Cell, Vol. 93, 617–625, May 15, 1998, Copyright ©1998 by Cell Press

Crystal Structure of the Catalytic Domain of the Human Cell Cycle Control Phosphatase, Cdc25A

Eric B. Fauman,[1,6] John P. Cogswell,[2] Brett Lovejoy,[3,7] Warren J. Rocque,[4] William Holmes,[4] Valerie G. Montana,[3] Helen Piwnica-Worms,[5] Martin J. Rink,[4] and Mark A. Saper[1,6]

[1]Biophysics Research Division
and the Department of Biological Chemistry
The University of Michigan
Ann Arbor, Michigan 48109-1055
[2]Department of Functional Genetics
[3]Department of Structural Chemistry
[4]Department of Molecular Sciences
Glaxo Wellcome, Inc.
Research Triangle Park, North Carolina 27709
[5]Howard Hughes Medical Institute
Department of Cell Biology and Physiology
Washington University School of Medicine
St. Louis, Missouri 63110

Summary

Cdc25 phosphatases activate the cell division kinases throughout the cell cycle. The 2.3 Å structure of the human Cdc25A catalytic domain reveals a small α/β domain with a fold unlike previously described phosphatase structures but identical to rhodanese, a sulfur-transfer protein. Only the active-site loop, containing the Cys-$(X)_5$-Arg motif, shows similarity to the tyrosine phosphatases. In some crystals, the catalytic Cys-430 forms a disulfide bond with the invariant Cys-384, suggesting that Cdc25 may be self-inhibited during oxidative stress. Asp-383, previously proposed to be the general acid, instead serves a structural role, forming a conserved buried salt-bridge. We propose that Glu-431 may act as a general acid. Structure-based alignments suggest that the noncatalytic domain of the MAP kinase phosphatases will share this topology, as will ACR2, a eukaryotic arsenical resistance protein.

Introduction

The cyclin-dependent serine/threonine protein kinases (Cdks) ensure an orderly progression through the eukaryotic cell cycle (Morgan, 1995; Fisher, 1997). They are positively regulated through association with cyclin proteins and phosphorylation at a conserved threonine, Thr-161. They are kept inactive until needed by phosphorylation at position Tyr-15 in yeast or both Tyr-15 and Thr-14 in higher eukaryotes (Gould and Nurse, 1989; Krek and Nigg, 1991). Final activation of the kinase depends on the dual-specificity phosphatase Cdc25, which dephosphorylates both Tyr-15 and Thr-14 (Strausfeld et al., 1991; Honda et al., 1993).

[6]To whom correspondence should be addressed.
[7]Present address: Roche Bioscience, Molecular Structure Department, Palo Alto, California 94304-1397.

The human genome encodes at least three Cdc25 enzymes, identified as A, B, and C (Galaktionov and Beach, 1991), which appear to have specificity for different Cdk/cyclin complexes. Cdc25C dephosphorylates Cdk2/cyclin B, which in turn phosphorylates and activates Cdc25C, creating a positive feedback loop regulating entry into M phase (Hoffmann et al., 1993). Cdc25A is expressed during the G1→S transition and is presumed to promote entry into S phase by acting on Cdk2/cyclin A and Cdk2/cyclin E (Hoffmann et al., 1994; Jinno et al., 1994). Cdc25A is a substrate for Cdk2/cyclin E, suggesting the involvement of another positive feedback loop at this stage of the cell cycle.

A variety of protein tyrosine and dual-specificity (tyrosine and serine/threonine) phosphatases have been identified (Fauman and Saper, 1996). All employ a similar two-step hydrolysis reaction proceeding through a phosphocysteine intermediate, unlike the serine/threonine phosphatases, which use a one-step, metal-dependent hydration. The vaccinia virus H1 phosphatase (VH1) was the first phosphatase reported to display dual specificity (Guan et al., 1991). The structure of VHR, a human homolog of VH1, is very similar to the structure of the tyrosine-specific phosphatases (e.g., *Yersinia enterocolitica* YopH, human PTP1B, and PTPα), and there are conserved residues throughout their sequences (Figure 1) (Yuvaniyama et al., 1996). As deduced from these four crystal structures, the phosphate moiety of the substrate is cradled in an active-site loop formed by the residues of the conserved His-Cys-$(X)_5$-Arg sequence. A conserved aspartic acid from an adjacent loop facilitates catalysis by protonating the leaving group. The low molecular weight phosphatases, while distinct in sequence and structure, also possess an active-site Cys-$(X)_5$-Arg sequence and catalytic aspartic acid (Su et al., 1994; Zhang et al., 1994a).

Aside from the sequence His-Cys-$(X)_5$-Arg, Cdc25 has no sequence homology with the catalytic domains of the other protein tyrosine or dual-specificity phosphatases (Strausfeld et al., 1991). Two earlier, and widely cited, reports suggested more extensive sequence similarity between *Schizosaccharomyces pombe* Cdc25 and VH1 (Gautier et al., 1991; Moreno and Nurse, 1991). However, as additional sequences for Cdc25s and VH1-like dual-specificity enzymes have become available, it has become apparent that Cdc25 is a distinct family. Of the 19 residues found in common between the two original proteins, only the 3 from His-Cys-$(X)_5$-Arg are conserved in both families.

Despite limited sequence similarity, indications are that Cdc25 uses the same catalytic mechanism as the other protein tyrosine and dual-specificity phosphatases. Mutation of either the cysteine or the arginine of the Cys-$(X)_5$-Arg motif inactivates the enzyme (Gautier et al., 1991). Further, Cdc25 exhibits a two-step mechanism (Gottlin et al., 1996), presumably proceeding through a phosphocysteine intermediate, as seen in the other Cys-$(X)_5$-Arg phosphatases (Guan and Dixon, 1991).

VHR and VH1 share sequence similarities with the catalytic domains of the MAP kinase phosphatases

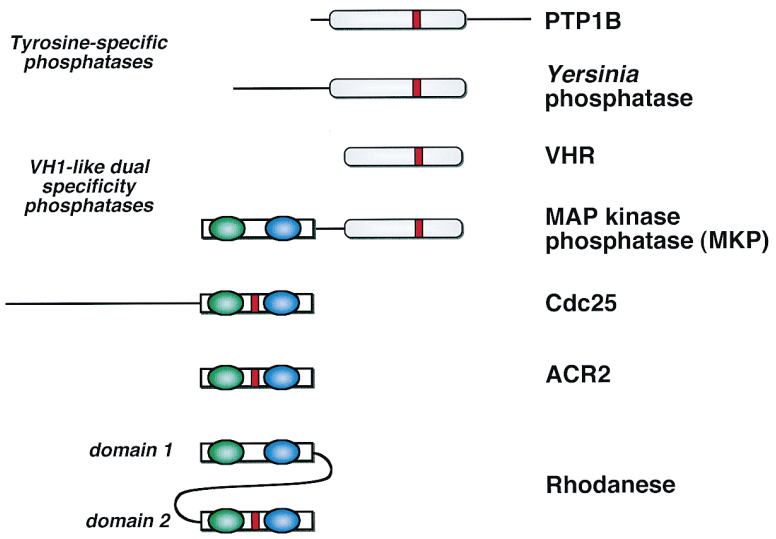

Figure 1. Domain Organization of Proteins Structurally and Functionally Related to Cdc25

Domains for MKP and ACR2 are defined from sequence comparisons alone; the rest all include results of crystal structure determinations. Gray, rounded rectangle represents the prototypical protein tyrosine phosphatase domain. White rectangle represents a Cdc25-like domain, including CH2A (green oval) and CH2B (blue oval) motifs. Small, red rectangle indicates position of a Cys-$(X)_5$-Arg motif or, in the case of rhodanese, the isolated catalytic cysteine.

(MKPs), such as PAC1 and CL100 (Figure 1). Surprisingly, two regions of homology (termed CH2A and CH2B) were identified between sequences in the N-terminal, noncatalytic domain of the MKPs and sequences flanking the active site of Cdc25 (Figure 1) (Keyse and Ginsburg, 1993).

The human Cdc25s are 470–530 residues in length and are composed of two domains: a variable N-terminal region, suspected to serve regulatory functions (Peng et al., 1997), and a highly conserved C-terminal catalytic domain, containing the CH2A, Cys-$(X)_5$-Arg, and CH2B motifs. Deletion experiments localized the catalytic activity to the last 175 residues of human Cdc25B (Xu and Burke, 1996). This abbreviated construct binds Cdk2/cyclin A as well as the full-length Cdc25B.

To elucidate the relationship between the Cdc25s and the other protein tyrosine phosphatases, the functions of the CH2 domains, and the substrate-specificity determinants of Cdc25, we determined the crystal structure of the catalytic domain of human Cdc25A.

Results and Discussion

The Overall Structure

We determined the X-ray crystal structure of the catalytic domain of human Cdc25A (residues 336–523) by

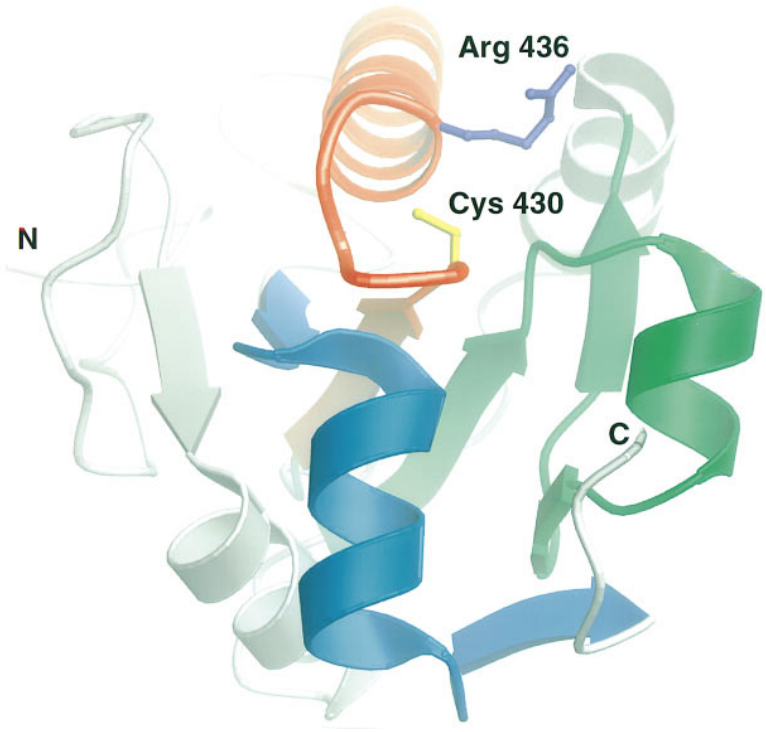

Figure 2. Ribbon Drawing of the Catalytic Domain of Human Cdc25A

The CH2A, active site, and CH2B motifs are indicated in green, red, and blue, respectively. The side chain of the Cys-430 nucleophile is shown in yellow. The N terminus (residue 335) is indicated on the left (N), while the C terminus extends toward the viewer and is clipped in this figure. Figure created with MOLSCRIPT (Kraulis, 1991) and the ray-tracing program VORT (University of Melbourne).

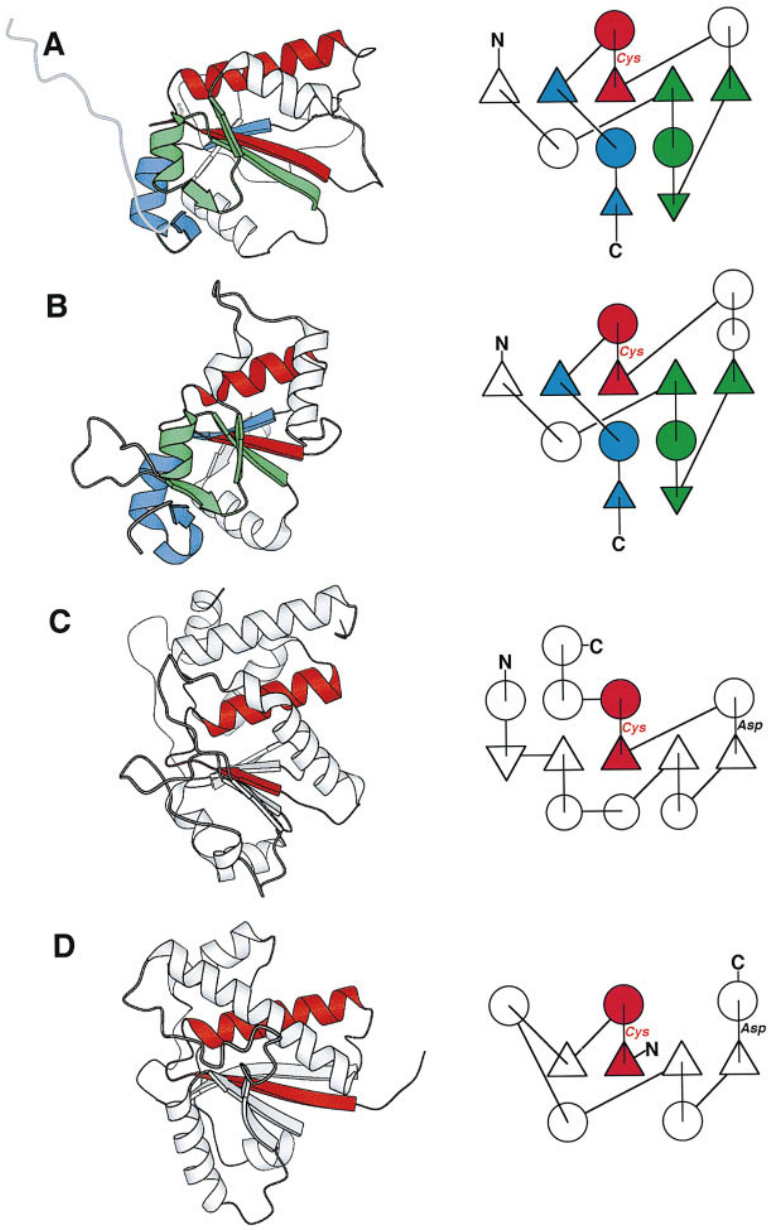

Figure 3. Topological Comparison of Cdc25 and Related Proteins

Depicted are (A) Cdc25A, (B) rhodanese (PDB entry 2ORA; Gliubich et al., 1996), (C) the VH1-like dual specificity phosphatase, VHR (PDB entry 1VHR; Yuvaniyama et al., 1996), and (D) the bovine low molecular weight phosphatase (PDB entry 1PNT; Zhang et al., 1994a).

While all the proteins have an α/β fold consisting of a β sheet surrounded by α helices, only Cdc25 and rhodanese are topologically equivalent. The CH2A, active site, and CH2B motifs are indicated in green, red, and blue, respectively. Helices are shown as circles; strands, as triangles. N and C indicate the N and C termini of each domain. Red "Cys" indicates position of the nucleophilic cysteine, while "Asp" indicates the general acid in VHR and the bovine low molecular weight phosphatase. Ribbon drawings created with MOLSCRIPT (Kraulis, 1991).

the multiple isomorphous replacement method (Table 1) and refined the model to a crystallographic R-factor of 22.7% at a resolution of 2.3 Å.

The topology of the catalytic domain is unlike that of any other phosphatase (Figures 2 and 3). Cdc25A is a small α/β domain with a central five-stranded parallel β sheet in the order 15423. Three helices lie below the sheet (as viewed in Figure 2), and two longer helices lie over the sheet. The dual-specificity phosphatases and the core of the tyrosine-specific phosphatases are also α/β structures of similar size but possess a mixed, five-stranded β sheet in the order 12534. The low molecular weight phosphatases have a four-stranded parallel β sheet in the order 2134.

Surprisingly, the topology of Cdc25A is identical to that of both halves of the sulfur transfer protein, rhodanese (Figures 3 and 4). Rhodanese is highly conserved and is found in bacteria and mitochondria, although the exact biological role is still unknown. Rhodanese is composed of two domains of identical topology, but low sequence similarity, believed to be the result of a gene duplication (Ploegman et al., 1978). Only the second domain contains the active site with a nucleophilic cysteine. We previously noted that rhodanese shared with the *Yersinia* phosphatase an unusual loop conformation that directs eight sequential main chain nitrogens in toward an anion-binding site (Stuckey et al., 1994), a structure which has now been seen in all protein tyrosine phosphatases. However, while the analogy between rhodanese and the other tyrosine phosphatases

encompasses only the active-site strand-loop-helix region, the similarity between Cdc25A and each half of rhodanese includes all the secondary structures.

A comparison with a database of representative protein folds (Holm and Sander, 1994) confirmed that rhodanese is the only other known structure to share Cdc25A's topology, with a Z score of 10.3 (standard deviations above the mean), a root-mean-square deviation (rmsd) of 2.8 Å over 109 α carbon positions and 12 sequence identities. Eighteen unrelated α/β proteins have Z scores of between 3.5 and 2.3. The *Yersinia* phosphatase and VHR each receive Z scores no higher than 2.3 (rmsd more than 4.3 Å over about 80 α carbons, 8 or fewer sequence identities). For comparison, the *Yersinia* phosphatase and VHR structures share a Z score of 12.9, with an rmsd of 3.1 Å over 160 α carbons and 25 identities, and the two halves of rhodanese score 13.6 with an rmsd of 2.1 Å over 119 α carbons and 15 identities.

While the low molecular weight phosphatase structure is even more dissimilar (Z score = 1.9), it can be made more similar through a simple circular permutation. The Cys-$(X)_5$-Arg motif in the low molecular weight phosphatase occurs near the N terminus (Figure 3D). A hypothetical structure, in which the N and C termini of the low molecular phosphatase are fused and new termini are introduced between the second strand and helix, receives a Z score of 5.5 (rmsd 2.9 over 94 α carbons, 14 sequence identities) when compared to the Cdc25A structure. Although this circular permutation aligns the active-site Cys-$(X)_5$-Arg motifs, it is doubtful that this cryptic topological similarity implies any evolutionary relationship, since more extensive sequence similarities are absent.

The Active Site

In common with the other Cys-$(X)_5$-Arg phosphatases, the Cys-$(X)_5$-Arg motif in Cdc25A lies on a loop between a β strand and a helix (red, in Figures 2, 3, and 4). The catalytic cysteine, Cys-430, is the last residue of the strand and projects into the center of the loop (Figure 5). The Glu-431 side chain points above the loop. Phe-432 and Glu-435 point away from the loop, while Ser-433 and Ser-434 point into the loop. The Oγ of Ser-433 is within hydrogen bonding distance of the Sγ of Cys-430. In this position, it could stabilize the thiolate ion of Cys-430, assisting in the second step of catalysis in a manner analogous to Thr-410 of the *Yersinia* phosphatase or Ser-131 of VHR (Denu and Dixon, 1995). Arg-436 sits far above the loop and does not appear to be well positioned to assist in catalysis. The side chain of His-429 forms a hydrogen bond with the carbonyl of Cys-430, supporting the strained main chain dihedral angles for that residue.

In contrast to the other Cys-$(X)_5$-Arg phosphatases, the active site of Cdc25 is extremely shallow (Figure 5). There are no auxiliary loops, which extend over the active site, as there are, for example, in the *Yersinia* phosphatase or VHR. The active-site Phe-432 may help to recognize tyrosine substrates, analogous to Tyr-128 of VHR or the conserved Phe-229 of the *Yersinia* phosphatase. Recognition of the Cdk substrate probably relies

Table 1. Summary of X-Ray Crystal Structure Determination

Crystal[a]	Resolution (Å)	Observations	Unique Reflections	R_{sym}[b] (%)	$<I/\sigma_I>$	Completeness (%)	Phasing Resolution (Å)	R_{cullis}[c]	PP[d]	Number of Sites
Native	2.1	39,374	12,551	6.7 (37.2)	18 (3)	98.3 (100)	—	—	—	—
PCMP	2.2	34,271	10,714	11.5 (53.2)	13 (2)	97.0 (96)	2.2	0.58	2.58	5 Hg
SMET	2.0	44,839	14,243	10.2 (49.0)	14 (3)	96.9 (94)	3.5	0.76	1.31	6 Se
Hg/Se	2.3	46,180	9,644	14.0 (66.2)	12 (3)	99.1 (99)	2.3	0.58	2.43	5 Hg, 6 Se
TMLA	2.4	40,720	8,500	14.1 (56.2)	14 (3)	98.1 (97)	2.4	0.82	1.15	6 Pb
PPC	2.6	23,164	6,714	7.7 (46.2)	22 (3)	99.5 (100)	4.5	0.81	1.24	5 Hg
UrAc	2.1	45,167	12,610	9.6 (49.1)	16 (4)	98.8 (98)	2.2	0.87	0.89	3 U
AuCN	2.4	26,342	8,353	9.9 (43.6)	11 (4)	97.4 (95)	2.5	0.87	0.89	1 Au
HgAc	2.8	8,810	4,165	9.5 (38.6)	11 (3)	72.3 (84)	3.5	0.72	1.64	4 Hg
HgCl	2.9	12,027	4,730	12.2 (58.1)	8 (3)	96.4 (97)	2.9	0.78	1.37	4 Hg
Thimersol	2.8	29,045	5,436	4.3 (6.8)	41 (27)	98.9 (99)	2.8	0.53	2.73	2 Hg
Hg/Pb	2.6	17,123	6,656	5.2 (9.5)	30 (21)	97.7 (99)	2.6	0.55	2.65	1 Hg, 1 Pb

[a] Abbreviations for heavy atom derivatives are as follows: PCMP, *p*-chloromercuriphenol; SMET, seleno-L-methionine–labeled protein; Hg/Se, PCMP soaked into seleno-L-methionine–containing crystals; TMLA, trimethyllead acetate; PPC, potassium platinum chloride; UrAc, uranyl acetate; AuCN, gold cyanide; HgAc, mercury acetate; HgCl, mercury chloride; Hg/Pb, double soak with thimersol and trimethyl lead acetate.

[b] $R_{sym} = \Sigma_j|I_j - <I>|/\Sigma_j I_j$, where $<I>$ is the average intensity of reflection j for its symmetry equivalents; values in parentheses calculated from the 5% of data in the highest resolution shell.

[c] $R_{cullis} = \Sigma_j||F_{PH}| - |F_P + F_H||/\Sigma_j F_H$, where F_{PH} and F_P are the measured derivative and native structure factor and F_H is the calculated heavy atom structure factor.

[d] Phasing power = $<F_H>/E$, where $<F_H>$ is the root-mean-square heavy atom structure factor and E is the residual lack of closure error.

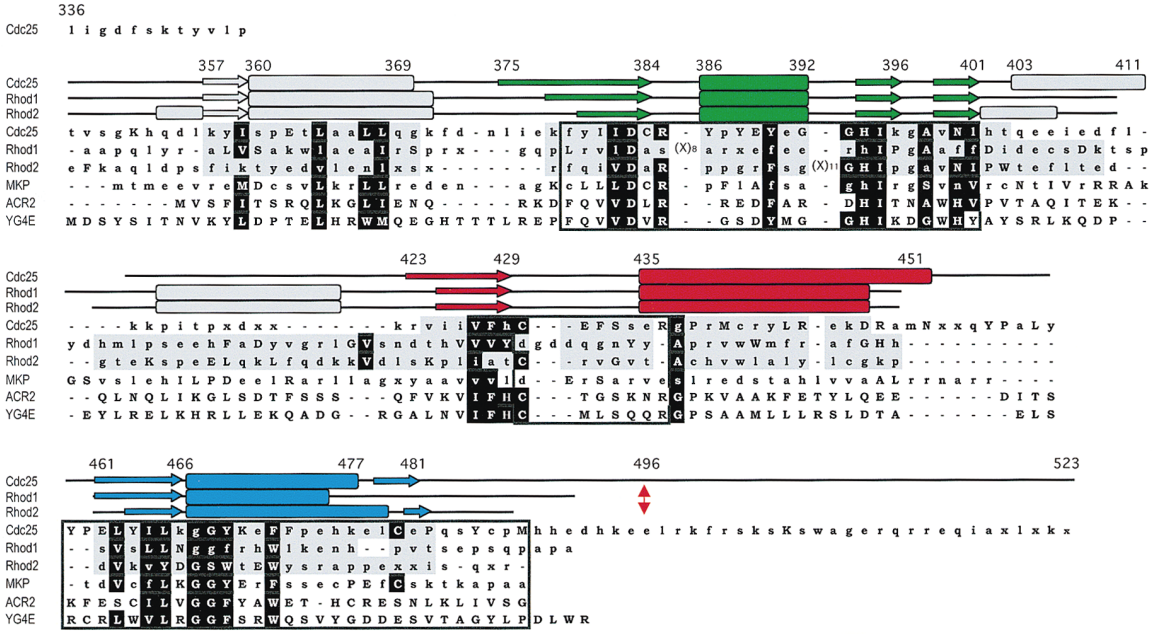

Figure 4. Sequence Alignment Comparing Cdc25 to Rhodanese, MAP Kinase Phosphatase, and Two *S. cerevisiae* Genes, ACR2 and YG4E

Sequences listed for Cdc25, Rhod1, Rhod2, and MKP are consensus sequences constructed by taking all available sequences with at least 25% sequence identity with Cdc25A, bovine rhodanese domain 1, bovine rhodanese domain 2, or CL100, respectively. Capital letters indicate highly conserved positions in each consensus sequence. Light shading indicates regions of structural equivalence between Cdc25A and both rhodanese domains, as determined by the LSQ_IMPROVE routine of O (Jones et al., 1991). Dark shading indicates sequence similarities between highly conserved positions. Secondary structure elements are indicated above the sequences (arrows, β strands; rounded rectangles, α helices). Sequence numbers are for Cdc25A. The CH2A, active site, and CH2B motifs are indicated in green, red, and blue, respectively.

on an extensive protein–protein interface, as opposed to a few select residues right at the active site. There is one report that Cdc25 hydrolyzes Thr-14 more readily than Tyr-15 (Borgne and Meijer, 1996). We believe this is primarily due to the greater accessibility of Thr-14 as revealed in the crystal structure of Cdk2/cyclin A (Russo et al., 1996).

The catalytic Cys-430 can form a disulfide bridge with the absolutely conserved Cys-384 on the adjacent β strand (Figures 5 and 6). In the refined native structure, the two thiol groups are 5.9 Å apart, and there is no evidence for a disulfide. However, in some crystals, including one grown in the presence of 3 mM sodium vanadate, difference electron density consistent with a disulfide is clearly visible (Figure 6B). This observation may explain Cdc25's sensitivity to oxidation and loss of activity in the absence of reducing agents, which hampered its early identification as a phosphatase (Dunphy and Kumagai, 1991; Kumagai and Dunphy, 1991). Whether the potential to form this disulfide reflects a physiological mechanism for enzyme inhibition during oxidative stress in vivo is not known. However, it was recently reported that natural killer cells subjected to thiol deprivation show Cdk hyperphosphorylation and cell cycle arrest, presumably due to inhibition of Cdc25 (Yamauchi and Bloom, 1997).

The CH2A Motif

The first 23 residues of the catalytic domain (residues 335–357) meander with no regular secondary structure.

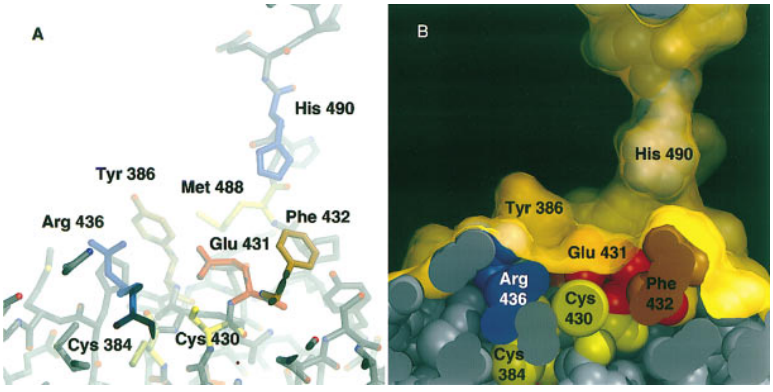

Figure 5. The Active Site of Cdc25A

(A) Ball-and-stick model.

(B) CPK model, in the same orientation, with a molecular surface (Nicholls et al., 1991) in gold. Unlabeled residues are gray; all others are colored according to chemical property. The C-terminal tail of Cdc25A can be seen behind the active site, extending away from the protein. Graphics rendered with O (Jones et al., 1991) and VORT (University of Melbourne).

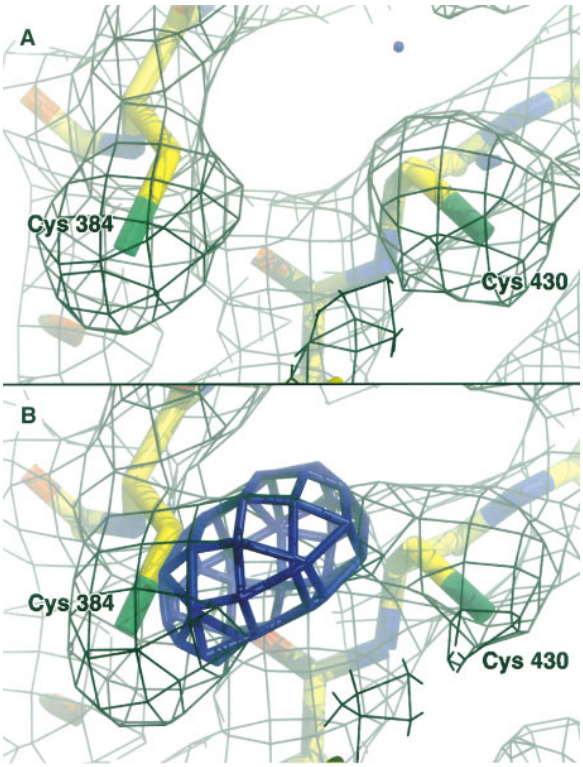

Figure 6. Electron Density Maps Showing Formation of a Disulfide Bond between Absolutely Conserved Cysteines

(A) In the native refined model, there is no positive difference electron density. (B) However, when oxidized, Cys-384 appears to form a disulfide with Cys-430. Calculated structure factors for a $2F_o$–F_c map (0.8 σ, black) and an F_o–F_c map (3.0 σ, blue) are from the native model, which is shown with carbon, nitrogen, oxygen, and sulfur atoms in yellow, blue, red, and green, respectively. Observed amplitudes are from either (A) the native crystal or (B) a crystal grown in the presence of 3 mM vanadate. However, no bound vanadate was located in the electron density maps from this crystal. Graphics were rendered with O (Jones et al., 1991) and VORT (University of Melbourne).

The initial Leu-Ile-Gly-Asp sequence, characteristic of the start of the Cdc25 catalytic domain (Eckstein et al., 1996), forms a short loop extending away from the protein, presumably leading to the N-terminal domain of Cdc25. After the irregular region, a short strand, helix, and loop extend behind the β sheet, leading into the CH2A motif. Conserved residues in this region (Ile-359, Met-364, Val-367, Leu-368) contribute to the protein's hydrophobic cores.

The CH2A motif (residues 379–401; green in Figures 3, 4, and 5) is the first of two regions identified because of significant sequence identity with the noncatalytic domain of the MKPs. CH2A is largely a self-contained strand-helix-strand domain, with many of the highly conserved residues (Asp-383, Arg-385, Tyr-390, His-394, Asn-400) forming hydrogen bonds to side chain and main chain atoms within the subdomain (Figure 7). Between the helix and the second strand is another short strand (residues 394–396), which forms a small antiparallel β ribbon with a segment of the CH2B motif.

The absolutely conserved Asp-383 had been proposed to be a general acid in the reaction (Eckstein et al., 1996), assisting in catalysis by donating a proton to the leaving group hydroxyl, by analogy to the *Yersinia* phosphatase reaction (Zhang et al., 1994b). Replacement of Asp-383 with Asn reduced k_{cat} by at least 150-fold; however, the analogous experiment in the *Yersinia* phosphatase reduces k_{cat}/k_M by 1000-fold. In the crystal structure, Asp-383 points away from the active site and cannot be considered to participate in catalysis (Figure 7). The rate reduction observed likely owes to the disruption of the salt bridge with the buried Arg-385, which is also absolutely conserved. That Asp-383 is also conserved in the CH2A motifs of domains with no phosphatase activity (Figure 4) further supports the notion that this residue is conserved for structural, rather than mechanistic, reasons.

The only acidic residue close enough to the active site to serve as a general acid is the absolutely conserved Glu-431, immediately following the nucleophilic cysteine (Figure 5). This residue has never been independently mutated. However, Gautier and colleagues observed that replacing the five residues following the active-site cysteine (Glu-Phe-Ser-Ser-Glu) with the sequence (Ser-Ala-Gly-Val-Gly) resulted in complete loss of activity (Gautier et al., 1991). There is, however, no requirement that Cdc25 contain a general acid. For example, the maximum turnover rate for the *Yersinia* phosphatase with the general acid removed is 0.9 s^{-1} against pNPP (Zhang et al., 1994b), compared to 0.35 s^{-1} for the intact catalytic domain of Cdc25B (Gottlin et al., 1996).

The CH2B Motif

The CH2B motif (residues 459–488; blue in Figures 2, 3, and 4) consists of a strand followed by a helix, which is situated under the active-site loop such that its N-terminal end points towards the active site (Figure 2). A final short strand forms the second half of a β ribbon with the CH2A motif. Many of the highly conserved residues contribute to a small hydrophobic core within the CH2B motif (Tyr-469, Phe-473, Cys-480, Tyr-485). The absolutely conserved Pro-482 contains a *cis*-peptide bond.

The Cdc25 sequences show less similarity in the region following the CH2B motif (Figure 4). In the crystal structure, residues 489–496 form an extended structure projecting away from the rest of the protein, interacting with neighboring proteins in the crystal lattice. Residues 497–523 are disordered and are not visible in the electron density maps, although there is room for the additional 27 residues in the crystal lattice. The penultimate highly conserved residue of Cdc25, Met-488, packs between the conserved Tyr-386 and Tyr-388 of the CH2A motif (Figures 5 and 7). In human Cdc25C, the last 38 residues, up to this methionine, can be removed without affecting pNPP or Cdc2 hydrolase activity (Lee et al., 1992).

A Recurring Fold

The structures of Cdc25A and rhodanese identify a constellation of approximately 24 residues that together determine the topology of this 140-residue domain (Figure 4). These 24 residues also include nearly all of the

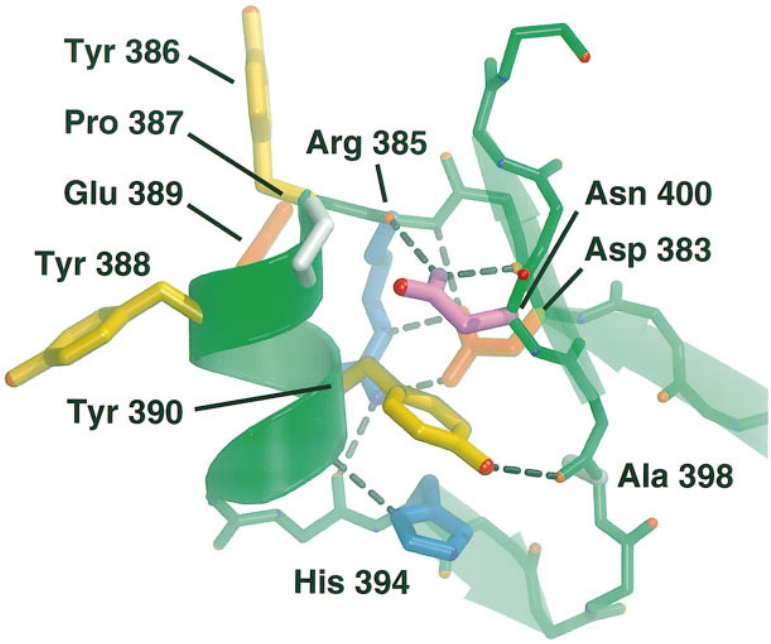

Figure 7. Conserved Residues in the CH2A Region in Cdc25A

Asp-383, Arg-385, Tyr-390, His-394, and Asn-400 all make intradomain hydrogen bonds. The small side chain of Ala-398 makes room for Tyr-390. Tyr-386, Pro-387, Tyr-388, and Glu-389, which point away from the CH2A motif, are conserved in all Cdc25s but not in rhodanese and MKPs. Figure generated with Molscript (Kraulis, 1991) and VORT (University of Melbourne).

sequence identities seen between Cdc25s from different species, between Cdc25 and the N-terminal region of the MKPs, between rhodaneses from different sources, and between the two halves of rhodanese. Many of these residues lie within the CH2A and CH2B regions, motifs which were defined originally based solely on identities between Cdc25 and the MKPs. Remarkably, the rhodanese PROSITE (Bairoch et al., 1997) motifs, which distill the salient features of rhodanese sequences, also lie within the CH2A and CH2B regions, even though these PROSITE motifs were constructed with reference to rhodanese sequences alone.

The sequence and structural similarity between Cdc25 and rhodanese suggests they be classed together as proteins of identical topology containing an active-site nucleophilic cysteine. Other proteins that show the same constellation of conserved amino acids probably also possess this same topology. One example would be ACR2, a unique eukaryotic arsenate resistance enzyme in *Saccharomyces cerevisiae* (Bobrowicz et al., 1997; Wysocki et al., 1997), which contains a Cys-$(X)_5$-Arg motif (Figures 1 and 4). By analogy with the bacterial ArsC protein (Gladysheva et al., 1996) and based on the structure of Cdc25, we propose that ACR2 uses its Cys-76 as a nucleophile to reduce arsenate to arsenite. Another *S. cerevisiae* protein, YG4E, is 27% identical in sequence to ACR2, including the Cys-$(X)_5$-Arg motif (Figure 4). YG4E has no known function but also likely adopts the Cdc25 fold.

Based on the pattern of conserved residues, the MKPs probably also have this same topology in their N-terminal, noncatalytic domains, even though the Cys-$(X)_5$-Arg motif is absent there (Figures 1 and 4). These sequence similarities may reflect similar biological targets. The substrates of Cdc25 and the MKPs are Cdk2 and the MAP kinases, homologous serine/threonine kinases (Goldsmith and Cobb, 1994). But while Cdc25 acts on residues in the N-terminal lobe of Cdc2, the MKPs dephosphorylate residues in the larger, C-terminal lobe of MAP kinases. Perhaps the N-terminal domain of the MKPs confers binding between MKP and the N-terminal lobe of the MAP kinases.

Residues that are highly conserved among the Cdc25s but not conserved in the other sequences (Figure 4) are likely to be serving functional instead of structural roles. One example is the sequence Tyr-Pro-Tyr-Glu (residues 386–389), which projects out from the CH2A helix (Figure 7). These residues may be involved in specific interactions between Cdc25 and its Cdk substrate.

Conclusion

Much of the literature for the past six years has relied on an assumption of homology between Cdc25 and the other known dual-specificity and tyrosine-specific phosphatases (e.g., Eckstein et al., 1996). The structure of the catalytic domain of Cdc25A, compared to those of the other Cys-$(X)_5$-Arg phosphatases, suggests otherwise. The *Yersinia* phosphatase (Stuckey et al., 1994), PTP1B (Barford et al., 1994), PTPα (Bilwes et al., 1996), and VHR (Yuvaniyama et al., 1996) all share a common topology and possess sequence similarity in regions away from the active site. Cdc25 has its own topology and groups of highly conserved residues and most likely represents the results of convergent evolution toward the Cys-$(X)_5$-Arg solution of phosphomonoester hydrolysis. In this light, the low molecular weight phosphatases may be seen as yet a third case of convergent evolution toward this solution (Ramponi and Stefani, 1997).

Experimental Procedures

Sample Preparation

Human Cdc25A cDNA was obtained from David Beach (Cold Spring Harbor Laboratory). DNA encoding amino acids 336-523, plus an

initiating methionine, was amplified by PCR, confirmed by sequencing, and cloned into a pET21a vector (Novagen). Cdc25A (336-523) was overexpressed in *Escherichia coli* BL21 (DE3) cells and purified to homogeneity in a three-column procedure on a Pharmacia BioPilot system at 4°C. Cation exchange (Poros HS 50-micron media) followed by a concentration step over the same media and then a size exclusion column (Pharmacia S-75 Superdex) generated about 0.5 g of pure protein per 10 l of *E. coli*. Seleno-L-methionine–labeled protein was similarly purified from *E. coli* B834 (DE3) cells, grown in 2× M9 medium supplemented with 0.1 g/l ampicillin, 2.0 g/l glucose, 246 mg/l Mg_2SO_4–$7H_2O$, 11 mg/l $CaCl_2$, 33.7 mg/l thiamine-HCl, 30 mg/l cysteine, and 30 mg/l seleno-L-methionine (Sigma). A glutathione-S-transferase fusion protein of the full-length enzyme, Cdc25A (1-523) was purified from *E. coli* BL21 (DE3) by binding and elution from Glutathione Sepharose 4B. Purified Cdc25A (336-523) appears to be 16-fold more active than GST-Cdc25A (1-523) in dephosphorylating Cdk2 labeled by the protein kinase Wee1 (data not shown).

Crystallization

Crystals of Cdc25A (336-523) and of the seleno-L-methionine–containing protein were grown in hanging drops by the vapor diffusion method at 4°C. Four microliters of protein (10 mg/ml in 200 mM NaCl, 1 mM DTT, 20 mM HEPES [pH 7.4]) plus 4 μl of precipitant (18%–20% PEG 3350 [Hampton], 0.025% β-octylglucoside, 0.11 M sodium citrate [pH 5.8]) were equilibrated against 1 ml of precipitant. Large ($1.0 \times 0.2 \times 0.2$ mm^3) square pyramidal crystals appeared in 2–3 days. Drops that did not spontaneously nucleate were seeded with microcrystals diluted in precipitant.

Data Collection

Prior to data collection, cryosolvent (25% PEG 400, 75% precipitant solution) was slowly added to the hanging drop. For heavy atom derivatives, crystals were next soaked in cryosolvent containing 10 mM of the desired heavy atom solution for 24–48 hr. Crystals were flash frozen and maintained at 120 K with nitrogen gas cooled by the *X-stream* (Molecular Structure Corporation) system. Diffraction intensities for the native crystal and most of the derivatives were collected on an R-AXIS IV/Rigaku RU-200 system running at 50 kV 100 mA; intensities for the thimersol derivatives were collected on an R-AXIS II. Intensities were integrated, scaled, and merged with DENZO (Otwinowski and Minor, 1997) and were converted to amplitudes by the method of French and Wilson (1978). Crystals were of the space group $P4_1$ (or $P4_3$), with $a = b = 43.5$ Å, $c = 117.1$ Å.

Phasing and Refinement

Sites for the *p*-chloromercuriphenol derivative were found by Patterson techniques. Sites for the remaining derivatives were located in difference Fourier maps and confirmed in Patterson Harker sections. Heavy atom sites were refined and phases calculated to 2.2 Å with MLPHARE (Collaborative Computational Project No. 4, 1994), giving an overall figure of merit of 0.64. The initial trace into solvent-flattened 3.0 Å maps (Cowtan, 1994) located all major secondary structures and identified $P4_1$ as the correct enantiomorph. Assignment of sequence and connectivity was aided by positions of selenium peaks in methionine side chains and a relatively accurate secondary structure prediction (Rost and Sander, 1994). The final model, after cycles of fitting (Jones et al., 1991) and refinement (Brünger, 1992), contains 162 residues and 69 water molecules. The first residue in the structure is Met-335, which was introduced during cloning. The C terminus is disordered, and residues Leu-497 through Leu-523 cannot be located. The final model, including individual atomic B factors and a bulk solvent model, yields an R-factor of 22.7% for all data to 2.3 Å and a free R-factor of 29.6%. The average protein B factor of 36 $Å^2$ agrees well with the value of 34 $Å^2$ obtained from a Wilson plot. The rmsd from ideality are 0.013 Å for bonds, 1.59° for angles, 24.9° for dihedral angles, and 1.54° for improper angles.

Acknowledgments

We thank William Clay for expert baculovirus preparation, Annie Hassell and Steve Jordan for help with crystallography, and Rebecca Schutt for excellent crystal growth. We also thank Michael Byrnes and Mary Stephenson for preliminary recombinant protein preparations. This work was supported in part by the National Institutes of Health grants (AI34095 to M. A. S. and GM47017 to H. P.-W.) and a grant from the John and Suzanne Munn Endowed Research Fund of the University of Michigan Comprehensive Cancer Center. H. P.-W. is an investigator of the Howard Hughes Medical Institute.

Received March 10, 1998; revised April 13, 1998.

References

Bairoch, A., Bucher, P., and Hofmann, K. (1997). The PROSITE database, its status in 1997. Nucleic Acids Res. *25*, 217–221.

Barford, D., Flint, A.J., and Tonks, N.K. (1994). Crystal structure of human protein tyrosine phosphatase 1B. Science *263*, 1397–1404.

Bilwes, A.M., den Hertog, J., Hunter, T., and Noel, J.P. (1996). Structural basis for inhibition of receptor protein-tyrosine phosphatase alpha by dimerization. Nature *382*, 555–559.

Bobrowicz, P., Wysocki, R., Owsianik, G., Goffeau, A., and Ulaszewski, S. (1997). Isolation of three contiguous genes, ACR1, ACR2 and ACR3, involved in resistance to arsenic compounds in the yeast *Saccharomyces cerevisiae*. Yeast *13*, 819–828.

Borgne, A., and Meijer, L. (1996). Sequential dephosphorylation of p34(cdc2) on Thr-14 and Tyr-15 at the prophase/metaphase transition. J. Biol. Chem. *271*, 27847–27854.

Brünger, A.T. (1992). Slow-cooling protocols for crystallographic refinement by simulated annealing. Acta Cryst. *A46*, 585–593.

Collaborative Computational Project, No. 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Cryst. *D50*, 760–763.

Cowtan, K. (1994). "dm": an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography *31*, 34–38.

Denu, J.M., and Dixon, J.E. (1995). A catalytic mechanism for the dual-specific phosphatases. Proc. Natl. Acad. Sci. USA *92*, 5910–5914.

Dunphy, W.G., and Kumagai, A. (1991). The cdc25 protein contains an intrinsic phosphatase activity. Cell *67*, 189–196.

Eckstein, J.W., Beer-Romero, P.B., and Berdo, I. (1996). Identification of an essential acidic residue in Cdc25 protein phosphatase and a general three-dimensional model for a core region in protein phosphatases. Protein Sci. *5*, 5–12.

Fauman, E.B., and Saper, M.A. (1996). Structure and function of the protein tyrosine phosphatases. Trends Biochem. Sci. *21*, 413–417.

Fisher, R.P. (1997). CDKs and cyclins in transition(s). Curr. Opin. Cell Biol. Genet. *7*, 32–38.

French, G.S., and Wilson, K.S. (1978). On the treatment of negative intensity observations. Acta Cryst. *A34*, 517–525.

Galaktionov, K., and Beach, D. (1991). Specific activation of cdc25 tyrosine phosphatase by B-type cyclins: evidence for multiple roles of mitotic cyclins. Cell *67*, 1181–1194.

Gautier, J., Solomon, M.J., Boojer, R.N., Bazan, J.F., and Kirschner, M.W. (1991). cdc25 is a specific tyrosine phosphatase that directly activates p34cdc2. Cell *67*, 197–211.

Gladysheva, T., Liu, J., and Rosen, B.P. (1996). His-8 lowers the pKa of the essential Cys-12 residue of the ArsC arsenate reductase of plasmid R773. J. Biol. Chem. *271*, 33256–33260.

Gliubich, F., Gazerro, M., Zanotti, G., Delbono, S., Bombieri, G., and Berni, R. (1996). Active site structural features for chemically modified forms of rhodanese. J. Biol. Chem. *271*, 21054–21061.

Goldsmith, E.J., and Cobb, M.H. (1994). Protein-kinases. Curr. Opin. Struct. Biol. *4*, 833–840.

Gottlin, E.B., Xu, X., Epstein, D., Burke, S., Eckstein, J.W., Ballou, D.P., and Dixon, J. (1996). Kinetic analysis of the catalytic domain of human Cdc25B. J. Biol. Chem. *271*, 27445–27449.

Gould, K.L., and Nurse, P. (1989). Tyrosine phosphorylation of the fission yeast cdc2+ protein kinase regulates entry into mitosis. Nature *342*, 39–45.

Guan, K.L., and Dixon, J.E. (1991). Evidence for protein-tyrosine-phosphatase catalysis proceeding via a cysteine-phosphate intermediate. J. Biol. Chem. *266*, 17026–17030.

Guan, K., Broyles, S.S., and Dixon, J.E. (1991). A Tyr/Ser protein phosphatase encoded by a vaccinia virus. Nature *350*, 359–362.

Hoffmann, I., Clarke, P.R., Marcote, M.J., Karsenti, E., and Draetta, G. (1993). Phosphorylation and activation of human cdc25-C by cdc2-cyclin B and its involvement in the self-amplification of MPF at mitosis. EMBO J. *12*, 53–63.

Hoffmann, I., Draetta, G., and Karsenti, E. (1994). Activation of the phosphatase activity of human cdc25A by a cdk2-cyclin E dependent phosphorylation at the G1/S transition. EMBO J. *13*, 4302–4310.

Holm, L., and Sander, C. (1994). The FSSP database of structurally aligned protein fold families. Nucleic Acids Res. *22*, 3600–3609.

Honda, R., Ohba, Y., Nagata, A., Okayama, H., and Yasuda, H. (1993). Dephosphorylation of human p34cdc2 kinase on both Thr-14 and Tyr-15 by human cdc25B phosphatase. FEBS Lett. *318*, 331–334.

Jinno, S., Suto, K., Nagata, A., Igarashi, M., Kanaoka, Y., Nojima, H., and Okayam, H. (1994). Cdc25A is a novel phosphatase functioning early in the cell cycle. EMBO J. *13*, 1549–1556.

Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Cryst. *A47*, 110–119.

Keyse, S.M., and Ginsburg, M. (1993). Amino acid sequence similarity between CL100, a dual-specificity MAP kinase phosphatase and cdc25. Trends Biochem. Sci. *18*, 377–378.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic figures. J. Appl. Cryst. *24*, 946–950.

Krek, W., and Nigg, E. (1991). Mutations of p34cdc2 phosphorylation sites induce premature mitotic events in HeLa cells: evidence for a double block to p34cdc2 kinase activation in vertebrates. EMBO J. *10*, 3331–3341.

Kumagai, A., and Dunphy, W. (1991). The cdc25 protein controls tyrosine dephosphorylation of the cdc2 protein in a cell-free system. Cell *64*, 903–914.

Lee, M.S., Ogg, S., Xu, M., Parker, L., Donoghue, D., Maller, J., and Piwnica-Worms, H. (1992). cdc25+ encodes a protein phosphatase that dephosphorylates p34 cdc2. Mol. Biol. Cell *3*, 73–84.

Moreno, S., and Nurse, P. (1991). Clues to action of cdc25 protein. Nature *351*, 194.

Morgan, D.O. (1995). Principles of CDK regulation. Nature *374*, 131–134.

Nicholls, A., Sharp, K., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In Methods in Enzymology, Volume 276: Macromolecular Crystallography, part A, C.W. Carter, Jr. and R.M. Sweet, eds. (New York: Academic Press), pp. 307–326.

Peng, C.Y., Graves, P.R., Thoma, R.S., Wu, Z., Shaw, A.S., and Piwnica-Worms, H. (1997). Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216. Science *277*, 1501–1505.

Ploegman, J.H., Drent, G., Kalk, K.H., and Hol, W.G. (1978). The covalent and tertiary structure of bovine liver rhodanese. Nature *273*, 124–129.

Ramponi, G., and Stefani, M. (1997). Structure and function of the low M-r phosphotyrosine protein phosphatases. Biochim. Biophys. Acta *1341*, 137–156.

Rost, B., and Sander, C. (1994). Combining evolutionary information and neural networks to predict protein secondary structure. Proteins *19*, 55–72.

Russo, A., Jeffrey, P., and Pavletich, N. (1996). Structural basis of cyclin-dependent kinase activation by phosphorylation. Nat. Struct. Biol. *3*, 696–700.

Strausfeld, U., Labbe, J.C., Fesquet, D., Cavadore, J.C., Picard, A., Sadhu, K., Russell, P., and Doree, M. (1991). Dephosphorylation and activation of a p34cdc2/cyclin B complex in vitro by human CDC25 protein. Nature *351*, 242–245.

Stuckey, J.A., Schubert, H.L., Fauman, E.B., Zhang, Z.-Y., Dixon, J.E., and Saper, M.A. (1994). Crystal structure of *Yersinia* protein tyrosine phosphatase at 2.5 Å and the complex with tungstate. Nature *370*, 571–575.

Su, X.-D., Taddel, N., Stefani, M., Ramponi, G., and Nordlund, P. (1994). The crystal structure of a low-molecular-weight phosphotyrosine protein phosphatase. Nature *370*, 575–579.

Wysocki, R., Bobrowicz, P., and Ulaszewski, S. (1997). The *Saccharomyces cerivisiae* ACR3 gene encodes a putative membrane protein involved in transport. J. Biol. Chem. *272*, 30061–30066.

Xu, X., and Burke, S.P. (1996). Roles of active site residues and the NH2-terminal domain in the catalysis and substrate binding of human Cdc25. J. Biol. Chem. *271*, 5118–5124.

Yamauchi, A., and Bloom, E.T. (1997). Control of cell cycle progression in human natural killer cells through redox regulation of expression and phosphorylation of retinoblastoma gene product protein. Blood *89*, 4092–4099.

Yuvaniyama, J., Denu, J.M., Dixon, J.E., and Saper, M.A. (1996). Crystal structure of the dual-specificity protein phosphatase VHR. Science *272*, 1328–1331.

Zhang, M., Van Etten, R.L., and Stauffacher, C.V. (1994a). Crystal structure of bovine heart phosphotyrosyl phosphatase at 2.2-Å resolution. Biochemistry *33*, 11097–11105.

Zhang, Z.-Y., Wang, Y., and Dixon, J.E. (1994b). Dissecting the catalytic mechanism of protein-tyrosine phosphatases. Proc. Natl. Acad. Sci. USA *91*, 1624–1627.

Brookhaven Protein Data Bank

Coordinates have been deposited with the Brookhaven Protein Data Bank (accession code 1c25) and may also be obtained from the authors (e-mail: saper@umich.edu or fauman@umich.edu).