JCTC Journal of Chemical Theory and Computation

Binding of an RNA pol II Ligand to the WW Domain of Pin1 Using Molecular Dynamics Docking Simulations

Chai Ann Ng,[†] Daniel P. Oehme,[†] Yusuke Kato,[‡] Masaru Tanokura,[‡] and Robert T. C. Brownlee*,[†]

Department of Chemistry, La Trobe University, VIC 3086, Australia, and Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

Received April 20, 2009

Abstract: A novel docking protocol using a long, all atom molecular dynamics (MD) simulation, in an explicit solvent medium, without using any distance constraints is presented. This MD docking protocol is able to dock ligands, based on the C-terminal domain (CTD) of RNA polymerase II, into the tryptophan-tryptophan (WW) domain of Pin1. In this docking process, a significant loop-bending event occurs in order to encircle the ligand into its solvent exposed binding site, which cannot be simulated using current protocols. The simulations were validated structurally and energetically against an X-ray structure to confirm correct sampling of conformational space. Based on these simulations, and justification of the starting structure as a valid intermediate structure, a potential molecular basis for binding was predicted as well as confirming the key residues involved in the formation of the final strong and stable Pin1 WW domain-ligand complex.

Introduction

Human Pin1 (Pin1), a peptidyl prolyl cis/trans isomerase, was first isolated from a yeast two-hybrid screen of interactors of the known mitotic regulator NIMA from *Aspergillus nidulans* and shown to be essential for normal growth in some organisms.[1,2] Pin1 has been found to be involved in many cell processes such as mitosis, transcription, and in response to DNA damage. It has been seen that Pin1 is overexpressed in cancers such as breast cancer,[3] and conversely it is down-regulated in degenerative neurons in patients who have Alzheimer's.[4] Because of its involvement in these major medical issues, Pin1 has been identified as a major target for drug discovery.[5] However, Pin1 is an extremely flexible protein,[6] and this challenges the standard drug design protocol, based on the notion that binding occurs in a lock and key fashion.

Pin1 and its homologues such as PinA[7] and Ess1[8] are composed of two domains: a C-terminal peptidyl-prolyl cis/trans isomerase (PPIase) domain and an N-terminal WW domain, which are linked by a long linker region.[9,10] It is currently hypothesized that Pin1 functions through the WW domain first acting as a targeting domain and binding the substrate before the PPIase domain isomerizes the substrate at specific phosphorylated serine/threonine-proline (pSer/Thr-Pro) motifs, either with the substrate still bound to the WW domain or, possibly, after the substrate has completely dissociated from the WW domain.[11] Named due to a pair of highly conserved tryptophan residues (at position 11 and 34 in Pin1), WW domains normally consist of 38−40 residues and fold into triple-stranded antiparallel beta sheets.[12] WW domains bind proline rich proteins and are classified into different groups depending on which motif they bind. Group IV WW domains, to which Pin1 belongs, bind to Pro residues preceded by a pSer/Thr.[13,14] Pin1 and its homologues are unique as they are the only parvulin-like proteins (a family of prolyl isomerase) with WW domains.[15] Assays on the WW domain and PPIase domain individually show that the WW domain interacts with pSer/Thr-Pro motif containing substrates with a higher affinity than the PPIase domain and of nearly equal affinity to that of the whole protein.[10,14] However in isolation, the WW domain shows no isomerase activity. It is therefore critical that the WW domain interacts

* Corresponding author e-mail: R.Brownlee@latrobe.edu.au.
[†] La Trobe University.
[‡] University of Tokyo.

10.1021/ct900190n CCC: $40.75
Published on Web 08/27/2009

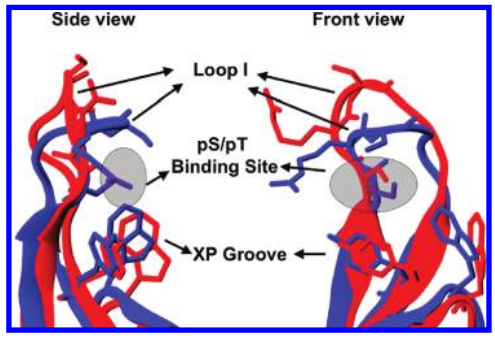

Figure 1. Front and side view of the superimposition of the X-ray structures of ligand free (red) and bound (blue) Pin1 WW domain. In the ligand free structure, termed the "open" form, loop I points upward. In the ligand bound structure, termed the "closed" form, the loop I bends down to cover the binding site.

with these pSer/Thr-Pro motifs containing peptides prior to the substrate being passed onto the PPIase domain for isomerization.[16]

A number of three-dimensional structures of Pin1 and its homologues are available through the Protein Data Bank[5,7−10,17,18] of which, two are extensively used in this work. The first is the X-ray structure of Pin1 bound to an AlaPro dipeptide at the PPIase domain but without a ligand bound to the WW domain (1PIN).[9] The second X-ray structure has a ligand comprising of a 7 amino acid sequence of the C-Terminal Domain (CTD) of RNA polymerase II bound to the Pin1 WW domain (1F8A).[10] This CTD ligand, (CTD-S2/S5) with sequence YpSPTpSPS, has three serines, two of which are phosphorylated, one at position 2 and one at position 5. It is of interest to note that all peptides bound to the WW domain bind with the proline in a trans conformation and that they all bind in the same unique binding mode.[11,17]

From the X-ray structures of Pin1, two important subdomains of the Group-IV WW domain were identified for ligand-binding (Figure 1). First the XP groove,[19] which is formed by residues Tyr23 and Trp34 which interacts with the proline residue of the pS/pT-P motif from the substrates in a hydrophobic fashion. Second loop I, the loop between the first and second β-sheets and composed of residues 16−21, is very flexible and binds the phosphate group immediately upstream of the proline. The flexibility of loop I is highlighted by comparing the ligand free (1PIN) and ligand bound (1F8A) X-ray structures of the Pin1 WW domain where a major difference is seen with loop I closing over the binding site in the ligand bound form (Figure 1).

A key aspect of the work in this paper is the ability to dock a ligand into a binding site. A number of programs have been developed with the premise of being able to dock a ligand into a receptors binding site. The algorithms used in these programs can be grouped according to the level of flexibility of the receptor and ligand that the algorithm allows.[20] In order of increasing complexity, the three groups are as follows: (a) rigid body docking, where both the receptor and small molecule are treated as rigid; (b) flexible ligand docking, where the receptor is held rigid, but the ligand is treated as flexible; and (c) flexible docking, where both receptor and ligand flexibility is considered. Thus far, the most commonly used docking algorithms use the rigid receptor/flexible ligand model where the receptor site is unable to change its conformation upon ligand binding. This is mainly because treatment of backbone flexibility in protein−protein docking is quite challenging, and thus few docking programs deal with the backbone flexibility.[21] Those programs that can deal with ligand and receptor flexibility however are normally only used for screening purposes.[20] Thus they are not suitable to use in studies such as these, where a ligand is to be docked into an open form of a protein.

In this work, we were guided by Hornak et al. who used all atom molecular dynamics (MD) simulations on a HIV protease complex.[22] These long simulations, produced with an implicit solvent model and no distance constraints, were able to sample the conformational changes that take place when a ligand is placed in the active site of HIV protease, and the protease moves from an open to a closed conformation. In contrast to the HIV protease simulations, an explicit solvent model was used in this study since it is known that water molecules play an important role in the stability of the ligand in the solvent exposed binding site.[10]

Our aim in this paper was to develop a molecular dynamics protocol to simulate the large conformational change of loop I in Pin1 moving from an open form, to a closed form when a pS/pT-P motif containing ligand (CTD-S2/S5) is bound. The protocol required validation which was achieved by producing multiple simulations. The first two were duplicate simulations of the CTD-S2/S5 ligand docked into Pin1. The third, also using the MD docking protocol, was a CTD-S5 simulation where the open form of Pin1 is docked with the singularly phosphorylated CTD-S5 ligand. Finally a standard MD simulation was run on the X-ray complex of Pin1 in its native bound conformation with the CTD-S2/S5 ligand to provide standard reference structures to compare with all three simulations run with the MD docking protocol. This reference simulation allowed us to validate the protocol, confirming that the MD docking protocol was able to produce complexes that sampled correct conformational space.

The second aim of this work was to use the simulations from the MD docking protocol to propose a potential binding mechanism for the CTD ligand to the Pin1 WW domain. By justifying that the starting structure used is a valid intermediate structure, key stages in the binding process were identified in the multiple simulations and the role of key residues in the binding process and stabilization of the complex, as identified by previous biological experimental studies, were explained. Most of the investigation of Pin1 in the literature has focused on its cellular role, yet there has been no investigation into the detailed ligand binding process. The importance of Pin1 in its many cellular processes means that the results from this study will be important in the future production of ligands to inhibit Pin1.

Materials and Method

Structure Creation. The Pin1 WW domain-CTD-S2/S5 complex used as the starting structure for the CTD-S2/S5 simulations was created using the scheme seen in Figure

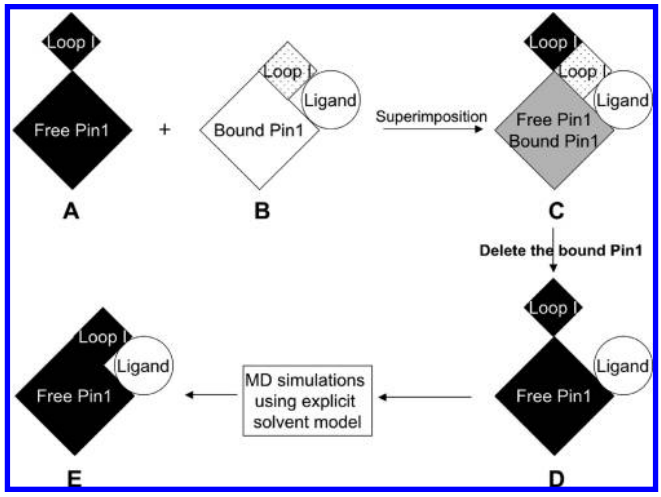

Figure 2. MD docking protocol. A→D show the production of the initial starting structure for the MD docking simulation. D→E show the initial structure moving into the closed form as seen in the "bound" X-ray structure as a result of the MD simulation.

2a−d. The "bound" structure of the Pin1 WW domain (b) (PDB ID: 1F8A), with the CTD-S2/S5 ligand (YpSPTpSPS), was superimposed onto the "free" structure of Pin1 (a) (residues 6−37 in PDB ID: 1PIN) using the Maestro program from Schrodinger.[23] From this superimposed structure (c), the "bound" Pin1 structure was deleted, and then the "free" Pin1 and the CTD ligand were merged together (d).

For the starting structure of the CTD-S5 simulation, where S2 is unphosphorylated, the Pin1-CTD-S2/S5 complex used as the starting structure for the CTD-S2-S5 simulations was modified by removing the phosphate group on S2 of the CTD-S2/S5 ligand to convert the phosphorylated serine to a standard serine residue. The crystal structure of the "bound" form of Pin1 (1F8A) was used as the starting structure for the X-ray complex simulation.

Parameters for the standard amino acids in the Pin1 WW domain and the ligand were taken from the AMBER ff03 force field,[24] while parameters for the phosphorylated serines (phosphoserine with deprotonated phosphate group, SER-PO3) were taken from the AMBER Parameter Database (http://pharmacy.man.ac.uk/amber/).[25] Hydrogen atoms were added such that all residues were in their standard protonation states at pH 7. Each complex was placed in a octahedral box of TIP3P water molecules[26] such that the box extended at least 10 Å from any atom of the Pin1-CTD ligand complexes. The systems were neutralized and made up to a salt concentration of 0.2 M by adding appropriate numbers of Na^+/Cl^- ions. These final steps of molecular system creation were achieved using the Leap module of AMBER9.[27]

MD Simulation Protocol. All simulations involved in the MD docking protocol were performed using the AMBER9 package of programs.[27] A four-step equilibration scheme was performed with the first step being a minimization where the Pin1/ligand complex was held fixed with a restraint of 500 kcal mol^{-1} $Å^{-1}$, such that the position of water molecules and ions could be relaxed. A second minimization followed, whereby all atoms were allowed to move for 1000 steps of steepest descent and then 1000 steps of conjugate gradient to relax the system as a whole. Once sufficiently relaxed, the system was carefully heated to 300 K over 10 ps using a Langevin dynamics temperature regulation scheme[28] with a collision frequency of 1 ps^{-1}, with everything except water held under a 10 kcal mol^{-1} $Å^{-1}$ restraint. Volume was kept constant, and SHAKE[29] was used to constrain bonds involving hydrogen such that a time step of 2 fs could be used. The final step of the equilibration scheme, to equilibrate the physical parameters such as temperature, pressure, and density of the system, was run over 10 ps using NPT conditions. Temperature and pressure were held constant using Langevin and isotropic scaling schemes, respectively. Production phase was run in 10 ns blocks under the same NPT conditions used for the final step of equilibration. The random seeds used to generate velocities were changed for each different block of a simulation to stop the possibility of synchronization of trajectories.[30] SHAKE was used, and long-range electrostatic interactions were calculated using the PME method with a cutoff of 10 Å.[31] Coordinates were printed to the trajectory file every 500 steps (every 1 ps) such that each trajectory held 10000 snapshots and a restart file was written every 1000 steps. Each simulation was continued on for a further 10 ns after a stable complex was formed for validation purposes.

Structural Analysis. Visual, distance, torsion, and rmsd analyses were completed using VMD.[32] All quantitative data used in the validation of the MD docking simulations were taken from the last 10 ns of the simulations. Over this time it was confirmed that the simulations were sampling a stable conformational space through hydrogen bond (H-bond) analysis using the ptraj module of AMBER9 (data not included).

Distances between the S5 phosphate of the CTD ligand, and key functional groups in loop I which are shown to H-bond to S5, were calculated between the phosphorus of the S5 and the hydrogen connected to the H-bond donor in loop I. Distance data for the Pin1 1F8A structure were calculated using the final structure from the first minimization of the X-ray complex from the X-ray complex simulation. This minimized structure was used instead of the initial X-ray structure as the X-ray structure did not include hydrogen atoms in optimized positions. The phi/psi angles on either side of the amide bond between Arg17 and Ser18 were calculated to quantify amide bond flipping.

The X-ray structure of the Pin1 complex (PDB ID: 1F8A) was used as the template structure for all rmsd calculations. All trajectories were aligned onto the backbone of this structure, and rmsds were calculated for each snapshot over the entire production phase. Snapshot intervals were changed to 5 ps to ensure that snapshots would be uncorrelated when compared to the next/previous snapshot.[33] Heavy atom and backbone rmsds were calculated for the entire Pin1 WW domain (residues 6−37), the loop I residues (16−21), and the XP groove (23 and 34) and were then averaged for each 10 ns portion of production phase.

Root mean square fluctuations (RMSF) for each residue were calculated using ptraj (see the Supporting Information). Each 10 ns portion of a simulation was rms fitted, for residues

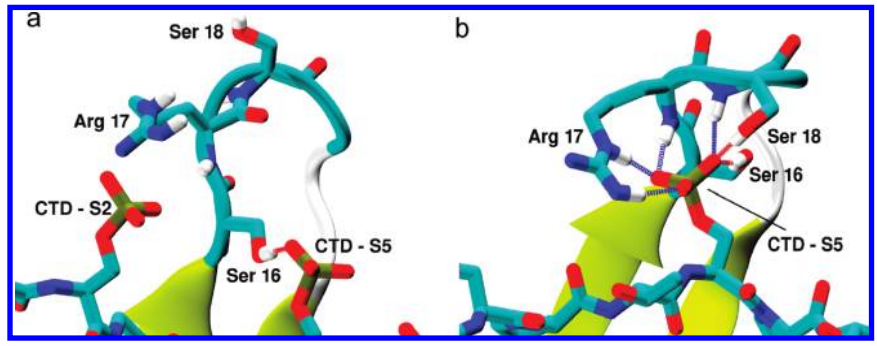

Figure 3. Interactions between loop I of the Pin1 WW domain and the CTD S2/S5 ligand in (a) the open and (b) closed form of the Pin1 complex from the MD docking protocol simulation. Specific hydrogens are shown to highlight the 6 key hydrogen bonds shown in (b) that stabilize the bound form of the complex.

6−37 from Pin1, onto the first structure of that portion of the simulation. RMSF were calculated for both Pin1 and CTD ligand residues.

Binding Energy Calculations. Binding energies were calculated using the MM-PBSA approach from AMBER9.[34,35] The single trajectory approach for calculating MM-PBSA energies was used with the ensemble of structures obtained from the final 10 ns of each simulation. Snapshots were taken every 5 ps from these trajectories such that 2000 snapshots were used in the calculations. For each snapshot, water molecules and counterions were removed, and estimates of the binding energies were calculated using

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{ligand}) \quad (1)$$

where the free energy of each part of the system was calculated using the formula

$$G = <E_{MM}> + <G_{solv}> - <TS_{solute}> \quad (2)$$

with <> signifying average values over the simulation. The E_{MM} term incorporates the molecular mechanical (MM) energies of the system. Using the single trajectory approach, this term was calculated using just the nonbonded, electrostatic and van der Waals components. The solvation energy (G_{solv}) term was divided into two parts

$$G_{solv} = G_{pol} + G_{np} \quad (3)$$

The nonpolar contribution (G_{np}) to G_{solv} was calculated using

$$G_{np} = \gamma SASA + \beta \quad (4)$$

where $\gamma = 0.0072$ kcal mol^{-1} Å^{-2}; $\beta = 0.0$ kcal/mol; and SASA is the solvent accessible surface area. The polar contribution (G_{pol}) to G_{solv} was calculated using both the Poisson−Boltzmann (PB) and Generalized-Born (GB) methods. PB estimates of G_{solv} were calculated by solving the Poisson−Boltzmann equation using the PBSA module of AMBER 9. GB estimates of G_{solv} were calculated using the GB model developed by Onufriev et al. and implemented in SANDER from AMBER 9.[36] The interior and exterior dielectric constants were set to 1 and 80, respectively, with a grid spacing set to 0.5 Å and 1000 linear iterations were performed. Bondi radii and a probe radius of 1.4 Å were used for both G_{pol} and G_{np} calculations. The entropy component (TS_{solute}) of the system was not calculated as discussed in the Results section of this paper.

Results and Discussion

The MD docking protocol was tested by running three separate simulations. The first test involved running on a simulation on the Pin1 CTD-S2/S5 complex. Further testing of the MD docking protocol was achieved by performing an independent duplicate of the CTD-S2/S5 simulation. The independence of the duplicate simulation was achieved by changing the random number seeds which initiate velocities and the forces on atoms at each stage of the simulation. This guaranteed that although the simulations started from the same structure, the velocities generated would be independent and therefore the simulations also. It has also been shown that if the random number seed values are not changed, simulations can synchronize.[30]

Binding energy studies have shown that there is no significant difference in binding affinity between Pin1 binding to the CTD-S2/S5 and CTD-S5 ligands,[10] suggesting that phosphorylation of S2 has a limited role in the binding process and therefore the bound structure would resemble the X-ray structure of the CTD-S2/S5 complex. The third and final test of the MD docking protocol was to run a simulation on the Pin1 CTD-S5 complex to determine whether the protocol was able to predict a bound conformation of this new complex, and, if so, did it have a similar conformation to the bound CTD-S2/S5 ligand complex.

In all three MD docking simulations performed, Pin1 was seen to fold from its open form (Figure 3a), into an intermediate "semi-open" form of the Pin1 WW domain after an 180° amide flip between Arg17 and Ser18 in loop I, to finally the closed form (Figure 3b), creating a ligand binding pocket as seen in the X-ray crystal structure 1F8A.

Validation of the MD Docking Simulations. The quality of the simulations, and in particular, the final structures in which the ligands were bound to Pin1, were validated through comparisons with the X-ray structure of the bound complex (1F8A). It should be noted that this X-ray structure was not used to generate the simulation. rmsd, distance, and binding energy comparisons were made from the structures from the final 10 ns of each simulation (2000 structures) to the X-ray structure.

The data from Table 1 show that for the backbone atoms of the whole protein, loop I, and the XP groove, the rmsds

Table 1. Rmsds over the Last 10 ns of Each Simulation against the X-ray Structure of the Pin1 Complex (1F8A)[a]

	all		loop I		XP groove	
	backbone	heavy	backbone	heavy	backbone	heavy
CTD-S2/S5	0.97 (0.07)	1.97 (0.09)	1.01 (0.16)	1.53 (0.11)	0.77 (0.09)	1.20 (0.03)
CTD-S2/S5 (duplicate)	1.02 (0.08)	2.17 (0.08)	0.87 (0.10)	1.61 (0.17)	0.72 (0.08)	1.23 (0.05)
CTD-S5	1.04 (0.11)	2.00 (0.12)	0.99 (0.13)	1.37 (0.19)	0.69 (0.06)	1.18 (0.02)
X-ray complex	0.78 (0.11)	2.13 (0.25)	1.10 (0.31)	1.43 (0.20)	0.62 (0.12)	0.67 (0.08)

[a] Standard deviation in parentheses.

Table 2. Distances (Å) from the Phosphorus of S5 to the Hydrogens of Loop I Functional Groups over the Last 10 ns[c]

	Ser16 OH	Arg17 N	Arg17 NE	Arg17 NH	Ser18 N	Ser18 OH
Pin1 1F8A[a]	2.93	2.98	2.88	2.47	3.32	5.57
CTD-S2/S5	2.85 (0.03)	2.74 (0.06)	2.85 (0.03)	2.66 (0.06)	3.27 (0.26)	3.17 (0.75)
CTD-S2/S5 (duplicate)	2.82 (0.03)	2.72 (0.05)	2.86 (0.05)	2.72 (0.08)	3.10 (0.05)	2.81 (0.18)
CTD-S5	2.84 (0.03)	2.69 (0.07)	2.79 (0.04)	2.70 (0.09)	3.03 (0.13)	2.90 (0.16)
X-ray complex	2.81 (0.03)	2.67 (0.01)	2.78 (0.01)	2.65 (0.01)	3.13 (0.02)	2.81 (0.02)
initial[b]	2.16	5.78	10.41	11.83	9.00	10.51

[a] Pin1 1F8A distances are taken from the final structure of the first stage of minimization from the X-ray complex simulation. [b] Initial distances are taken from the structure used to start all the CTD simulations. [c] Standard deviation in parentheses.

Table 3. Binding Energies from Last 10 ns of Simulations[a]

method	CTD-S2/S5	CTD-S2/S5 (duplicate)	CTD-S5	X-ray complex
PB	−64.18 (1.05)	−64.05 (0.92)	−63.81 (0.89)	−66.23 (0.99)
GB	−52.42 (0.80)	−53.57 (0.75)	−53.65 (0.73)	−55.92 (0.78)

[a] Energies in kcal mol^{-1} and standard errors in parentheses.

are all about 1 Å. This is well below the standard 2 Å limit with which structures are said to be "nativelike".[37] In order to ascertain the quality of this data, a MD simulation of the X-ray structure (1F8A) with hydrogens added and in a solvent box was run and the last 10 ns of the simulation compared back to the single X-ray structure (shown in bottom row of Table 1). These data are almost identical to the data seen for the simulations run with the MD docking protocol suggesting the simulations are sampling a conformational space similar to the X-ray structure.

Interactions between the loop I residues in Pin1 and the S5 phosphate of the CTD ligand are compared by calculating the distances of six important hydrogen bonds from the last 10 ns of the simulations against the distances seen in the X-ray complex. Table 2 shows that five of the six hydrogen bonds between the loop and the ligand are directly comparable with the X-ray structure. For Ser18 OH however, there is a considerable difference between the distance in the simulation (3.17 Å) and the distance in the X-ray structure (5.6 Å). Interestingly, the X-ray complex simulation shows this considerable difference to the single X-ray complex structure as well and is discussed below. For completeness, Table 2 also shows the six H-bond distances from the initial structures used for the CTD-S2/S5 simulations where loop I was in its open form. These highlight the dramatic changes in structure that took place in the simulation. Overall the six H-bond distances are internally self-consistent and comparable with the X-ray structure confirming that our simulations were sampling the correct conformational space.

The final tool used to validate the MD docking simulations was the calculation of "binding energies" using the MM-PBSA protocol within the AMBER software.[34,35] The binding energies of the MD docking simulations and the X-ray complex simulation are very similar, differing by on average 2.2 and 2.7 kcal/mol for the Poisson−Boltzmann (PB) and Generalized Born (GB) methods, respectively (Table 3). The results from the PB and GB methods are expected to be different but show internal self-consistency[38] which is what is seen in these results. It is important to note that these energies are purely relative and are not absolute estimates of the binding energy. MMPBSA/GBSA has consistently been shown to be a good method for comparing binding energies of similar complexes.[34,38,39] Entropy calculations have not been included as it has been documented in many papers that it is difficult to calculate, and with similar ligands it can be assumed that differences in entropy will be small.[34,40] Given that the simulations started from structures very different to the X-ray structure, and that such close agreement is seen, the binding energy data again confirms that the simulations are sampling a very similar conformational space.

Importance of an Explicit Solvent Model. The use of an explicit water model was essential to simulate the transformation of the Pin1 WW domain from the "open" to "closed" form. Preliminary simulations performed on the CTD-S2/S5 complex following the protocol of Hornak et al., using an implicit solvent model failed consistently with the CTD ligand quickly flying away from the binding site. These simulations failed due to the binding site being on the surface of the protein and, hence, solvent exposed, which contrasts to the HIV protease work where the ligand is placed in a cavity which is solvent excluded. By modeling the Pin1/CTD ligand simulations in explicit solvent, the movement of the ligand was significantly restrained in the binding site which allowed loop bending and, eventually, the formation of the strong and stable complex.

The crystal structure of the Pin1 complex (1F8A) showed that there is a key water mediated interaction between Tyr23 and the S5 phosphate of the ligand. Due to the solvent exposure of the binding site, it is expected that many different water molecules will mediate this interaction. This is seen in the MD docking simulations with the water molecules continually exchanging with the bulk solvent. This continual

exchange of water molecules allows for a consistent interaction between Tyr23 and S5, helping to stabilize the complex at early stages and increase strength of binding during the latter stages. This again illustrates the importance of performing these simulations in explicit solvent as an artificial restraint would need to be placed on the structural water molecule for it to stay in position in an implicit solvent simulation.

Successful simulation using the MD docking protocol relied upon placement of proline 6 of the ligand in the XP groove of Pin1, with the proline sandwiched between Try23 and Trp34. It could be suggested that this is a limitation of the MD docking protocol in that we have used the bound X-ray structure as a basis for this work. However we view the work in this article is a proof of concept showing that we can position a ligand with one key component near the correct position and run a simulation to produce a known bound complex.

Overall it has been shown that the MD docking protocol has been able to dock ligands into a binding site and simulate the movement of the complex from an "open" free state to a "closed" bound state. It has also predicted that the CTD-S2 ligand, of which no crystal structure has been produced, binds in an identical manner to the CTD-S2/S5 ligand, as highlighted by similar rmsd, distances and pleasingly having a similar binding energy as seen previously in experimental studies.

Proposed Binding Mechanism

Justification of Initial Ligand Placement. If the starting structure used in the MD docking simulations is a valid intermediate in the binding process, mechanistic information from the MD simulations can provide a detailed description of the binding process. The simulations presented in this paper start with the ligand proline (Pro6) positioned close to the XP groove. Given that the predominant ligand interactions with Pin 1 in the bound structure are (i) the Pro6/XP groove interactions and (ii) the phosphorylated serine (pSer) interactions with the flexible loop I, it might be argued however that an alternative starting geometry based on the pSer/loop I interaction could be plausible.

There is significant evidence supporting the starting geometry used in these simulations as a valid intermediate. The first reason is by docking Pro6 into the XP groove, significant hydrophobic burial occurs, in which the placement of the hydrophobic proline in XP groove reduces the unfavorable solvent contacts these hydrophobic residues experience. Second, by docking Pro6 into the XP groove, the pSer is positioned in the vicinity of the loop binding site as a result of the geometric consequences of Pin1 only binding the trans form of the CTD ligand.[11,17] Placing pSer near the binding site allows subsequent bending of loop I to generate a number of H-bonds to pSer. Third, mutational analysis of the XP groove residues shows no binding is seen when either residue, Tyr23 or Trp34, is mutated to alanine.[14] This suggests that binding can only occur when a fully functional XP groove is available. Finally the fact that our MD docking simulations have been able to correctly produce

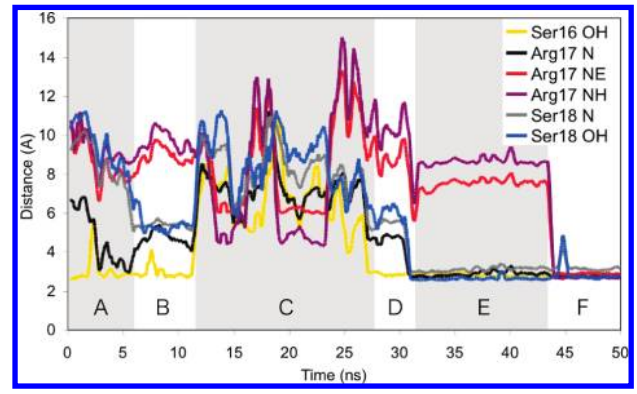

Figure 4. Distances from key functional groups in loop I residues to the phosphorus atom of S5 of the CTD ligand for the first 50 ns. The timeline is shaded to define the six stages of the simulation. Distances were calculated as running averages over 100 snapshots. After 50 ns, only insignificant changes are seen and have therefore been excluded from the figure for clarity.

structures that sample a similar conformational space to the X-ray structure suggest this could be a valid intermediate structure.

The alternative starting geometry which relies upon pSer first interacting with loop I of Pin1 is not favored as a valid intermediate for a number of reasons. The first is that pSer is a polar residue which would result in a significant desolvation penalty which pSer would have to overcome upon binding by forming interactions with a number of residues simultaneously.[41−44] This is unlikely since loop I is not optimized to interact in its open form with pSer as: Ser16 is hidden in the binding site and is not accessible; Ser18 points out toward the solvent; while Arg17 is flapping around in space, far from the binding site. Another reason is that any interactions between pSer and loop I would significantly decrease the flexibility of loop I,[45−47] resulting in significant barriers to allow loop I to bend, and to bring the CTD ligand along with it. Finally, mutational analysis of key loop I residues to alanine show that although a reduction of binding strength is seen, binding occurs none the less.[10,14,48] Therefore by removing potential H-bond partners, binding still occurs suggesting initial binding may not take place via this approach.

CTD-S2/S5 Docking Simulation. Given that the starting structure has been shown to be a valid intermediate, the MD docking simulations could be investigated to identify a potential binding mechanism. The final structures from the initial CTD-S2/S5 MD docking simulation show that there were 6 key hydrogen bonds (H-bonds) between residues in loop I and the CTD ligand (as shown in Figure 3b). The progress of binding was followed by monitoring these distances throughout the simulation, in which 6 distinct stages are clearly identified (Figure 4). This MD docking simulation is much more readily understood by visualization and therefore a QuickTime movie (15MB) of this simulation is available in the Supporting Information for this paper. Each stage is characterized by a geometric change in the structure of the complex. The stages, A through F, of the simulation are described below and give a fascinating insight into the proposed binding mechanism.

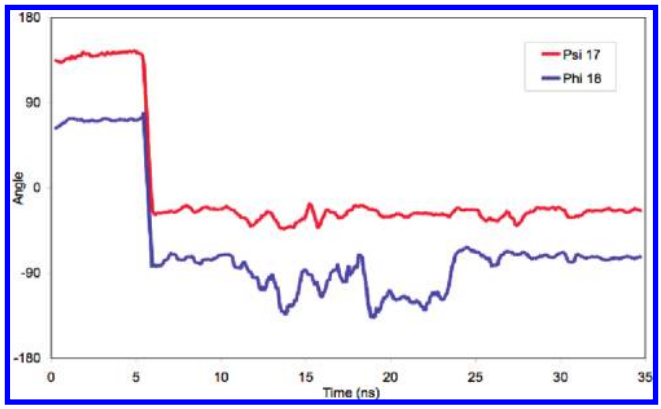

Figure 5. Plot of the psi 17 (17N - 17C-alpha - 17C - 18N) and phi (17C - 18N - 18C-alpha - 18C) torsion angles from the first 35 ns of the simulation. Angles were calculated as running averages over 100 snapshots.

A) Starting Conformation. At the start of the simulation the proline at position 6 of the CTD-S2/S5 ligand is positioned into the XP groove of Pin1 according to the bound X-ray structure. Throughout the whole simulation there is very little movement in this proline, with this group functioning as an anchor, holding the ligand in place while other motions takes place. With the proline sandwiched in between Tyr23 and Trp34, the S5 phosphate points up toward loop I and forms a stable hydrogen bond with the Ser16 OH. Loop I is open at this stage of the simulation, and no other hydrogen bonds are seen between the S5 phosphate and the rest of loop I.

B) Amide Flip. After about 6 ns the amide bond between Arg17/Ser18 flips, as shown by the 180° change in the psi-17/phi-18 angles in Figure 5. This flip of the amide bond does not significantly alter the backbone of adjacent residues nor does it change the omega angle of the amide bond (data not shown). The carbonyl of Arg17 now points out toward the solvent while the amine (N−H) of Ser18 points toward the ligand, and loop I bends into a "semiopen" conformation. Thus the distances between both the backbone Ser18 NH (Ser18 N) and the Ser18 OH and the S5 phosphorus decrease from about 8 Å to just above 5 Å (Figure 4). Following the flip, a short period of stability (from 8−12 ns) is seen in which a number of water molecules position themselves between loop I and the S5 phosphate, forming water mediated H-bonds. These water molecules effectively block S5 from interacting directly with the loop I residues, in turn blocking the transition to the "closed" form. As a result of this, instead of moving closer toward loop I, at about 12 ns the S5 phosphate moves away from the binding site and a period of chaos ensues.

C) Chaos. During this period, no hydrogen bonds are seen between loop I and the S5 phosphate, and a high degree of flexibility is seen throughout the complex. From RMSF data (Supporting Information) it can be seen that the backbone of S5 is quite stable with the side chain being flexible. The flexibility of the S5 side chain is highlighted in this "chaos" period where it rotates about 60° out into the solvent, as shown in Figure 6, allowing layers of water molecules to position themselves between S5 and loop I. During this chaos period water mediated interactions are seen between S2 and S5, interactions between S2 and the side chain of Arg17, and interactions between S5 and the side chain of Arg17. After 16 ns of chaos, S5 rotates back into the binding site, as shown by the decreases in length between Ser16 OH and S5, and chaos finally comes to an end when a stable H-bond again forms between these two residues at about 28 ns.

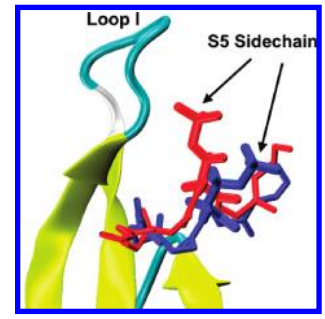

Figure 6. Pin1 WW domain in its semiopen form with CTD-S2/S5 ligand having its S5 side chain positioned in the pS/pT binding site (red) and after 60° rotation away from the binding site in the "chaos" period (blue).

D) Bending Cascade. The reformation of the S5 to Ser16 OH H-bond starts a cascade of events which allow the loop to fully bend over into the "closed" conformation. The first step of this cascade occurs at 30 ns where the water mediating the interaction between the Arg17 backbone NH (Arg17 N) and the S5 phosphate moves from the binding site allowing a direct H-bond. This causes a small downward movement of loop I which in turn removes the waters mediating interactions between Ser18 N/Ser18 OH and the S5 phosphate. This causes another downward movement of the loop and two direct H-bonds between Ser18 and S5 are formed. Thus at 32 ns we have 4 direct H-bonds to S5 phosphate with all distances being consistently 3 Å, and therefore the complex is in a stable conformation.

E) Arg17 Side Chain Optimization. During the time that the rest of loop I was stable, the side chain of Arg17 alternately interacts with both the solvent and with the S2 phosphate. At ∼44 ns, the side chain of Arg17 finally forms two H-bonds with the S5 phosphate. After this time in the simulation, the Arg17 side chain remains at the standard 3 Å distance from the S5 phosphate.

F) Strong Stable Complex. Throughout the remaining 25 ns (through to 70 ns) a predominantly strong and stable complex is seen. Occasionally loop I tries to move back to the semiopen form, as seen by the Ser18 distance increasing to ∼5 Å. This is quickly rectified, and the Ser18 distances decrease back to 3 Å almost immediately, presumably because of the stability caused from the rest of the H-bonding interactions, such that the strong stable complex is the predominant conformation seen in the last 25 ns of the simulation.

CTD-S2/S5 Duplicate Docking Simulation. Although the same conformational space is being sampled at the end of the duplicate simulation, Figure 7 shows there are considerable differences in the time evolution of the two simulations, in particular that there is no "chaos" period (stage 'C') in the duplicate simulation. This is because S5 does not move outside the binding site and is shown in Figure 7 by the Ser16

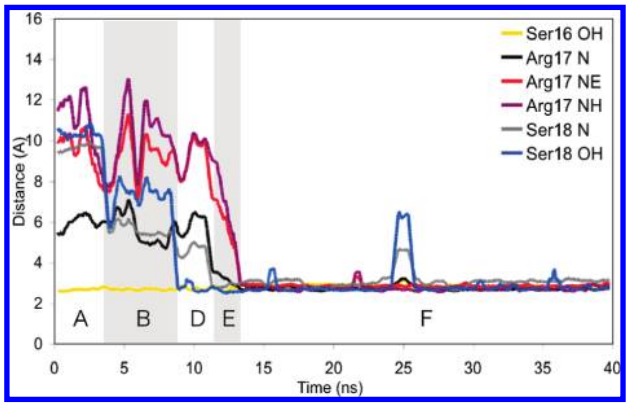

Figure 7. Distances from key functional groups in loop I residues to the phosphorus atom of S5 of the CTD ligand for the duplicate simulation. The timeline is shaded to define five out of the six stages of the simulation as defined for the initial CTD-S2/S5 simulation. Distances were calculated as running averages over 100 snapshots.

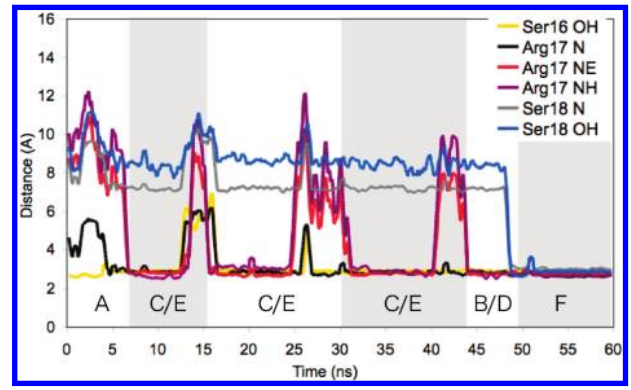

Figure 8. Distances from key functional groups in loop I residues to the phosphorus atom of S5 of the CTD ligand for the CTD-S5 simulation. The timeline is shaded to define the six stages of the simulation. Distances were calculated as running averages over 100 snapshots.

OH distance never deviating from around the 2.7 Å distance. After the amide flip, which occurs on a similar time scale to the initial simulation, S5 stays in the binding site, and the bending cascade occurs much earlier. In the duplicate we also see that side chain optimization occurs at a much earlier stage than the initial simulation. Overall this means that a stable complex is formed after only ~15 ns compared to the 45 ns it took in the initial simulation. Although no "chaos" period is seen in the duplicate simulation, all other five stages of the binding can be identified.

An interesting aspect of the duplicate simulation is that the bending cascade occurs in the opposite order to the initial simulation. Instead of interacting with Arg17 N first and then interacting with Ser18 N and Ser18 OH, S5 H-bonds to Ser18 OH first and then to Ser18 N followed by Arg17 N. Visual analysis shows that this is because for the first part of the simulation, S5 interacts with Arg21 of loop I. This causes S5 to sit in the far right-hand side of the binding site (Figure 1-front view). In this position, S5 is unable to interact with Arg17, and it is only when Ser18 OH starts to interact with S5 (~9 ns) that we see a shift in the position of S5 back to the center of the binding site. This movement allows Ser18 N to H-bond to S5 which in turn allows Arg17 N to interact. With S5 centrally located in the binding pocket, side chain optimization can then occur. It is important to see the movement of S5 from interacting with Arg21 at one end of loop I to eventually interacting with Arg17 as it shows that although there are some favorable interactions between Arg21 and S5, it is more favorable for S5 to sit in the left-hand side of the binding site where multiple interactions can occur which will allow for a more stable complex.

As this duplicate has been started with a different seed compared to the initial simulation, the velocities of each atom at the start of the simulations will be different. Because of this, the different simulation will follow different trajectories. It is therefore to be expected that the simulations follow different paths and are not identical. However, although the simulations do not follow the same path, they do both show the key steps of binding: amide flip; a loop I bend; and Arg17 side chain optimization to derive bound structures as close to identical as possible to the X-ray structure, thus highlighting the quality of the MD docking protocol.

CTD-S5 Docking Simulation. The MD docking simulation of this mutated ligand showed that despite the absence of phosphorylation of Ser2 of the CTD ligand, the system was able to sample a conformational space very close to that of the X-ray structure of the CTD-S2/S5 complex. The binding process differed however from that seen for the initial MD docking simulation and its duplicate, although all six stages can be seen in the trajectory (Figure 8). A striking feature of this "unphosphorylated" simulation is that the amide flipping is the last conformational change to occur and does not occur until ~48 ns which is more than three times longer than any other simulation.

As it takes so long for the amide to flip, a number of conformations of the complex are seen in the preceding time. It can be seen from the distance graph (Figure 8) that after ~5 ns, both Ser16 OH and Arg17 N are H-bonded to S5. This indicates that Arg17 N can bind to S5 at any stage of the whole binding process and is not dependent on loop bending as suggested from the initial simulations. Unexpected interactions between the Arg17 side chain and S5 (stage 'E') are also seen before loop I bending. Due to the amide flip and loop I bending having not occurred yet, these interactions are not stable, and, thus, the system is able to move into small stages of chaos. It is thought that these stages of chaos are shorter than seen in the initial simulations due to S2 no longer being phosphorylated and therefore not being able to stabilize the Arg17 side chain or S5 in their "chaotic" forms, allowing them to position themselves back to the binding site relatively quickly.

In the end it appears coincidental that at the time that the amide flip occurs, all other important residues are H-bonded to S5 which allows immediate formation of the strong and stable complex. This is different than the other simulations where the amide flip occurred before S5 was in its "bound" position. It can therefore be concluded that the bending cascade (stage 'D') can only occur once the amide has flipped and the S5 is positioned correctly in the binding site. This shows the pronounced importance of the amide flip and how it is a rate limiting step of the proposed binding mechanism.

Key Residues in Simulations. MD docking simulations in this work have provided significant insight into a potential binding mechanism of CTD ligands with the pS/pT-P motif to the Pin1 WW domain. In particular it highlights the importance of the structure of the complex around the XP groove; the amide flip; and the interactions between S5 and the loop I residues.

The MD simulations produce in this work show that the structure in the vicinity of the XP groove is very stable, as seen from the backbone data in Table 1 with an average rmsd of ~0.7 Å. This compares favorably with the rmsd from the X-ray complex simulations and given the low backbone rmsd to start with, it can be concluded that the backbone of the XP groove must have started in a conformation very similar to the final bound conformation and does not deviate significantly from this position. It is appropriate to suggest that the stability of the XP groove is due to its hydrophobicity and it being unfavorable for it to interact more significantly with the surrounding polar environment.

Other than interacting with proline, Tyr23 plays another key role in binding by being able to also form water mediated H-bonds to S5. Interaction energy calculations show that there is a significantly high electrostatic interaction between Tyr23 and the S5 phosphate. These energy calculations are unable however to include the contribution that the water mediated H-bonds play as they are transient (different water molecules play the role of the mediator). It can also be seen from these calculations that there is a significant VDW contribution to the interaction energy between proline for both Tyr23 and Trp34. No doubt this is in part due to their hydrophobic interactions to the proline from the ligand. Together with the data presented earlier, this work has shown the critical role that the XP groove has in binding and the production of a stable complex.

A significant difference in the orientation of the amide bond between Arg17 and Ser18 is observed between the open (1PIN) and closed (1F8A) forms of Pin1 WW domain. In the open form, the NH points away from where the ligand binds, while, in the closed form, the NH points toward the ligand. In all the MD docking simulations produced in this work, this amide 17/18 flip was seen allowing a conformational change of loop I to occur unhindered.

This amide bond flipping is a very important step in the binding process of the ligand to Pin1, as the loop cannot bend until this amide flip has taken place. Without the amide flip, interactions between S5 and Ser18 N or Ser18 OH would not be seen which would severely limit the ability of a strong and stable complex to be formed. To consolidate this hypothesis, a further simulation was run in which the amide bond in the free/open structure was manually flipped such that it showed no structural difference to its closed form. From this simulation, conformational change of loop I was seen after just 1 ns, which gave strong evidence to suggest that the NH of the amide between Arg17 and Ser18 must be pointing toward the ligand for conformational change to occur. Therefore the timing of the amide flip is critical to the timing of conformational change and therefore binding.

Initial comparison of the Pin1 free and bound complexes showed the greatest difference between the two structures was seen in the loop I region. This has been highlighted in our simulations and through rmsd, distance, and RMSF calculations, these changes have been quantified. All this is understandable as this is where the important loop bending occurs which locks the ligand into its bound position.

Of the loop I residues, the first residue of the loop, serine 16, is the most stable. The reduced movement seen in Ser16 relative to the other residues suggests that Ser16 may initially be in an optimal position for binding. Our simulations show that Ser16 consistently interacts with S5, and the interaction is necessary for subsequent stages of the binding process to occur. Without interaction between Ser16 and S5, no other interactions are seen between S5 and the loop except for the Arg17 side chain for small time scales in the chaos period of the initial CTD-S2/S5 simulation. It therefore seems that Ser16 is crucial to the whole binding process, which has been suggested in other work.[10,19,48,49]

Previous work, based on the crystal structure of the bound form of Pin1, suggests that Ser18 plays a minimal role in the binding to the CTD ligand.[10] In that X-ray structure (which does not include hydrogens), the Ser18 backbone NH is slightly outside the H-bonding distance, and the oxygen in the Ser18 side chain points away from the ligand, thereby preventing any H-bonding to S5. It is because of these distances, and some binding energy calculations,[48] that Ser18 was predicted to be unimportant. However Ser18 is one of only two residues from loop I that are conserved in the analogues of Pin1, PinA, and Ess1, suggesting a more significant role in the binding process. In addition to this, our MD simulation of the X-ray complex, which started from the initial X-ray structure, actually shows the Ser18 side chain rotating such that the distance from the S5 phosphorus to Ser18 OH decreases from 5.57 Å to within 3 Å. The Ser18 backbone NH also sits at a similar distance, ultimately allowing two H-bonds to form to S5. The MD docking simulations carried out also show this conformation where the Ser18 NH/OH are within H-bonding distance of S5 (Table 2). Given that this form of the complex is the predominant one seen, the simulations indicate that Ser18 has a much more significant role to play in binding than previously thought.

Since the X-ray structure of Pin1 bound to the CTD ligand was published it has been thought that Arg17 played a very important role in binding.[10] Due to arginine being a positively charged amino acid and the S5 phosphate from the ligand being negatively charged, it is easy to understand why this has been thought. This has been supported by binding studies which have shown that by mutating Arg17 to Ala, the binding strength decreases 6-fold,[10] and that a homologue of Pin1, PinA, where Arg17 is replaced by Asn, also has a binding strength ~6-fold less.[19] Our simulations also support the importance of Arg17 to the binding strength by showing that Arg17 contributes 3 H-bonds to the final complex, two from the side chains and one from the backbone NH.

Arginine is a highly charged, flexible residue that is regularly seen on the surface of proteins. The ability of our simulations to show this flexibility, yet also to show the remarkable stability when the Arg17 side chain interacts with S5 is extremely important to the successful simulation of

the binding process. The fact that different time intervals are seen for when this stability occurs also highlights the fact that the Arg17 side chain optimization step is an independent process and another rate-limiting process in binding of the Pin1 WW domain to a ligand.

As important as Arg17 has been shown to be to the overall binding strength of Pin1 to the CTD ligand, our simulations show that Arg17 is not critical for binding to occur. There are numerous stages in the simulations where a stable complex is formed without interaction between S5 and the Arg17 side chain and where Arg17 interactions occur yet stability is not seen. The binding studies discussed above also suggest that Arg17 is not critical as complexes are still formed with acceptable binding energies when Arg17 is not present. Sequence alignment of the three analogous WW domain containing proteins, Pin1, PinA, and Ess1, show that the key residues in loop I at positions 16, 17, and 18 are conserved with residue 17 either an arginine, asparagine, or lysine. This suggests that the binding motif that pSer looks for may be S-X-S, where X is an amino acid containing at least one amine group in its side chain.

Conclusion

The development of an MD docking protocol using all atom unrestrained MD simulations in this work has allowed us to simulate the binding of a ligand into a solvent exposed binding site. Validation of the MD docking simulations has shown that our MD docking protocol has produced structures that sample the same conformational space as the X-ray structure of the bound complex. The simulations also highlight the importance of using an explicit solvent model for docking a ligand into a solvent exposed binding site.

The results of the simulations, given that the starting structure is a valid intermediate as has been justified, show that the binding of a ligand containing the pS/pT-P motif to Pin1 is a multistage process. First, the proline of the ligand docks itself in between the XP groove residues Tyr23 and Trp34 of Pin1 where it is stabilized by hydrophobic interactions. The next steps in the binding mechanism are the flip of the amide bond between Arg17 and Ser18 and Arg17 side chain optimization. The flipping of the amide bond allows loop I to move into a semiopen conformation, from which it eventually moves into a closed conformation when S5 is correctly positioned in the binding site. Optimization of the Arg17 side chain occurs when H-bonds form between the NHs of the Arg17 side chain and the oxygens from the S5 phosphate. It is only when all stages of the binding mechanism have been completed do we see the strong and stable complex.

Our work using a MD docking protocol for solvent exposed binding sites has shown that a number of interactions between the pS-P motif in the ligand and the XP groove and loop I residues are vital for the formation and strength of the bound complex. These results confirm that for a ligand to bind into the Pin1 WW domain, it must have a hydrophobic fragment to bind into the XP groove, connected to a fragment containing a phosphate or some very electronegative atoms in the correct spatial relationship to interact with loop I as highlighted by recently published work involving phosphorylated serine mimics.[50]

This work has shown that the MD docking protocol can be used as a refining tool after docking. To become a truly predictive tool, the protocol should be able to place the CTD ligand in any conformation in the environment surrounding the binding site and, following a MD docking simulation, have the complex settle into the bound conformation. It would then be possible to use the protocol to identify potential ligands designed to inhibit phospho-peptide binding.

Acknowledgment. The Victorian Partnership for Advanced Computing (VPAC) is thanked for access to their computational resources. D.O. is supported by an Australian Postgraduate Award (APA) scholarship.

Supporting Information Available: A QuickTime movie of the CTD-S2/S5 MD docking simulation; additional data from the MM-PBSA calculations; figures detailing the change of RMSF over time for residues in Pin1 and in the CTD-ligand; and the complete ref 27. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Joseph, J. D.; Daigle, S. N.; Means, A. R. PINA is essential for growth and positively influences NIMA function in Aspergillus nidulans. *J. Biol. Chem.* **2004**, *279*, 32373–32384.

(2) Lu, K. P.; Hanes, D.; Hunter, T. A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **1996**, *380*, 544–547.

(3) Wulf, G. M.; Ryo, A.; Wulf, G. G.; Lee, S. W.; Niu, T.; Petkova, V.; Lu, K. P. Pin1 is overexpressed in breast cancer and cooperates with Ras signaling in increasing the transcriptional activity of c-Jun towards cyclin D1. *EMBO J.* **2001**, *20*, 3459–3472.

(4) Lu, K. P. Pinning down cell signalling, cancer and Alzheimer's disease. *Trends Biochem. Sci.* **2004**, *29*, 200–209.

(5) Bayer, E.; Goettsch, S.; Mueller, J. W.; Griewel, B.; Guiberman, E.; Mayr, L. M.; Bayer, P. Structural analysis of the mitotic regulator hPin1 in solution: Insights into domain architecture and substrate binding. *J. Biol. Chem.* **2003**, *278*, 26183–26193.

(6) Namanja, A. T.; Peng, T.; Zintsmaster, J. S.; Elson, A. C.; Shakour, M. G.; Peng, J. W. Substrate recognition reduces side-chain flexibility for conserved hydrophobic residues in human Pin1. *Structure* **2007**, *15*, 313–327.

(7) Ng, C. A.; Kato, Y.; Tanokura, M.; Brownlee, R. T. C. Structural characterisation of PinA WW Domain and comparison with other Group IV WW Domains, Pin1 and Ess1. *Biochim. Biophys. Acta* **2008**, *1784*, 1208–1214.

(8) Li, Z.; Li, H.; Devasahayam, G.; Gemmill, T.; Chaturvedi, V.; Hanes, S. D.; Van Roey, P. The structure of the Candida albicans Ess1 prolyl isomerase reveals a well-ordered linker that restricts domain mobility. *Biochemistry* **2005**, *44*, 6180–6189.

(9) Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P. Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* **1997**, *89*, 875–886.

(10) Verdecia, M. A.; Bowman, M. E.; Lu, K. P.; Hunter, T.; Noel, J. P. Structural basis for phosphoserine-proline recognition

by group IV WW domains. *Nat. Struct. Biol.* **2000**, *7*, 639–643.

(11) Lu, K. P.; Liou, Y. C.; Zhou, X. Z. Pinning down proline-directed phosphorylation signaling. *Trends Cell Biol.* **2002**, *12*, 164–172.

(12) Macias, M. J.; Hyvönen, M.; Baraldi, E.; Schultz, J.; Sudol, M.; Saraste, M.; Oschkinat, H. Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* **1996**, *382*, 646–649.

(13) Espanel, X.; Sudol, M. A single point mutation in a group I WW domain shifts its specificity to that of group II WW domains. *J. Biol. Chem.* **1999**, *274*, 17284–17289.

(14) Lu, P. J.; Zhou, X. Z.; Shen, M.; Lu, K. P. Function of WW domains as phosphoserine- or phosphothreonine-binding modules. *Science* **1999b**, *283*, 1325–1328.

(15) Sekerina, E.; Rahfeld, J. U.; Müller, J.; Fangḧ; Anel, J.; Rascher, C.; Fischer, G.; Bayer, P. NMR solution structure of hPar14 reveals similarity to the peptidyl prolyl cis/trans isomerase domain of the mitotic regulator hPin1 but indicates a different functionality of the protein. *J. Mol. Biol.* **2000**, *301*, 1003–1017.

(16) Joseph, J. D.; Yeh, E. S.; Swenson, K. I.; Means, A. R.; Winkler, K. E. The peptidyl-prolyl isomerase Pin1. *Prog. Cell Cycle Res.* **2003**, *5*, 477–487.

(17) Wintjens, R.; Wieruxzeski, J.; Drobecq, H.; Rousselot-Pailley, P.; Buee, L.; Lippens, G.; Landrieu, I. 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. *J. Biol. Chem.* **2001**, *276*, 25150–25156.

(18) Jäger, M.; Zhang, Y.; Bieschke, J.; Nguyen, H.; Dendel, G.; Bowman, M. E.; Noel, J. P.; Gruebele, M.; Kelly, J. W. Structure-function-folding relationship in a WW domain. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10648–10653.

(19) Kato, Y.; Ito, M.; Kawai, K.; Nagata, K.; Tanokura, M. Determinants of ligand specificity in groups I and IV WW domains as studied by surface plasmon resonance and model building. *J. Biol. Chem.* **2002**, *277*, 10173–10177.

(20) May, A.; Sieker, F.; Zacharias, M. How to efficiently include receptor flexibility during computational docking. *Curr. Comput.-Aided Drug Des.* **2008**, *4*, 143–153.

(21) Lee, K.; Lee, J. W. Computational approaches to protein-protein docking. *Curr. Proteomics* **2008**, *5*, 10–19.

(22) Hornak, V.; Okur, A.; Rizzo, R. C.; Simmerling, C. HIV-1 protease flaps spontaneously close to the correct structure in simulations following manual placement of an inhibitor into the open state. *J. Am. Chem. Soc.* **2006**, *128*, 2812–2813.

(23) *Maestro, version 8*; Schrodinger, LLC: New York, NY, 2007.

(24) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, W.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.

(25) Homeyer, N.; Horn, A. H. C.; Lanig, H.; Sticht, H. AMBER force field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine and phosphohistidine. *J. Mol. Model.* **2006**, *12*, 281–289.

(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(27) Case, D. A. et al. *AMBER 9*; University of California: San Francisco, CA, 2006.

(28) Chandrasekhar, S. Stochastic problems in Physics and Astronomy. *Rev. Mod. Phys* **1943**, *15*, 1–89.

(29) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular Dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(30) Cerutti, D. S.; Duke, R.; Freddolino, P. L.; Fan, H.; Lybrand, T. P. A. Vulnerability in popular molecular dynamics packages concerning Langevin and Andersen dynamics. *J. Chem. Theory Comput.* **2008**, *4*, 1669–1680.

(31) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: an Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(32) Humphrey, W.; Dalke, A.; Schulten, K. VMD - visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38, 27−28.

(33) Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein-Protein Binding by binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *J. Mol. Biol.* **2003**, *330*, 891–913.

(34) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham III, T. E. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(35) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **1998**, *120* (37), 9401–9409.

(36) Onufirev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem.* **2000**, *104*, 3712–3720.

(37) Lei, H.; Duan, Y. Two-stage Folding of HP-35 from Ab Initio Simulations. *J. Mol. Biol.* **2007**, *370*, 196–206.

(38) Gohlke, H.; Case, D. A. Converging Free Energy Estimates: MM-PB(GB)SA Studies on the Protein-Protein Complex Ras-Raf. *J. Comput. Chem.* **2004**, *25*, 238–250.

(39) Kuhn, B.; Kollman, P. A. Binding of a Diverse Set of Ligands to Avidin and Streptavidin: An Accurate Quantitative Prediction of Their Relative Affinities by a Combination of Molecular Mechanics and Continuum Solvent Models. *J. Med. Chem.* **2000**, *43*, 3786–3791.

(40) Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. Use of MM-PBSA in Reproducing the Binding Free Energies to HIV-1 RT of TIBO Derivatives and Predicting the Binding Mode to HIV-1 RT of Efavirenz by Docking and MM-PBSA. *J. Am. Chem. Soc.* **2001**, *123*, 5221–5230.

(41) Hendsch, Z. S.; Tidor, B. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* **1994**, *3* (2), 211–226.

(42) Hendsch, Z. S.; Tidor, B. Electrostatic interactions in the GCN4 leucine zipper: substantial contributions arise from intramolecular interactions enhanced on binding. *Protein Sci.* **1999**, *8* (7), 1381–1392.

(43) Sheinerman, F. B.; Honig, B. On the Role of Electrostatic Interactions in the Design of Protein-Protein Interfaces. *J. Mol. Biol.* **2002**, *318*, 161–177.

(44) Schapira, M.; Totrov, M.; Abagyan, R. Prediction of the binding energy for small molecules, peptides and proteins. *J. Mol. Recognit.* **1999**, *12* (3), 177–190.

(45) Borchert, T. V.; Kishan, K. V. R.; Zeelen, J. P.; Schliebs, W.; Thanki, N.; Abagyan, R.; Jaenicke, R.; Wierenga, R. K. Three new crystal structures of point mutation variants of mono TIM: conformational flexibility of loop-1, loop-4 and loop-8. *Structure* **1995**, *3* (7), 669–679.

(46) Gershenson, A.; Schauerte, J. A.; Giver, L.; Arnold, F. H. Tryptophan Phosphorescence Study of Enzyme Flexibility and Unfolding in Laboratory-Evolved Thermostable Esterases. *Biochemistry* **2000**, *39*, 4658–4665.

(47) Muralidhara, B. K.; Chen, M.; Ma, J.; Wittung-Stafshede, P. Effect of Inorganic Phosphate on FMN Binding and Loop Flexibility in Desulfovibrio desulfuricans Apo-flavodoxin. *J. Mol. Biol.* **2005**, *349*, 89–97.

(48) Lu, P. J.; Zhou, X. Z.; Liou, Y. C.; Noel, J. P.; Lu, K. P. Critical role of WW domain phosphorylation in regulating phosphoserine binding activity and Pin1 function. *J. Biol. Chem.* **2002**, *277*, 2381–2384.

(49) Zarrinpar, A.; Lim, W. A. Converging on proline: the mechanism of WW domain peptide recognition. *Nat. Struct. Biol.* **2000**, *7*, 611–613.

(50) Wu, B.; F., R. M.; Wei, J.; Yuan, H.; Dahl, R.; Zhang, Z.; Pellecchia, M. Discovery and Binding Studies on a Series of Novel Pin1 Ligands. *Chem. Biol. Drug. Des.* **2009**, *73*, 369–379.

CT900190N