

Structural Analysis of the Protein Phosphatase 1 Docking Motif: Molecular Description of Binding Specificities Identifies Interacting Proteins

Heike Meiselbach,¹ Heinrich Sticht,¹ and Ralf Enz^{1,*}

¹Emil-Fischer-Zentrum

Institut für Biochemie

Friedrich-Alexander-Universität Erlangen-Nürnberg

Fahrstrasse 17

91054 Erlangen

Germany

Summary

The interplay between kinases and phosphatases represents a fundamental regulatory mechanism in biological systems. Being less numerous than kinases, phosphatases increase their diversity by the acquisition of a variety of binding partners, thereby forming a large number of holoenzymes. Proteins interacting with protein phosphatase 1 (PP1) often bind via a so-called docking motif to regulate its enzymatic activity, substrate specificity, and subcellular localization. Here, we systematically determined structural elements that mediate the binding specificity of PP1 interacting proteins, and propose a refined consensus sequence for high-affinity PP1 ligands. Applying this pattern to database searches, we predicted and experimentally confirmed several previously unknown PP1 interactors. Thus, the suggested PP1 docking motif enables a highly specific prediction of PP1 binding partners, thereby facilitating the genome-wide identification of PP1 interactors.

Introduction

Signaling processes in cells are controlled by a delicate balance of kinase and phosphatase activity. While the number of serine/threonine kinases is predicted to be in the range of 300, only about 20 serine/threonine phosphatases are known [1]. This underrepresentation is balanced by the acquisition of a large number of interaction partners that convert serine/threonine phosphatases into many different forms, which increases diversity and substrate specificity [2].

Protein phosphatase 1 (PP1) is a major serine/threonine phosphatase in eukaryotes that is ubiquitously expressed and regulates a large number of cellular activities, including neurotransmission, synaptic plasticity and memory, apoptosis, protein synthesis, muscle contraction, and the cell cycle [3]. The enzyme consists of one catalytic subunit (PP1 α , PP1 β/δ , or PP1 γ) that binds to over 50 different interacting subunits in a mutually exclusive manner, thereby forming a large number of holoenzymes [2]. The broad substrate specificity of PP1 leads to the idea that the enzymatic specificity is mainly dictated by these interacting subunits. PP1 interaction partners have diverse functions, such as targeting and anchoring of the catalytic subunit in specific subcellular compartments, regulating activity, or defin-

ing substrate specificity of the enzyme, and they may also serve as substrates themselves. Thus, identification and characterization of PP1 binding partners is crucial to understand the enzymatic function of this phosphatase.

Binding to PP1 catalytic subunits is largely mediated by a short sequence stretch referred to as a docking motif. It interacts, remote from the active site, with a conserved patch on the surface of PP1 catalytic subunits [4]. This docking motif might serve as an anchor for the initial binding of interaction partners to PP1 catalytic subunits, and thereby facilitate binding of additional secondary interaction sites that have lower affinity and are able to influence the enzymatic activity and substrate specificity of the phosphatase [5]. Ligand binding via the docking motif has only a minor effect on the PP1 structure itself, but might nevertheless modulate phosphatase activity.

Recently, we identified a conserved stretch of 5 amino acids in the intracellular C termini of the metabotropic glutamate receptors mGluR1a, mGluR5a, mGluR5b, and mGluR7b, binding to the catalytic isoforms PP1 γ 1 and PP1 γ 2 (KSV[ST]W [6]). Deletion of amino acids on both sides of this motif did not affect the interaction with PP1 γ 1, while deletion of the KSV[ST]W motif prevented binding, indicating that these amino acid sequences represent the only PP1 γ 1 binding sites within the mGluR C termini [6–8]. Although the overall architecture of this pattern is similar to predicted PP1 docking motifs of the form [KR][X]_{0–1}[V][P][FW] [9], we observed important differences: in PP1 γ binding metabotropic glutamate receptors, deletion of serine +2 or replacement of serine/threonine at position +4 with alanine prevented interaction with the enzyme [6]. Existing consensus motifs in the literature fail to predict these effects, instead showing that distinct sequence preferences exist in the PP1 docking motif, which are not specified by the current available patterns.

The ligand binding surface region of PP1 γ 1 exhibits a high excess of negative charges (D166, E167, D240, D242, E287) and contains two hydrophobic pockets, mainly formed by the side chains of residues I169, L243, F257, M283, L289, and F293 [4]. These different properties of the PP1 γ 1 surface determine the amino acid types tolerated at the different ligand positions. Analysis of the three-dimensional structure of the mGluR docking motifs in complex with PP1 γ 1 showed that the ligands bind in an extended backbone conformation to the PP1 γ 1 surface [6].

A detailed knowledge of the nature and binding affinities of amino acids allowed within the PP1 docking motif could facilitate the identification of proteins interacting with the phosphatases by selecting candidates from existing databases. Here, we used an integrative approach, combining structure analysis, molecular modeling, and molecular dynamics simulations that provide a detailed insight into the PP1-ligand interactions, in order to guide and supplement experimental studies. Our studies resulted in an experimentally tested consensus motif for PP1 interacting proteins, which then was

*Correspondence: ralf.enz@biochem.uni-erlangen.de

used to identify previously unknown PP1 binding partners from databases. Upon validating their PP1 binding capabilities, we found that our suggested amino acid pattern predicts PP1 interacting proteins more specifically than previous motifs.

Results

Sequence Preferences at the First Two Positions of the PP1 Docking Motif

To specify amino acids that are tolerated in the PP1 docking motifs, we used the C-terminal domain of mGluR7b as a model system. The individual contribution of each of the five positions of the mGluR7b wild-type sequence KSVTW in binding to PP1 γ 1 was investigated by replacing each amino acid by a range of residues covering different physical and chemical properties. The type of mutations tested was guided by the three-dimensional modeled structure of the KSVTW motif interacting with the PP1 γ 1 binding pocket [6].

Lysine (K911) at position +1 of the PP1 docking motif interacts with a negatively charged surface patch formed by residues D166, E167, and E287 of PP1 γ 1 (Figure 1A). Exchange of this residue with other basic amino acids (histidine, arginine) retained the binding between the glutamate receptor and the phosphatase (Figure 1B). Thus, position +1 shows a strict preference for charged basic residues, while polar amino acids carrying a side chain hydroxyl group (serine, threonine), acidic residues (glutamate), and hydrophobic side chains of different volumes (alanine, phenylalanine, leucine, methionine, proline) were not tolerated.

For serine (S912) at position +2, the modeled complex between PP1 γ 1 and the KSVTW docking motif did not show any specific side chain interactions, suggesting that any amino acid might be tolerated at this position [6]. To test this hypothesis, S912 was mutated into all naturally occurring amino acids. However, in contrast to the prediction, a very diverse picture emerged (Figure 1C). Very small or large hydrophobic side chains (glycine, leucine, isoleucine) were not tolerated, nor were aromatic residues (phenylalanine, tryptophan, tyrosine), negatively charged residues (aspartate, glutamate), or proline. In contrast, hydrophobic residues with intermediate side chain volumes (alanine, methionine, valine) were allowed, but showed a reduced binding affinity for PP1 γ 1. Higher binding strengths were observed for polar and uncharged side chains (cysteine, asparagine, glutamine, serine, threonine) and basic amino acids (histidine, arginine, lysine; Figure 1C). The observed specificity pattern is rather surprising, since neither serine +2 in the model of the KSVTW-PP1 γ 1 complex [6], nor arginine +2 in an RRVSF-PP1 γ 1 crystal structure [4], formed any specific electrostatic interaction with the enzyme, despite their location in the spatial vicinity of D242 of PP1 γ 1.

Therefore, we performed molecular dynamics simulations to gain insight into the dynamic and conformational variability of the interactions formed by the residues at positions +1 and +2 of the RRVSF and KSVTW motifs bound to PP1 γ 1 in water. Analysis of the RRVSF-PP1 γ 1 complex [4] showed that the arginines at positions +1 and +2 readily form numerous polar interactions (hydrogen bonds, salt bridges) with negatively

charged side chains and backbone carbonyl groups of PP1 γ 1 during the simulation. The types of polar interactions are shown in detail in Figures 2A and 2B, which depict the intermolecular contacts after 0.3 ns of simulation. Short donor-acceptor distances of less than 3.5 Å, indicative of the formation of hydrogen bonds and salt bridges, are found between the guanidino group of arginine +1 and the side chains of D166 and E167, as well as on the backbone carbonyl of E287. In addition, the guanidino group of arginine +2 forms salt bridges with the side chain of D242 of PP1 γ 1 (Figures 2A and 2B).

From this snapshot, it is evident that arginine at position +1 forms a larger number of polar interactions than arginine +2, suggesting a more dominant role for ligand binding. Therefore, we tested whether the interactions formed by arginine +1 are more stable than those formed by arginine +2 over the simulation time. Figure 2C shows that arginine +1 forms numerous stabilizing polar interactions with residues D166, E167, and E287 of PP1 γ 1. At each time point of the simulation, there are at least four polar interactions present, as indicated by short donor-acceptor distances of less than 3.5 Å (arrow in Figure 2C), showing that arginine +1 is efficiently tightened to the negatively charged cluster of the PP1 γ 1 surface.

Regarding the side chain of the adjacent arginine at position +2, polar interactions form with residues D240, L241, and D242 of PP1 γ 1 during the simulation (Figure 2D). While the interactions with D240 and L241 are formed with the uncharged carbonyl oxygens of the PP1 γ 1 backbone, the side chain of D242 forms a salt bridge with the arginine +2 side chain of the ligand. The fluctuations around 0.15–0.30 ns and 0.70–0.75 ns of the simulation reflected by donor-acceptor distances larger than 3.5 Å (Figure 2D) suggest that the corresponding interactions are less stable compared with those formed by arginine at position +1. This is in agreement with the stricter sequence preferences observed for position +1 in the binding studies described in Figure 1.

These differences in the relative importance of the first two ligand positions in PP1 γ 1 binding are further emphasized by simulations performed for the KSVTW-PP1 γ 1 complex. In this structure, lysine +1 forms similar salt bridges, like arginine at the same position in the RRVSF-PP1 γ 1 complex: namely, to the side chains of D166 and E167 (Figures 3A and 3B). Lysine +1 also forms contacts with the backbone carbonyl of E287; however this interaction was not present at 0.3 ns. In contrast to arginine +1 in the RRVSF motif, the fluctuations of the hydrogen bonds and salt bridges of lysine +1 are slightly larger (see Figure 3C, around 0.5 and 0.6 ns), which can most likely be attributed to differences in the ligand sequences, particularly at position +2.

Compared to arginine +2 of the RRVSF sequence, serine +2 of the KSVTW motif can only form a subset of the interactions (compare Figure 2B with 3B and Figure 2D with 3D). This is plausible, since serine has a polar but not a charged side chain, and thus cannot form salt bridges to D242. Instead, a hydrogen bond between the side chain hydroxyl group of serine +2 and the side chain carbonyl of D242 forms during the simulation, and is stable over 30% of the total simulation time (Figure 3D). Although this interaction can be considered to be weaker than that formed by arginine +2, this finding

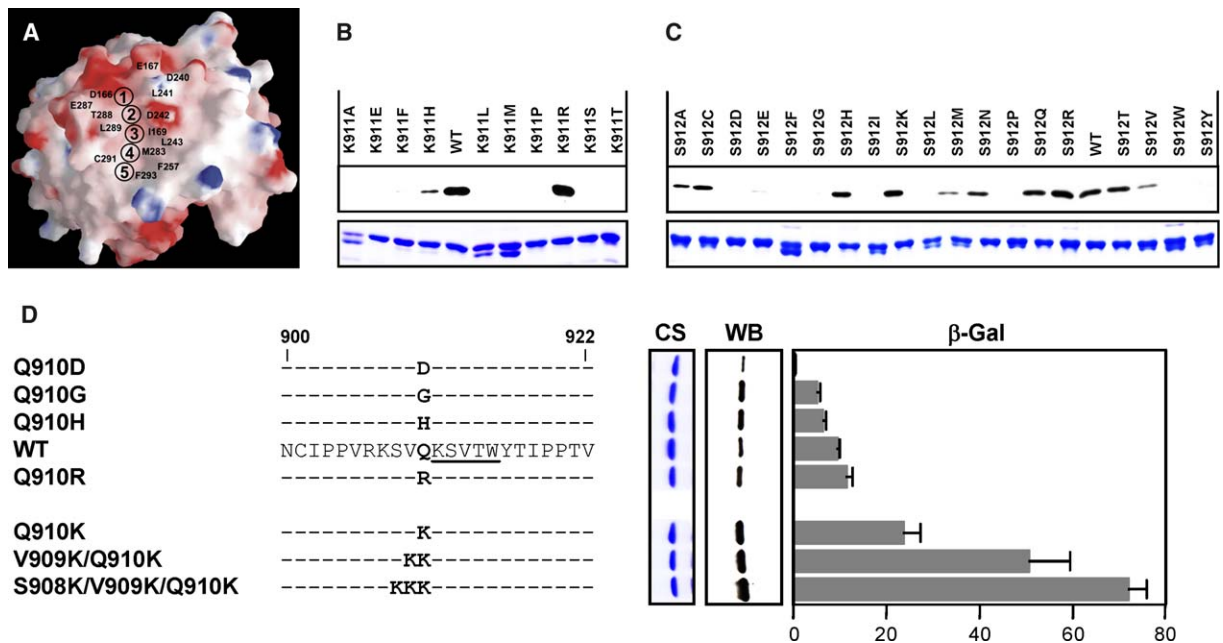


Figure 1. Mutational Analysis of the N-Terminal Amino Acids of the PP1 Docking Motif in mGluR7b

(A) Solvent-accessible surface area and electrostatic potential of PP1 γ 1, calculated based on the coordinates from [4]. Regions of positive and negative charges are colored in blue and red, and residues important for ligand binding are labeled with their sequence positions. The location of the five amino acids forming the PP1 docking motif in interacting proteins are indicated by circles.

(B and C) Mutations were introduced at position +1 (B) and position +2 (C) of the PP1 γ 1 docking motif KSVTW of mGluR7b, as indicated. Constructs were expressed as GST-fusion proteins, immobilized on glutathione Sepharose, and incubated with T7-tagged PP1 γ 1. Bound PP1 γ 1 was detected on Western blots using an anti-T7 immune serum (upper panels). Protein concentrations of coated Sepharose beads are shown on Coomassie-stained SDS-PAGE (lower panels).

(D) Amino acids located N-terminal to the KSVTW motif were mutated as indicated and resulting constructs were tested for binding to PP1 γ 1, as described above (CS, Coomassie stained SDS-PAGE; WB, Western blot). The C-terminal sequence of mGluR7b is shown for the wild-type (WT), and the PP1 docking motif is underlined. In vivo protein interactions were analyzed in yeast cells, and relative binding affinities were quantified and visualized as arbitrary β -galactosidase units (β -Gal, horizontal columns). Each value represents the mean of three yeast clones. Error bars are SEM.

clearly shows that a serine is also able to form stabilizing side chain interactions. Therefore, distinct sequence preferences exist for the residue at position +2 of the PP1 docking motif, consistent with our experiments described in Figure 1C.

The N-Terminally Adjacent Region of the PP1 Docking Motif Can Modulate Binding Affinity

Theoretical predictions assumed that amino acids located N-terminal to the PP1 docking motif could increase the binding affinity due to interactions with a negative electrostatic surface potential located on the enzyme surface (Figure 1A; [4, 10]). To test this prediction experimentally, we introduced a series of mutations at position -1 of the PP1 docking motif of mGluR7b, and analyzed the ability of the mutants to interact with the enzyme (Figure 1D). All introduced mutations were tolerated at this position, consistent with the presence of different amino acids at the corresponding location in more than 50 PP1 interacting proteins [2]. However, relative binding affinities varied largely, being lower for acidic side chains (aspartate) and higher for the basic residues histidine, arginine, and lysine. This can be seen best upon analyzing the binding strengths in vivo using yeast cells (right panel of Figure 1D). To test if the presence of additional basic amino acids would still increase ligand affinity, we introduced successive lysines N-terminal to

the PP1 docking motif, which resulted in higher binding affinities, most evident when the interactions were monitored in vivo (Figure 1D). Thus, we conclude that residues located N-terminal to the PP1 docking motif are able to modulate the binding affinity of ligands.

Preference for Valine at Position +3

The residue at position +3 of the PP1 docking motif packs into a hydrophobic pocket of the PP1 γ 1 surface (red circle in Figure 4A), meeting strict steric requirements, in contrast to positions +1 and +2. Mutation of valine (V913) at the third position of the PP1 docking motif into other hydrophobic residues containing smaller (alanine, glycine) or larger (leucine) side chains completely disrupted the binding (Figure 4B). Only the mutation of valine into isoleucine resulted in a weak interaction with PP1 γ 1. Exchanging one of the valine methyl groups with the hydroxyl group of threonine prevented binding, indicating that, due to the hydrophobic character of the pocket, polar amino acids are not tolerated at this position.

While several PP1 binding proteins contain an isoleucine at position +3 (e.g., Inhibitor-1 [11]), isoleucine showed an extremely weak affinity for PP1 γ 1 in the context of the mGluR7b C terminus (Figure 4B). Therefore, molecular modeling was used to investigate the steric consequences of a valine to isoleucine replacement in

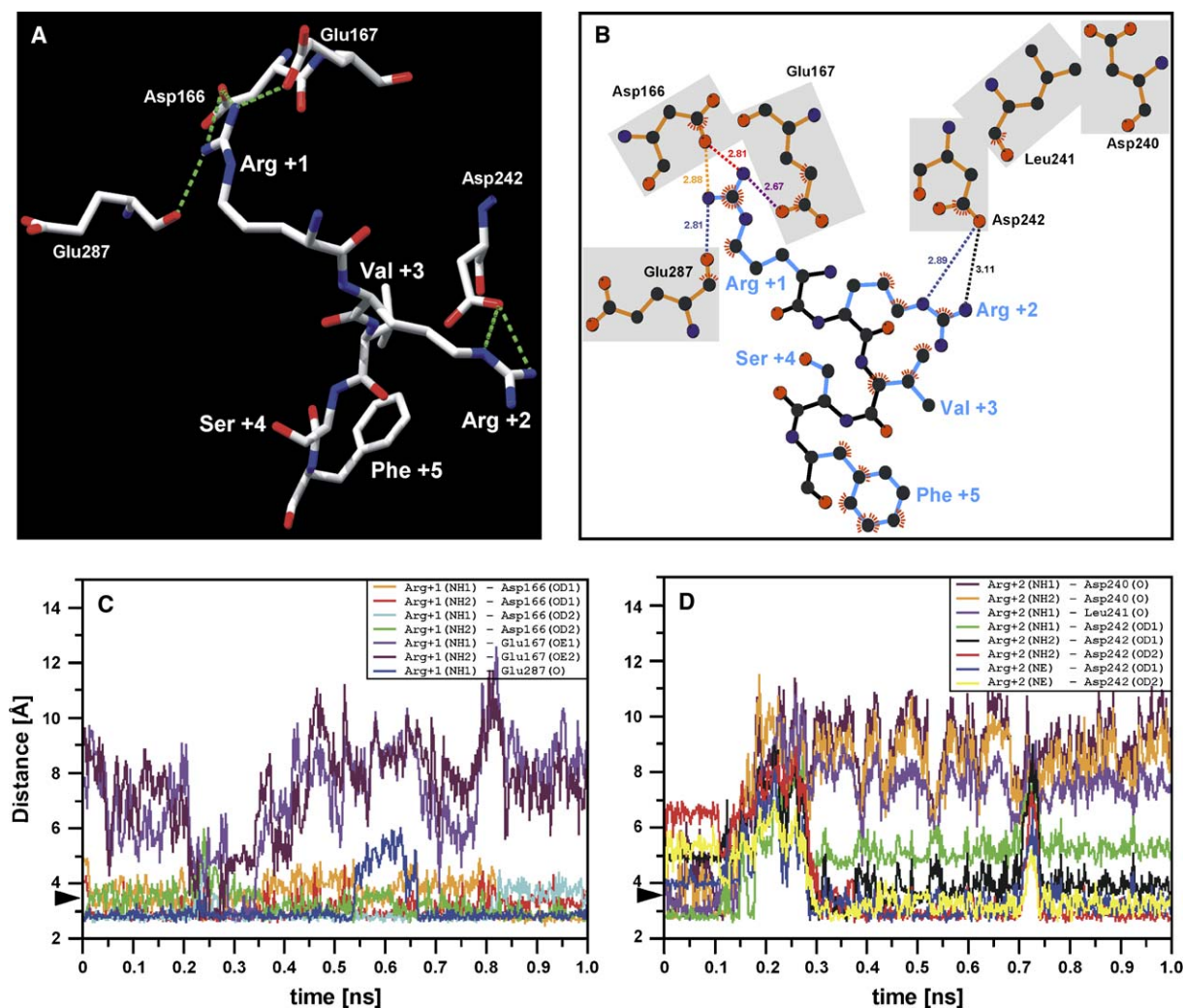


Figure 2. Interactions between PP1 γ 1 and the RRVSF Ligand in the Time Course of a Molecular Dynamics Simulation
 (A) Hydrogen bonds and salt bridges (green dotted lines) formed by arginine +1 and arginine +2 after 0.3 ns of MD simulation. Functional groups are colored according to their atom types, and are labeled with their sequence position, where a “+” denotes a residue of the ligand.
 (B) Two-dimensional amino acid representation specifying in detail the interactions shown in (A). The ligand is shown in ball-and-stick presentation with carbon, oxygen, and nitrogen atoms colored in black, red, and blue, respectively. Bonds of the protein backbone are shown in black, and those of the side chains in blue. Residues of PP1 γ 1 are marked by gray rectangles. For each of the polar interactions, the donor-acceptor distance is given in Å, and the color of the dotted lines was chosen to match the color coding used in (C) and (D). Hydrophobic interactions are indicated by red dashes.
 (C) Electrostatic interactions of arginine +1 as a function of simulation time. Interactions with different groups of PP1 γ 1 are colored according to the legend given in the inset. Distances of less than 3.5 Å (arrowhead) indicate the presence of a hydrogen bond or salt bridge. Those interactions that are present at 0.3 ns are shown in (A) and (B) in detail.
 (D) Electrostatic interactions of arginine +2 as a function of simulation time, visualized as in (C).

the RRVSF-PP1 γ 1 crystal structure. In order to allow a proper fit into the pocket, different side chain rotamers for isoleucine were tested individually. Even for the lowest-energy rotamer, some clashes were detected between two hydrogen atoms of the isoleucine +3 side chain and the side chains of I169 and L243 of PP1 γ 1 that flank the hydrophobic pocket (red arrows in Figure 4C). These clashes are better seen in a space-filled model of the ligand atoms (Figure 4D). Energy minimization shows that removal of these clashes forces the ligand backbone to adopt a less favorable geometry, offering a plausible explanation for the weaker binding strength of isoleucine compared with valine, which does not form such clashes (Figure 4E).

Interestingly, the two clashes are located at the side of the PP1 γ 1 hydrophobic pocket that is, in principle, sufficiently deep to accommodate an isoleucine. Indeed, an isoleucine is present in the docking motifs of three well-studied PP1 interacting proteins (Inhibitor-1, DARPP-32, and Neurabin I). These ligands contain two basic amino acids at the first two motif positions (Figure 4F), which represents a strong interaction sequence, as deduced from the experiments described above. To investigate whether these high-affinity interactions would be able to partially compensate for the presence of a less favorable isoleucine at position +3, we introduced the complete docking motifs of Inhibitor-1, DARPP-32, and Neurabin I in the corresponding location of the

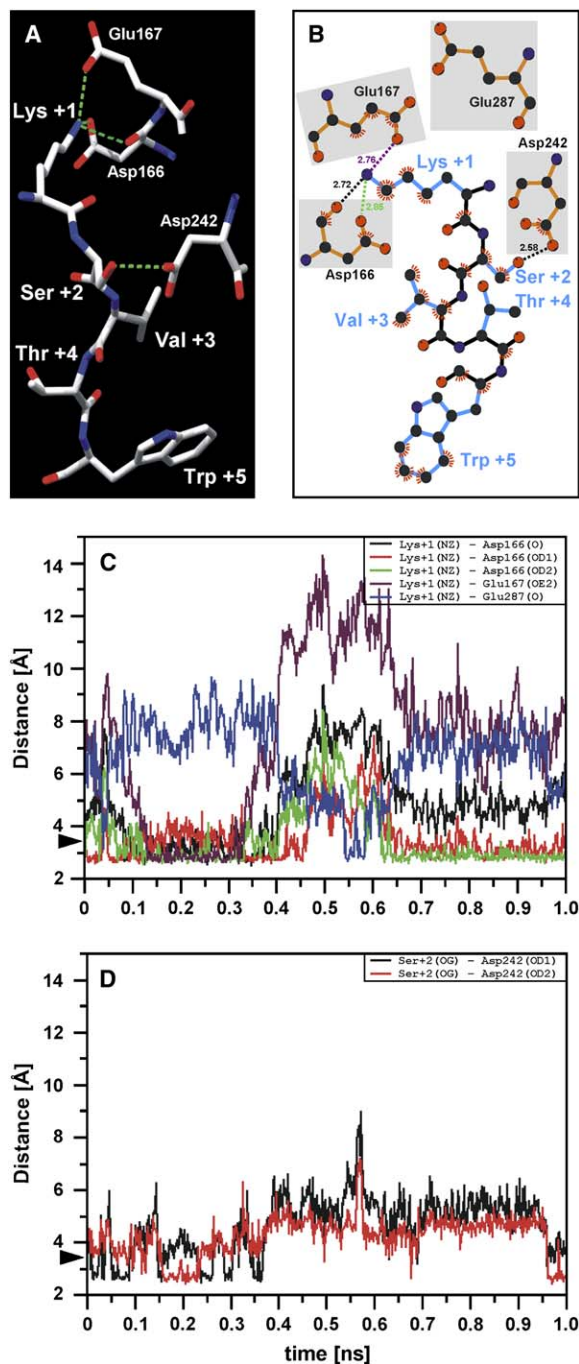


Figure 3. Contacts between PP1 γ 1 and the KSVTW Sequence during a Molecular Dynamics Simulation

Hydrogen bonds and salt bridges (green dotted lines) formed by lysine +1 and serine +2 after 0.3 ns of MD simulation are visualized in three dimensions (A), and are shown in more detail in a two-dimensional representation (B). Electrostatic interactions of lysine +1 (C) and serine +2 (D) as a function of simulation time. Only interactions present at 0.3 ns are visualized in (A) and (B). See Figure 2 for details.

mGluR7b C terminus. Analysis of their binding behavior revealed a weak affinity for PP1 γ 1 comparable to that of the mGluR7b-V913I mutant (Figure 4F), indicating that basic amino acids at positions +1 and +2 of the docking motif cannot compensate for the low binding affinity of isoleucine at position +3.

Amino Acids Tolerated at Position +4 Are Restricted to Basic and Small Polar Residues

To analyze which residues are compatible with position +4 of the PP1 γ 1 binding motif, we tested 10 amino acid types with different side chain properties (Figure 4G). The polar side chain of threonine +4 (T914) could be replaced by polar side chains of basic amino acids (histidine, lysine, arginine), which increased the binding strength. In contrast, acidic (glutamate) and hydrophobic (alanine, leucine, valine) residues were not tolerated. While threonine could be replaced by glutamine and serine [6], a mutation to tyrosine disrupted the interaction (Figure 4G), showing that the presence of a hydroxyl group alone is not sufficient for binding. Thus, position +4 reveals a preference for basic and uncharged polar aliphatic side chains, which can be explained by the solvent-exposed location of this position and the negative charge of the opposing PP1 γ 1 surface. Modeling studies suggest that the uncharged polar aliphatic side chains of cysteine and asparagine might also be tolerated at position +4 (data not shown), and were therefore also included in the pattern.

High Affinity for Unpolar Aromatic Amino Acids at Position +5

Empirically derived consensus sequences for the PP1 docking motif suggest that position +5 might tolerate phenylalanine, tryptophan, and tyrosine [5]. The fact that tryptophan is actually tolerated is evident from its presence in the mGluR7b wild-type sequence and from molecular modeling, which has shown that the hydrophobic pocket, indicated in Figure 4A by a yellow circle, is sufficiently large to accommodate tryptophan [6]. In contrast, nonaromatic amino acids result in a poor packing in the hydrophobic pocket, as investigated by an extended computational analysis (data not shown). For tyrosine, no significant steric clashes were detected, but the presence of a hydrophilic hydroxyl group that is deeply buried in the hydrophobic pocket being in close contact with the side chains of L243 and F257 of PP1 γ 1 (yellow arrow in Figure 4H) is expected to be energetically highly unfavorable. These predictions were verified by mutating tryptophan of the KSVTW sequence into alanine, phenylalanine, and tyrosine (Figure 4I). While the exchange of tryptophan with phenylalanine was tolerated, the presence of a hydrophobic nonaromatic alanine, or a more hydrophilic aromatic tyrosine, completely prevented binding to PP1 γ 1. Therefore, we conclude that hydrophobic aromatic side chains are required at this position.

Definition of a Consensus Sequence for the PP1 Docking Motif

Our results propose the degenerated consensus sequence [HKR][ACHKMNQRSTV][V][CHKNQRST][FW] as a high-affinity PP1 docking motif. This motif favors basic residues at positions +1 and +2, valine at position +3, polar amino acids, with amino or hydroxyl groups that are able to accept or donate protons, at position +4, and hydrophobic aromatic residues at position +5. In addition, uncharged residues are allowed at position +2 if their side chain volume is between that of alanine and methionine, with the exception of proline, which probably breaks the extended conformation of the protein

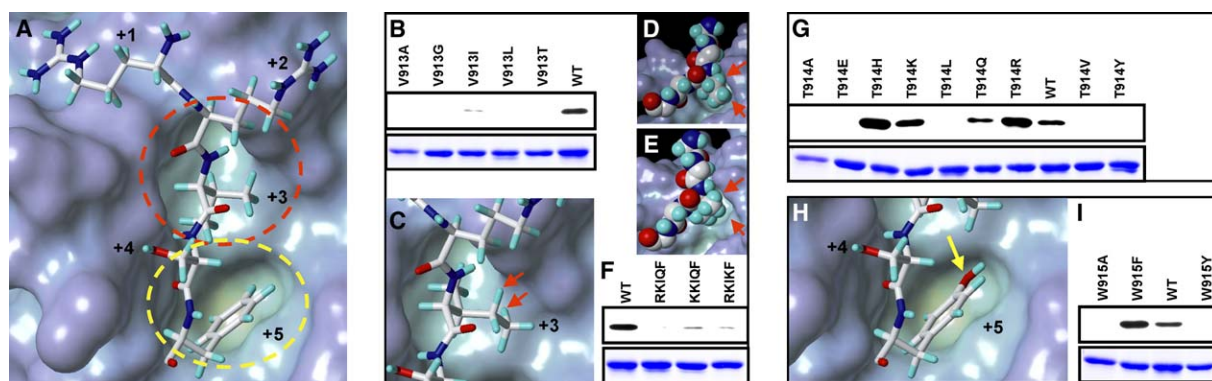


Figure 4. Amino Acid Preferences at Ligand Positions +3 to +5 of the PP1 Docking Motif

- (A) Spatial view of the two hydrophobic pockets (red and yellow circles) in the PP1 γ 1 crystal structure [4] that accommodate residues +3 and +5 of the G_M ligand sequence RRVSF, while residue +4 is located in a polar and partially solvent-exposed environment.
- (B) Mutations were introduced at position +3 of the KSVTW motif, and binding to PP1 γ 1 was analyzed as described in Figure 1. Bound PP1 γ 1 was detected on Western blots (upper panel), and the protein concentrations of coated Sepharose beads are shown on Coomassie stained SDS-PAGE (lower panel; WT, wild-type).
- (C) Detailed view of a modeled isoleucine present at position +3 using the crystal structure of PP1 γ 1 in complex with the RRVSF sequence as template [4]. The lowest-energy side chain rotamer for isoleucine +3 is shown in stick presentation, and the two clashes with residues I169 and L243 of PP1 γ 1 are indicated by red arrows.
- (D) The space-filled presentation of isoleucine +3 emphasizes the location of the clashes between ligand and enzyme (red arrows). For better visualization of the pocket depth, the view is rotated approximately 90° around the vertical axis with respect to (C).
- (E) Same view as in (D), but showing a valine instead of an isoleucine side chain at position +3. The red arrows denote the position of the clashes detected for isoleucine, which are absent for valine.
- (F) PP1 docking motifs of Inhibitor-1 (RKIQF), DARPP-32 (KKIQF), and Neurabin I (RKIKF) were introduced in the corresponding position of the mGluR7b C terminus and tested for their binding capabilities, as described above.
- (G) Threonine +4 of the KSVTW motif was exchanged with other amino acids as indicated, and binding was analyzed as described above.
- (H) Unfavorable location of the side chain hydroxyl group that deeply penetrates into the hydrophobic pocket (yellow arrow) in a modeled ligand containing a tyrosine at position +5.
- (I) Mutations were introduced at position +5 of the KSVTW sequence and binding of resulting mutants was analyzed as above.

backbone. The extremely low binding affinity of isoleucine at position +3 (Figure 4B) suggests that well-known PP1 ligands, like Inhibitor-1, DARPP-32, and Neurabin I, which carry an isoleucine at the respective location, might use additional interaction sites to compensate for the low affinity of their docking motifs (Figure 4F). Indeed, the idea that multiple PP1 binding sites act cooperatively has been demonstrated [12, 13]. To further test the role of isoleucine +3 in our consensus sequence proposed above, we analyzed the affinity of seven additional proteins of yet unknown PP1 binding properties. Each protein carries an isoleucine at position +3 in their potential PP1 docking motifs (cystatinB: KAISF; GABA_AR π : RKISF; moesin: RNISF; NPY2R: KRISF; ORP150: KVITF; ROMK1: KTITF; sodium pump 3: RNITF). Because none of these proteins interacted with PP1 γ 1 in glutathione-S-transferase (GST) pull-down experiments (data not shown), we classified isolated PP1 docking motifs carrying isoleucine at position +3 as noninteractors, and consequently did not include isoleucine at position +3 of our consensus sequence.

Sensitivity and Specificity of the Novel Consensus Sequence

The quality of the novel pattern for the PP1 docking motif was assessed by calculating its sensitivity and specificity, and by comparing these values to the performance of previous patterns. We calculated the sensitivity as the fraction of PP1 interactors identified by the patterns using a set of 56 experimentally tested PP1 ligands reported in the literature. Only those 38 proteins selected by the classical [KR][X]₀₋₁[V][P][FW] pattern

are listed in Table 1, while the other 18 known PP1 ligands use different binding mechanisms. A total of 18 of these 56 proteins were also selected by the newly defined [HKR][ACHKMNQRSTV][V][CHKNQRST][FW] pattern. Based on these numbers, the sensitivity of the classical and new patterns was calculated to be 0.68 and 0.32, stating that 68% or 32% of the known PP1 interactors were identified.

While this lower sensitivity might be considered a disadvantage of our novel consensus motif, this pattern should have the advantage of being more specific, leading to a reduction in the number of proteins erroneously predicted to bind PP1 (false-positive hits). Because the number of proteins that contain a PP1 docking motif, but do not interact with PP1, is unknown, we created our own test set consisting of proteins that contain the classical or new pattern and were not known a priori to bind to PP1.

To estimate the number of potential PP1 interactors that would be predicted by the novel [HKR][ACHKMNQRSTV][V][CHKNQRST][FW] pattern, we performed a UniProt-SwissProt database search, which resulted in 186 rat candidate proteins, which were further filtered by applying three criteria. First, we assumed that the regulation of PP1 would be conserved between mammals, implying that the PP1 docking motif should be present in at least three mammalian orthologs. Second, from the resulting 131 proteins, we disregarded candidates that contain the docking motif in extracellular domains, transmembrane regions, or signal peptides, leaving 89 hits. Third, proteins listed in a database for PP1 interactors [14] were omitted, because we intended to identify *new*

Table 1. Sensitivity of the Newly Suggested Consensus Sequence

Gene Name	Protein Accession No.	Motif	Selection
<i>PPP1R1A</i>	Q13522	NSP RKIQF TVP	–
<i>PPP1R1B</i>	Q9UD71	KDR KKIQF SVP	–
<i>PPP1R1C</i>	AAH17943	NSP KKIQF AVP	–
<i>PPP1R3A</i>	Q16821	SGT RRVSF ADS	+
<i>PPP1R3B</i>	AAH43388	KVK KRVSF ADN	+
<i>PPP1R3C</i>	CAH69995	QAK KRVSF ADS	–
<i>PPP1R3D</i>	O95685	ELG SRVHF AVR	–
<i>PPP1R3E</i>	BAB15779	DTR KRVSF ADA	+
<i>PPP1R3F</i>	NP_149992	VAP RRVLF ADE	–
<i>PPP1R3G</i>	XP_371796	KCK KRVSF ADT	+
<i>PPP1R8</i>	Q12972	RKN SRVSF SED	–
<i>PPP1R9A</i>	Q9ULJ8	PAN RKIKF SSA	–
<i>PPP1R9B</i>	NP_115984	APS RKIHF STA	–
<i>PPP1R10</i>	Q96QC0	RKR KSVTW PEE	+
<i>PPP1R11</i>	AAI04751	KPE KKVEW TSD	–
<i>PPP1R12A</i>	Q14974	RQK TKVKF DDG	–
<i>PPP1R12B</i>	O60237	RGS PRVSF EDG	–
<i>PPP1R12C</i>	NP_060077	RRA RTVSF ERA	+
<i>PPP1R13A</i>	AAH58918	AHG MRVKF NPL	–
<i>PPP1R13B</i>	Q96KQ4	GHG LRVRF NPL	–
<i>PPP1R14B</i>	Q96C90	GPG PRVSF QSP	–
<i>PPP1R14C</i>	Q8TAE6	GGG ARVSF QSP	–
<i>PPP1R14D</i>	Q9NXH3	NPC KKVHW ASG	+
<i>PPP1R15A</i>	CAG33540	LKA RRVRF SEK	+
<i>PPP1R15B</i>	CAI16570	VKR KKVTF LEE	+
<i>Ppp1r16a</i>	XP_343275	KPQ KQVHF PPS	+
<i>PPP1R16B</i>	Q96T49	GRR KVSF EAS	+
<i>AKAP1</i>	Q92667	SSP KGVLF SSK	–
<i>AKAP11</i>	Q9UKA4	HSG KKVQF AEA	+
<i>BCL2</i>	AAA51813	NWG RIVAF FEF	–
<i>GRM1</i>	Q13255	SNG KSVSW SEP	+
<i>GRM5</i>	P41594	SNG KSVTW AQN	+
<i>GRM7</i>	NP_870989	SVQ KSVTW YTI	+
<i>SLC12A2</i>	P55011	KGR FRVNF VDP	–
<i>NEK2</i>	P51955	VIK KKVHF SGE	+
<i>Pic1</i>	BAA08351	GRK KTVSF SSM	+
<i>PCDH7</i>	BAA25196	QPF RRVTF SVV	+
<i>SFPQ</i>	Q15233	GKQ LRVRF ACH	–

Known PP1 binding proteins and corresponding amino acid sequences selected by the [KR][X]₀₋₁[VI][P][FW] pattern. The five residues forming the docking motifs in rat or human sequences are shown in bold, and are extended by three N- and C-terminal residues. The upper part of the table shows PP1 interactors according to their official nomenclature, while additional PP1 binding proteins in the lower part of the table were taken from [2, 6]. In addition to the listed proteins, 18 PP1 interactors use binding sites other than the [KR][X]₀₋₁[VI][P][FW] pattern [2], and are not shown. Positive or negative selection of the docking motifs by the suggested [HKR][ACHKMQRSTV][V][CHKNQRST][FW] pattern is indicated by “+” and “–”.

PP1 ligands. From the remaining 84 putative PP1 interactors, 8 proteins were selected to test the binding capability of the respective docking motifs (Figure 5A). We applied the same filtering procedure to select four additional proteins that contain the classical [KR][X]₀₋₁[VI][P][FW] pattern, but not the novel motif.

Testing the PP1 binding properties of these proteins confirmed seven of them as PP1 interactors (Figure 5B). While both types of motifs succeeded in identifying these seven binders, the classical pattern produced five false-positive hits, which is a significantly higher number than the single false-positive hit produced by the novel pattern. This indicated that the new pattern is more

specific in identifying true PP1 interactors than the previously suggested motif.

Finally, we investigated whether the docking motifs of newly identified PP1 interactors were indeed responsible for the observed interactions, or if binding was mediated by additional sites within the ligand sequences. We used the mGluR7b C terminus, containing the KSVTW docking motif, to compete for the interaction of PP1 γ 1 with the newly identified PP1 ligands. First, PP1 γ 1 was bound to these ligands, which were immobilized on glutathione Sepharose (Figure 6A). These beads were then incubated with the His-tagged C terminus of mGluR7b, or with the non-PP1 γ 1 interacting His-mGluR7a [6] as a control. Analyses of the remaining amount of PP1 γ 1 bound to its immobilized ligands on Western blots showed that only mGluR7b, but not mGluR7a, displaced the phosphatase into the supernatant (Figure 6B). In addition, a peptide containing the motif KRVTW also competed with the newly identified PP1 ligands for binding to the enzyme (Figure 6C). Therefore, we conclude that the newly identified PP1 ligands use their docking motifs to bind to the phosphatase.

Discussion

Properties of the Novel Consensus Sequence for the PP1 Docking Motif

Cells use short linear peptide motifs for diverse processes, such as protein phosphorylation, protein targeting, and protein-protein interactions. In this study, we propose an experimentally tested consensus sequence [HKR][ACHKMQRSTV][V][CHKNQRST][FW] for proteins interacting with a conserved patch on the PP1 surface. Position +1 represents the key anchor point of the N-terminal part of the motif by forming a network of salt bridges and hydrogen bonds that proved to be stable in a simulation of the protein dynamics. The dominant role of the salt bridges explains why only positively charged amino acids are tolerated at this motif position. In contrast, the salt bridges formed by arginine +2 in the RRVSF motif are less stable than those formed by arginine +1, indicating that salt bridges at position +2 are less important. Indeed, these interactions can partly be replaced by hydrogen bonds formed by the polar but uncharged serine +2 in the KSVTW sequence. The different number and stability of the interactions formed during the molecular dynamics simulation gives a plausible explanation for arginine binding with higher affinity than serine, and for serine nevertheless exhibiting higher affinity than those amino acids without polar groups in their side chains. This difference in importance of the first two basic residues in the interaction motif is consistent with mutational studies showing that deletion of serine +2 in the KSVTW sequence, which moves an uncharged glutamine to position +1 and the positively charged lysine to the second position of the resulting QKVTW pattern, impedes PP1 γ 1 binding [6].

In contrast to the preference for positively charged and polar residues at the first two motif positions, position +3 favors the hydrophobic amino acid valine, and position +5 only tolerates the uncharged and aromatic side chains of phenylalanine and tryptophan. The sequence variability observed at position +4 is again larger, allowing basic and uncharged polar aliphatic

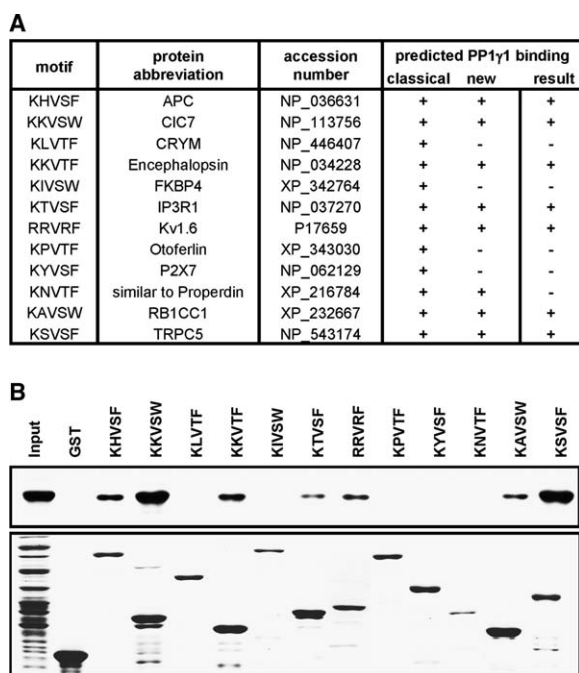


Figure 5. Specificity of the Novel Consensus Sequence and Identification of New PP1 Binding Partners

(A) Selected proteins from a database search predicted to be PP1 interactors (“+”) or noninteractors (“-”) based on the classical [KR][X]₀₋₁[VI][P][FW] or new [HKR][ACHKMNQRSTV][V][CHKNQRST][FW] consensus sequence for the PP1 docking motif.

(B) To evaluate these predictions, the PP1 γ 1 binding capability of soluble proteins or intracellular domains of transmembrane proteins was tested as described in Figure 1. Bound PP1 γ 1 was detected on Western blots (upper panel); protein concentrations of coated Sepharose beads are shown in the lower panel.

side chains. Thus, our analysis allows the definition of an experimentally validated docking motif for PP1 regulatory proteins.

In a recent study, the PP1 binding protein NIPP1 tolerated any amino acid, except proline, that was exchanged for the wild-type threonine at position +4 of its docking motif [9]. This finding conflicts with the PP1 binding characteristics of the glycogen-targeting subunit, G_M, as well as that of mGluR types, in which mutation of a serine or threonine at position +4 into alanine or valine prevented PP1 binding (this study and [6–8, 15]). The discrepancy might be explained by different methods being used to monitor the protein interactions (enzymatic activity and Western blots), or by the presence of phosphatase binding regions distinct from the docking motif that would compensate for a low-affinity sequence (see below). However, the presence of additional secondary binding sites was excluded for NIPP1, as well as for the metabotropic glutamate receptors [6, 9]. In the case of, for example, mGluR7b, an alanine scan covering the complete isoform-specific C-terminal region that was sufficient to interact with PP1 demonstrated that only those five amino acids forming the docking motif interact with the phosphatase [8]. Furthermore, deletion of amino acids on both sides of this motif did not affect the interaction, while deletion of the docking motif itself prevented binding. Finally, inser-

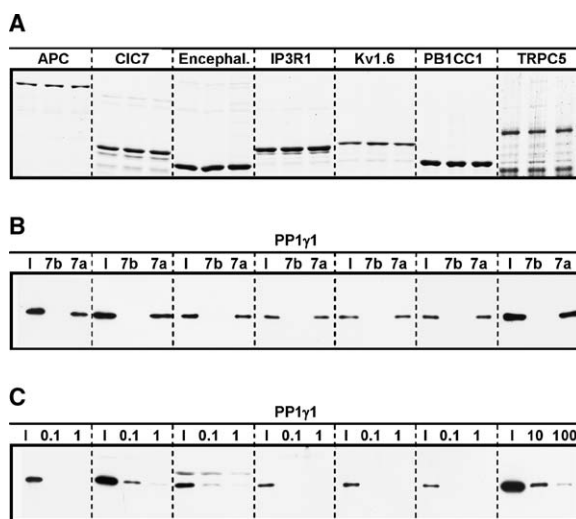


Figure 6. The Newly Identified PP1 Binding Partners Use Their Docking Motifs for Interaction with the Phosphatase

(A) PP1 γ 1 was bound to the seven newly identified ligands immobilized on glutathione Sepharose. Similar amounts of each type of coated beads, as analyzed on Coomassie-stained SDS-PAGE, were then incubated with His-mGluR7b (7b), containing a KSVTW motif, or with His-mGluR7a (7a) for control.

(B) Competition between PP1 γ 1 ligands and mGluR7b, but not mGluR7a, displaced PP1 γ 1 from the ligand-coated beads into the supernatant. Remaining amounts of PP1 γ 1 bound to the PP1 ligands were analyzed by pulling down the glutathione Sepharose, and subsequent detection of PP1 γ 1 on Western blots, as described in Figure 1.

(C) The coated beads shown in (A) were also used to compete for the binding between PP1 γ 1 and its ligands with increasing concentrations of a KRVTW motif containing peptide. Remaining amounts of bound PP1 γ 1 were analyzed as above. Numbers indicate the concentration of the peptide in μ M. (I, input; Encephal., encephalopsin).

tion of the mGluR7b docking motif in a corresponding position of the noninteracting mGluR7a isoform resulted in binding of this construct to PP1 [6]. Therefore, a more likely explanation for the observed differences could be the presence of three arginines and three lysines immediately N-terminal to the PP1 docking motif of NIPP1. N-terminal basic residues are known to dramatically increase the binding affinity of PP1 docking motifs (see Figure 1D) [4, 10], which could explain the high sequence variability tolerated at position +4 of the NIPP1 docking motif.

Predictive Power of the Novel Consensus Sequence

Our results revealed that the proposed [HKR][ACHKMNQRSTV][V][CHKNQRST][FW] pattern exhibits a lower sensitivity and a higher specificity compared to the previously suggested [KR][X]₀₋₁[VI][P][FW] pattern. Less stringent consensus sequences cover the majority off all PP1 interactors, but at the expense of producing more false-positive hits. Alternatively, our newly defined consensus sequence selects a subpopulation of PP1 binding proteins that is more enriched in true interactors.

A higher specificity is mainly advantageous for genome-wide searches that aim to identify a relatively small number of proteins suitable for subsequent experimental testing. Database searches revealed that our

novel pattern occurs approximately 8.5-times less frequently than the classical [KR][X]₀₋₁[VI][P][FW] motif in rat proteins (1578 hits by the classical pattern/186 hits by the novel pattern). In conjunction with the filtering procedures outlined in the results, the number of 186 proteins is reduced to 84 hits, which is a size suitable for experimental testing. Based on the specificity of the novel pattern assessed from the data in Figure 5, one can estimate that approximately 90% of these predicted interactors actually bind to PP1. In contrast, even after application of the filtering procedure, the classical pattern produces more than 700 hits in the rat proteome—far too many for experimental testing. In addition, this number is supposed to contain a large proportion of false-positive hits, as estimated from the binding studies shown in Figure 5.

The problem of false-positive hits can generally be further reduced by including additional filters that analyze the cellular location of the proteins and exclude regions not accessible for the phosphatase (e.g., globular domains, transmembrane, and extracellular regions). Such filters are already implemented in recent bioinformatics search tools, like the eukaryotic linear motif searcher (<http://elm.eu.org>). The benefit of such filters becomes evident from analysis of the only sequence mispredicted to bind PP1 by the novel pattern in our test set (Figure 5). In this protein, the sequence KNVTF is buried in a globular thrombospondin type 1 domain, and is therefore not available for PP1 interaction.

PP1 Binding Sites Not Yet Covered by Any Motif

Aside from the PP1 docking motif present in most interacting proteins, the existence of additional binding sites was described [5, 14]. In the recently determined crystal structure of PP1 δ in complex with the myosin phosphatase-targeting subunit, MYPT1, an additional interaction site is formed by a spatially distant region of the ligand [13]. In the PP1 docking motif of MYPT1 (TKVKF), threonine +1 and lysine +2 do not form specific polar side chain interactions, suggesting that these residues play no dominant role in binding. Indeed, the sequence TKVKF is not selected by our consensus sequence, suggesting that this motif alone is not capable of mediating the formation of a tight MYPT1-PP1 δ protein complex. Rather, it is suggested that the cooperative binding of this low-affinity docking motif, together with the additional interaction site observed in the MYPT1-PP1 δ co-crystal, increases the binding strength.

The nature of additional interaction sites is poorly characterized as yet, and cannot, therefore, be easily converted into a reliable profile for database searches. Instead, our proposed [HKR][ACHKMNQRSTV][V][CHK NQRST][FW] pattern was optimized for specificity, thus reducing the number of false-positive hits. Although this pattern cannot detect all PP1 interactors, proteins that use a PP1 docking motif as a high-affinity anchor will be predicted with a high specificity.

Significance

The correct targeting and localization of proteins to subcellular compartments represent an important biological mechanism for regulating cellular function.

Increasing evidence underlines the importance of macromolecular signaling complexes, where functionally related proteins are arranged in close proximity. Therefore, elucidating the molecular composition of these signaling complexes represents a fundamental step toward understanding the function of biological systems. Kinases and phosphatases are key players in the regulation of numerous cellular processes. PP1 is a major serine/threonine phosphatase, and regulates its enzymatic activity, substrate specificity, and subcellular localization by acquisition of different binding partners that mostly use a five amino acid long peptide sequence for PP1 binding. A detailed knowledge of the nature of amino acids allowed within this PP1 docking motif would greatly facilitate the identification of new PP1 ligands. Therefore, we systematically analyzed each individual position of the PP1 docking motif, and propose the new consensus motif [HKR][ACHKMNQRSTV][V][CHK NQRST][FW] for PP1 interactors. We also present a dynamic visualization of PP1 in complex with different ligand sequences. Our results show that position +1 represents the key anchor point of the motif by forming a network of salt bridges and hydrogen bonds. While positions +1, +2, and +4 prefer positively charged and/or polar residues, position +3 favors the hydrophobic amino acid valine, and position +5 only tolerates the unpolar and aromatic side chains of phenylalanine and tryptophan. The benefit derived from our consensus sequences is evident upon searching PP1 binding proteins. We identified 84 conserved mammalian proteins that potentially bind to the phosphatase, and could experimentally verify 7 previously unknown PP1 interactors out of 8 predicted candidates. Thus, the new pattern allows a specific prediction of PP1 binding partners, and will facilitate the genome-wide identification of such proteins.

Experimental Procedures

Molecular Biology

Mutations were introduced into the C-terminal domain of mGluR7b using standard PCR cloning techniques. cDNA sequences of proteins tested for interaction with PP1 γ 1 were amplified from rat brain cDNA libraries or from reverse-transcribed mRNA isolated from rat kidney (ROMK1) or rat lung (GABA_AR π), as previously described [7]. Resulting PCR products were fused to the coding sequence of GST in pET41 (Novagen, Madison, WI), or cloned in the yeast two-hybrid bait vector, pBTM116. The complete coding region of rat PP1 γ 1 was also generated by PCR, fused to the T7-epitope of pET21 (Novagen), or ligated in the yeast two-hybrid prey vector pVP16. All constructs were sequenced to check for PCR errors.

GST Pull-Down Assays

To test protein-protein interactions in vitro, pET21 and pET41 constructs were transformed in *Escherichia coli* BL21(DE3)pLysS, and protein expression was induced by adding 1 mM isopropyl- β -D-thiogalactoside (Sigma-Aldrich, Deisenhofen, Germany). Fusion proteins were purified under native conditions, immobilized to glutathione Sepharose beads, and incubated with the soluble fraction of *E. coli* expressing PP1 γ 1, as previously described [6–8]. Competition experiments were performed as previously described [8], using recombinantly expressed proteins and a peptide (SVQKRVTYTI; Coring, Gernsheim, Germany) with blocked end groups. A control peptide (SVQLVTWYTI) was not soluble in aqueous solutions. To obtain comparable conditions in competition experiments, *E. coli* protein extracts of similar protein concentrations, as measured at 280 nm, were used. Bound proteins were separated on SDS-PAGE

and visualized with Coomassie Brilliant Blue R-250 (Serva, Heidelberg, Germany), or detected by Western blotting using a monoclonal anti-T7 antibody and the enhanced chemiluminescence system (ECL; Amersham Biosciences, Freiburg, Germany). Unless otherwise stated, all reagents were purchased from Novagen.

Yeast Two-Hybrid Techniques

To test protein-protein interactions *in vivo*, the L40 yeast strain (Invitrogen, Karlsruhe, Germany) was used. In brief, yeast strains expressing PP1 γ 1 were transformed with bait constructs, and protein interactions were monitored by the activation of *His3* and *LacZ* reporter genes on selection plates, as previously described [7]. Binding affinities were calculated according to the "Yeast Protocols Handbook" (Clontech) using *o*-nitrophenyl- β -D-galactopyranoside (Sigma-Aldrich) as substrate.

Computational Techniques

The crystal structure of PP1 γ 1 in complex with the RRVSF peptide [4] was used as template to model PP1 γ 1 binding sequences investigated in this study. Models were generated by substituting side chains using the Sybyl 6.5 [16] and InsightII [17] program packages, as described previously [6].

Molecular dynamics (MD) calculations were done with the AMBER 7 program [18] using the parm99 force field [19, 20] and the TIP3P water model [21]. Simulations were performed in a periodic box filled with water molecules with at least 10 Å of solvent around every atom of the solute. An appropriate number of counter ions were added to neutralize the charges of the systems, and the particle mesh Ewald summation method [22] was employed to calculate the long-range electrostatic interactions.

All structures were minimized in a three-step procedure by using the SANDER module of AMBER. In the first step, the solvent was allowed to relax while restraining the protein atoms in their original position with a force constant of 500 kcal mol⁻¹ Å⁻². Afterwards, additional relaxation of the protein side chains was allowed by restraining only the backbone atoms. In the last step, all restraints were removed. The minimization consisted of 250 steps of steepest descent, followed by 250 or 7250 steps of conjugated gradient minimization.

MD simulations were performed using the SHAKE procedure [23] to constrain all bonds involving hydrogen atoms. The integration time step of the simulation was 2 fs and an 8.5 Å cutoff was used for the nonbonded interactions, which were updated every 15 steps. The temperature of the system was raised gradually from 50 to 298 K, followed by 20 ps of dynamics at 298 K. In order to increase the accuracy of the modeled structure, the equilibration period was extended by 1 ns for this system. A subsequent 1 ns MD simulation was performed for data collection, and 500 snapshots were saved for the analysis.

Protein-ligand contacts were analyzed with LIGPLOT [24] using default settings. The detailed hydrogen bond analysis, as a function of the simulation time, was restricted to the ligand side chains, since backbone interactions cannot explain sequence preferences of the ligand. For visualization and structural analysis, the programs Sybyl [16], MolMol [25], Grasp [26], and Deep View [27] were used.

Acknowledgments

We thank Nadja Schröder for excellent technical assistance and Jens Wilke and Pamela Strissel for critically reading the manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft to R.E. (SFB 539) and H.S. (SFB 473).

Received: July 12, 2005

Revised: October 6, 2005

Accepted: October 20, 2005

Published: January 20, 2006

References

1. Ceulemans, H., Stalmans, W., and Bollen, M. (2002). Regulator-driven functional diversification of protein phosphatase-1 in eukaryotic evolution. *BioEssays* 24, 371–381.
2. Cohen, P.T. (2002). Protein phosphatase 1: targeted in many directions. *J. Cell Sci.* 115, 241–256.
3. Ceulemans, H., and Bollen, M. (2004). Functional diversity of protein phosphatase-1, a cellular economizer and reset button. *Physiol. Rev.* 84, 1–39.
4. Egloff, M.P., Johnson, D.F., Moorhead, G., Cohen, P.T., Cohen, P., and Barford, D. (1997). Structural basis for the recognition of regulatory subunits by the catalytic subunit of protein phosphatase 1. *EMBO J.* 16, 1876–1887.
5. Bollen, M. (2001). Combinatorial control of protein phosphatase-1. *Trends Biochem. Sci.* 26, 426–431.
6. Croci, C., Sticht, H., Brandstätter, J.H., and Enz, R. (2003). Group I metabotropic glutamate receptors bind to protein phosphatase 1C. Mapping and modeling of interacting sequences. *J. Biol. Chem.* 278, 50682–50690.
7. Enz, R. (2002). The metabotropic glutamate receptor mGluR7b binds to the catalytic gamma-subunit of protein phosphatase 1. *J. Neurochem.* 81, 1130–1140.
8. Enz, R., and Croci, C. (2003). Different binding motifs in metabotropic glutamate receptor type 7b for filamin A, protein phosphatase 1C, protein interacting with protein kinase C (PICK) 1 and syntenin allow the formation of multimeric protein complexes. *Biochem. J.* 372, 183–191.
9. Wakula, P., Beullens, M., Ceulemans, H., Stalmans, W., and Bollen, M. (2003). Degeneracy and function of the ubiquitous RVXF motif that mediates binding to protein phosphatase-1. *J. Biol. Chem.* 278, 18817–18823.
10. Zhao, S., and Lee, E.Y. (1997). A protein phosphatase-1-binding motif identified by the panning of a random peptide display library. *J. Biol. Chem.* 272, 28368–28372.
11. Huang, F.L., and Gliismann, W.H. (1976). Separation and characterization of two phosphorylase phosphatase inhibitors from rabbit skeletal muscle. *Eur. J. Biochem.* 70, 419–426.
12. Beullens, M., Vulsteke, V., Van Eynde, A., Jagiello, I., Stalmans, W., and Bollen, M. (2000). The C-terminus of NIPP1 (nuclear inhibitor of protein phosphatase-1) contains a novel binding site for protein phosphatase-1 that is controlled by tyrosine phosphorylation and RNA binding. *Biochem. J.* 352, 651–658.
13. Terrak, M., Kerff, F., Langsetmo, K., Tao, T., and Dominguez, R. (2004). Structural basis of protein phosphatase 1 regulation. *Nature* 429, 780–784.
14. Garcia, A., Cayla, X., Caudron, B., Deveaud, E., Roncal, F., and Rebollo, A. (2004). New insights in protein phosphorylation: a signature for protein phosphatase 1 interacting proteins. *C. R. Biol.* 327, 93–97.
15. Liu, J., Wu, J., Oliver, C., Shenolikar, S., and Brautigan, D.L. (2000). Mutations of the serine phosphorylated in the protein phosphatase-1-binding motif in the skeletal muscle glycogen-targeting subunit. *Biochem. J.* 346, 77–82.
16. Tripos (2002). Sybyl 6.9, Release 7.0A (computer program). Tripos, St. Louis, MO.
17. Accelrys (2000). Insight II, First Edition (computer program). Accelrys, San Diego, CA.
18. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham, T.E., 3rd, DeBolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* 91, 1–41.
19. Cheatham, T.E., 3rd, Cieplak, P., and Kollman, P.A. (1999). A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* 16, 845–862.
20. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M.J., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995). A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.* 117, 5179–5197.
21. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935.

22. Darden, T.A., York, D.M., and Pedersen, L.G. (1993). Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.
23. Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341.
24. Wallace, A.C., Laskowski, R.A., and Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134.
25. Koradi, R., Billeter, M., and Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.
26. Nichollas, A., Sharp, K., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.
27. Guex, N., and Peitsch, M.C. (1997). SWISS-MODEL, and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723.