^{\eta}$ group interacts with a terminal phosphate oxygen and the His βD4 imidazole ring, whereas the other $N^{\eta}-H^{\eta}$ group interacts with the pTyr ring (Table 6). The third interaction between Arg αA2 N^{η} and the pY−1 carbonyl oxygen is not seen in the PLCC·pY MD/NMR structure, consistent with the lack of NOEs between the two residues.[19] However, it is not clear whether the absence of NOEs is truly a reflection of absence of interactions or the result of flexibility in this region.

Overall, the recognition of the pTyr is remarkably similar among the crystal structures and the MD structure. The strictly conserved arginine at βB5 forms bidentate ion-pairing interactions with two phosphate oxygens. The highly conserved arginine at αA2 interacts simultaneously with a phosphate oxygen and the pTyr ring. Finally, the least conserved of the three key recognition residues, either βB7 (Src, SypN, PLCC) and/or βD6 (Lck), is involved in hydrogen-bonding interactions with a phosphate oxygen and/or amino−aromatic/aliphatic−aromatic interactions with the pTyr ring.

Interpretation of NMR Chemical Shifts. In addition to the PLCC·pY MD structure and *ab initio* chemical shift calculations, the free and PLCC·P_i structures complement the NMR study[23] by providing an atomic picture and plausible interpretations for the observed chemical shift changes in different environments (see Introduction).

P_i−Imidazole N^{ϵ} and H^{ϵ} Chemical Shifts. The changes in hydrogen-bonding interactions of the H^{ϵ} protons of Arg 18, 37, 39, and 59 in the free structure upon binding isolated phosphate ion, seem to account for the observed N^{ϵ} and H^{ϵ} chemical shift differences upon substitution of imidazole with phosphate ion buffer. In the free structure, a water molecule (Bul 1061) forms a hydrogen bonding bridge between Arg 37 H^{ϵ} and His 57 $N^{\delta 1}$, both of which are relatively solvent inaccessible (Table 5), so that their interactions with Bul 1061 are likely to be long-lived. On the other hand, Arg 37 H^{ϵ} interacts with Ser 48 $O^{\gamma 1}$ during the last 60 ps in the PLCC·P_i simulation (see Results). Thus, the loss of the Arg 37 $H^{\epsilon}-$Bul 1061 hydrogen bond in the free structure cannot account for the observed *upfield* shift of the Arg 37 $N^{\epsilon}-H^{\epsilon}$ crosspeak. Instead, the latter may result from the different electron distributions between Arg 37 side chains in the free and PLCC·P_i structures. Isolated arginine bears one positive charge, which is likely to be equally distributed on its amino (η position) and imino (ϵ position) side chain groups. In the PLCC·P_i structure however, both Arg 37 amino groups form

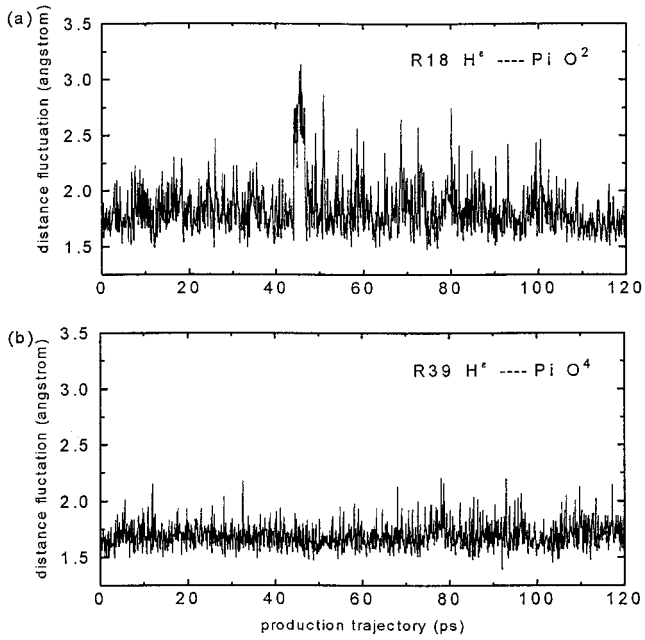

Figure 5. Time series of distance fluctuations for (a) Arg 18 $H^{\epsilon}-O^2$ and (b) Arg 39 $H^{\epsilon}-O^4$ in PLCC·P_i.

two strong hydrogen bonds with two phosphate oxygens. Thus, the electron density is expected to be more localized on the imino group in the PLCC·P_i structure relative to free PLCC, shielding the Arg 37 H^{ϵ} proton more in the PLCC·P_i structure than in the free one so as to account for the observed upfield shift.

The *downfield* shifts of the Arg 39 and Arg 18 N^{ϵ} and H^{ϵ} peaks are caused by deshielding effects of a hydrogen bond with a phosphate oxygen in the PLCC·P_i structure. The smaller downfield shift observed for Arg 18 relative to Arg 39 is probably due to (i) extra shielding of the Arg 18 H^{ϵ} proton in the free MD structure, which not only interacts with water molecules but also forms a weak hydrogen bond with Glu 22 $O^{\epsilon 2}$; (ii) the weaker Arg 18 $H^{\epsilon}-O^2$ hydrogen bond compared to the Arg 39 $H^{\epsilon}-O^4$ hydrogen bond in the PLCC·P_i structure, as indicated by their relative distance fluctuations (Table 2 and Figure 5); (iii) ring current effects stemming from the axial position of Arg 18 H^{ϵ} with respect to the His 57 ring plane in PLCC·P_i ($R_c = 4.4$ Å, $\theta = 9°$ compared to $R_c = 3.8$ Å, $\theta = 47°$ in the free structure).

The downfield shift of the Arg 59 H^{ϵ} resonance upon binding of phosphate ion was initially thought to reflect weak or exchanging hydrogen bonds to the phosphate ion.[23] During the PLCC·P_i simulation Arg 59 H^{ϵ} does not interact with the phosphate but forms weak, transient hydrogen bonds with several water molecules. This is consistent with the negligible shift of the Arg 59 N^{ϵ} resonance upon binding of phosphate ion, the low $N^{\epsilon}-H^{\epsilon}$ order parameter (0.31 ± 0.01) for Arg 59 in the PLCC·pY complex which indicates unrestricted mobility of the $N^{\epsilon}-H^{\epsilon}$ bond on the picosecond/nanosecond time scale and relatively high H^{ϵ} exchange rates,[23] implying substantial solvent exposure (see also Table 5).

pY1021−P_i N^{ϵ} and H^{ϵ} Chemical Shifts. Upon replacing phosphate ion with phosphopeptide-containing buffer the upfield shifts of the Arg 39 and 59 $N^{\epsilon}-H^{\epsilon}$ resonances suggest that the position of these groups is above or below the plane of the ring.[23] For Arg 59, the MD PLCC·pY structure supports this interpretation (see Table 7) whereas in the NMR structure, the Arg 59 N^{ϵ} and H^{ϵ} atoms are too far from the pTyr ring centroid (distances greater than 6.8 Å) to form amino−aromatic interactions. However, for Arg 39, both MD and NMR PLCC·pY structures show the $N^{\epsilon}-H^{\epsilon}$ group located equatorial to the pTyr

ring plane (Table 6). This discrepancy may be due to the positional imprecision of the Arg 39 and pTyr side chains relative to each other in the NMR and MD structures since only two NOE and two hydrogen-bonding restraints between pTyr and Arg 39 were imposed during NMR structure refinement,[19] and only one of the NOE restraints involves side chain−side chain interactions.

Upon replacing phosphate ion with peptide-containing imidazole buffer, the Arg 18 N^{ϵ} peak shifted *downfield* whereas its H^{ϵ} resonance shifted slightly *upfield*. The latter appears to be consistent with the loss of the Arg 18 $H^{\epsilon}-O^2$ hydrogen bond in the MD PLCC·pY structure. However, the observed *downfield* shift of Arg 18 N^{ϵ}, the discrepancy between the computed and experimental Arg 18 H^{ϵ} shifts (Table 4) as well as the X-ray structures of Src and Lck bound SH2 domains suggest the presence of a hydrogen bond to Arg 18 H^{ϵ}. The latter may be absent from both NMR and MD structures due to the lack of NOEs between Arg 18 and pTyr.[14,19]

The almost unperturbed shift of the Arg 37 $N^{\epsilon}-H^{\epsilon}$ crosspeak upon substitution of phosphate ion with peptide-containing imidazole buffer appears to be consistent with the similar hydrogen-bonding networks involving Arg 37 H^{ϵ} found in the PLCC·P_i and PLCC·pY structures (see Results, Tables 1 and 2).

Arg 37 N^{η} and H^{η} Chemical Shifts. In contrast to phosphopeptide-containing imidazole buffer, only two crosspeaks corresponding to two $N^{\eta}-H^{\eta}$ of one of the two Arg 37 NH_2 groups were observed at 30 °C in the phosphate ion buffer.[23] The other two peaks corresponding to two $N^{\eta}-H^{\eta}$ of the other Arg 37 NH_2 group can only be measured at a lower temperature (10 °C), suggesting the presence of intermediate exchange motional processes (microsecond-to-millisecond time scale) causing resonance broadening. The similarity of the hydrogen bonding distributions in the PLCC·P_i and PLCC·pY MD structures agrees with the similar $N^{\eta}-H^{\eta}$ resonance patterns of PLCC in these two environments. However, the bound phosphate ion may not exist solely as a dianion but rather a dianion in equilibrium with a monoanion if its pK_a value is around the pH of the chemical shift experiments. This may diminish the charge-charge interaction between phosphate ion and the Arg 37 amino side chain and may explain why only two Arg 37 $N^{\eta}-H^{\eta}$ crosspeaks were seen at 30 °C. Lowering the temperature would slow down the $C^{\zeta}-N^{\eta}$ bond rotation and would favor the existence of dianionic phosphate, consistent with the observation of all four Arg 37 $N^{\eta}-H^{\eta}$ crosspeaks at a lower temperature (10 °C).

On replacement of phosphate ion buffer with peptide-containing imidazole buffer all four crosspeaks corresponding to the four Arg 37 $N^{\eta}-H^{\eta}$ bonds appeared at 30 °C, indicating that rotation around both $C^{\zeta}-N^{\eta}$ bonds had been considerably slowed down in the PLCC·pY complex. However, the pair of peaks that was observable at 30 °C in the phosphate ion buffer experiment was still more intense than the other pair, implying a persisting difference in the motional restriction around the two $C^{\zeta}-N^{\eta}$ bonds. In the PLCC·pY MD structure (Figure 3), the Arg 37 $N^{\eta 1}$ amino group participates in three hydrogen bonds with two phosphate oxygens and His 57, whereas the Arg 37 $N^{\eta 2}$ amino group is involved in one hydrogen bond to a phosphate oxygen and one to a water molecule. The multiplicity of stable hydrogen bonds involving one of the two NH_2 groups relative to the other in the MD structure could explain the experimentally observed resonance intensity differences. Simultaneous breaking and reorganization of the three hydrogen bonds involving the $N^{\eta 1}$ amino group, and thus rotation around the $C^{\zeta}-N^{\eta 1}$ bond, would be expected to be far more inhibited than dynamical exchange of the single protein−peptide hydrogen bond involving the $N^{\eta 2}$ amino group. Therefore, the current PLCC·pY MD structure not only suggests an unambiguous assignment of NMR $N^{\eta}-H^{\eta}$ resonances to distinct protons of Arg 37 side chain (see above and Table 4), but also provides an explanation for the different relative intensities of the Arg 37 amino resonances observed experimentally.

Arg 39 N^{η} and H^{η} Chemical Shifts. Substitution of imidazole with phosphate ion buffer caused no observable effect on the arginine $N^{\eta}-H^{\eta}$ resonances except for those of Arg 37.[23] This may be due to intermediate exchange broadening causing certain resonances to be undetectable in the spectra (see below). However, upon substitution of phosphate ion with phosphopeptide-containing imidazole buffer, an "apparent" downfield shift (in the ^1H dimension) of one of the two observable Arg 39 $N^{\eta}-H^{\eta}$ crosspeaks was found. It was interpreted as arising from an equatorial position of the $N^{\eta}-H^{\eta}$ group with respect to the pTyr ring.[23] This is consistent with $N^{\eta 1}$ and $H^{\eta 12}$ distances to the ring centroid and inclination angles of 5.4 Å, 72° and 4.9 Å, 81°, respectively in the PLCC·pY MD structure. Compared to the NMR structure (Table 6), the $H^{\eta 12}$ atom is now closer by 2.7 Å to the aromatic ring centroid in the MD structure. However, Table 4 shows that the contribution of ring current effects is no more than 0.2 ppm: the average shifts of $H^{\eta 11}$ and $H^{\eta 12}$ are 7.35 and 7.20 ppm in the presence and absence of a benzyl ring, respectively, very close to the observed downfield shift at 7.3 ppm in phosphopeptide-containing imidazole buffer. It is possible that in phosphate ion buffer, there is also a $N^{\eta}-H^{\eta}$ resonance around 7.3 ppm, reflecting the interaction of $H^{\eta 11}$ with phosphate oxygen O^3 found in the PLCC·P_i structure, but it is unobservable due to the resonance broadening.

Arg 18 N^{η} and H^{η} Chemical Shifts. Despite an Arg 18($H^{\eta 11}$)$-O^2$ hydrogen bond in the PLCC·P_i structure, no significant changes in the Arg 18 $N^{\eta}-H^{\eta}$ chemical shifts were detected upon substitution of imidazole with phosphate ion buffer. In analogy to Arg 39, this may be due to intermediate exchange broadening causing certain resonances to be undetectable in the spectra. In contrast to Arg 39 however, the $N^{\eta}-H^{\eta}$ crosspeak ($\delta N^{\eta} = 71.0$ ppm; $\delta H^{\eta} = 6.1$ ppm)[22] of Arg 18 was shifted upfield upon replacing phosphate ion with phosphopeptide. This is consistent with $H^{\eta 11}$ positioned axial to the pTyr ring plane ($R_c = 5.1$ Å, $\theta = 21°$) as well as to the His 57 ring plane in the PLCC·pY structure ($R_c = 3.4$ Å, $\theta = 14°$). The second crosspeak ($\delta N^{\eta} = 73.2$ ppm)[22] of Arg 18 cannot be detected at 30 °C due to broadening of the H^{η} resonances. The unobservable crosspeak with the more downfield shift in the ^{15}N dimension (73.2 ppm) relative to the one at 71.0 ppm is consistent with an Arg 18 $H^{\eta 21}$ proton hydrogen bonding to a phosphate oxygen in the PLCC·pY MD structure. This interaction may slow down the rotation rate of the Arg 18 $N^{\epsilon}-C^{\zeta}$ bond to a rate comparable to the time scale of the NMR chemical shift measurement, thus explaining why the second peak could not be observed at 30 °C, and furthermore predicting its appearance at lower temperatures.

Arg 59 N^{η} and H^{η} Chemical Shifts. There are no hydrogen bonds between the side chain atoms of Arg 59 and the pTyr in the PLCC·pY MD and NMR structures. This is in accord with the lack of chemical shift changes in the $N^{\eta}-H^{\eta}$ resonances upon binding phosphate or phosphopeptide as well as the higher $^{13}C^{\zeta}$ T_1 value of Arg 59 relative to Arg 18, 37, and 39, indicating that the motions of the C^{ζ} carbons of the latter three arginines are more constrained than those of Arg 59.[23] Furthermore, the simulation preserves the Arg 59 aliphatic side chain−aromatic ring interactions seen in the NMR structure, as well as in the crystal structures of the closely related Src and SypN SH2 domains.[17,18]

Conclusions

The present simulations appear to have served in the refinement of a protein NMR structure. The interactions at the pTyr−binding site of the MD PLCC·pY complex are on the whole compatible with the observed NOEs and NMR chemical shift data,[23,19] despite the fact that most of these contacts are absent in the NMR structure used as the simulation starting point. In particular, the good agreement between predicted chemical shifts based on the MD-derived structure and experimental ones, the striking similarity between the present model of the pTyr binding site and crystal structures of related SH2 domains complexed with high affinity phosphopeptides, as well as the overall ability to explain the NMR chemical shift data indicate that the MD PLCC·pY structure is a realistic representation of (at least) the dominant pTyr-binding interactions. Specifically, the close agreement of the theoretical chemical shifts involving Arg 37 with experiment is particularly reassuring, since this is the only absolutely conserved residue at the binding site among various SH2 domains[40] and is therefore likely to account for a significant portion of the phosphopeptide binding free energy. All these results, in turn, are evidence for the adequacy of the newly-developed pTyr parameters and the current CHARMM force field.[30]

The present simulation permits identification of the primary pTyr recognition elements, viz., Arg 37 (βB5), 39 (βB7), and 18 (αA2), as well as interactions stabilizing them. Specifically, the hydrogen-bond network among Arg 37, Ser 48, and His 57, contributing to stabilization of the pTyr residue within the pocket, may explain the very high conservation of these residues among various SH2 domains.[13] In the same regard, the particularly low pKa of His 57[24] and thus the neutrality of the residue at the experimental and physiological pH, crucial for formation of the observed interactions with Arg 37 and Ser 48, appears to be of biological relevance.

The MD structure of the pTyr binding pocket is characterized by a substantially increased number of hydrogen bonds involving phosphate oxygens (10) relative to the NMR structure (one). In particular, Arg 37 forms bidentate ion-pairing interactions with two phosphate oxygens, while Arg 39 and 18 interact with three and one phosphate oxygens, respectively. There is a dearth of hydrogen bonds in the NMR structure despite the fact that hydrogen-bond restraints involving both N^{η} groups of Arg 37 and Arg 39 $N^{\epsilon}-H^{\epsilon}$, and the phosphate group were imposed during the NMR structure refinement procedure.[19] This is not surprising as the Coulombic and Lennard-Jones terms were absent in the potential energy functions used in the NMR refinement, and the hydrogen-bond restraints (based on chemical shift information) by necessity constitute nonspecific interactions involving *any* phosphate oxygen.

The simulations and chemical shift calculations provide an assignment of the observed $^1H^{\epsilon}$ and H^{η} resonances to individual proton(s) of each of the pTyr recognition arginines as well as interpretations of the chemical shift changes observed experimentally upon substitution of imidazole with phosphate ion buffer and of the latter with peptide-containing imidazole buffer.[23] In particular, in the case of Arg 37 the current structural model has offered an unambiguous one-to-one assignment of distinct NMR resonances to individual Arg 37 H^{η} protons, and it also accounts for the differential $N^{\eta}-H^{\eta}$ crosspeak intensities observed in the spectra.[23] Furthermore, the experimentally predicted amino−aromatic and aliphatic−aromatic interactions involving pTyr and arginines 18 and 59 are well accounted for by the simulations.

In conclusion, the present findings suggest a strategy for enhancement of the information content of hydrogen-bonding, amino−aromatic, and aliphatic−aromatic interactions in protein NMR structures through a combination of NOE and chemical shift measurements and molecular dynamics simulation and chemical shift calculation approaches. A similar strategy can be applied to other systems whose structures cannot be obtained at high resolution.

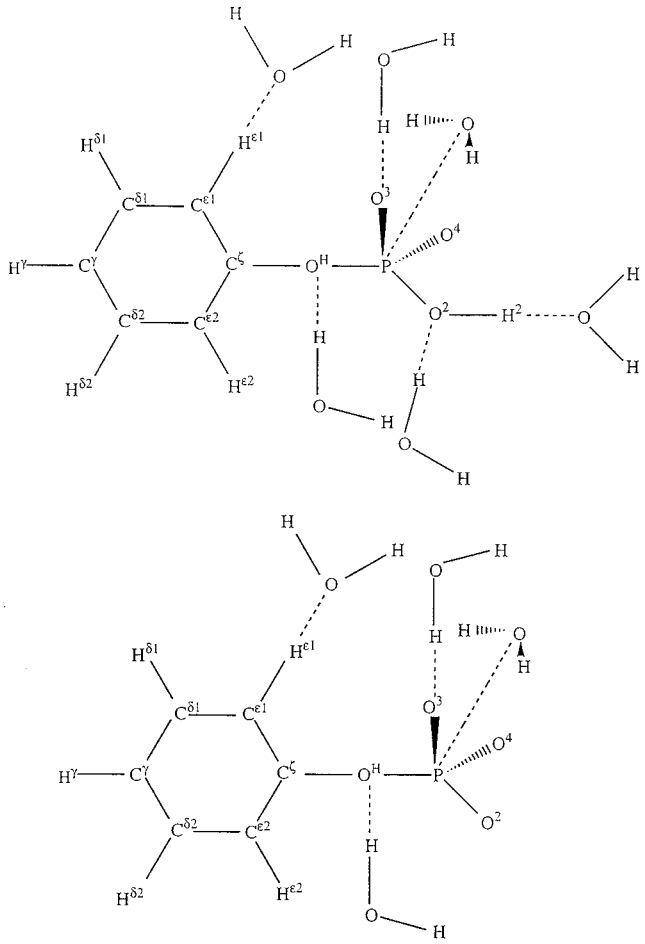

Figure 6. Phenyl phosphate−water interactions. Geometries for the anionic (HPP) and dianionic species (PP). Atom names represent the nomenclature used in the text.

Acknowledgment. We are indebted to S. M. Pascal, T. Yamazaki, A. U. Singer, L. E. Kay, and J. D. Forman-Kay for providing us with a preprint of their NMR work[23] prior to publication and for their refined solution structure, the list of experimental restraints, and helpful discussions. We also thank T. Yamazaki and A. U. Singer for providing us with information regarding the histidine and phosphotyrosine ionization states. We are grateful to Prof. M. Karplus for the CHARMM program. M.-H.F. and M.P. were supported by the Institute of Biomedical Sciences, Academia Sinica, Taiwan. M.P. also acknowledges a grant from the Medical Research Council of Canada. This work was supported by the Institute of Biomedical Sciences, Academia Sinica, and the National Center for High Performance Computing, Republic of China. M.-H.F. and M.P. contributed equally to this work.

Appendix: Parameterization Procedure

Parameterization of monoanionic and dianionic states of tyrosine phosphate was based on the model compounds, hydrogen phenyl phosphate (HPP) and phenyl phosphate (PP), respectively (Figure 6). The parameters were optimized to reproduce both experimental data and *ab initio* calculations. The experimental data included a survey of the Cambridge Crystal Data Bank (CCDB)[42] for all phosphate-containing compounds in which one of the ester linkages involved a phenyl moiety,

Table 9. Phenyl Phosphate to Water Minimum Interaction Energies and Geometries from the Empirical and *Ab Initio* Calculations

	ab initio		empirical	
atom that interacts with water	E_{min} (kcal/mol)	R_{min} (Å)	E_{min} (kcal/mol)	R_{min} (Å)
HPP(O^H)	−10.12	2.04	−9.84	1.92
HPP(O^2)	−7.33	2.04	−7.30	1.85
HPP(H^2)	−0.59	2.04	−0.49	1.86
HPP(P)	−8.95	3.84	−9.79	3.61
HPP(O^3)	−13.03	1.86	−13.05	1.66
HPP($H^{\epsilon 1}$)	2.43	3.07	2.37	2.74
PP(O^H)	−19.01	1.87	−18.72	1.84
PP(P)	−19.17	3.25	−23.51	3.28
PP(O^3)	−21.26	1.75	−21.18	1.60
PP($H^{\epsilon 1}$)[a]	4.57	4.58	4.76	4.58

[a] No minima was present in the *ab initio* or empirical calculations.

Table 10. Partial Atomic Charges and Atom Types for the Phenyl Phosphate Atoms

	anionic HPP		dianionic PP	
atom	type	charge	type	charge
$C^{\epsilon 1}$	CA	−0.21	CA	−0.21
$H^{\epsilon 1}$	HP	0.21	HP	0.21
$C^{\epsilon 2}$	CA	−0.21	CA	−0.21
$H^{\epsilon 2}$	HP	0.21	HP	0.21
C^{ζ}	CA	−0.16	CA	−0.28
O^H	ON2B	−0.36	ON2B	−0.28
P	P	1.40	P	1.44
O^2	ON4	−0.76	ON3	−0.96
H^2	HN4	0.40		
O^3	ON3	−0.76	ON3	−0.96
O^4	ON3	−0.76	ON3	−0.96

two of the phosphate oxygens were anionic and resolved to a *R*-factor of 0.08 or better. *Ab-initio*-based data included geometries and rotation surfaces about the $C^{\zeta}-O^H$ bonds of HPP and PP, as well as minimum interaction energies and distances between the model compounds and waters.[43]

The *ab initio* calculations were performed with the program Gaussian 92.[44] The geometries and $C^{\zeta}-O^H$ rotation surfaces of HPP and PP were optimized with the $C^{\zeta}-O^H-P-O^2$ and $O^H-P-O^2-H^2$ dihedrals constrained to 180° at the HF/6-31+G* level, while the phenyl phosphate−water interaction energies and distances were optimized with the monomer geometries fixed using the HF/6-31G* basis set. Use of the latter is consistent with the version 22 CHARMM phosphate parameterization[30] and previous studies on a variety of systems.[46] The empirical energy minimizations were performed with version 22 of the program CHARMM[26] using adopted-basis Newton−Raphson and Newton−Raphson minimizers to root-mean-square forces of 10^{-6} kcal/(mol/Å). No truncation of the nonbonded interactions was performed. The CHARMM-modified TIP3P water model[27] was employed. Initial parameters for the phosphate group of phenyl phosphate were obtained from monoanionic and dianionic methyl phosphate parameters in the version 22 CHARMM all-hydrogen nucleic acid parameters,[45] whereas starting parameters for the aromatic part of phenyl phosphate came from phenol parameters in the version 22 CHARMM all-hydrogen protein parameter set.[30]

The partial atomic charges were optimized to reproduce the minimum interaction energies and geometries of the model compounds with individual water molecules (Figure 6 and Table 9); the optimized charges and atom types are listed in Table 10. Table 9 shows that for the monoanionic species, the *ab initio* and empirical interaction energies are in good agreement, while the empirical distances are generally shorter than the *ab initio* values in order to reproduce the macroscopic properties.[47,48] The more negative monoanionic HPP(P)−water empirical interaction energy relative to the *ab initio* value, which also occured in the case of methyl phosphate,[45] is required to maintain good agreement for the other interaction geometries studied. For the dianionic species, the *ab initio* and empirical interaction energies are in good agreement, except the DPP-(P)−water empirical interaction energy, which is overestimated as in the monoanion case. Altering the VDW parameters of the phosphorous atom led to better agreement for this particular interaction; however, such a change was not incorporated to maintain consistency with the existing nucleic acid parameters. The excellent agreement between the *ab initio* and empirical interaction energy of water with the $C^{\epsilon 1}-H^{\epsilon 1}$ site in both species suggests that relative to the standard charge of −0.115 on an aromatic carbon and 0.115 on an aromatic hydrogen, the increased charge separation at $C^{\epsilon 1}$ and $C^{\epsilon 2}$, as shown in Table 10, is appropriate for the present system.

Table 11. Comparison of the Empirical, CCDB Survey, and *ab Initio* Monoanionic Phenyl Phosphate Geometries

parameter	empirical[a]	survey	*ab initio*[a]
$C^{\zeta}-O^H$	1.392	1.390 ± 0.013	1.340
$P-O^H$	1.614	1.610 ± 0.017	1.653
$P-O^2$	1.566	1.595 ± 0.034	1.628
$P=O^3/O^4$	1.474(1.476)	1.481 ± 0.011	1.477(1.474)
$C^{\epsilon 1}/C^{\epsilon 2}/C^{\zeta}$	1.400(1.412)	1.380 ± 0.008	1.393(1.394)
$P-O^H-C^{\zeta}$	125.8	124.2 ± 2.4	125.5
$O^2-P=O^3/O^4$	104.3(104.3)	108.8 ± 1.8	109.9
$O^H-P=O^3/O^4$	108.1(108.1)	107.9 ± 1.4	109.4
$O^3=P=O^4$	121.0	119.9 ± 2.7	121.2
$P-O^2-H^2$	106.2		109.4
$O^H-C^{\zeta}-C^{\epsilon 1}/C^{\epsilon 2}$	117.1(124.2)	119.4 ± 0.4	117.0
$C^{\epsilon 1}-C^{\zeta}-C^{\epsilon 2}$	118.7	121.1 ± 0.8	119.4

[a] Minimum with the $C^{\zeta}-O^H-P-O^2$ and $O^H-P-O^2-H^2$ dihedrals constrained to 180°. The values in parentheses correspond to the minimum not being symmetric. Distances are in angstroms and angles in degrees.

Table 12. Comparison of the Empirical and *ab Initio* Dianionic Phenyl Phosphate Geometries

parameter	empirical[a]	*ab initio*[a]
$C^{\zeta}-O^H$	1.407	1.291
$P-O^H$	1.642	1.826
P=O	1.485	1.503(1.496)
$C^{\epsilon 1}/C^{\epsilon 2}-C^{\zeta}$	1.400 (1.410)	1.416 (1.413)
$P-O^H-C^{\zeta}$	129.0	126.9
$O^H-P=O$	103.6 (100.7)	101.4 (99.1)
O=P=O	115.6 (115.2)	115.5 (117.3)
$O^H-C^{\zeta}-C^{\epsilon 1}/C^{\epsilon 2}$	124.2 (117.3)	124.8 (118.4)
$C^{\epsilon 1}-C^{\zeta}-C^{\epsilon 2}$	118.6	116.6
$C^{\epsilon 1}-C^{\zeta}-O^H-P$	0.0	0.0

[a] Minimum with the $C^{\zeta}-O^H-P-O^2$ dihedrals constrained to 180°. The values in parentheses correspond to the identical nonsymmetric internal coordinate. Distances are in angstroms and angles in degrees.

The covalent bond lengths, angles, and dihedral angles for the monoanionic and dianionic species were optimized to reproduce the CCDB[42] and *ab initio* geometries (Tables 11 and 12). To obtain the agreement in Tables 11 and 12, it was necessary to introduce a new atom type, ON2B (Table 10), for the O^H atom. This is suggested to be due to conjugation between the phenyl ring and the phosphate group (leading to significant differences compared to methyl phosphate), as evidenced by a significant decrease in the C−O bond length upon going from methyl phosphate (1.443 ± 0.009 Å) to phenyl phosphate (1.390 ± 0.013 Å) from a survey of CCDB. The new atom type also aided in optimization of the surrounding angle parameters and the dihedral parameters (see below). For

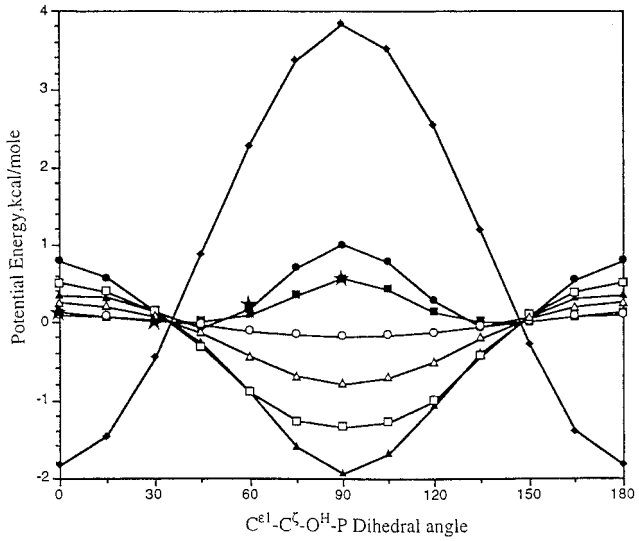

Figure 7. Potential energy (kcal/mol) as a function of the $C^{\epsilon 1}-C^{\zeta}-O^{H}-P$ dihedral angle from the HF/6-31+G* ab initio (*) and empirical calculations, which include the total (filled squares), electrostatic (filled circles), VDW (filled triangles), bond (open circles), angle(open squares), Urey−Bradley (open triangles), and dihedral (filled diamonds) contributions.

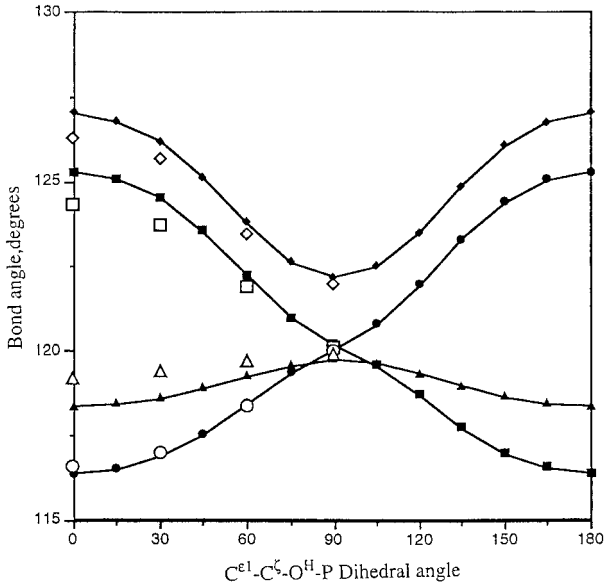

Figure 8. Angles as a function of the $C^{\epsilon 1}-C^{\zeta}-O^{H}-P$ dihedral angle from the HF/6-31+G* *ab initio* (ai) and empirical (exp) calculations for the $C^{\epsilon 1}-C^{\zeta}-O^{H}$ (ai (open squares); emp (filled squares)), the $C^{\epsilon 2}-C^{\zeta}-O^{H}$ (ai (open circles); emp (filled circles)), the $C^{\epsilon 1}-C^{\zeta}-O^{\epsilon 2}$ (ai (open triangles); emp (filled triangles)), and the $C^{\zeta}-O^{H}-P$ (ai (open diamonds); emp (filled diamonds)) angles.

the monoanion, agreement is within the error reported in the CCDB, whereas for the dianion, the distances and angles agree to within ±0.02 Å and ±2°, respectively. The discrepancy in $C^{\zeta}-O^{H}$ (0.12 Å) and $P-O^{H}$ (−0.18 Å) is deemed acceptable based on limitations in the Hartree−Fock approximation, which has been reported to result in the favoring of alternating single and double bonds in conjugated systems.[49]

Force constants were optimized to reproduce the energetics and structural aspects of the $C^{\epsilon 1}-C^{\zeta}-O^{H}-P$ dihedral angle rotation; the results of the monoanion are shown in Figures 7 and 8, while those of the dianion are listed in Table 13. Figure 7 shows the *ab initio* and empirical potential energies as a function of the $C^{\epsilon 1}-C^{\zeta}-O^{H}-P$ dihedral angle rotation as well as a breakdown of the empirical results into the individual energy terms for the monoanionic species. All of the terms in the empirical energy function make a significant contribution to the empirical potential energy surface, the largest contribution stemming from the dihedral term due to the conjugated nature of the ester linkage in phenylphosphate. The electrostatic and VDW contributions are due to interactions of the phosphate with the ring; the magnitude of these terms emphasizes the need for an iterative approach to the parameterization process as changing the intermolecular terms will affect the intramolecular ones. To check that the change in geometry as a function of the $C^{\epsilon 1}-C^{\zeta}-O^{H}-P$ dihedral angle was being properly reproduced, the $C^{\epsilon 1}-C^{\zeta}-O^{H}$, $C^{\epsilon 2}-C^{\zeta}-O^{H}$, $C^{\epsilon 1}-C^{\zeta}-C^{\epsilon 2}$, and $C^{\zeta}-O^{H}-P$ angles were obtained from both the *ab initio* and empirical surfaces, as shown in Figure 8. The change in the angles as a function

Table 13. Dianionic Phenyl Phosphate Conformational Energies and Geometries of Selected Angles

	empirical		ab initio	
parameter $C^{\epsilon 1}-C^{\zeta}-O^{H}-P$	0.0	90.0	0.0	90.0
energy (kcal/mol)	0.0	2.86	0.0	3.40
$O^{H}-C^{\zeta}-C^{\epsilon 1}$	124.2	120.1	124.8	121.4
$O^{H}-C^{\zeta}-C^{\epsilon 2}$	117.2	120.1	118.4	121.3
$C^{\epsilon 1}-C^{\zeta}-C^{\epsilon 2}$	118.6	119.8	116.7	117.2
$C^{\zeta}-O^{H}-P$	129.0	125.4	126.9	123.4

of the $C^{\epsilon 1}-C^{\zeta}-O^{H}-P$ dihedral angle from the *ab initio* calculations is adequately reproduced by the empirical force field, indicating that the balance between the intermolecular and intramolecular portions of the force field is reasonable. The force constants from the monoanionic species were directly transferred to the dianionic species. The resulting empirical energies and bond angles of the dianionic species for $C^{\epsilon 1}-C^{\zeta}-O^{H}-P = 0°$ (minimum) and 90° (maximum) are compared with the corresponding *ab initio* values in Table 13. The empirical force field underestimates the barrier height but the difference is not large enough to warrant the introduction of new atom types, whereas the *ab initio* angles are reproduced by the empirical force field to within 2.6°.

JA961530R