

A Computational Study of the Phosphorylation Mechanism of the Insulin Receptor Tyrosine Kinase

Baojing Zhou and Chung F. Wong*

Department of Chemistry and Biochemistry, and Center for Nanoscience, University of Missouri—Saint Louis, One University Boulevard, Saint Louis, Missouri 63121

Received: December 9, 2008; Revised Manuscript Received: February 13, 2009

Although various groups have studied the phosphorylation mechanism of the insulin receptor tyrosine kinase (IRK), an unanimous picture has not yet emerged. In this work, we performed a computational study to gain further insights. We first built a structural model of the reactant complex with the guide of several crystal structures and previous computational studies of the cyclic AMP-dependent protein kinase. We then optimized the structure by performing geometry optimization using a quantum mechanical model containing nearly 300 atoms. A reaction path was then traced between the reactant and the product by using a multiple coordinate-driven method. The calculations mapped out a sequence of structural changes depicting the conversion of the reactant to the product. Analysis of the structural changes revealed the formation of a dissociative transition state and the involvement of a proton transfer from the hydroxyl group of the tyrosyl residue of the peptide substrate to a conserved aspartate in the active site of the enzyme. The proton transfer began well before the transition state was reached and finished only shortly before the product was completely formed. In addition, the formation of a hydrogen bonding network among Arg1136, Asp1132, the γ -phosphate of ATP, and the tyrosine residue of the substrate appeared to hold the latter two in a near-attack position for reaction. The model estimated a reaction barrier of 14 kcal/mol, semiquantitatively in accord with experiment.

I. Introduction

Protein phosphorylation plays an important role in signal transduction and cellular regulation.^{1–3} Although it is well-known that protein kinases catalyze these reactions, much work remains to understand the complex reaction mechanisms in atomic detail. The whole process includes substrate recognition, phosphoryl transfer, and product release. Large conformational change beyond the catalytic core could also play a role. In the chemical step, protein kinases catalyze the transfer of the γ -phosphate of ATP to the hydroxyl group of serine, threonine, or tyrosine of the targeted protein substrates. Their extraordinary catalytic power can reach 10^9 to 10^{11} times the uncatalyzed reactions.^{3,4}

Tyrosine kinases, with 90 members, form one large subgroup of protein kinases in humans⁵ and play important roles in physiological and pathophysiological processes.⁶ Because the insulin receptor tyrosine kinase (IRK) influences cellular metabolism and growth and its defect can lead to insulin resistance and other diseases, it is among one of the most studied tyrosine kinases in humans. As many other protein kinases, the catalytic domain of IRK contains about 300 residues, forming a small N-terminal lobe and a larger C-terminal lobe. The N-terminal lobe consists of a five-stranded β sheet and a single α helix, while the C-terminal lobe is largely helical.⁷ A flexible hinge region, spanning several residues in length, connects the N- and C-terminal lobes and facilitates relative domain movement. Between the two lobes sits the binding site for ATP, which serves as the phosphate source for the substrate.

Different protein kinases have different substrate specificities. Many currently known substrates of IRK contain the consensus sequence YMXM.⁸ The basal form of IRK has low kinase

activity. Structural study suggests that the basal form is autoinhibited by the activation loop (A-loop), which transverses the N- and C-terminal lobes and blocks the binding sites, preventing ATP and the substrate from interacting with the enzyme.⁹ Upon binding of insulin to the extracellular α domain, three tyrosine residues—Y1158, Y1162, and Y1163—are phosphorylated,^{10,11} causing the A-loop to open up and the binding sites for ATP and the substrate to be exposed.

The crystal structure of triphosphorylated IRK (PDB code 1IR3)¹² showed that the ATP analogue AMP-PNP, two Mg^{2+} ions, the tyrosine in the substrate, and important catalytic residues are well ordered (see Figure 1). The amide nitrogen of the conserved Ser1006 on the glycine-rich loop [RELG¹⁰⁰³QG¹⁰⁰⁵SFG¹⁰⁰⁸MV] forms a hydrogen bond with an oxygen on the β -phosphate of AMP-PNP. In addition, the conserved Lys1030 hydrogen bonds to an oxygen on the α -phosphate. These interactions between the phosphate groups and IRK are less extensive than those observed in the cAMP-dependent protein kinase (cAPK).^{12,13} The Mg^{2+} ion (Mg1) that chelates the γ - and β -phosphate groups is hexagonally coordinated. Unlike some other protein kinases, the second Mg^{2+} ion (Mg2) does not form as many direct interactions with ATP; it is coordinated to the ATP analogue only through an oxygen on the β -phosphate. Instead, several water molecules and an aspartate provide the other ligands for this metal ion. Asp1150 of the DFG motif and Asn1137 are two important conserved residues anchoring the two Mg^{2+} ions. The substrate sits further away from the N-terminal lobe in comparison to the serine/threonine protein kinase cAPK, which is not surprising because the side chain of tyrosine is larger than that of serine or threonine.^{12,13} The Asp at position P-1 of the substrate interacts with Lys1085 of the D helix indirectly via a structural water. On the other hand, residues (P+1) to (P+3) of the peptide substrate form a short β -strand that pairs up with β 11 at the

* To whom correspondence should be addressed. E-mail: wongch@umsl.edu.

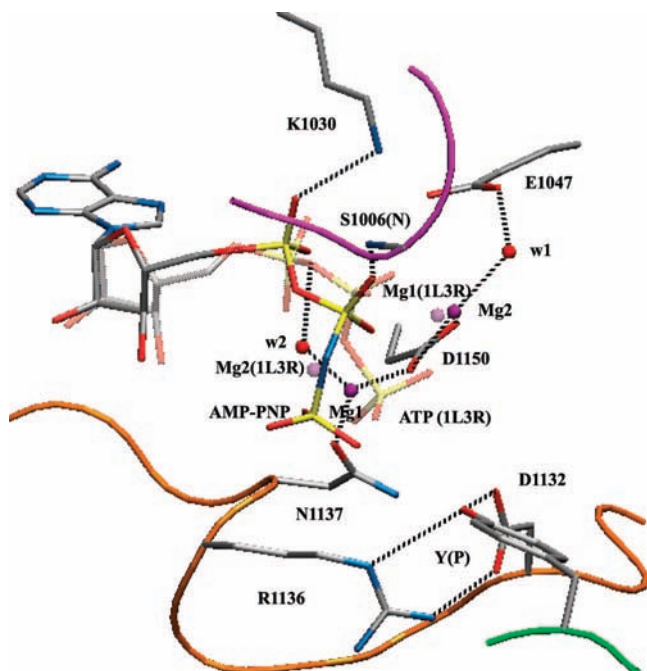


Figure 1. A model derived from the X-ray structure of the active site of the insulin receptor protein kinase (pdb code 1IR3). The phosphate arm of AMP-PNP was aligned with that of ATP in cyclic AMP-dependent protein kinase (PDB code 1L3R). Selected hydrogen bonds and bonds to Mg^{2+} ions are shown in dotted black lines. Also see the X-ray data column in Table 1 for additional geometrical features. Carbon atoms are colored silver, nitrogen atoms blue, oxygen atoms red, phosphorus atoms yellow, and magnesium atoms purple. The glycine-rich loop, the catalytic loop, and the substrate backbone are colored magenta, orange, and green, respectively.

end of the A-loop in an antiparallel fashion. The hydroxyl group on the tyrosine [Tyr(P)] of the peptide substrate forms a hydrogen bond with Asp1132, which might act as a catalytic base for the phosphorylation reaction. Arg1136, which forms hydrogen bonds with Asp1132 and the Tyr(P) in the crystal structure, might also play an important role in catalysis by holding the reactive groups together. Although this crystal structure has generated useful insights into how the enzyme might position ATP and the peptide substrate to promote catalysis, the replacement of the bridging oxygen between the γ - and β -phosphate groups by a nitrogen in AMP-PNP might not realistically reflect how ATP actually binds in the active site.¹² Therefore, in using this structure to build our computational model, as shown below, we also examined the crystal structures of several related tyrosine kinases and the results from previous calculations on cAPK.

Several kinetic studies investigated the mechanism of phosphorylation catalyzed by IRK and reached different conclusions.^{14–17} Graves et al. favored a reaction mechanism in which the hydroxyl group of the tyrosine in the substrate was deprotonated before the formation of the transition state (TS).^{14,15} The transition state was associative in which the entering and leaving groups were tightly bonded to the γ -phosphate. On the other hand, Ablooglu et al. suggested that the hydroxyl group only deprotonated after the formation of the TS, which was dissociative in nature.¹⁶ They also estimated the free energy barrier for the forward reaction to be 15.4 kcal/mol. (Note that their peptide substrate, pY-IRS939, contained an Asn instead of an Asp at the (*P*-1) position, i.e., differed from the peptide substrate used in determining the crystal structure: PDB entry 1IR3.)

TABLE 1: Distances (Å), from Experimental and Computational Structures, Characterizing Important Interactions in the Reactant, the Product, and the Transition State

distance	theoretical computations			X-ray data 1IR3
	reactant	transition state	product	
$O_{\beta\gamma}-P_{\gamma}$	1.82	2.41	3.10	n/a
$P_{\gamma}-O_{Tyr(P)}$	3.27	2.32	1.80	5.08
$O-H_{Tyr(P)}$	1.01	1.28	1.72	n/a
$D1132-H_{Tyr(P)}$	1.60	1.17	1.00	n/a
$D1132-O_{Tyr(P)}$	2.60	2.44	2.71	2.65
$R1136-O_{\gamma_2}$	2.54	2.86	2.82	3.91
$R1136-O_{Tyr(P)}$	4.20	3.98	4.01	3.08
$R1136-D1132$	2.88	2.87	2.98	2.87
$K1030-O_{\alpha_1}$	2.82	2.53	2.53	2.83
$1006N-O_{\beta_2}$	2.70	2.85	2.79	2.89
$Mg1-O_{\gamma_1}$	2.07	2.14	2.18	5.11
$Mg1-O_{\beta_1}$	2.11	2.01	2.00	2.50
$Mg1-D1150-1$	2.27	2.34	2.43	2.61
$Mg1-D1150-2$	2.22	2.25	2.25	2.44
$Mg1-w1$	2.19	2.20	2.21	2.31
$Mg1-w2$	2.07	2.09	2.11	n/a
$Mg2-O_{\gamma_2}$	1.99	2.05	2.09	n/a
$Mg2-O_{\beta\gamma}$	2.25	2.04	2.01	n/a
$Mg2-D1150-2$	2.03	2.04	2.05	2.05
$Mg2-N1137$	2.08	2.10	2.16	2.17
$Mg2-w3$	2.10	2.09	2.14	n/a
$Mg2-w4$	2.22	2.24	2.28	2.04
$O_{\alpha_2}-w3$	2.54	2.54	2.55	2.71

On the other hand, the theoretical study on the phosphorylation mechanism of IRK lags behind experimental work, although it can, in principle, provide detailed atomistic insights not easily obtainable from experiment. Although one theoretical study has appeared, the calculations were performed only with a small active-site model containing two phosphate groups to mimic ATP and a phenol group to represent the tyrosine residue of the peptide substrate.¹⁸ The rest of the protein environment was simply modeled by using a dielectric constant of 4 in the calculations. The authors proposed a four-step mechanism consisting of ATP protonation, shifting of a proton from the nonbridging oxygen on the γ -phosphate to the bridging oxygen between the β - and γ -phosphates of ATP, detachment of the γ -phosphate from ATP, and the attachment of this phosphate group to Tyr(P). This mechanism differed from the experimental proposals of Ablooglu et al.¹⁷ and Graves et al.¹⁵

In this work, we performed calculations by building a much larger active-site model containing nearly 300 atoms. We also used larger basis sets: 6-31G* and 6-31+G* instead of 3-21G* in the earlier work.¹⁸ We made these calculations practical by using a two-layered ONIOM model^{19–26} in which we described only the most important region by ab initio density functional theory (DFT)²⁷ with the rest by a cheaper semiempirical PM3 model. Although QM/MM methods are also commonly used to study large biomolecular systems, the B3LYP/PM3 approach provides a useful alternative. For example, Cui et al. recently chose such an approach because it described charge-transfer and nonadditive many-body effects that molecular mechanics could not.²⁸ Moreover, the ONIOM B3LYP/PM3 approach has been successfully applied to study other enzymatic reactions. For example, Sklenak et al. applied this approach to study the catalytic mechanism of yeast cytosine deaminase.²⁹ Pawlak et al., on the other hand, used this method to study enoyl-CoA hydratase.³⁰

The crystal structures of several tyrosine kinases in addition to the one for IRK (PDB code 1IR3) were used, along with results from previous calculations of cAPK, to guide the construction of an initial reactant model. In addition, by using an efficient multiple coordinate-driven method (MCD),³¹ we searched for a near optimal reaction path with fewer human biases. Among several possible mechanisms proposed above, our calculations were most consistent with the dissociative mechanism proposed by Ablooglu et al.¹⁷ Our calculations also provided atomistic detail not directly obtainable from experiment. We found the reaction mechanism to involve a proton transfer from the hydroxyl group of Tyr(P) to the conserved ASP1132 in the active site, but the proton transfer occurred in concert with the phosphoryl transfer rather than happening before or after the transition state. In addition, the calculations suggested that Arg1136 formed a hydrogen bonding network with several residues in the active site, helping to position the γ -phosphate of ATP to react with Tyr(P).³² Although the calculational model employed various approximations, it produced an activation barrier not far from the experimental value by Ablooglu et al.¹⁷

II. Construction of the Structural Model of the Active Site

The lack of a reliable structure of the ternary complex of IRK, in which the γ -phosphate of ATP and the hydroxyl group of Tyr(P) are in near-attack position,³² has posed difficulty for the theoretical study of the reaction mechanism of the enzyme. Nevertheless, the crystal structure of IRK (pdb code 1IR3) in complex with a nonhydrozable ATP analogue, AMP-PNP, and a peptide substrate determined at 1.9 Å provides a reasonable starting point for building a structural model to begin computational study. The main drawback of this structure is that AMP-PNP is only an analogue of ATP and might not mimic ATP well. Fortunately, one can gain some insights by also examining the crystal structures of similar enzymes. By comparing the structure of IRK (pdb code 1IR3) with that of cAPK (pdb codes 1L3R and 1ATP), we found that the γ -phosphate of AMP-PNP was not aligned in a near-attack position with the hydroxyl group of the to-be-phosphorylated tyrosine residue of the substrate. On the other hand, by comparing the active sites of three tyrosine kinases, IRK (pdb code 1IR3),¹² insulin-like growth factor 1 receptor (pdb code 1K3A),³³ and epidermal growth factor receptor (EGFR, pdb code 2GS6),³⁴ containing different transition-state analogues, we found that although the γ -phosphate could adopt somewhat different orientations in these complexes, the rest of the active-site residues had a similar structure; the rmsd deviation of the four catalytic residues—R1136, N1137, D1132, and D1150—between IRK and those of the other two tyrosine kinases amounted to only 0.7 Å. Moreover, the crystal structure of a binary complex of IRK with a tetrafluorotyrosyl peptide substrate and that of the ternary complex of IRK in PDB entry 1IR3 shows similar positioning of Tyr(P) and key residues in its immediate neighborhood.¹⁶ Therefore, it is reasonable to build a better reacting structure of IRK by simply repositioning the γ -phosphate according to other suitable crystal structures but keeping everything else the same as in PDB entry 1IR3.

To suitably reposition the γ -phosphate, we started from two structural models constructed in Henkelman et al.'s earlier ab initio DFT study of the cAPK-catalyzed phosphoryl transfer reaction.³⁵ They constructed these models based on two crystal structures of cAPK from PDB entries 1ATP and 1L3R. The model started from 1ATP gave significantly overestimated activation barrier whereas the latter gave a lower activation

barrier in line with experiment. By comparing the cAPK structures in Henkelman et al.'s work with the IRK crystal structure (PDB code 1IR3) containing the ATP mimic AMP-PNP, the γ -phosphate of the latter structure was clearly not aligned well for attacking Tyr(P). Therefore, we realigned the phosphate arm of ATP in our model to match those of Henkelman et al.'s two starting structures. It turned out that AMP-PNP in the crystal structure aligned better with the ATP in the structure derived from 1L3R of cAPK. The α - and β -phosphate atoms deviated by only 0.52 and 0.61 Å, respectively, between the two structures (see Figure 1). The only major difference occurred in the γ -phosphate, with its phosphorus atom differing by 2.17 Å between the two structures. We therefore replaced the coordinates of the triphosphate in the crystal structure of IRK by those from Henkelman et al.'s structure deriving from 1L3R. The phosphate arm of ATP was now well positioned to react with the hydroxyl group of the tyrosine in the peptide substrate. The ATP-substrate alignment angle, $O_{\beta,\gamma}-P_{\gamma}-O_{\text{Tyr(P)}}$, was 169.47° and the distance between P_{γ} and $O_{\text{Tyr(P)}}$ measured 3.93 Å. In this reaction model, the second Mg^{2+} ion ($\text{Mg}2$ of Figure 1) in IRK became the primary Mg^{2+} ion, which now chelated the γ - and β -phosphates. The other Mg^{2+} ion ($\text{Mg}1$ in Figure 1) was now situated between the γ - and α -phosphates.

Our final calculational model of the active site included an ATP molecule modified from the AMP-PNP molecule, two Mg^{2+} ions, the tyrosine residue of the substrate, and eight residues in the active site: G1005, S1006, K1030, E1047, D1132, R1136, N1137, and D1150. We also added four water molecules to coordinate the Mg^{2+} ions hexagonally. Because we needed to cut some peptide bonds to separate the active site region from the rest of the protein, we capped the exposed N-termini with acetyl groups (CH_3CO) and blocked the exposed C-termini by methylamide groups (NHCH_3). The net charge of Arg and Lys was set to +1, while that of Asp and Glu was set to -1. Because the crystal structure did not contain hydrogens, we used the "hbuild" module of CHARMM³⁶ to add them and then optimize their positions by 200 steps of steepest decent energy minimization, holding all other atoms fixed. The resulting reaction model was composed of 279 atoms.

III. Calculation Details

A. Structural Optimization. We optimized the geometry of the IRK active site model by using the ONIOM model^{19,20,22–26} in GAUSSIAN03.³⁷ Our ONIOM model consisted of two layers in which we treated the most important region at the B3LYP(6-31G*) level and the rest of the system at the semiempirical PM3 level. The high-level region contained 95 atoms including the phosphate arm of ATP, two Mg^{2+} ions, four water molecules, the peptide linkage between G1005 and S1006, and the side chains, partial or whole, of K1030, E1047, D1132, R1136, N1137, D1150, and Tyr(P). Assuming that the crystal structures realistically describe the protein structure well except at its active site (because of the unnatural AMP-PNP analogue), we only allowed 66 atoms to move during the geometry optimization. These atoms included the phosphate arm of ATP, the two Mg^{2+} ions, the side chains of R1136, D1132, and Tyr(P), and the four water molecules. After obtaining an optimized reactant state, we used the MCD method³¹ that we modified to use with the ONIOM model to drive the reactant to the product (see the details below). We then further refined the resulting product state by geometry optimization within the same ONIOM model. The default convergence criteria of GAUSSIAN03³⁷ were used in these calculations.

B. Reaction Path. We obtained the reaction path connecting the reactant and the product states by our modified version of the MCD method.³¹ This method speeds up the path search by using the adiabatic approximation. One first chooses by chemical intuition a set of active coordinates to drive a reaction from the reactant to the product. The remaining coordinates only change adiabatically, i.e., they are energy minimized at each fixed value of the reaction coordinate. To identify the reaction coordinate, defined in terms of the active coordinates, the method looks for a path associated with minimum work. Starting with the reactant, it first calculates the forces and an approximate Hessian matrix by quantum mechanics and then uses them in a Taylor's series expansion to compute quickly the energy change when the reactant is moved in many different directions toward the product. The direction that leads to the slowest energy rise (when the reaction is going uphill) or to the quickest energy drop (when the reaction is going downhill) is chosen. The method repeats this until the product is reached. We have extended this MCD method to use the ONIOM approach so that the most important region can be treated with a higher level theory while the less important regions are described with lower level theories (quantum or classical). This has allowed us to construct larger active site models.

To use this method effectively, one needs to choose a good set of active coordinates such that the adiabatic approximation holds. One does not know beforehand which degrees of freedom are best to include but we have found a useful strategy. We first started with the minimum number of degrees of freedom based on chemical intuition. For example, in the phosphorylation reaction of protein kinases, the hydroxyl oxygen of the side chain of a serine, threonine, or tyrosine residue attacks the γ -phosphate of ATP to bond with it eventually. Simultaneously, the bond between the γ -phosphorus and its bridging oxygen to the β -phosphate breaks. One can first start with these two clearly essential degrees of freedom to perform a reaction path search. When we did this, the energy curve was not smooth but with a jump, suggesting that other degrees of freedom coupled tightly with the above two such that the adiabatic approximation did not work well when they were not included in the active coordinates.

To find additional essential degrees of freedom to include, we identified those—excluding the soft degrees of freedom, such as reorientation of side chains, which could still be reasonably treated within the adiabatic approximation—that changed the most in converting from the reactant to the product or vice versa. From this analysis, we added two more degrees of freedom to describe the observed proton transfer between the tyrosine OH group and the conserved aspartate side chain (see the reactant and product columns in Table 1). These were the distance between the hydrogen of the tyrosine OH group and the aspartate side chain and the distance associated with the tyrosine O—H bond which significantly increased beyond the bonding value as the result of proton transfer. Recent studies of cAPK predicted a similar mechanism.^{35,38–40} After adding these two degrees of freedom, we obtained a smoother energy profile with a lower activation barrier (see Section IV.C). Thus, a strategy to identify reaction path with this method is to first start with the minimum number of active coordinates based on chemical intuition— $\text{O}_{\beta,\gamma}-\text{P}_{\gamma}$ and $\text{P}_{\gamma}-\text{O}_{\text{Tyr(P)}}$ distances in this example—then identify other hard degrees of freedom that change significantly from the reactant to the product—such as those involving bond breaking and bond formation—add them to the active coordinates, and redo the path search.

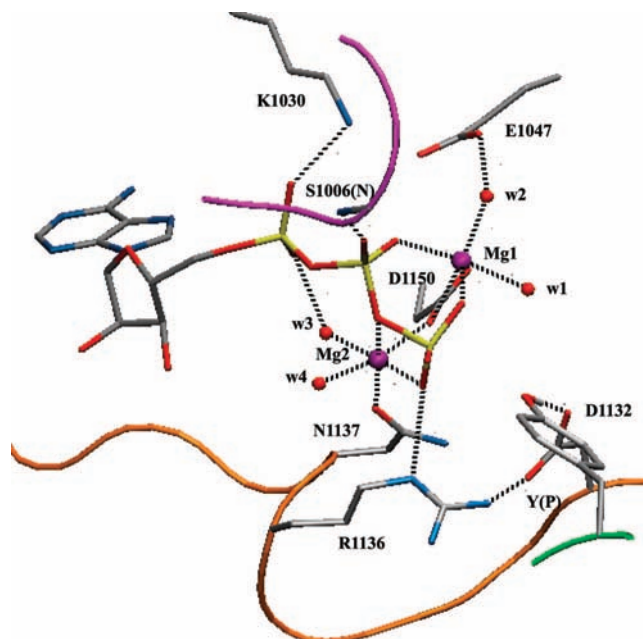


Figure 2. Optimized structure of the reactant. Selected hydrogen bonds and bonds to Mg^{2+} are shown as dotted black lines. See the reactant column in Table 1 for further geometric features. The coloring scheme is the same as that in Figure 1.

During the final reaction path search with the MCD method including the four active coordinates, the adiabatic optimization was stopped when the maximum force fell below 0.01 au and the maximum displacement below 0.3 au, as recommended by ref 31. We limited the total change of the four active coordinates to less than 0.15 Å in each search step to reduce numerical errors.

IV. Results and Discussions

A. Reactant Structure. ATP and Tyr(P) on the substrate were in near-attack position in the optimized reactant state (see Figure 2). The $\text{P}_{\gamma}-\text{O}_{\beta,\gamma}$ bond length measured 1.82 Å, which matched well with Pauling's estimate of 1.73 Å.⁴¹ The ATP–substrate alignment angle, $\text{O}_{\beta,\gamma}-\text{P}_{\gamma}-\text{O}_{\text{Tyr(P)}}$, was 159.84° and the distance between P_{γ} and $\text{O}_{\text{Tyr(P)}}$ spanned 3.27 Å, shortened from its value of 3.93 Å in the initial reactant model before geometry optimization. The near-attack position between ATP and Tyr(P) also resulted from further extension of the phosphate arm and the rotation of the phenolic ring during geometry optimization.

The conserved Asp1132 in the active site formed a hydrogen bond with the hydroxyl group of Tyr(P), suggesting its potential role as a catalytic base. As the hydroxyl group of Tyr(P) approached the γ -phosphorus of ATP, the CA—CB—CG—OD1 dihedral angle of Asp1132 changed significantly from 16.93° in the crystal structure to -31.28° in our optimized model of the reactant state.

The conserved Arg1136 in the active site of tyrosine kinases is believed to play a critical role in holding ATP and Tyr(P) together in a near-attack position for reaction.³² We observed this in the optimized reactant structure, which appeared to be stabilized by a network of hydrogen bonds among Arg1136, Asp1132, the γ -phosphate of ATP, and the tyrosine residue of the substrate.

The two Mg^{2+} ions are thought to play important roles in phosphoryl transfer by screening the negative charges on ATP and orienting the γ -phosphate to better interact with Tyr(P) of

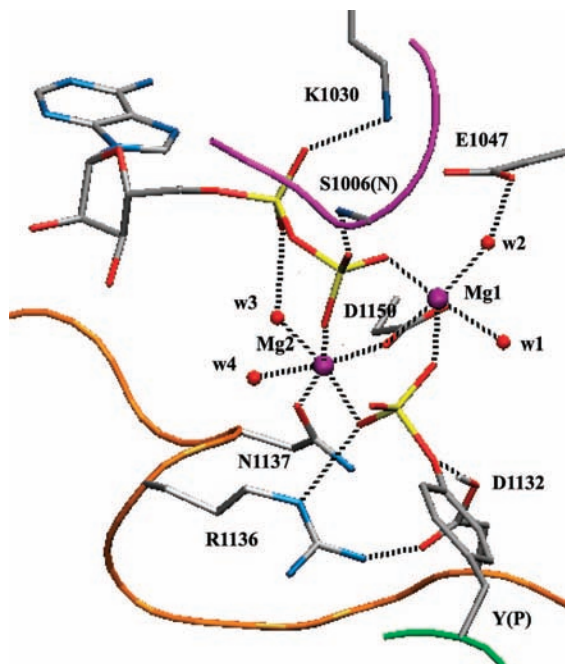


Figure 3. Optimized structure of the product. Selected hydrogen bonds and bonds to Mg^{2+} are shown as dotted black lines. See the product column in Table 1 for further geometric features. The coloring scheme is the same as that in Figure 1.

the substrate. In the optimized reactant structure (Figure 2), they were hexagonally coordinated to the oxygens on the phosphate arm of ATP, on the side chains of Asn1137 and Asp1150, and on the four water molecules. The primary Mg^{2+} ion (Mg1) that chelated the β - and γ -phosphates was coordinated to the two oxygen atoms on Asp1150 and to the oxygens of two water molecules ($\text{Mg}\cdots\text{O} = 2.0\text{--}2.3$ Å). These two water molecules ($w1$ and $w2$) also hydrogen bonded to the oxygens of the γ -phosphate and on the side chain of Glu1047, respectively. The latter water ($w2$) was observed in the crystal structure of IRK. This magnesium ion situated a bit further away from Asp1150 than from its other four ligands (see Table 1). The secondary Mg^{2+} ion (Mg2) coordinated to a terminal oxygen of the γ -phosphate, the bridging oxygen between the β - and γ -phosphates, the oxygens on the side chain of Asn1137 and Asp1150, and to two water molecules ($\text{Mg}\cdots\text{O} = 2.0\text{--}2.2$ Å). Instead of chelating to the α -phosphate directly, Mg2 interacted with the α -phosphate via a water ($w3$) molecule, which was also observed in the crystal structure of IRK.

B. Product Structure. In the active site of IRK, both the γ -phosphate of ATP and Asp1136 sat next to the hydroxyl group of Tyr(P). Thus, protonation of either one of them by the hydroxyl group was possible, leading to different intermediate or product states. The preferred reaction scheme should be one that gives a lower reaction barrier. Our MCD calculations using two and then four active coordinates (see Section III.B) showed that the proton of the hydroxyl group of Tyr(P) transferred more easily to a side-chain oxygen of Asp1136 than to a phosphate oxygen.

As shown in Figure 3, Tyr(P) on the substrate was phosphorylated in the product state. The $\text{O}_{\beta\gamma}\text{--P}_{\gamma}$ and $\text{P}_{\gamma}\text{--O}_{\text{Tyr(P)}}$ distances were 3.10 and 1.80 Å, respectively (see Table 1). The oxygen (OD2) on Asp1136 was protonated and the $\text{CA}\text{--CB}\text{--CG}\text{--OD1}$ dihedral angle of Asp1136 measured -10.94° , which differed from its value of -17.61° in the reactant. The hydrogen bond interactions between Asp1132 and Tyr(P) were retained,

TABLE 2: Displacement (Å) of Non-Hydrogen Atoms in the Active Site Resulting from the Phosphoryl Transfer Reaction

residue	atom	displacement	residue	atom	displacement
ATP	P_{γ}	1.30	NE		0.25
	$\text{O}_{\gamma 3}$	0.90	CD		0.24
	$\text{O}_{\gamma 2}$	0.77	Tyr(P)	OH	0.35
	$\text{O}_{\gamma 1}$	0.74		CE2	0.33
	$\text{O}_{\beta 2}$	0.50		CZ	0.26
	$\text{O}_{\alpha,\beta}$	0.30		CD2	0.24
	$\text{O}_{\beta 1}$	0.30	MG2	MG	0.24
	$\text{O}_{\alpha 2}$	0.29	D1132	OD1	0.23
	P_{β}	0.28	w1	OH2	0.60
	$\text{O}_{\beta,\gamma}$	0.24	w4	OH2	0.47
R1136	NH1	0.58	w3	OH2	0.36
	NH2	0.36	w2	OH2	0.25
	CZ	0.34	MG1	MG	0.20

although the $\text{O}_{\text{D1132}}\text{--O}_{\text{Tyr(P)}}$ distance increased by about 0.1 Å in comparison to that in the reactant.

Arg1136 still hydrogen bonded to the γ -phosphate with a distance of 2.82 Å, similar to that in the reactant. The hydrogen bond to Asp1132 also remained but the donor–acceptor distance increased by 0.15 Å from the reactant to 2.81 Å in the product.

Both Mg^{2+} ions remained hexagonally coordinated in the product state with similar metal–ligand distances as in the reactant. The bridging oxygen between β - and γ -phosphates, $\text{O}_{\beta\gamma}$, moved closer to Mg2 after the release of the γ -phosphate and their separation distance decreased by 0.24 Å in comparison to that in the reactant. The distances between Mg2 and its other oxygen ligands experienced less pronounced changes. Again, we observed the indirect interactions between Mg2 and an oxygen on the α -phosphate via a water molecule ($w3$).

Table 2 lists active site heavy atoms that displaced more than 0.2 Å in converting from the reactant to the product. The γ -phosphorus atom showed the largest movement of 1.30 Å. The oxygen atoms on the γ -phosphate changed a bit less, about 0.74–0.90 Å. The side chain of Arg1136 followed the motion of the γ -phosphate and its nitrogen atoms shifted by more than 0.25 Å. Atoms on the β -phosphate only displaced by 0.25–0.40 Å. The α -phosphate experienced significantly less movement as it was further away from the reaction center. The side chain of Tyr(P), which gained a phosphate group, moved quite a bit. The hydroxyl oxygen moved by about 0.35 Å. One oxygen (OD1) on Asp1132, which formed a hydrogen bond with Arg1136, moved by 0.23 Å. The other oxygen (OD2), which abstracted a proton from Tyr(P), moved only 0.06 Å (not shown in Table 2). The secondary Mg^{2+} ion moved by 0.24 Å. In contrast, the primary Mg^{2+} ion moved by only 0.20 Å. Following the movement of the magnesium ions and the phosphate arm of ATP, four structural water molecules experienced displacement of the order of 0.25–0.6 Å.

When we re-evaluated the total energy of the reactant and the product with the larger 6-31+G* basis set, we found the reactant to be more stable by 2.16 kcal/mol. From kinetic measurements, in which a substrate peptide containing an Asn instead of an Asp at position (P-1) was used, Ablooglu et al. estimated a free energy difference between the reactant and the product to be 1.4 kcal/mol favoring the product.¹⁷ The difference between our calculation and experiment might result from the different substrates used or/and from the uncertainty of the calculations.

C. Reaction Path and Transition State. With the optimized reactant and the product structures at hand, we used the MCD method³¹ to find a low-barrier reaction path connecting the reactant and the product, in order to understand the sequence

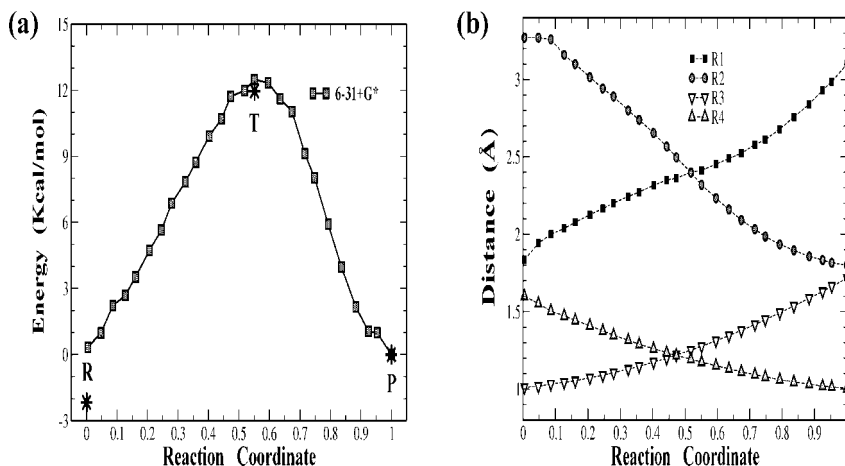


Figure 4. The reaction path obtained from the MCD method using four active coordinates. (Reactant at left) (a) Energy profile from B3LYP/6-31+G* calculations. The structures (opaque squares) produced by the MCD calculations were only partially optimized for the nonactive coordinates. Stars (*) denote the reactant (R), the transition state (T), and the product (P) whose nonactive coordinates were fully optimized at the B3LYP/6-31G* level. (b) Progress of the four active coordinates: $O_{\beta\gamma}-P_{\gamma}$ (R1), $P_{\gamma}-O_{Tyr(P)}$ (R2), $O_{Tyr(P)}-H_{Tyr(P)}$ (R3), and $H_{Tyr(P)}-O_{D1132}$ (R4) as a function of the reaction coordinate.

of structural changes occurring in the phosphate transfer reaction. As long as the minimum energy path is smooth and continuous, the structure corresponding to the maximum energy along the reaction path provides a good approximation to the TS, and its energy relative to the reactant or the product states offers an estimate to the associated reaction barrier.^{42,43} As mentioned earlier, we started off with two active coordinates but found that four were more appropriate, i.e., $O_{\beta\gamma}-P_{\gamma}$ (R1), $P_{\gamma}-O_{Tyr(P)}$ (R2), $O-H_{Tyr(P)}$ (R3), and $H_{Tyr(P)}-O_{D1132}$ (R4) distances. We drove the reaction forward (from the reactant to the product) and backward (from the product to the reactant) and found the energy profile of the backward reaction to be smoother and gave a lower reaction barrier as depicted in Figure 4a. In the figure, the reactant and product corresponded to $rc = 0$ and $rc = 1$, respectively. The maximum energy along the reaction path shown in Figure 4a occurred at $rc \approx 0.55$ and the associated structure approximately characterized the TS. We fully optimized the nonactive coordinates to obtain the structure of the TS (see Table 1) and found its energy to be higher than that of the reactant by ~ 14.2 kcal/mol when calculated at the B3LYP/6-31+G* level—this gave an estimate of the reaction barrier.

The progress of the four active coordinates illustrated in Figure 4b provided atomistic insights into the reaction mechanism. As the reaction proceeded from the reactant to the product, the $O_{\beta\gamma}-P_{\gamma}$ (R1) and $O-H_{Tyr(P)}$ (R3) distances increased, while the $P_{\gamma}-O_{Tyr(P)}$ (R2) and $H_{Tyr(P)}-O_{D1132}$ (R4) distances decreased. The progress of R3 and R4 suggested that proton transfer from the hydroxyl group of Tyr(P) to Asp1132 was synchronous with the phosphotransfer. The structure of the TS, occurring at the peak of the reaction path (see Table 1), was nearly symmetric at the γ -phosphate with respect to the leaving terminal oxygen of ADP and the entering hydroxyl group of Tyr(P). The $O_{\beta\gamma}-P_{\gamma}$ and $P_{\gamma}-O_{Tyr(P)}$ distances amounted to about 2.41 and 2.32 Å, respectively. On the basis of Pauling's formula suggested by Mildvan,^{4,41} this TS was 93% and 90% dissociative with respect to the leaving (ADP) and the entering groups [hydroxyl group of Tyr(P)], respectively. For comparison, a previous theoretical investigation on the cAPK catalytic system yielded a TS that was about 87% dissociative with respect to the entering hydroxyl group of Ser(P).⁴⁰ Finally, we found the hydrogen bond between Arg1136 and the γ -phosphate to remain intact in the TS but appeared weakened a bit as

reflected by their larger separation of 2.86 Å in comparison to those of the reactant and the product (see Table 1).

V. Conclusions

In summary, we have built a model for the active site of the IRK system for studying its catalytic mechanism of phosphate transfer. Because the crystal structure of the transition state analogue (pdb code 1IR3) did not appear to mimic the transition state well, we used the crystal and computational structures from other protein kinases as a guide to obtain a better structural model. The main modification was the repositioning of the γ -phosphate of ATP to adopt a near-attack position for reaction. We then obtained refined structures of the reactant complex and the product by geometry optimization using a quantum mechanical model containing nearly 300 atoms. When tracing the reaction path between the reactant and the product using a multiple coordinate-driven method, we found a reaction barrier of about 14 kcal/mol. Although we were not aiming to obtain fully quantitative agreement with experiment because of various approximations employed in the calculations, this activation barrier is in line with the experimental estimate of 15 kcal/mol. (Using a smaller 3-21G basis set in the calculations gave a similar activation barrier, within 1 kcal/mol, and comparable structural changes (data not shown)). The reaction mechanism suggested by our computational study was similar to that found in the cyclic AMP-dependent serine/threonine kinase, indicating that some serine/threonine and tyrosine protein kinases might share a similar mechanism. Specifically, analyzing the structural change along our calculated reaction pathway suggested that the reaction went through a dissociative transition state and the proton on the hydroxyl group of Tyr(P) on the peptide substrate transferred to the nearby conserved Asp1131 in concert with the phosphoryl transfer reaction. In addition, Arg1136 appeared to stabilize the active site by being part of a hydrogen bonding network involving the γ -phosphate of ATP, the conserved Asp1136, and the hydroxyl group of Tyr(P).

Acknowledgment. The National Institutes of Health and a Research Award from the University of Missouri—Saint Louis provided partial support for this work. We also thank the University of Missouri Bioinformatics Consortium for supplying computational resources. In addition, we appreciate useful discussions with Dr. Zunnan Huang.

References and Notes

- (1) Hunter, T. *Cell* **1995**, *80*, 225.
- (2) Graves, J. D.; Krebs, E. G. *Pharmacol. Ther.* **1999**, *82*, 111.
- (3) Adams, J. A. *Chem. Rev.* **2001**, *101*, 2271.
- (4) Mildvan, A. S. *Proteins: Struct. Funct. Genet.* **1997**, *29*, 401.
- (5) Manning, G.; White, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002**, *298*, 1912.
- (6) Pawson, T.; Hunter, T. *Curr. Opin. Genet. Dev.* **1994**, *4*, 1.
- (7) Cowan-Jacob, S. W. *Cell. Mol. Life Sci.* **2006**, *63*, 2608.
- (8) Shoelson, S. E.; Chatterjee, S.; Chaudhuri, M.; White, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2027.
- (9) Hubbard, S. R.; Wei, L.; Ellis, L.; Hendrickson, W. A. *Nature* **1994**, *372*, 746.
- (10) Rosen, O. M.; Herrera, R.; Olowe, Y.; Petruzzelli, L. M.; Cobb, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3237.
- (11) Ellis, L.; Clauser, E.; Morgan, D. O.; Edery, M.; Roth, R. A.; Rutter, W. J. *Cell* **1986**, *45*, 721.
- (12) Hubbard, S. R. *EMBO J.* **1997**, *16*, 5572.
- (13) Bossemeyer, D.; Engh, R. A.; Kinzel, V.; Ponstingl, H.; Huber, R. *EMBO J.* **1993**, *12*, 849.
- (14) Martin, B. L.; Wu, D.; Jakes, S.; Graves, D. J. *J. Biol. Chem.* **1990**, *265*, 7108.
- (15) Yuan, C.-J.; Jakes, S.; Elliott, S.; Graves, D. J. *J. Biol. Chem.* **1990**, *265*, 16205.
- (16) Ablooglu, A. J.; Till, J. H.; Kim, K.; Parang, K.; Cole, P. A.; Hubbard, S. R.; Kohanski, R. A. *J. Biol. Chem.* **2000**, *275*, 30394.
- (17) Ablooglu, A. J.; Kohanski, R. A. *Biochemistry* **2001**, *40*, 504.
- (18) Pichierrì, F.; Matsuo, Y. *THEOCHEM* **2003**, *622*, 257.
- (19) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
- (20) Dapprich, S.; Komaromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *THEOCHEM* **1999**, *462*, 1.
- (21) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419.
- (22) Vreven, T.; Morokuma, K. *J. Chem. Phys.* **2000**, *113*, 2969.
- (23) Vreven, T.; Mennucci, B.; da Silva, C. O.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62.
- (24) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760.
- (25) Morokuma, K. *Bull. Korean Chem. Soc.* **2003**, *24*, 797.
- (26) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theor. Comput.* **2006**, *2*, 815.
- (27) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Clarendon: New York, 1989. Dreizler, R. M.; Gross, E. K. U. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*; Springer-Verlag: Berlin, Germany, 1990.
- (28) Cui, Q.; Guo, H.; Karplus, M. *J. Chem. Phys.* **2002**, *117*, 5617.
- (29) Sklenak, S.; Yao, L.; Cukier, R. I.; Yan, H. *J. Am. Chem. Soc.* **2004**, *126*, 14879.
- (30) Pawlak, J.; Bahnson, B. J.; Anderson, V. E. *Nukleonika* **2002**, *47*, S33.
- (31) Berente, I.; Náráy-Szabó, G. *J. Phys. Chem. A* **2006**, *110*, 772.
- (32) Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 139.
- (33) Favelyukis, S.; Till, J. H.; Hubbard, S. R.; Miller, W. T. *Nat. Struct. Biol.* **2001**, *8*, 1058.
- (34) Zhang, X.; Gureasko, J.; Shen, K.; Cole, P. A.; Kuriyan, J. *Cell* **2006**, *125*, 1137.
- (35) Henkelman, G.; LaBute, M. X.; Tung, C.-S.; Fenimore, P. W.; McMahon, B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15347.
- (36) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
- (37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B. et al. *Gaussian 03*, Revision C.02; Gaussian, Inc., Wallingford, CT, 2004.
- (38) Díaz, N.; Field, M. J. *J. Am. Chem. Soc.* **2004**, *126*, 529.
- (39) Cheng, Y.; Zhang, Y.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 1553.
- (40) Valiev, M.; Yang, J.; Adams, J. A.; Taylor, S. S.; Weare, J. H. *J. Phys. Chem. B* **2007**, *111*, 13455.
- (41) Pauling, L. *The Nature of the Chemical Bond*; Cornell University Press: Ithaca, NY, 1960.
- (42) Rothman, M. J.; Lohr, L. L. *Chem. Phys. Lett.* **1980**, *70*, 405.
- (43) Scharfenberg, P. *Chem. Phys. Lett.* **1981**, *79*, 115.

JP810827W