Cell, Vol. 112, 243–256, January 24, 2003, Copyright ©2003 by Cell Press

Structural Basis for Phosphodependent Substrate Selection and Orientation by the SCF^{Cdc4} Ubiquitin Ligase

Stephen Orlicky,[1] Xiaojing Tang,[1] Andrew Willems,[1] Mike Tyers,[1,2,*] and Frank Sicheri[1,2,*]
[1]Program in Molecular Biology and Cancer
Samuel Lunenfeld Research Institute
Mount Sinai Hospital
600 University Avenue
Toronto, Ontario M5G 1X5
[2]Department of Molecular and Medical Genetics
University of Toronto
Toronto, Ontario M5S 1A8
Canada

Summary

Cell cycle progression depends on precise elimination of cyclins and cyclin-dependent kinase (CDK) inhibitors by the ubiquitin system. Elimination of the CDK inhibitor Sic1 by the SCF^{Cdc4} ubiquitin ligase at the onset of S phase requires phosphorylation of Sic1 on at least six of its nine Cdc4-phosphodegron (CPD) sites. A 2.7 Å X-ray crystal structure of a Skp1-Cdc4 complex bound to a high-affinity CPD phosphopeptide from human cyclin E reveals a core CPD motif, Leu-Leu-pThr-Pro, bound to an eight-bladed WD40 propeller domain in Cdc4. The low affinity of each CPD motif in Sic1 reflects structural discordance with one or more elements of the Cdc4 binding site. Reengineering of Cdc4 to reduce selection against Sic1 sequences allows ubiquitination of lower phosphorylated forms of Sic1. These features account for the observed phosphorylation threshold in Sic1 recognition and suggest an equilibrium binding mode between a single receptor site in Cdc4 and multiple low-affinity CPD sites in Sic1.

Introduction

The ubiquitin proteolytic system controls the abundance of regulatory proteins in signaling, development, and cell cycle progression. Substrate ubiquitination is catalyzed by a cascade of enzymes, termed E1, E2, and E3, which activate and then conjugate ubiquitin to the substrate (Hershko and Ciechanover, 1998). E3 enzymes, also known as ubiquitin ligases, contain substrate-specific recognition domains and catalyze the final step in ubiquitin transfer. Recognition is mediated by elements in the substrate termed degrons (Varshavsky, 1991). Control of the E3-substrate interaction forms the basis for regulated proteolysis. Posttranslational modification often serves to target substrates to their cognate E3 enzymes. Two main classes of E3 enzyme are now evident, as characterized by the presence of either a HECT domain or a RING domain. The HECT domain forms a catalytically essential thioester with ubiquitin, whereas the RING domain provides a docking site for E2 enzymes, which provide catalytic activity (Pickart, 2001).

* Correspondence: sicheri@mshri.on.ca (F.S.), tyers@mshri.on.ca (M.T.)

Phosphorylation-dependent degrons direct many substrates to multisubunit E3 enzymes termed Skp1-Cdc53/Cullin-F box protein (SCF) complexes. SCF complexes are built from an invariant core machinery comprised of the adaptor protein Skp1, the scaffold protein Cdc53 (called Cul1 in metazoans), and the RING-H2 domain protein Rbx1 (also called Roc1 or Hrt1), which interacts with an E2 enzyme, usually Cdc34 (Pickart, 2001). Substrates are brought to the core SCF complex by one of a large family of variable adaptor subunits called F box proteins, each of which targets a limited number of specific substrates (Bai et al., 1996; Patton et al., 1998). F box proteins typically have a bipartite structure with an N-terminal ~40 amino acid F box motif and a C-terminal substrate interaction region, such as a WD40 repeat domain or a leucine rich repeat (LRR) domain (Bai et al., 1996; Feldman et al., 1997; Skowyra et al., 1997). The overall architecture of SCF complexes is conserved in several related ubiquitin-ligase complexes, including the Anaphase Promoting Complex/Cyclosome and the Von Hippel Lindau (VHL) tumor suppressor protein complex, each of which contain cullin family members, RING-H2 domains, and substrate recognition subunits (Pickart, 2001; Kaelin, 2002).

Cell cycle progression depends on the precisely timed destruction of cyclins and cyclin-dependent kinase (CDK) inhibitors by the ubiquitin system (Harper et al., 2002). In yeast, G1 cyclin CDK activity phosphorylates the CDK inhibitor Sic1, whose degradation is necessary for onset of B type cyclin CDK activity and DNA replication (Schwob et al., 1994). Phospho-Sic1 is specifically recognized by the WD40 domain of the F box protein Cdc4, which recruits Sic1 for ubiquitination by the Cdc34-SCF complex (Bai et al., 1996; Feldman et al., 1997; Skowyra et al., 1997). Stable forms of Sic1 that lack CDK phosphorylation sites cause a G1 phase arrest (Verma et al., 1997), whereas deletion of *SIC1* causes premature DNA replication and genome instability (Lengronne and Schwob, 2002). Cdc4 recruits several other substrates to the SCF core complex in a phosphorylation-dependent manner, including the Cln-Cdc28 inhibitor/cytoskeletal scaffold protein Far1, the replication protein Cdc6, and the transcription factor Gcn4 (Patton et al., 1998). The F box protein Grr1 functions in an analogous manner to render G1 cyclins unstable throughout the cell cycle in a manner that depends on recognition of phospho-epitopes by its LRR domain (Skowyra et al., 1997; Hsiung et al., 2001).

In the metazoan cell cycle, SCF complexes target phosphorylated forms of the CDK inhibitor p27^{Kip1} and cyclin E, among other substrates. Interestingly, F box protein specificity for these substrates is reversed compared to yeast in that the WD40 domain of hCdc4/Fbw7/Ago/SEL-10 recognizes cyclin E (Strohmaier et al., 2001; Koepp et al., 2001; Moberg et al., 2001), whereas the LRR domain of Skp2 recognizes p27^{Kip1} in conjunction with the CDK binding protein Cks1 (Harper, 2001). Both of these degradation pathways are perturbed in cancer cells. Many primary tumors express high levels of Skp2, which leads to premature degradation of p27^{Kip1} and cell

cycle entry (Harper, 2001). Conversely, loss of Cdc4 function causes deregulation of cyclin E-CDK2 activity, which leads to precocious S phase entry and genome instability (Spruck et al., 1999). Mutations in the *Drosophila* homolog of *CDC4*, called *ago*, were isolated as homozygous recessive alleles in a screen for excess cell proliferation, a defect attributed to ectopic cyclin E activity (Moberg et al., 2001). Mutations in *hCDC4* have been detected in several cancer cell lines that exhibit high levels of cyclin E (Moberg et al., 2001; Strohmaier et al., 2001), as well as in a significant fraction of primary endometrial cancers (Spruck et al., 2002). In addition, *hCDC4* is located in the 4q32 region, which is often deleted in various cancers (Spruck et al., 2002). Significantly, a high level of cyclin E correlates strongly with low survival rates in breast cancer (Keyomarsi et al., 2002). Other important substrates appear to be targeted for degradation by Cdc4 orthologs in a phosphorylation-dependent manner, including activated intracellular forms of the developmental regulator Notch and the presenilin proteins, which are implicated in familial early onset Alzheimer's disease (Lai, 2002; Selkoe, 2001). SCF-dependent proteolysis also mediates other signaling events including phosphorylation-dependent degradation of the NFκB inhibitor IκBα and the proto-oncogene product β-catenin by the F box protein β-TrCP (Pickart, 2001).

Several F box proteins can recognize short phosphopeptide motifs that correspond to substrate sequences. However, it is unknown whether such interactions are analogous to phosphorylation-dependent interactions of SH2, PTB, 14-3-3, WW, and FHA domains, each of which has been crystallized with its cognate phosphopeptide (Yaffe and Elia, 2001). For many SCF substrates, including Sic1, Cdc6, and Cln2, phosphorylation on multiple dispersed sites is required for recognition and degradation (Verma et al, 1997; Patton et al., 1998). We recently defined a high-affinity consensus phosphopeptide binding motif for Cdc4, termed the Cdc4 phosphodegron (CPD), which bears the consensus I/L-I/L/P-pT-P-$<$KR$>_4$, where $<$ $>$ indicates a disfavored residue (Nash et al., 2001). The P0 phospho-threonine residue or less favorably, a phospho-serine residue, and the P+1 proline are essential for interaction with Cdc4. Unexpectedly, the CPD consensus is at odds with the CDK phosphorylation site consensus: S/T-P-X-K/R (Endicott et al., 1999). Thus, substrate recognition by the targeting kinase is counter-balanced against the targeting component of the degradation machinery. All nine CPD sites in Sic1 have one or more suboptimal features: all lack consensus hydrophobic residues in the P-1 or P-2 positions, four have serine in place of threonine in the P0 position, and seven contain a disfavored basic residue in one of the +2 to +5 positions. Unexpectedly, Sic1 must be phosphorylated on at least six of its nine sites in order to allow recognition by Cdc4 (Nash et al., 2001). This requirement for multisite phosphorylation in principle renders the rate of Sic1 degradation proportional to the sixth power of G1 CDK concentration (Ferrell, 1996). The inherently ultrasensitive nature of the Sic1 degradation reaction appears critical for the coordinated initiation of DNA replication by S phase CDK activity (Nash et al., 2001; Lengronne and Schwob, 2002).

The mechanism of the ubiquitin conjugation reaction is not well understood. The ability of E2-E3 enzyme complexes to form polymers of ubiquitin, itself an 8 kDa protein, on a protein substrate presumably demands a large catalytic cradle to accommodate the initial reactants (Pickart, 2001). The sequential addition of ubiquitin moieties onto the substrate must also entail considerable flexibility of the substrate and/or the enzyme complex in order to extend the ubiquitin chain. Recent structure determination and modeling of three E2-E3 complexes has provided insight into these issues. A complex of the E2 enzyme UbcH7 and the HECT domain enzyme E6AP reveals a distance of ~50Å between the E2 and E3 active sites, suggesting that catalytic transfer of ubiquitin requires large-scale movements in an as yet undefined process (Huang et al., 1999). Similarly, a complex between UbcH7 and the RING domain E3 c-Cbl contains a substantial gap between the E2 active site and the substrate binding site on c-Cbl (Zheng et al., 2000). Structures of the SOCS-box adaptor protein VHL in complex with a hydroxylated substrate peptide have recently been solved (Kaelin, 2002), but the orientation of the substrate binding site with respect to the E2 enzyme is unknown. Finally, structure determination and molecular modeling of the holo-SCF^{Skp2} complex again suggests a distance of ~50Å between the substrate binding LRR domain in Skp2 and the E2 active site (Zheng et al., 2002). Notably, the extensive interdigitation of the Skp1-Skp2 interface and the Skp2 interdomain interface rigidly fixes the orientation of the LRRs of Skp2, suggesting that the F box protein might hold the substrate in a very precise orientation with respect to the E2 enzyme (Schulman et al., 2000). However, because the substrate binding site on Skp2 has not been determined by either mutation or by cocrystallization with substrate peptide, it is not possible to deduce how SCF substrates might be positioned with respect to the E2 catalytic site. To understand the basis for phospho-epitope recognition by Cdc4 and to gain insight into the nature of substrate orientation by E3 enzymes, we have determined the structure of a Skp1-Cdc4-CPD phosphopeptide cocomplex.

Results

Alignment of Skp1 and Cdc4 homologs from various species and limited proteolysis of full-length recombinant proteins were used to deduce loop regions in *Saccharomyces cerevisiae* Skp1 and Cdc4 that might interfere with protein crystallization (Figure 1). Crystals of a ternary complex of ScSkp1 bound to Cdc4 and a CPD phosphopeptide were obtained that diffracted to a resolution of 2.7Å (Table 1). For Skp1, a nonessential loop spanning residues 37–64 was removed. The Cdc4 fragment used extends from residues 263–744, which encompasses the F box motif to the end of the WD40 domain, and was engineered to remove two predicted loop regions (Figure 1B). This Cdc4 construct lacks an essential ~40 residue domain that precedes the F box in different WD40 domain-containing F box protein family members (Wolf et al., 1999). The high-affinity CPD phosphopeptide corresponds to nine residues of human cyclin E, Gly-Leu-Leu-pThr-Pro-Pro-Gln-Ser-Gly, which binds Cdc4 with a K_d of 1 μM (Nash et al., 2001).

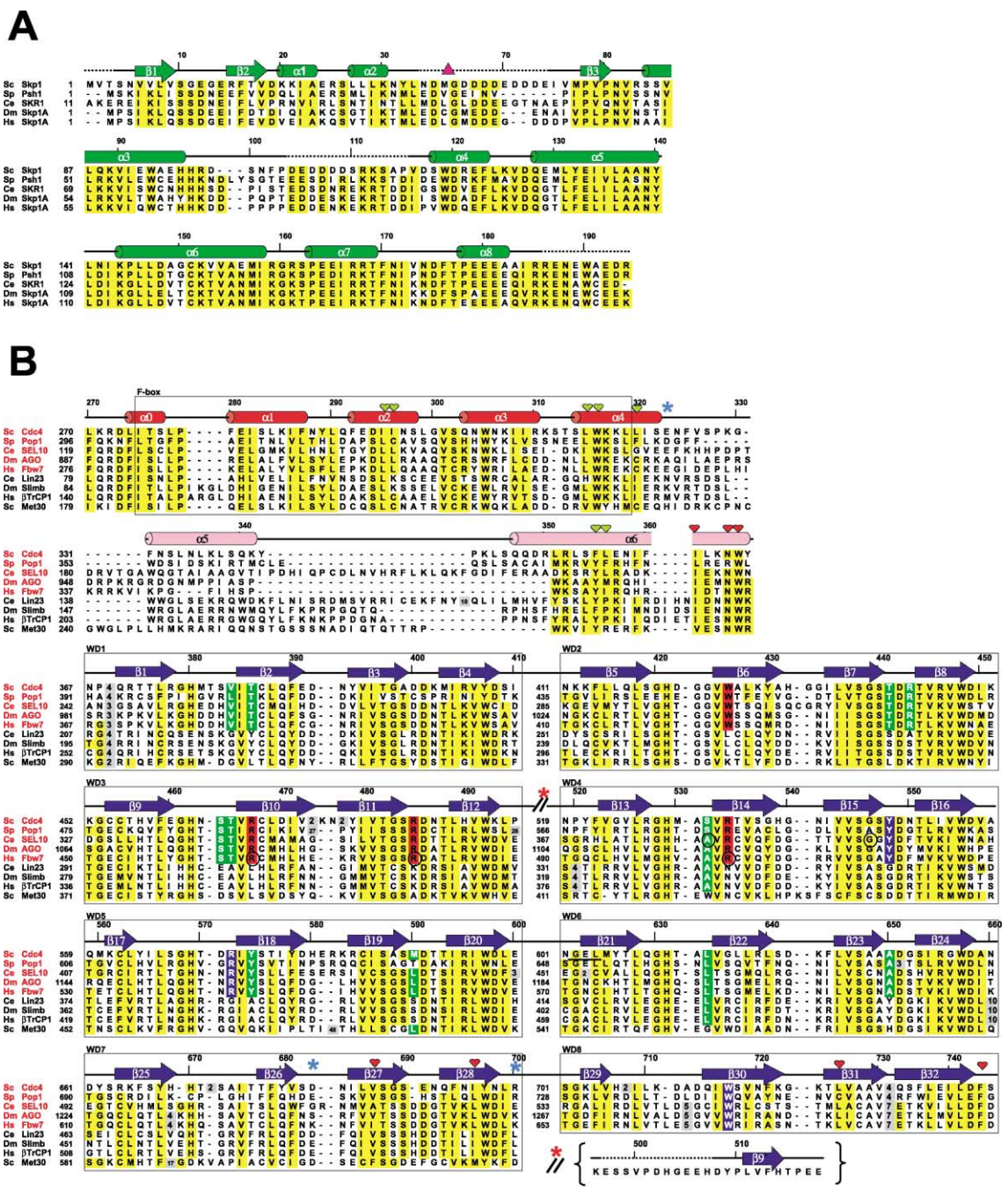

Figure 1. Structure-Based Sequence Alignments

(A) Skp1 orthologs. (B) Cdc4 orthologs (red) and paralogs (black). Human Fbw7 and β-TrCP1 are isoforms 1 and 2, respectively. Secondary structure elements are colored as in Figure 2A. Disordered regions in the crystal structure are shown as dashed lines. Red residues are essential for the Cdc4 function; blue residues strongly influence, but do not abrogate, function; green residues are nonessential but conserved around the binding pocket; and yellow residues are conserved elsewhere. Circles indicate mutations associated with excessive cell proliferation in flies and/or cancer in humans (see text). Deletion of residues 37–64 in Skp1 is denoted by a triangle, and a replacement of two closely placed loops from residues 602–605 and 609–624 is denoted by the underline of the short interloop sequence Gly-Glu-Leu. Insertions to optimize sequence alignments are indicated by number of residues inserted in gray. The nonstandard β strand element 9′ in ScCdc4 is marked by the red asterisk and is shown in full at the bottom of the alignment. Residues that anchor helix α6 to the F box domain are marked by green hearts, those that anchor helix α6 to the WD40 domain by red hearts, and those that make direct contact between the WD40 domain and F box domain by blue asterisks.

The F Box Interface

Yeast Skp1 forms an elongated structure with a mixed α/β topology identical to that reported for human Skp1 (Schulman et al., 2000) and consists of a three-strand β sheet, denoted β1 to β3, and eight α helices, denoted α1 to α8 (Figure 2A). The structure of Cdc4 consists of

Table 1. Data Collection, Structure Determination, and Refinement Statistics

Phasing Statistics			
	Peak	Inflection	Remote
Wavelength (Å)	0.9798	0.9800	0.90000
Resolution (Å)[1]	2.8 (2.9–2.8)	2.9 (3.0–2.9)	2.7 (2.8–2.7)
R_{sym} (%)	5.9 (37.2)	6.1 (36.1)	5.0 (28.9)
Total reflections	311,509	187,010	298,371
Unique reflections	107,167	96,027	116,218
Completeness (%)	99.8 (99.1)	99.3 (98.3)	97.7 (93.6)
I/σ	9.9 (2.7)	7.4 (2.1)	10.1 (2.9)
Phasing power (ISO/ANO)[2]	5.2/1.3	4.0/0.94	0/0.91

Refinement Statistics (Remote Wavelength)	
Resolution range (Å): 20–2.7	Number of protein atoms: 9,364
Reflections: 113,960	Number of water molecules: 72
R_{factor}/R_{free} (%)[3]: 23.8/27.3	
Rms deviations	
Bonds (Å): 0.0089	
Angles (°): 1.42	
Number of Skp1-Cdc4-CPD complexes per asymmetric unit = 2.	
Space group $P3_2$: a = b = 107.7 Å, c = 168.3 Å; a = b = 90°, c = 120°.	

Numbers given in parentheses refer to data for the highest resolution shell.
[1] Rsym = $100 \times \Sigma|I-\langle I\rangle|/\Sigma\langle I\rangle$, where I is the observed intensity and $\langle I\rangle$ is the average intensity from multiple observations of symmetry-related reflections.
[2] Phasing power for isomorphous and anomalous acentric reflections = $\langle[|F_h(calc)/\text{phase-integrated lack of closure}]\rangle$.
[3] R_{free} was calculated with 8.8% of the data.

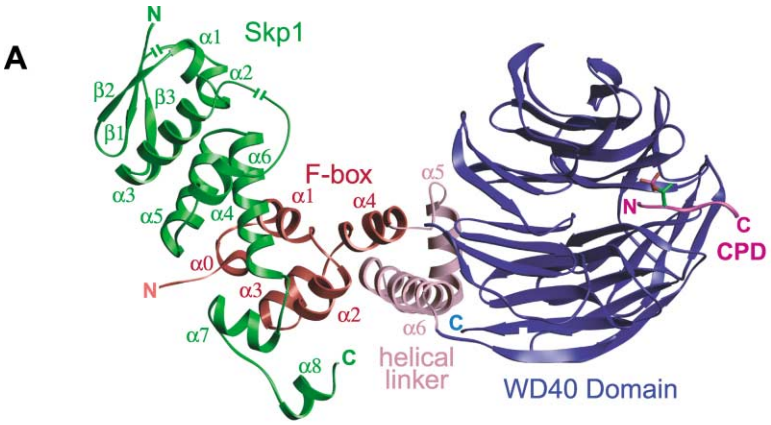

Figure 2. The Skp1-Cdc4-CPD Complex

(A) Ribbons representation of Skp1 (green), the F box domain (residues 274–319 in red), the helical linker region (residues 331–366 in pink), and the WD40 domain of Cdc4 (residues 367–744 in blue). The bound cyclin-E-derived CPD peptide is in purple with the phospho-Thr moiety shown in ball and stick representation. Positions of disordered loop regions are shown as ribbon breaks.

(B) The WD40 domain of Cdc4. β propeller blades are denoted PB1 to PB8. Ribbons and CPD peptide are colored as in (A). Position of the WD40 domain is identical to that in Figures 4A–4C.

(C) The structured linkage between the WD40 domain and the F box domain of Cdc4.

an F box domain, an α-helical linker, and a WD40 domain (Figures 2A–2C). The F box domain is comprised of five α helices, denoted α0 to α4. This topology differs slightly from that reported for the F box domain of hSkp2, which consists of a loop region, L1, and three helices, denoted α1 to α3 (Schulman et al., 2000). Helix α0 in Cdc4 corresponds most closely in sequence and position to the loop region L1 of hSkp2, while a half-turn remnant of Cdc4 helix α4 is discernable in the transition sequence between the hSkp2 F box and the LRR domain. As observed in the hSkp1-hSkp2 complex, Skp1 and the F box domain of Cdc4 associate by interdigitation of helices α0 to α3 of Cdc4 with helices α5 to α8 of Skp1, with the interface itself comprised of an interprotein four-helix bundle. This mode of association gives rise to a contiguous hydrophobic core that spans Skp1 and the F box domain of Cdc4. Superposition of the yeast and human structures reveals that Skp1 helix α8 and F box helix α4 deviate significantly in that only the first half of helix α8 is ordered in ScSkp1 and only a half-turn fragment of the F box helix α4 is apparent in hSkp2 (Figure 3A). The difference in position and length of F box helix α4 and Skp1 helix α8 reflects the different roles these secondary structure elements play in the linkage between their respective F box and ligand binding domains, as described below.

The WD40 Domain

Eight copies of the WD40 repeat motif in Cdc4 form an eight-blade β-propeller structure (Figure 2B). The WD40 repeat motif of approximately 40 residues composes the outer β strand of one propeller blade and the inner three strands of the adjacent blade in a continuous circular arrangement (Fulop and Jones, 1999). The actual Cdc4 structure contrasts to the seven-blade β-propeller predicted for Cdc4 and its orthologs based on previously solved WD40 domain structures, all of which contain only seven blades (Koepp et al., 2001; Nash et al., 2001). This discrepancy is attributable to the cryptic nature of the eigth WD40 repeat motif. Structure-based sequence alignment suggests that the WD40 domains of the F box proteins Met30 and β-TRCP will form canonical seven-blade β-propeller structures (Figure 1B). A variant 5-β strand structure occurs in blade two, in which a large insert in the β12-β13 linker allows the outermost β9[1] strand to run parallel to the β9 strand. This five-strand composition is unique to the fungal Cdc4 orthologs. In terms of overall structural dimensions, the WD40 domain resembles a conical frustum of 40 Å diameter top surface and 50 Å bottom surface, an overall thickness of 30 Å, and a central pore of 6 Å diameter. The CPD binding site resides on the top surface of the frustum and runs across the edge of the pore, while the bottom surface of the frustum links to the F box domain.

The F Box to WD40 Domain Linker

The F box domain of Cdc4 is followed by a helical extension that forms a structured bridge to the WD40 domain. The bridge consists of two α helices, α5 and α6, that together with helices α3 and α4 of the F box domain form a platform and stalk-like structure that positions the WD40 domain well away from the F box domain (Figures 2A and 2C). The relative orientation of the F box domain and WD40 domain is imposed almost entirely through the integrity of the stalk-like helix α6, which is 30 Å in length. The N-terminal end of helix α6 is anchored into the hydrophobic core of the F box domain through interactions involving α6 residues Phe355 and Leu356 and F box residues Ile295, Ile296, Leu315, Trp316, and Leu319 (Figure 2C). Helix α5 packs along side the base of helix α6 opposite to the F box domain through hydrophobic interactions involving Tyr342, Leu338 and

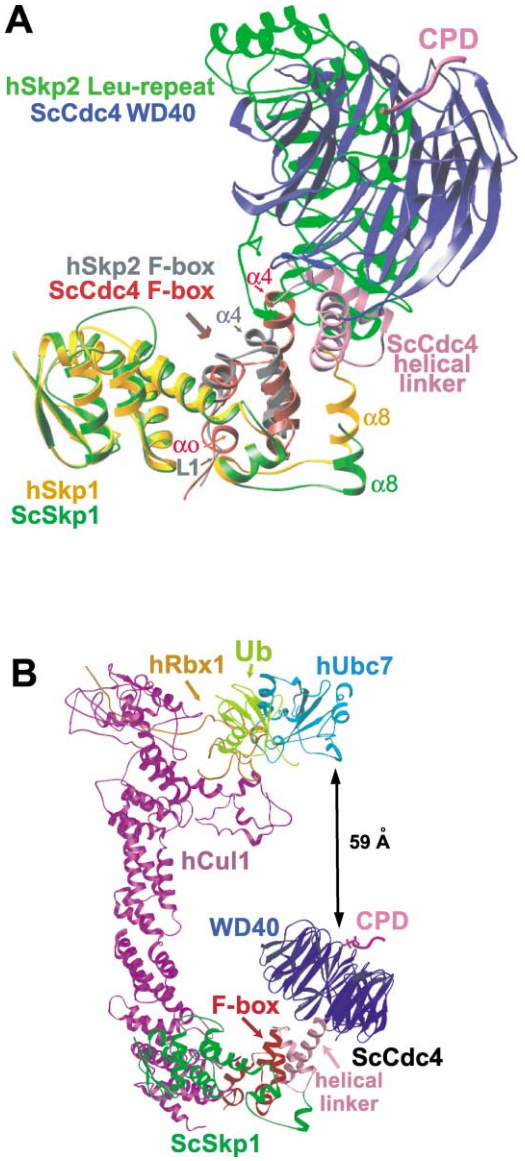

Figure 3. Substrate Orientation within the Skp1-Cdc4-CPD Complex

(A) Comparison of the ScSkp1-ScCdc4-CPD complex and the hSkp1-hSkp2 complex. Complexes were superimposed through a least squares optimization of Skp1 β strands 1 to 3 and α helices 1 to 6 (RMSD Cα = 0.74Å). Skp1 and F box secondary structure elements that deviate significantly in size and position between the two structures are labeled.

(B) Model of the ubiquitin-E2-SCF^{Cdc4}-CPD complex. The arrow indicates the 59 Å distance separating the phosphate group of the CPD and the active site cysteine of the E2.

Leu334. The C-terminal end of helix α6 inserts obliquely between propeller blades β7 and β8 of the WD40 domain through van der Wals and hydrophobic interactions involving residues Trp365 and Ile361 with WD40 domain residues Val687, Ile696, Leu726, and Phe743 in β propeller blades seven and eight. Asn364 of helix α6 also forms a tight hydrogen bond with the backbone carbonyl group of Phe743 in propeller blade eight. The conservation of many of these residues, with the possible exception of those within helix α5, suggests that a structured linkage between the WD40 and F box domains may be a common feature of the WD40 family F box proteins.

The interdomain connection between the F box and the WD40 domains of Cdc4 appears less rigid than the corresponding region in hSkp2 (Figure 3A). Outside of the stalk helix α6, only two close contacts (<3.5 Å) are observed between the WD40 domain and other regions of Cdc4 (Figure 2C). These contacts consist of hydrogen bonds between Asn684 and Arg700 in two loop regions of propeller blade seven with Glu323 in the α4-α5 linker of the helical extension. Both hydrogen bonds are maintained in the two Cdc4 molecules of the crystal asymmetric unit, but all three residues are poorly conserved amongst Cdc4 orthologs (Figure 1B). The lack of additional stabilizing interactions suggests that the F box to WD40 domain linker is not exceedingly rigid, and indeed, the WD40 domain in the two Cdc4 molecules of the crystal asymmetric unit differ relative to their F box domains by a 5° rotation about the long axis of helix α6. In contrast, in hSkp2, the F box domain helix α4 terminates abruptly in an immediate transition to the LRR domain fold such that the adjoined domains form a rigid hydrophobic core (Schulman et al., 2000). Although Skp2 and Cdc4 employ structurally divergent F box interfaces, the general position of the WD40 and LRR domains are nonetheless similar (Figure 3A).

Model of the SCF^{Cdc4} E2 Complex

The structure of the Skp1-Cdc4-CPD complex sheds light on how substrates are presented by the F box protein to the E2 for ubiquitin transfer. A complete model of the E2-SCF^{Cdc4}-substrate complex, consisting of ubiquitin, hUbc7, hCul1, hRbx1, ScCdc4, ScSkp1, and the CPD peptide, is shown in Figure 3B. This model is based on the reconstructed E2-SCF^{Skp2} complex derived by Pavletich and colleagues (Zheng et al., 2002) in conjunction with an NMR-based ubiquitin-E2 thioester model (Hamilton et al., 2001). Two interesting features are apparent. First, the distance between the E2 active site cysteine and the phosphate group of the bound CPD peptide is approximately 59 Å, which is similar to the spacing reported between the substrate interaction site and the E3 catalytic site in the hUbc7-Cbl structure (Zheng et al., 2000). Second, the WD40 domain presents the CPD peptide in a direct line of sight to the E2. Although the ligand binding site on hSkp2 has not been determined, mutagenesis studies on the LRR-containing F box protein Grr1 in yeast suggest that substrates bind to the inner side of the curved repeat surface (Hsiung et al., 2001). If the position of this site is maintained in hSkp2, then the LRR domain of Skp2 is predicted to project substrates in an orthogonal direction to that of the Cdc4 WD40 domain (Figure 3A).

Phosphopeptide Recognition

The CPD binding surface represents the most conserved part of the WD40 repeat domain structure (Figures 4A–4D). The central CPD sequence Leu-pThr-Pro-Pro was modeled unambiguously in unbiased experimental electron density maps in both Skp1-Cdc4-CPD complexes of the crystal asymmetric unit (see Supplemental Figure S1 online at http://www.cell.com/cgi/content/full/112/2/243/DC1). Interpretable electron density is also apparent for the P-2 Leu, P+3 Gln, P+4 Ser, and P+5 Gly positions, but only in one complex of the crystal asymmetric unit. The CPD peptide binds in an extended manner across β propeller blade two with the N terminus oriented toward the central pore of the WD40 domain and the C terminus oriented toward the outer rim. We observed identical substrate peptide orientations and contacts for an independent Skp1-Cdc4-CPD structure with a phosphopeptide derived from the transcription factor Gcn4, which is a physiological substrate of Cdc4 in yeast (Meimoun et al., 2000; Chi et al., 2001). However, of the Gcn4 peptide sequence Phe-Leu-Pro-pThr-Pro-Val-Leu-Glu-Asp, only the core residues Pro-pThr-Pro had discernable electron density (data not shown).

The CPD sequence requirements for interaction with Cdc4 are fully accounted for by structural elements in the WD40 domain. An absolute requirement for phosphorylation at Ser or Thr at the P-0 position of the CPD derives from a network of electrostatic interactions and hydrogen bonds that coordinate the P0 pThr phosphate group (Figures 4C and 4D). This interaction is mediated by residues that are conserved across all Cdc4 orthologs (Figure 1B). The P0 phosphate group forms direct electrostatic interactions with the guanidinium groups of Arg485, Arg467, and Arg534, and a direct hydrogen bond with the side chain of Tyr548. The side chain of Tyr548 is coordinated by stacking interactions with the guanidinium group of Arg572, which in turn is coordinated by a hydrogen bond to the side chain of Tyr574. Although Cdc4 shows a 6-fold preference for pThr over pSer (Nash et al., 2001), the structural basis for this selectivity is not obvious since the Cγ methyl group of Thr is directed toward solvent and does not make contact with the WD40 domain surface.

A second absolute requirement for CPD-Cdc4 interaction rests on the P+1 proline, the side chain of which projects into a three-sided pocket on the WD40 surface. One side of this pocket is formed by the side chain of Trp426, which packs in a coplanar manner with the P+1 proline side chain. The opposite side of this binding pocket is formed by the side chain of Arg485 via coordination of the proline side chain and backbone carbonyl group through van der Waals and hydrogen bonding interactions, respectively. The side chains of Thr441 and Thr465 define the remaining side of the P+1 proline binding pocket, with Cγ side chain groups composing a hydrophobic surface. The hydroxyl groups of Thr441 and Thr465 orient away from the P+1 binding pocket, where they are well placed to influence binding specificity for CPD residues C-terminal to the P+1 position. Unlike Trp426 and Arg485, which are invariant amongst the Cdc4 orthologs, Thr441 and Thr465 are both substituted with Ile in the *S. pombe* Cdc4 ortholog Pop1 (Figure 1B). This substitution might restrict CPD sequences able to bind Pop1 through steric or hydrophobic con-

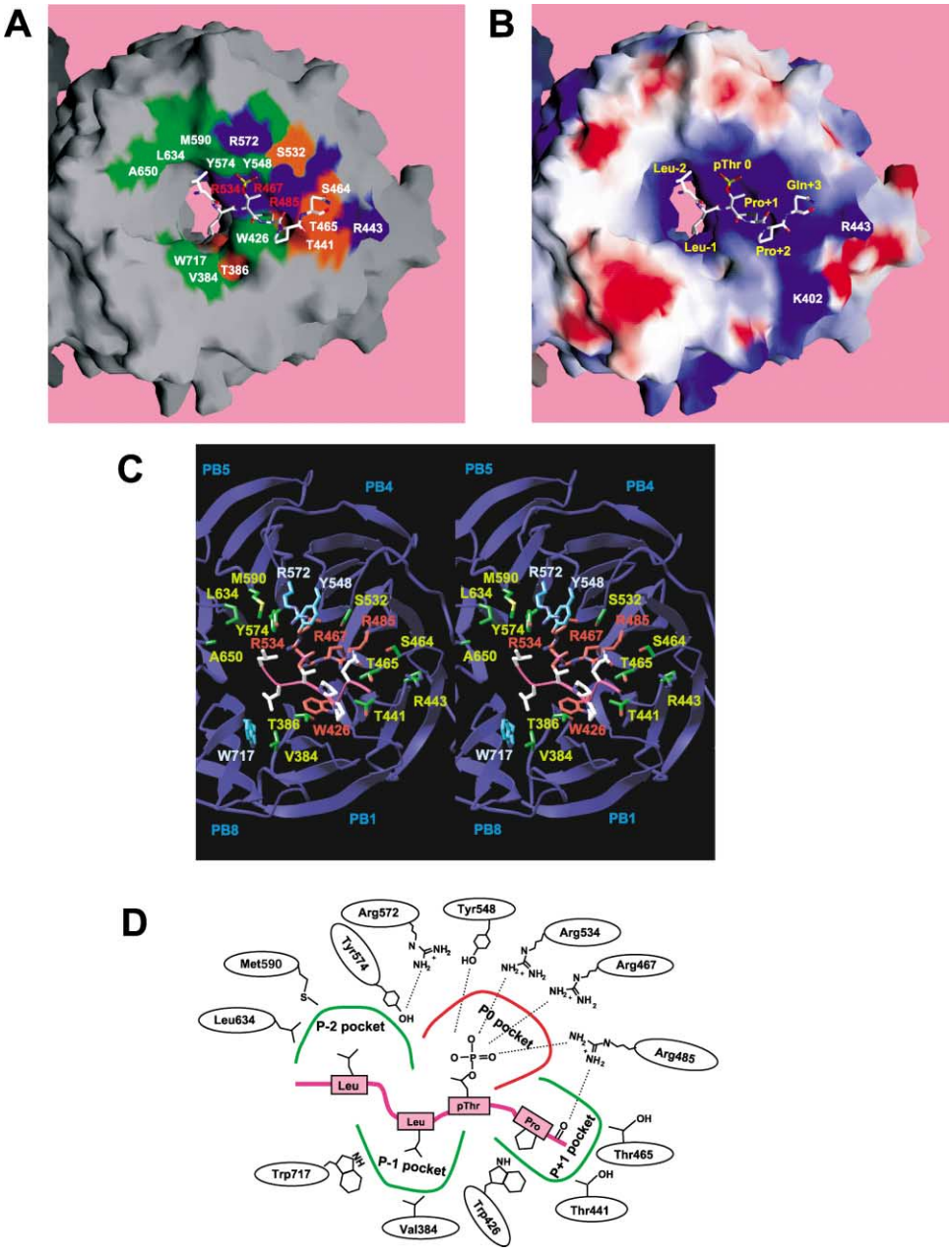

Figure 4. The CPD Binding Pocket of the WD40 Domain

(A) Surface representation of the CPD binding pocket, indicating invariant and highly conserved residues. Basic (blue), hydrophobic (green), and small polar residues (orange) are shown. The bound CPD is in ball and stick representation with carbon (white), nitrogen (blue), oxygen (red), and phosphorous (yellow) atoms shown.

(B) Surface representation of CPD binding region indicating electrostatic potential. Blue and red indicate regions of positive and negative potential, respectively, over the range 10 to -10 k_BT.

(C) Stereo ribbons representation of side chains and molecular interactions in the CPD binding pocket. Highly conserved and invariant side chains of Cdc4 and the CPD are displayed in ball and stick representation. Sites of mutation that give rise to severe and intermediate loss of function (see Figure 5) are colored red and blue, respectively; nonessential residues are colored green.

(D) Schematic of CPD binding pocket interactions with the CPD peptide.

straints on residues C-terminal to the P+1 proline position.

Cdc4 displays a strong preference for the hydrophobic residues Leu/Ile/Pro at the P-1 and Leu/Ile at the P-2 positions of the CPD. In the crystal structure, the P-1 Leucine side chain fits into a hydrophobic pocket composed of invariant residues Trp426, Trp717, and Thr386, and the conserved hydrophobic residue Val384. While less precisely modeled, the main chain position of Leu$-$2 lies in close proximity to a third hydrophobic pocket composed of the invariant residue Tyr574, and the conserved hydrophobic residues Met590 and Leu634. The

hydrophobic character of the P-1 and P-2 pockets is manifest as selection against both charged and small polar residues at these positions in the CPD consensus (Nash et al., 2001).

The WD40 phosphorecognition domain of Cdc4 is unusual in that it exhibits strong selectivity against either Arg or Lys residues in the P+2 to P+5 CPD positions but otherwise shows no sequence preference at these positions (Nash et al., 2001). In the crystal structure, the side chain of P+2 Pro is directed toward solvent, while the main chain conformation of Pro+1 and Pro+2 causes the CPD to kink away from the peptide binding surface from the Pro+2 position onward. As a result, only one additional contact with Cdc4 is made by the CPD following the Pro+1 position, namely a weak hydrogen bond with suboptimal geometry between the P+4 Gln side chain and the side chain of Arg485. Because the P+1 Pro main chain is forced away from the WD40 domain surface, the selection against basic residues in the P+2, +3, +4, and + 5 positions in the CPD consensus is almost certainly due to through-space electrostatic repulsion. This effect arises from a dominant-positive electrostatic potential generated by both the invariant triad of Arg residues that comprise the core pThr-Pro binding pocket and by a radial extension of the surface due to Arg572, Arg443, and Lys402, the former two of which are conserved amongst Cdc4 orthologs (Figures 1B and 4B).

A number of natural mutations detected in metazoan orthologs of Cdc4 corroborate the structure-based analysis. Two human ovarian cancer cell lines bear missense mutations at conserved Arg residues that correspond to Arg467 and Arg534 in yeast Cdc4 (Moberg et al., 2001). In the crystal structure, these residues make direct contact with the P0 phosphate group and are essential for function (Figures 4C and 4D). In a recent study of human primary endometrial tumors, mutations in phosphate binding Arg residues equivalent to Arg467 and Arg485 were detected in 2 of 13 tumor samples (Spruck et al., 2002). Other cancer-associated nonsense and frameshift mutations truncate hCdc4 within the WD40 domain (Moberg et al., 2001; Strohmaier et al., 2001; Spruck et al., 2002). Similarly, all three characterized mutations in the *Drosophila ago* gene that lead to excess cell proliferation affect the WD40 domain (Moberg et al., 2001). One of these mutations, Ala1118Val, corresponding to position Ser532 in ScCdc4, substitutes a conserved small residue with a bulkier residue at the center of the critical Arg485-Arg467-Arg534 triad (Figure 4C).

Mutational Analysis of the F Box to WD40 Domain Linker

To probe the importance of orientation and rigidity in the F box WD40 interdomain linker, we introduced point mutations, insertions, or deletions into the platform and stalk structure of Cdc4. None of these deletions affected the ability of the recombinant proteins to bind phospho-Sic1 in vitro or protein abundance in vivo (Figure 5A and data not shown). Introduction of the helix-destabilizing residues glycine and proline into helix α5 did not compromise Cdc4 function in vivo (Figure 5B), consistent with the poorly conserved nature of this region (Figure 1B). However, two different deletions of helix α5 eliminated Cdc4 function in vivo, indicating that the F box-WD40 domain interface is an essential structural component. Similarly, placement of helix-destabilizing residues at the center of helix α6 or the lengthening of this helix by the insertion of one, two, three, four, eight, or twelve amino acid residues disrupted Cdc4 function in vivo. Helix α6 is thus critical for productive orientation of the WD40 domain.

Mutational Analysis of the CPD Binding Surface

Previous mutational analysis based on sequence conservation in the Cdc4 family identified Arg467, Arg485, and Arg534 as essential for substrate binding and function in yeast (Nash et al., 2001). Two of the three corresponding residues in hCdc4, Arg417, and Arg457 are essential for the binding of phospho-cyclin E, while the third corresponding to Arg485 in yeast was not tested (Koepp et al., 2001). To systematically probe the role of residues that form the highly conserved peptide binding surface, we generated a panel of Cdc4 mutants and tested each for pSic1 binding in vitro, complementation of a *cdc4*Δ strain, and sensitivity to increased *SIC1* dosage. Four mutants, Arg467Ala, Arg485Ala, Arg534Ala, and Trp426Ala, were unable to bind phospho-Sic1 in vitro or complement a *cdc4*Δ strain but were fully competent for Skp1 binding (Figures 5A and 5B). The essential function of these residues is not confined to elimination of Sic1 because none of the corresponding mutant alleles were able to rescue a *cdc4*Δ *sic1*Δ strain (see Supplemental Figure S2 online at http://www.cell.com/cgi/content/full/112/2/243/DC1). These results reflect the critical structural role played by these residues in coordination of the P0 phosphate and the P+1 proline of the CPD. Mutation of the remaining phosphate-coordinating residue Tyr548 did not cause loss of viability but did result in dosage sensitivity to $SIC^{Thr33Val}$, which encodes a partially stabilized version of Sic1 (Figure 5C). Mutation of Arg572 had the same effect, as befits the observed stacking interaction between this residue and Tyr548. Although both mutants were severely impaired for binding to phospho-Sic1 in vitro, this effect may be exacerbated by the tendency of these recombinant proteins to aggregate (data not shown). In summary, the six residues that directly or indirectly coordinate the primary pThr-Pro core motif are critical for CPD recognition in vitro and Cdc4 function in vivo.

Disruption of residues that confer selection at the P-2, P-1, and P+2 to P+5 positions had only modest effects on the ability of Cdc4 to target pSic1. A Trp717Asn mutation predicted to disrupt the P-1 pocket conferred sensitivity to dosage of $SIC1^{Thr33Val}$, but did not overtly affect the pSic1-Cdc4 interaction in vitro. Individual mutations in all other residues that are well positioned to affect substrate selection, namely Arg443Ala, Arg443Asp, Lys402Ala, Tyr574Phe, and Val384Asn, were indistinguishable from wild-type in each of the assays used. Substrate selection residues on the WD40 surface thus contribute only modestly, if at all, to the essential function of Cdc4. As described below, however, these residues play a subtle but critical role in setting the phosphorylation threshold for the CPD-Cdc4 interaction.

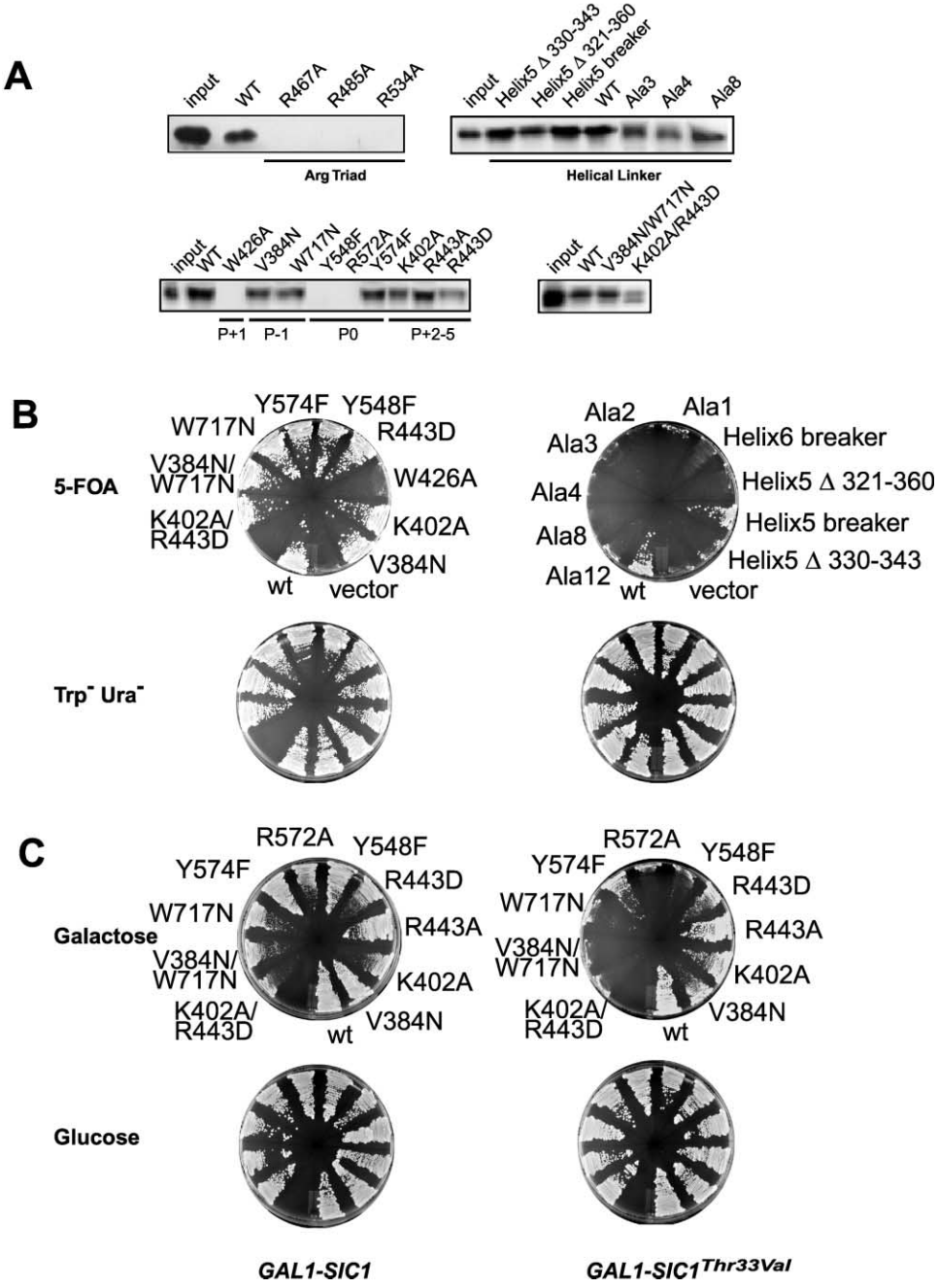

Figure 5. Structure-Guided Mutational Analysis of Cdc4

(A) Residues required for interaction of phospho-Sic1 and Cdc4 in vitro. Sic1 was phosphorylated with Cln2-Cdc28 kinase and captured onto resin loaded with either wild-type or the indicated mutant forms of the Skp1-Cdc4 complex (see text).

(B) Residues essential for Cdc4 function in vivo. Complementation of a *cdc4*Δ strain by the indicated alleles was assessed in a plasmid shuffle assay. The R485A, R467A, and R534A mutations in Cdc4 have been previously shown to disrupt function in vivo (Nash et al., 2001) and so are not shown.

(C) Effect of Cdc4 mutations on sensitivity to increased *SIC1* dosage. Strains bearing indicated *CDC4* alleles were tested for sensitivity to overexpression of wild-type *SIC1* and a partially stabilized version, $SIC1^{Thr33Val}$ from the *GAL1* promoter. Strains were incubated on galactose or glucose medium for 2 days at 30°C.

Modulation of CPD Substrate Selectivity

A critical feature of the Sic1-Cdc4 interaction is the requirement for phosphorylation of Sic1 on multiple sites. To enforce this requirement, each of the phosphorylation sites in the native Sic1 sequence are suboptimal in one or more respects (Figure 6A). The Cdc4-CPD structure suggests that selectivity against basic residues may be due to electrostatic repulsion generated from the conserved patch of basic residues in and around the CPD binding pocket, while selection for hy-

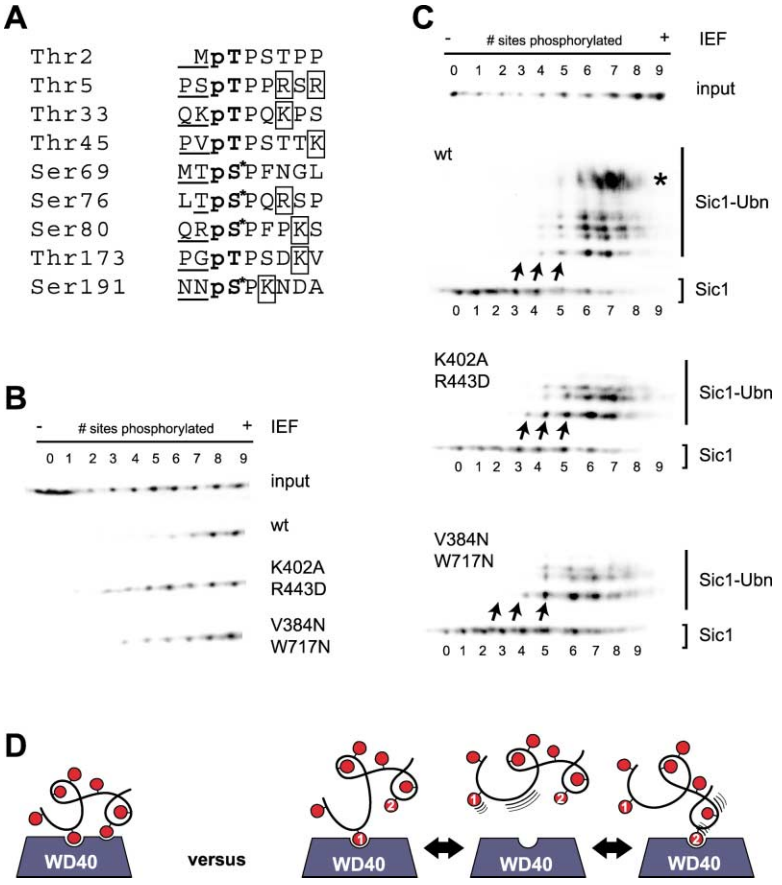

Figure 6. Modulation of the Multisite Requirement for Phospho-Sic1-Cdc4 Interaction

(A) All natural CPD sites in Sic1 deviate from the CPD consensus. Underlined residues indicate suboptimal residues at the P-1 and P-2 positions, boxed residues indicate suboptimal basic residues at the P+2 to P+5 positions, and asterisks indicate a suboptimal pSer at the P0 position.

(B) Capture of Sic1 phospho-isoforms by wild-type and mutant Cdc4. Pools of differentially phosphorylated Sic1 were captured on Skp1-Cdc4 resin, using either wild-type or the indicated mutant forms of Cdc4 compromised for selection at the P-1 position (V384N W717N) or the P+2 to P+5 positions (K402A R443D). The input and bound phospho-Sic1 isoform pools were resolved by denaturing IEF-2D gel electrophoresis and visualized by anti-Sic1 immunoblot.

(C) Ubiquitination of phospho-Sic1 isoforms by wild-type and mutant SCF^{Cdc4} complexes. Pools of differently phosphorylated Sic1 were incubated in solution with an equimolar amount of the indicated SCF^{Cdc4} complexes, Cdc34, ubiquitin, and ATP for 1 hr at 30°C. Input and reaction products were separated and visualized as in (B). Arrows indicate the less-phosphorylated forms of Sic1 captured by Cdc4 selection mutants. Asterisk indicates more extensively ubiquitinated species.

(D) Possible interaction mechanisms for single site and multisite dependent substrate binding to Cdc4. In a two-site cooperative interaction model (left), a primary high-affinity CPD binding site acts in conjunction with a secondary weak CPD binding site. The free energy for the two interactions is additive and so the overall K_d increases multiplicatively. In a single-site allovalent interaction (right), multiple low-affinity CPD sites engage a single CPD binding site on Cdc4 in equilibrium. The high local concentration of CPD sites increases the probability of binding such that Sic1 is unable to diffuse away from Cdc4 before rebinding occurs. The probability of rebinding increases as an exponential function of the number of CPD sites, thus accounting for the apparent cooperativity of the interaction.

drophobic residues arises from the P-1 pocket that is composed in part by Val384 and Trp717. To examine the basis for selection against suboptimal CPD motifs, we assessed the effects of mutations in nonessential residues in these two regions on the multisite phosphorylation requirement for Sic1 recognition.

We monitored the ability of Cdc4 to capture various phosphoisoforms of wild-type Sic1 from a pool of recombinant Sic1 that had been phosphorylated to various extents by Cln2-Cdc28. As resolved by isoelectric focusing, this pool contained roughly equal amounts of Sic1 phosphorylated on one, two, three, four, five, six, seven, eight, and nine sites. Wild-type Cdc4 was only able to capture Sic1 phosphorylated on six or more sites (Figure 6B). This result formally demonstrates the transition in binding affinity between five and six phosphorylation sites, as initially inferred from capture of a series of Sic1 phosphorylation site mutants by Cdc4 (Nash et al., 2001). The role of positive electrostatic potential in selecting against suboptimal CPD sequences with basic residues at C-terminal positions was tested with the Lys402Ala Arg443Asp double mutant. This mutant was able to select Sic1 phosphoisoforms that contained as few as three phosphorylation sites (Figure 6B). The ability of the Lys402Ala Arg443Asp double mutant to capture lower phosphorylated forms of Sic1 was also evident in one-dimensional SDS-PAGE (Figure 5A). Similarly, perturbation of the P-1 hydrophobic pocket with a Val384Asn Trp717Asn double mutation allowed capture of Sic1 phosphorylated on as few as four sites. These in vitro binding results were recapitulated in solution-based in vitro ubiquitination assays with wild-type and mutant forms of Cdc4. Both double mutant forms of Cdc4 were able to convert Sic1 phosphoryated on four or five sites to oligo-ubiquitinated species, whereas wild-type Cdc4 was unable to do so (Figure 6C). The double mutants were, however, less efficient than wild-type at elaborating fully ubiquitinated species of phospho-Sic1, perhaps because of protein stability effects or interference with catalytic steps after substrate binding. This interpretation is consistent with the sensitivity of strains bearing the double mutant alleles to $SIC1^{Thr33Val}$ dosage (Figure 5B). Overall, reengineering of negative selection residues in the Cdc4 WD40 domain supports the notion that the series of suboptimal CPD motifs in Sic1 sets a high phosphorylation threshold for its recognition by Cdc4.

Discussion

The structure of the Skp1-Cdc4-CPD complex provides direct visualization of substrate orientation within an SCF complex. Insights gained from the structure include

the unexpectedly frail interface between the F box and the WD40 repeat domain, the basis for dedicated pThr-Pro dipeptide recognition by a novel eight-blade WD40 propeller, and a detailed understanding of the basis for selection against natural CPD sequences. The latter feature appears to be tailored to enforce multisite phosphorylation-dependent degradation of Sic1, which in turn would help engender a highly cooperative onset of DNA replication (Nash et al., 2001). Similar principles may well operate for other Cdc4 substrates, including cyclin E, Notch^{IC}, and presenilin in mammalian cells (Strohmaier et al., 2001; Lai, 2002; Selkoe, 2001). Because yeast and human Cdc4 are structurally and functionally analogous (Nash et al., 2001; Strohmaier et al., 2001; Koepp et al., 2001), the structure of yeast Cdc4 affords obvious insights for pharmacological modulation of hCdc4 function in these pathways. Interestingly, a significant proportion of characterized human and fly *CDC4* mutations alter residues in the CPD binding pocket. Given the probable requirement for homodimerization in active SCF complexes (Wolf et al., 1999), such mutations might act in a partial dominant-negative manner to confer a growth advantage in the heterozygous state.

Phosphorecognition by Cdc4

The specificity of phosphorylation-dependent recognition by the WD40 domain of Cdc4 is governed by three main determinants: (1) a dedicated pThr-Pro binding pocket, (2) a deep hydrophobic pocket that selects hydrophobic residues N-terminal to the phosphorylation site, and (3) a through-space electrostatic selection against basic residues C-terminal to the phosphorylation site. As for all documented phosphodependent lipid/protein recognition modules, the Cdc4 WD40 domain employs arginine residues to directly contact the phosphate group of the ligand. However, unlike most domains in which adjacent residues impose subtle effects on specificity (Yaffe and Elia, 2001), the P+1 proline is an integral component of the core binding determinant (Nash et al., 2001). In the Cdc4-CPD cocrystal, ligand residues are locked in place by direct contact of the phosphate and proline carbonyl groups with three conserved and essential Arg residues, while the proline side chain inserts into a tight hydrophobic pocket formed by Trp426, Thr441, and Thr465. Because the phosphobinding pocket infrastructure has no obvious demarcation between the pThr and Pro binding sites, the Cdc4 WD40 domain is in effect a dedicated pThr-Pro binding module.

Comparison to Other Peptide Recognition Modules

Interesting parallels can be drawn between the Cdc4 WD40 domain, 14-3-3 domains, and the class IV WW domains, which all have the ability to recognize phospho-Ser/Thr epitopes in the context of adjacent proline residues (Yaffe and Elia, 2001). The interaction of the Pin1 class IV WW domain with a pSer-Pro peptide differs from Cdc4 in that it does not rely on an extensive network of Arg residues for phosphate coordination (Verdecia et al., 2000). However, a striking similarity between Pin1 and Cdc4 lies in the P+1 proline binding pocket, which in both cases depends on a highly conserved tryptophan side chain to engage the P+1 proline pyrrolidine ring through a coplanar interaction. In contrast to Cdc4, Pin1 actually displays a preference for Arg in the P+2 position, such that the binding specificity of the pSer-Pro recognition domain closely matches that of the targeting CDK enzymes.

14-3-3 domains bind pSer epitopes with a preference, but not an absolute requirement, for proline residues at the P+2 position (Yaffe et al., 1997). This less-stringent selection arises because the 14-3-3 proline binding pocket is able to accommodate other residues with propensity to form β turns. Interestingly, the proline preferences in both the 14-3-3 and Cdc4 WD40 domains give rise to the same qualitative effect: in each case the prolines terminate direct interactions between the peptide and the ligand binding domain by orienting the peptide away from the domain surface. In the case of Cdc4, biologically significant electrostatic effects operate in spite of the loss of direct peptide contact. As far as we are aware, physiologically relevant substrate antiselection mediated by charge repulsion is unique amongst known protein interaction modules.

The structure of the Cdc4 WD40 domain provides direct evidence that WD40-type repeats can assemble into propellers with more than seven blades (Fulop and Jones, 1999). WD40 domains are known to interact with other proteins in at least two different modes, either across the front face of the propeller, as in the case of Cdc4, or on the outer edge of the propeller, as in the case of clathrin (ter Haar et al., 2000). Modeling of the F box protein β-TrCP, which binds the doubly phosphorylated consensus motif DpSGXXpS in IκBα, β-catenin, and Vpu (Yaffe and Elia, 2001), reveals an extensive conserved basic region on the front face of the propeller, which may engage substrate phosphoepitopes in an analogous manner to Cdc4 (data not shown).

Spatial Orientation of SCF Substrates

A conserved feature between all E3 structures solved to date is the substantial distance between the substrate binding site and the catalytic site (Huang et al., 1999; Zheng et al., 2000, 2002). Superposition of the Skp1-Cdc4 complex onto a model of the Skp1-Cul1-Rbx1-E2-ubiquitin complex suggests that the substrate is positioned for direct frontal attack by the E2 catalytic site but that a gap of some 59Å between the two sites must be bridged, presumably by the substrate polypeptide. The disordered structure of Sic1 lends itself to this possibility (Nash et al., 2001). Intriguingly, overlay of the WD40 domain of Cdc4 with the LRR of Skp2 does not align the defined phosphopeptide binding pocket of Cdc4 with a potential phosphorecognition site on the concave face of the LRR repeats (Zheng et al., 2002), at least as defined by mutational analysis of the related F box protein Grr1 in yeast (Hsiung et al., 2001). If the relative position of substrates in the WD40 versus LRR class of F box proteins do in fact differ, spatial leeway in substrate presentation must be possible.

Based on the extensive Skp1-Skp2 interface and on the inactivation of Cul1 by insertion of a flexible linker, it has been proposed that SCF complexes and perhaps E3 enzymes in general must present substrates to the catalytic site in a rigidly defined fashion (Zheng et al., 2002). Unexpectedly, the WD40 domain and the F box of Cdc4 are linked only by a single α-helical stalk, with

very limited additional contacts. Despite relatively low sequence conservation in the α helix 6 structure that supports the WD40 domain, spatial constraints are nevertheless evident, as shown by the sensitivity of the structure to rotational and translational shifts caused by insertion of additional residues into the stalk. It is also possible that regions truncated from Cdc4 to enable crystallization may normally help stabilize the interdomain interface.

Cooperativity in Substrate Selection by Cdc4

The properties of the Cdc4 phosphopeptide binding module differ from those of other known modules in the important respect that the interaction with core recognition elements is partially offset by specific selection against basic residues in the substrate peptide. This feature establishes an intrinsic antagonism between the recognition mechanism and the targeting CDK kinases, which prefer Ser/Thr-Pro sites with C-terminal basic residues (Endicott et al., 1999). Significantly, all of the natural CPD motifs in Sic1 contain one or more mismatches to the optimal CPD consensus. This system, based on positive and negative ligand selection, may not only set an elevated threshold for kinase activity, but may also allow the threshold to be precisely tuned for any given substrate by varying the number, spacing, and properties of each site. Thus, Cdc4 is able to target numerous critical factors for phosphorylation-dependent degradation, including the Cdk inhibitor Sic1, the CDK inhibitor and polarization factor Far1, the replication initiator Cdc6, and the transcription factor Gcn4, all of which may be controlled with different kinetics and different phosphorylation thresholds (Patton et al., 1998). In one extreme, typified by Gcn4 and cyclin E, the substrate may contain a high-affinity site that is augmented by several minor low-affinity sites (Meimoun et al., 2000; Chi et al., 2001; Strohmaier et al., 2001). In the other, more akin to Sic1, a large number of weak sites may cooperate to drive high-affinity binding only when a phosphorylation threshold is reached. As shown here, mutation of either the distal basic selection region or the P-1 pocket in Cdc4 shifts the binding equilibrium to lower phosphorylated forms of Sic1, which, in the absence of other structural effects that may compromise Cdc4, would be predicted to cause premature DNA replication and genome stability (Nash et al., 2001). These features distinguish Cdc4 from other known phospho-peptide binding modules characterized to date that typically interact with dedicated sites on their substrates through a single high-affinity interaction.

The mechanism that underlies the cooperative binding transition of the phospho-Sic1-Cdc4 interaction between five and six phosphorylation sites remains to be determined. In principle, multiple interactions sites might increase binding by engaging more than one binding site on Cdc4 (Figure 6D). This type of cooperative interaction is common in biological systems, as in the avidity of antibodies for polyvalent ligands and pathogen-host interactions (Mammen et al., 1998). Analogous cooperative binding interactions occur in signaling pathways. For instance, the dual SH2 domain phosphatase SH-PTP2 and the 14-3-3ζ protein both engage two substrate binding sites on their respective ligands (Eck et al., 1996; Yaffe et al., 1997). However, inspection of the Cdc4 WD40 domain surface does not reveal any obvious ligand binding sites that might accommodate a second phosphorylated peptide motif, nor is there any biochemical evidence for secondary binding sites (Nash et al., 2001). In addition, the wide range of substrates and site spacing accommodated by Cdc4, including random concatamers of synthetic CPD sites (Nash et al., 2001), is a priori difficult to explain by two or more fixed binding sites on Cdc4.

Instead, we favor a model that requires only a single phosphodependent binding site on Cdc4 (Figure 6D). In this scheme, phosphorylation of multiple CPD sites on Sic1 increases the local concentration of sites around Cdc4 once the first CPD site is bound, to the point where diffusion-limited escape from the receptor is overwhelmed by the probability of rebinding of any one CPD site. In a sense, Sic1 becomes kinetically trapped in close proximity to Cdc4. Mathematical modeling of an idealized polyvalent ligand-monovalent receptor interaction indicates that the rate of ligand escape from the receptor exhibits a negative exponential dependence on the number of ligand sites (P. Klein, personal communication). We propose the term allovalent to describe the ability of multiple weak, spatially separated ligand sites to cooperatively interact with a single receptor site. The prevalence of multisite phosphorylation (Cohen, 2000) and indeed of polyvalent ligands in general (Mammen et al., 1998) suggests that this type of behavior may underlie many biological processes.

Experimental Procedures

Protein Expression and Purification

The Cdc4 fragment employed for crystallization was deleted for residues 1–262, 602–605, 609–624, and 745–779 to remove loop regions based on sequence alignments and limited proteolysis of the intact SCF^{Cdc4} complex. Skp1 was deleted for a nonconserved loop insertion spanning residues 37–64. A ^{GST}Skp1-^{His6}Cdc4 complex was coexpressed from plasmid pMT3169 in B934 (DE3) bacterial strain (Stratagene) cells grown in minimal media supplemented with a mixture of selenomethionine (40 μg/ml) and methionine (0.4 ug/ml) and purified by double affinity tag chromatography (Nash et al., 2001). All mutations were constructed by standard methods, using oligonucleotides listed in Supplemental Table S1 (online at http://www.cell.com/cgi/content/full/112/2/243/DC1) and sequence verified in their entirety. Mutants were subcloned into pMT3055 or pMT3217 for expression in bacteria or yeast, respectively, as listed in Supplemental Table S2 (online at http://www.cell.com/cgi/content/cull/112/2/243/DC1). Cdc4 mutant proteins Ala1, Ala2, Ala12, and helix α6 breaker could not be stably expressed in bacteria; the Ala12 mutant also could not be expressed in yeast.

Crystallization, Data Collection, Structure Determination, and Modeling

Hanging drops containing 1 μl of 20 mg/ml protein and 1.2 molar equivalents of the cyclin-E-derived CPD peptide (acetyl-Gly-Leu-Leu-pThr-Pro-Pro-Gln-Ser-Gly-amide) in buffer (10 mM HEPES [pH 7.5], 250 mM NaCl, and 1 mM DTT) were mixed with equal volume of reservoir buffer (0.1 M Tris [pH 8.5] and 1.5 M ammonium sulfate). Crystals of the space group $P3_2$, (a =107.7Å, b = 107.7Å, c = 168.3Å, $\alpha = \gamma = 90°$, β =120°), with two Cdc4-Skp1-CPD complexes in the asymmetric unit, were obtained at 20°C. A Multiple Anomalous Dispersion (MAD) experiment was performed on a frozen crystal at the Advanced Photon Source (Argonne, IL) beamline BM 14-C and BM 14-D ($\lambda1$ = 0.9798 Å, $\lambda2$ = 0.9800 Å, $\lambda3$ = 0.9000 Å), using a Quantum 4 ADSC CCD detector. Data processing and reduction were carried out with the HKL program suite (Otwinowski and Minor,

1997). The programs SHARP (de La Fortelle and Bricogne, 1997) and SnB (Miller et al., 1994) were used in combination to locate and refine 19 of the total 22 selenium sites. Following phasing and density modification, a model was built using O (Jones et al., 1991) and refined to 2.7 Å resolution with NCS restraints using CNS (Brunger et al., 1998) to a working R_{value} of 23.8% and R_{free} of 27.3%. Pertinent statistics for data collection and refinement are shown in Table 1. Amino acids 37–74, 104–115, and 185–194 of Skp1 and amino acids 263–269 and 497–507 of Cdc4 were disordered and could not be modeled. 89.1% of the residues occupy the most favored regions of the Ramachandran plot, 10.8% the additional allowed region, and 0.2% the generously allowed region.

Ribbons representations were generated using Ribbons (Carson, 1991), surface representations were generated using Grasp (Nicholls et al., 1991), and electron density maps were generated using O (Jones et al., 1991). A model of the ubiquitin-E2-SCF^{Cdc4}-CPD complex was generated by superposition of the Skp1 subunits of the Skp1-Cdc4-CPD structure and the Skp1-Cul1-Rbx1 structure (PDB ID 1LDK) (Zheng et al., 2002), the RING finger domains from Rbx1 in the same Skp1-Cul1-Rbx1 complex and from the Cbl subunit of the Cbl-UbcH7 structure (PDB ID 1FBV) (Zheng et al., 2000), and the E2 subunits of the Cbl-UbcH7 structure and an NMR-based Ubc1-ubiquitin model (PDB ID 1FXT) (Hamilton et al., 2001). The Skp1, RING domain, and E2 subunits overlapped with RMSD values of 1.01 Å, 2.09 Å, and 2.04 Å, respectively.

Cdc4 Functional Assays

CDC4 mutant alleles were assessed for complementation of a *cdc4Δ* strain in a plasmid shuffle assay (Nash et al., 2001). Sensitivity to *SIC1* dosage was determined by transformation with pMT837 (*GAL1-SIC1*) or pMT767 (*GAL1-SIC1^{Thr33Val}*) and plating on glucose medium or galactose medium. For in vitro capture of phospho-Sic1 by Cdc4, 0.5 μg of bacterially expressed ^{HIS6}Sic1 was phosphorylated with immobilized Cln2-Cdc28 kinase from baculovirus-infected Sf9 cells and then incubated with 1 μg of immobilized wild-type or mutant $^{His}Cdc4^{263-744\_GST}$-Skp1 at 4°C for 1 hr, washed four times, and visualized by anti-Sic1 immunoblot. For isoelectric focusing (IEF)-2D gel analyses, an evenly distributed pool of phospho-Sic1 isoforms was generated by combining different time points in a Sic1 phosphorylation reaction. 2.5 μg of the phospho-Sic1 pool was bound to 5 μg of immobilized wild-type or mutant $Cdc4^{1-744}$-GST-Skp1. Captured isoforms were separated by denaturing IEF-2D gel electrophoresis using pH3-10NL Immobiline gel strips (Amersham) and visualized by anti-Sic1 immunoblot. Alternatively, the pool of phospho-Sic1 isoforms was incubated in solution with a ubiquitination reaction mix containing ATP, ubiquitin, yeast E1, Cdc34, and either wild-type or mutant SCF^{Cdc4} complex, composed of a 1:1 ratio of bacterial Cdc4-GST-Skp1 and insect cell-produced Cdc53-Rbx1, at 30°C for 1 hr as previously described (Nash et al., 2001).

Acknowledgments

We thank Piers Nash for providing phosphopetides; Zhen Lin for assistance with protein production; and Tony Pawson, Peter Klein, Jim Ferrell, Dan Durocher, Angelika Amon, Wade Harper, Ray Deshaies, and Lea Harrington for stimulating discussions. We also thank the BioCars staff at the Advanced Photon Source at Argonne National Laboratories for assistance with collection of diffraction data. This work was supported by grants from the Canadian Institutes of Health Research, the National Cancer Institute of Canada (NCIC), the Ontario Research and Development Challenge Fund, and MDS-Sciex. M.T. holds a Canada Research Chair in Biochemistry and F.S. is a Research Scientist of the NCIC.

Received: November 1, 2002
Revised: December 16, 2002

References

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebl, M., Harper, J.W., and Elledge, S.J. (1996). SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. Cell *86*, 263–274.

Brunger, A.T., Adams, P.D., Clore, G.M., Gros, P., Grosse-Kuntsleve, Jiang, R.W., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Carson, M. (1991). Ribbons 2.0. J. Appl. Crystallgr. *24*, 958–961.

Chi, Y., Huddleston, M.J., Zhang, X., Young, R.A., Annan, R.S., Carr, S.A., and Deshaies, R.J. (2001). Negative regulation of Gcn4 and Msn2 transcription factors by Srb10 cyclin-dependent kinase. Genes Dev. *15*, 1078–1092.

Cohen, P. (2000). The regulation of protein function by multisite phosphorylation: a 25 year update. Trends Biochem. Sci. *25*, 596–601.

de La Fortelle, E., and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472–494.

Eck, M.J., Pluskey, S., Trub, T., Harrison, S.C., and Shoelson, S.E. (1996). Spatial constraints on the recognition of phosphoproteins by the tandem SH2 domains of the phosphatase SH-PTP2. Nature *379*, 277–280.

Endicott, J.A., Noble, M.E., and Tucker, J.A. (1999). Cyclin-dependent kinases: inhibition and substrate recognition. Curr. Opin. Struct. Biol. *9*, 738–744.

Feldman, R.M., Correll, C.C., Kaplan, K.B., and Deshaies, R.J. (1997). A complex of Cdc4p, Skp1p, and Cdc53p/cullin catalyzes ubiquitination of the phosphorylated CDK inhibitor Sic1p. Cell *91*, 221–230.

Ferrell, J.E. (1996). Tripping the switch fantastic: how a protein kinase cascade can convert graded inputs into switch-like outputs. Trends Biochem. Sci. *21*, 460–466.

Fulop, V., and Jones, D.T. (1999). Beta propellers: structural rigidity and functional diversity. Curr. Opin. Struct. Biol. *9*, 715–721.

Hamilton, K.S., Ellison, M.J., Barber, K.R., Williams, R.S., Huzil, J.T., McKenna, S., Ptak, C., Glover, M., and Shaw, G.S. (2001). Structure of a conjugating enzyme-ubiquitin thiolester intermediate reveals a novel role for the ubiquitin tail. Structure *9*, 897–904.

Harper, J.W. (2001). Protein destruction: adapting roles for Cks proteins. Curr. Biol. *11*, R431–R435.

Harper, J.W., Burton, J.L., and Solomon, M.J. (2002). The anaphase-promoting complex: it's not just for mitosis any more. Genes Dev. *16*, 2179–2206.

Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. Annu. Rev. Biochem. *67*, 425–479.

Hsiung, Y.G., Chang, H.C., Pellequer, J.L., La Valle, R., Lanker, S., and Wittenberg, C. (2001). F-box protein Grr1 interacts with phosphorylated targets via the cationic surface of its leucine-rich repeat. Mol. Cell. Biol. *21*, 2506–2520.

Huang, L., Kinnucan, E., Wang, G., Beaudenon, S., Howley, P.M., Huibregtse, J.M., and Pavletich, N.P. (1999). Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2–E3 enzyme cascade. Science *286*, 1321–1326.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the localization of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kaelin, W.G., Jr. (2002). Molecular basis of the VHL hereditary cancer syndrome. Nat. Rev. Cancer *2*, 673–682.

Keyomarsi, K., Tucker, S.L., Buchholz, T.A., Callister, M., Ding, Y., Hortobagyi, G.N., Bedrosian, I., Knickerbocker, C., Toyofuku, W., Lowe, M., et al. (2002). Cyclin E and survival in patients with breast cancer. N. Engl. J. Med. *347*, 1566–1575.

Koepp, D.M., Schaefer, L.K., Ye, X., Keyomarsi, K., Chu, C., Harper, J.W., and Elledge, S.J. (2001). Phosphorylation-dependent ubiquitination of cyclin E by the SCF^{Fbw7} ubiquitin ligase. Science *294*, 173–177.

Lai, E.C. (2002). Protein degradation: four E3s for the notch pathway. Curr. Biol. *12*, R74–R78.

Lengronne, A., and Schwob, E. (2002). The yeast CDK inhibitor Sic1

prevents genomic instability by promoting replication origin licensing in late G1. Mol. Cell *9*, 1067–1078.

Mammen, M., Choi, S.-K., and Whitesides, G.M. (1998). Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. Angew. Chem. *37*, 2754–2794.

Meimoun, A., Holtzman, T., Weissman, Z., McBride, H.J., Stillman, D.J., Fink, G.R., and Kornitzer, D. (2000). Degradation of the transcription factor Gcn4 requires the kinase Pho85 and the SCF^{CDC4} ubiquitin-ligase complex. Mol. Biol. Cell *11*, 915–927.

Miller, R., Gallo, S.M., Khalak, H.G., and Weeks, C.M. (1994). SnB: crystal structure determination via Shake-and-Bake. J. Appl. Crystallgr. *27*, 613–621.

Moberg, K.H., Bell, D.W., Wahrer, D.C., Haber, D.A., and Hariharan, I.K. (2001). Archipelago regulates Cyclin E levels in Drosophila and is mutated in human cancer cell lines. Nature *413*, 311–316.

Nash, P., Tang, X., Orlicky, S., Chen, Q., Gertler, F.B., Mendenhall, M.D., Sicheri, F., Pawson, T., and Tyers, M. (2001). Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. Nature *414*, 514–521.

Nicholls, A., Sharp, K.A., 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. Methods Enzymol. *276*, 307–326.

Patton, E.E., Willems, A.R., and Tyers, M. (1998). Combinatorial control in ubiquitin-dependent proteolysis: don't Skp the F box hypothesis. Trends Genet. *14*, 236–243.

Pickart, C.M. (2001). Mechanisms underlying ubiquitination. Annu. Rev. Biochem. *70*, 503–533.

Schulman, B.A., Carrano, A.C., Jeffrey, P.D., Bowen, Z., Kinnucan, E.R., Finnin, M.S., Elledge, S.J., Harper, J.W., Pagano, M., and Pavletich, N.P. (2000). Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. Nature *408*, 381–386.

Schwob, E., Bohm, T., Mendenhall, M.D., and Nasmyth, K. (1994). The B-type cyclin kinase inhibitor $p40^{SIC1}$ controls the G1 to S transition in *S. cerevisiae*. Cell *79*, 233–244.

Selkoe, D.J. (2001). Presenilin, Notch, and the genesis and treatment of Alzheimer's disease. Proc. Natl. Acad. Sci. USA *98*, 11039–11041.

Skowyra, D., Craig, K.L., Tyers, M., Elledge, S.J., and Harper, J.W. (1997). F box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. Cell *91*, 209–219.

Spruck, C.H., Won, K.A., and Reed, S.I. (1999). Deregulated cyclin E induces chromosome instability. Nature *401*, 297–300.

Spruck, C.H., Strohmaier, H., Sangfelt, O., Muller, H.M., Hubalek, M., Muller-Holzner, E., Marth, C., Widschwendter, M., and Reed, S.I. (2002). *hCDC4* gene mutations in endometrial cancer. Cancer Res. *62*, 4535–4539.

Strohmaier, H., Spruck, C.H., Kaiser, P., Won, K.A., Sangfelt, O., and Reed, S.I. (2001). Human F box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line. Nature *413*, 316–322.

ter Haar, E., Harrison, S.C., and Kirchhausen, T. (2000). Peptide-in-groove interactions link target proteins to the beta-propeller of clathrin. Proc. Natl. Acad. Sci. USA *97*, 1096–1100.

Varshavsky, A. (1991). Naming a targeting signal. Cell *64*, 13–15.

Verdecia, M.A., Bowman, M.E., Lu, K.P., Hunter, T., and Noel, J.P. (2000). Structural basis for phosphoserine-proline recognition by group IV WW domains. Nat. Struct. Biol. *7*, 639–643.

Verma, R., Annan, R.S., Huddleston, M.J., Carr, S.A., Reynard, G., and Deshaies, R.J. (1997). Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase. Science *278*, 455–460.

Wolf, D.A., McKeon, F., and Jackson, P.K. (1999). F box/WD-repeat proteins Pop1p and Sud1p/Pop2p form complexes that bind and direct the proteolysis of Cdc18p. Curr. Biol. *9*, 373–376.

Yaffe, M.B., and Elia, A.E. (2001). Phosphoserine/threonine-binding domains. Curr. Opin. Cell Biol. *13*, 131–138.

Yaffe, M.B., Rittinger, K., Volinia, S., Caron, P.R., Aitken, A., Leffers, H., Gamblin, S.J., Smerdon, S.J., and Cantley, L.C. (1997). The structural basis for 14-3-3:phosphopeptide binding specificity. Cell *91*, 961–971.

Zheng, N., Wang, P., Jeffrey, P.D., and Pavletich, N.P. (2000). Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. Cell *102*, 533–539.

Zheng, N., Schulman, B.A., Song, L., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M., et al. (2002). Structure of the Cul1-Rbx1-Skp1-Fbox-Skp2 SCF ubiquitin ligase complex. Nature *416*, 703–709.

Accession Numbers

Coordinates have been deposited in the Protein Data Bank (accession code 1NEX).