


RESEARCH ARTICLE

Structural study of the Cdc25 domain from Ral-specific guanine-nucleotide exchange factor RalGPS1a

Wei Peng*, Jiwei Xu*, Xiaotao Guan, Yao Sun, Xuejun C. Zhang, Xuemei Li, Zihe Rao 

National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

 Correspondence: raozh@xtal.tsinghua.edu.cn

Received March 10, 2011 Accepted March 31, 2011

ABSTRACT

The guanine-nucleotide exchange factor (GEF) RalGPS1a activates small GTPase Ral proteins such as RalA and RalB by stimulating the exchange of Ral bound GDP to GTP, thus regulating various downstream cellular processes. RalGPS1a is composed of an N-terminal Cdc25-like catalytic domain, followed by a PXXP motif and a C-terminal pleckstrin homology (PH) domain. The Cdc25 domain of RalGPS1a, which shares about 30% sequence identity with other Cdc25-domain proteins, is thought to be directly engaged in binding and activating the substrate Ral protein. Here we report the crystal structure of the Cdc25 domain of RalGPS1a. The bowl shaped structure is homologous to the Cdc25 domains of SOS and RasGRF1. The most remarkable difference between these three Cdc25 domains lies in their active sites, referred to as the helical hairpin region. Consistent with previous enzymological studies, the helical hairpin of RalGPS1a adopts a conformation favorable for substrate binding. A modeled RalGPS1a-RalA complex structure reveals an extensive binding surface similar to that of the SOS-Ras complex. However, analysis of the electrostatic surface potential suggests an interaction mode between the RalGPS1a active site helical hairpin and the switch 1 region of substrate RalA distinct from that of the SOS-Ras complex.

KEYWORDS RalGPS1a, RalA, cdc25 domain, crystal structure

INTRODUCTION

Ral proteins (RalA and RalB) are members of the small GTPase Ras superfamily (Chardin and Tavittian, 1986) and play an essential role in a variety of physiological and pathological processes in mammalian cells, including exocytosis (Moskalenko et al., 2002; Cascone et al., 2008), cell proliferation (Chien and White, 2003; Lim et al., 2005), and oncogenic transformation (Rangarajan et al., 2004). As with other members of the Ras family, Ral cycles between its activated GTP-bound form and inactivated GDP-bound form in the cytoplasm. The exchange of a Ral-bound GDP molecule with GTP, thus activating Ral, is catalyzed by Ral-specific guanine-nucleotide exchange factors (GEFs), which are subdivided into the RalGDS and RalGPS families. Members of the RalGDS family, including RalGDS, Rgl, and Rlf (Albright et al., 1993; Murai et al., 1997; Wolthuis et al., 1997), contain a Ras binding domain (RBD) in their C-terminal region and are proposed to be stimulated by GTP-bound Ras. In contrast, the RalGPS family lacks an RBD in their sequences and may respond to other upstream stimuli independent of Ras activation (de Bruyn et al., 2000; Rebhun et al., 2000; Ceriani et al., 2007).

Two members of the RalGPS family have been identified in humans, namely RalGPS1a and RalGPS1b. They are translated from two splicing variants of the same premature mRNA and share a similar domain organization (Rebhun et al., 2000; Quilliam, 2006). An additional homological protein named RalGPS2 was identified from mouse (Rebhun et al., 2000). RalGPS contains a 30 kDa N-terminal catalytic domain (known as the Cdc25 domain), which shares about 30%

*These authors contributed equally to this work.

sequence identity with the equivalent domain of other Ras GEFs and the *Saccharomyces cerevisiae* Cdc25 (Broek et al., 1987). A proline rich region (about 15 residues in length), termed the PXXP motif, exists in the central part of the full length peptide sequence. This PXXP motif has been shown to interact with SH3 domains of the adaptor proteins Grb2 and Nck, and may recruit RalGPS into the tyrosine kinase receptor-Grb2-GEF pathway (Rebhun et al., 2000). Moreover, a pleckstrin homology (PH) domain is located in the C-terminal region of RalGPS. Evidence has shown that this PH domain directly interacts with certain components of the plasma membrane, such as phosphatidylinositol 4,5-bisphosphate (PIP2) and phosphatidylinositol 1,4,5-trisphosphate (PIP3), and is required for the membrane association of RalGPS1a and its biological activities in cells (Rameh et al., 1997; de Bruyn et al., 2000). The REM (Ras exchange motif) domain, normally located on the N-terminal side of the Cdc25 domain in the RalGDS family and other Ras GEFs, such as SOS, RasGRF1 (Freedman et al., 2006), and Epac2 (de Rooij et al., 2000), is absent in RalGPS proteins. The REM domain has been shown to stabilize the Cdc25 domain and to regulate its activity (Boriack-Sjodin et al., 1998). The lack of an REM domain in RalGPS may imply a different regulatory mechanism for its Cdc25 domain. For example, the C-terminal PH domain is suggested to play a similar role as the REM domain in RalGPS (de Bruyn et al., 2000; Ceriani et al., 2007), but this hypothesis remains to be verified experimentally. Taken together, the domain organization of the RalGPS family indicates its similar catalytic activity to, but distinct regulatory mechanism from, other Ral-specific GEFs.

The three-dimensional (3D) structures of several Cdc25-containing GEF proteins have been reported. The crystal structures of Ras-GEF SOS and its complex with a nucleotide free Ras in the active site have illustrated the substrate binding site and the catalytic mechanism of the Cdc25 domain. Briefly, the binding of SOS to the substrate Ras dramatically disrupts the conformation of the Ras nucleotide binding site, including the P loop, the switch 1 and switch 2 regions, and thus catalyzes the nucleotide release (Boriack-Sjodin et al., 1998). A positive feedback mechanism of SOS was later proposed based on the structure of a Ras-GTP complex bound to a cleft between the cdc25 and REM domains of SOS, which is distal to the active site of SOS (Margarit et al., 2003). Subsequently, the crystal structures of Rap GEF Epac2 (Rehmann et al., 2006) and Ras GEF RasGRF1 (Freedman et al., 2006) revealed a similar fold in their Cdc25 domains to that of SOS. Nevertheless, with the exception of the SOS-Ras complex, there have been no reported structures of other Ras GEF family proteins in complex with their substrates, and thus the structural basis for the substrate specificity of the GEF proteins remains elusive. Here we report the crystal structure of the RalGPS1a Cdc25 domain, which is the first Cdc25 domain from the Ral specific

GEFs. While it shares high similarity with other known Cdc25 domain structures, as expected from their high sequence homology, the RalGPS1a cdc25 domain also possesses several unique properties that may imply a novel mechanism of substrate recognition.

RESULTS

Overall features of the RalGPS1a Cdc25 domain structure

In order to understand the structural basis for RalGPS activity and substrate specificity, we determined the crystal structure of the RalGPS1a Cdc25 domain at 2.2 Å resolution using the molecular replacement method. The full length RalGPS1a comprising the Cdc25 domain, PXXP motif, and PH domain could not be well expressed and purified as a recombinant protein using an *E. coli* expression system, and the structure of the RalGPS1a PH domain has been reported previously (PDB ID: 2DTC). We therefore focused on the RalGPS1a Cdc25 domain, which shares about 30% sequence identity with the equivalent domains of RasGRF1 and SOS (Fig. 1). Crystals of the RalGPS1a Cdc25 domain belong to the space group $P2_1$ with unit cell parameters $a = 36.5$ Å, $b = 102.0$ Å, $c = 82.1$ Å, and $\beta = 96.5^\circ$. As suggested by Matthews coefficient analysis (Matthews, 1968), two protein molecules were found in an asymmetric unit with a solvent content of 52.7% (corresponding to a Matthews coefficient $V_M = 2.6$ Å³ Da⁻¹). The final refined model has an R-factor of 19.7% (free R-factor of 24.0%) without applying non-crystallographic symmetry restraints during refinement. Each protein molecule includes 245 amino acid residues spanning residues 44–288 and containing 14 α -helices. Residues 24–43 were missing in the final model due to the lack of interpretable electron density in the corresponding region. There was only one Cys residue per protein molecule, and no intermolecular disulfide bonds were identified in the final model. The environments surrounding the two protein molecules in an asymmetric unit were not identical; however, their structures were generally similar with a root mean square deviation (RMSD) of 0.44 Å between 244 C α atom pairs. We therefore focus our discussion below based on the structure of one protein molecule (arbitrarily chain A), unless otherwise stated. As summarized in Table 1, the refined model was of excellent quality.

The RalGPS1a Cdc25 domain adopts a bowl-shaped, all α -helical structure (Fig. 2A). Six long α -helices, $\alpha 1$, $\alpha 4$, $\alpha 5$, $\alpha 6$, $\alpha 9$, and $\alpha 14$, assemble together to form the main body of the Cdc25 domain. A helical hairpin formed by two antiparallel helices, $\alpha 11$ and $\alpha 12$, protrudes from the main body. Since a similar helical hairpin directly interacts with the nucleotide binding site of the substrate Ras and facilitates nucleotide exchange in the SOS-Ras complex, the helical hairpin is considered to be the active site of canonical Cdc25 domains (Boriack-Sjodin et al., 1998).

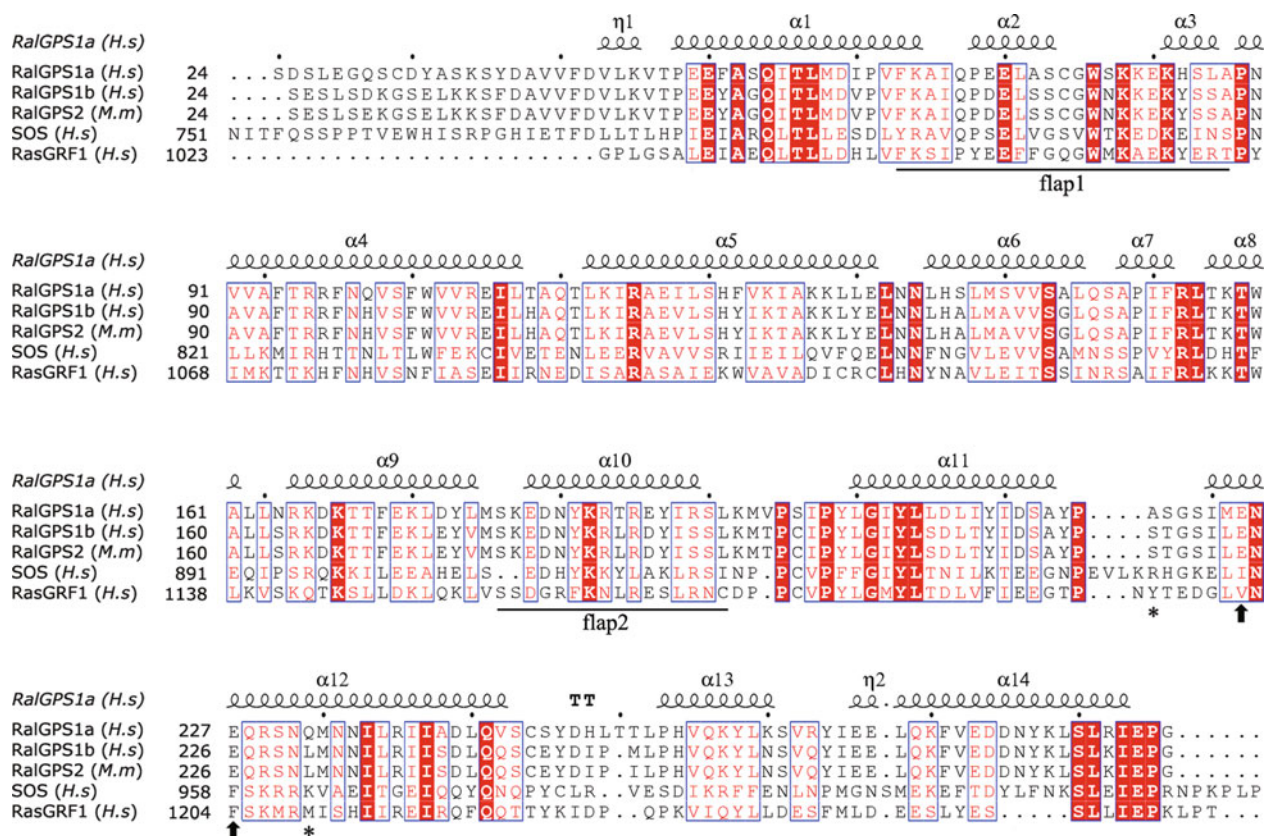


Figure 1. Sequence alignment of Cdc25 domains from RalGPS1a, RalGPS1b, RalGPS2, SOS, and RasGRF1. Amino acid sequences of Cdc25 domains from RalGPS1a (residues 24–289), RalGPS1b (24–287), RalGPS2 (24–287), SOS (751–1025), and RasGRF1 (1023–1262) were aligned by the program CLUSTALW2 (Wilbur and Lipman, 1983), and the alignment result was graphically displayed by the program ESPRIPT (Gouet et al., 1999). Secondary structure elements of the RalGPS1a Cdc25 domain are marked at the top of the sequence alignment. Residues involved in the interaction between the SOS helical hairpin and REM domain are marked with black arrows. Positive residues in SOS which interact with Ras switch 1 are marked with asterisks. Flap1 and Flap2 in RalGPS1a are indicated by black lines below the sequence.

Comparison with other known Cdc25 structures

We next compared the structure of the RalGPS1a Cdc25 domain with those of other Cdc25-containing GEFs. Structures of two other Cdc25 domains from Ras GEFs, SOS and RasGRF1, have been reported and share the same basic fold (Boriack-Sjodin et al., 1998; Freedman et al., 2006). The overall structure of our RalGPS1a Cdc25 domain exhibits high similarity with those of SOS and RasGRF1. Superposition of the RalGPS1 Cdc25 domain with the SOS Cdc25 domain (chain A of 2II0) resulted in an RMSD of 1.47 Å for 223 Cα atom pairs (using a 3 Å cutoff); and superposition with the RasGRF1 Cdc25 domain (chain S of 2IJE) yielded an RMSD of 1.16 Å for 213 Cα atom pairs (Fig. 2B). This structural architecture might be conserved among catalytic domains of other related Ras family GEFs, and the structural similarity suggests similar catalytic mechanisms for these Cdc25 domains.

The structure of the RalGPS1a Cdc25 domain also

exhibits a number of differences from those of SOS and RasGRF1. Major differences between the three available Cdc25 domain structures were located in the helical hairpin, especially the region connecting α11 and α12. Both SOS and RasGRF1 have a REM domain on the N terminal side of their Cdc25 domain. In the crystal structure of SOS, the region linking α11 and α12 contains two small antiparallel β-strands forming an inter-domain β-sheet with two other β-strands from the REM domain (Margarit et al., 2003). In addition, the hydrophobic side chains of two residues, Ile956 and Phe958, from the helical hairpin of the Cdc25 domain are inserted into a small hydrophobic groove on the surface of the REM domain. These inter-domain β-sheet and hydrophobic interaction were thought to stabilize the conformation of the helical hairpin in the SOS Cdc25 domain and also to regulate its GEF activity (Margarit et al., 2003). Although the 3D structure of the RasGRF1 REM domain has not been reported to date, it is hypothesized that a similar hydrophobic interaction between RasGRF1

Table 1 Data collection and refinement statistics

Parameters	RalGPS1a GEF domain
Data collection	
Space group	P2 ₁
Unit-cell parameters (Å, °)	$a = 36.5, b = 102.0, c = 82.1,$ $\alpha = \gamma = 90.0, \beta = 96.5$
Wave length (Å)	1.0000
Resolution (Å)	50–2.20 (2.32–2.20) ^a
Total observation	92,818 (5005)
Unique reflections	27,500 (2275)
Data completeness (%)	95.5 (79.2)
R_{merge}^b	0.139 (0.441)
$I/\sigma(I)$	8.34 (1.9)
Redundancy	3.4 (2.2)
Refinement	
Resolution range	40.8–2.2
No. of reflections in work set	26,070
No. of reflections in test set	1394 (5.08%)
No. of protein atoms	4016
No. of water molecules	124
R_{cryst} (work/free) ^c (%)	19.7/24.0
Average B factor (Å ²)	22.0 (19.4) ^d
r.m.s. deviation	
Bond lengths (Å)	0.006
Bond angles (°)	0.898
Ramachandran plot ^e	
Preferred region (%)	92.8
Allowed region (%)	7.2
Disallowed region (%)	0

^a Numbers in parentheses refer to the highest resolution shell.

^b $R_{\text{merge}} = \sum_h \sum_l |I_{hl} - \langle I_h \rangle| / \sum_h \sum_l \langle I_h \rangle$, where $\langle I_h \rangle$ is the mean intensity of the observations I_{hl} of reflection h .

^c $R_{\text{cryst}} = \sum_h |F_o - F_c| / \sum_h F_o$, where F_o and F_c are the observed and calculated structure factor amplitudes of reflection h .

^d Wilson B-factor.

^e Calculated by PROCHECK.

REM and Cdc25 domain may also exist because two hydrophobic residues, Val1202 and Phe1204, are located at the corresponding site in RasGRF1 (black arrows in Fig. 1 and 2C). In contrast, RalGPS1a lacks the REM domain, and our structure shows that the helical hairpin of its Cdc25 domain does not contain a pair of antiparallel β -strands. Moreover, three polar residues (Glu225, Asn226, and Glu227) occupy equivalent positions to these hydrophobic residues both at the amino acid sequence level and in the 3D structure of the RalGPS1a Cdc25 domain (Fig. 1 and 2C). Without a REM domain, the activity of the RalGPS1a Cdc25 domain should be regulated by a different mechanism from that for SOS and RasGRF1.

A model for the substrate binding mode of the RalGPS1a Cdc25 domain

Our RalGPS1a Cdc25 crystal structure provides a new opportunity to study the structural basis of RalGPS1a specificity towards Ral. RalGPS1a specifically activates the small GTPase Ral *in vivo* (de Bruyn et al., 2000; Rebhun et al., 2000; Ceriani et al., 2007), but there is no evidence to show that RalGPS1a has the ability to catalyze nucleotide exchange for any other Ras-like small GTPases either *in vitro* or *in vivo*. Although our attempts to crystallize RalGPS1a Cdc25 domain in complex with its substrate RalA failed, we could still model their interactions based on the homologous SOS-Ras complex structure (Margarit et al., 2003) and high structural conservation among Cdc25 domains and among Ras/Ral family members. We modeled the complex structure of the RalGPS1a cdc25 domain and RalA in three steps. First, we superimposed our structure of the RalGPS1a Cdc25 domain onto the SOS Cdc25 domain of the SOS-Ras complex (PDB ID: 1NVW) using a sequence homology based alignment protocol in the program PyMol. Next, we modeled the 3D structure of RalA using the on-line program Swiss-Model (Arnold et al., 2006) and the SOS-bound Ras structure as the homologous model, and the predicted RalA 3D structure was superimposed onto the Ras molecule of the SOS-Ras complex. Finally, the RalGPS1a-RalA complex model was refined by RosettaDock (Lyskov and Gray, 2008) to identify a low-energy conformation with optimized rigid body orientation and side chain conformation. This final RalGPS1a-RalA complex model (Fig. 3A and 3B) formed the basis for our analysis of the interaction between RalGPS1a and Ral.

The modeled RalGPS1a-RalA complex structure indicates an extensive interface encompassing a buried surface area of $\sim 3700 \text{ \AA}^2$. Such an extensive interface is close to that observed in the SOS/Ras complex structure (Boriack-Sjodin et al., 1998) and significantly larger than average protein-protein interaction found in Protein Data Bank (PDB) (Lo Conte et al., 1999). In addition, analysis of the electrostatic surface potential revealed a negatively charged surface area on helix $\alpha 11$ of Cdc25, which could insert into a positively charged site in RalA that usually binds GTP/GDP phosphate groups (Fig. 3C). In our modeled complex structure, $\alpha 11$ occupies the normal position of RalA switch 1 and therefore would facilitate the dissociation of GTP/GDP. Such an activation mechanism was illustrated clearly by the SOS-Ras complex structure (Boriack-Sjodin et al., 1998). However, a notable difference between the helical hairpins of RalGPS1a and SOS is observed in helix $\alpha 12$ and the junction connecting $\alpha 11$ and $\alpha 12$. In SOS, several basic residues including Lys949, Arg950, and the region Lys960-Arg961-Arg962-Lys963 confer positive charge on the surface of the helical hairpin and form electrostatic interactions with acidic residues such as Asp30, Glu31, and Asp33 of the Ras switch

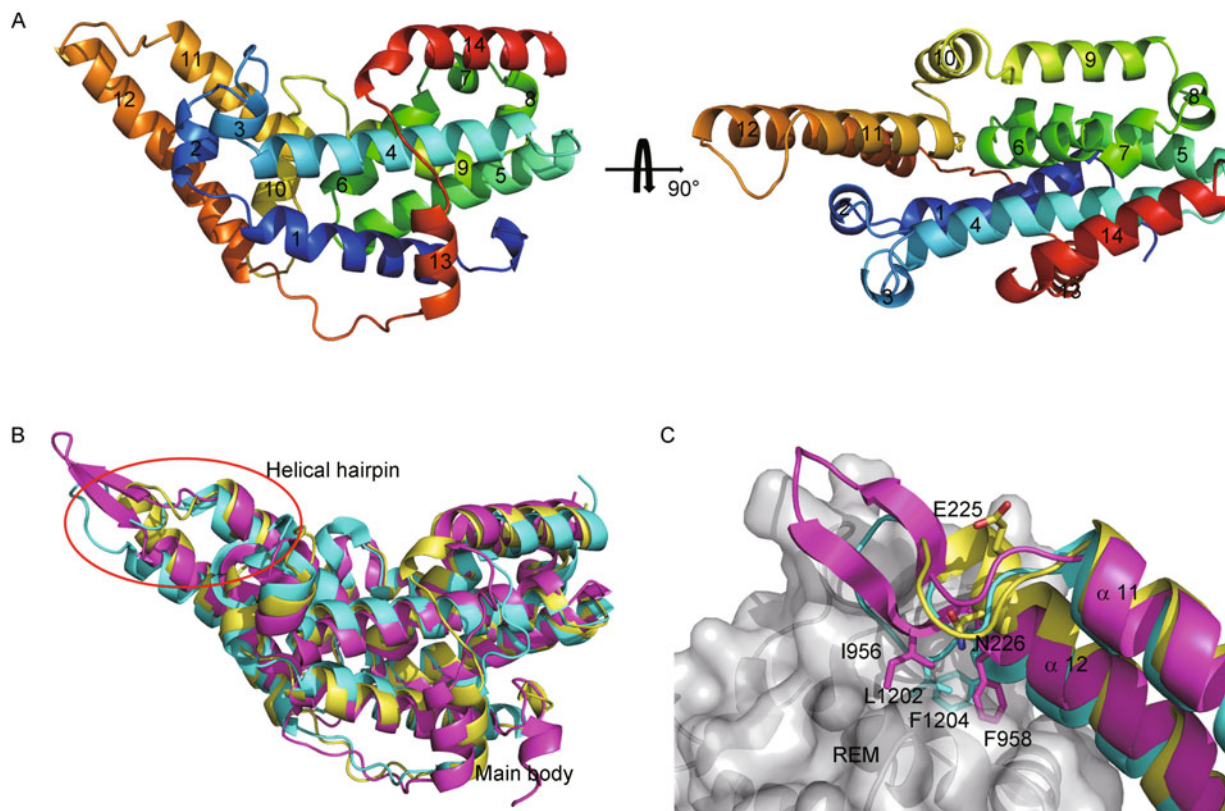


Figure 2. Overall structure of RalGPS1a Cdc25 domain, and comparison with the Cdc25 domains from SOS and RasGRF1. (A) Overall structure of the RalGPS1a Cdc25 domain in two views. 14 α -helices are numbered from the N-terminal to C-terminal. (B) Structural superposition of Cdc25 domains from RalGPS1a (yellow), SOS (magenta), and RasGRF1 (cyan). The red oval highlights major conformational differences between the three Cdc25 domains. (C) Residues involved in interactions with the REM domain. The helical hairpin of RalGPS1a (yellow), SOS (magenta), and RasGRF1 (cyan) are shown by ribbon representation. The REM domain (grey) of SOS is shown by a surface representation. Ile956 and Phe958 are buried into a hydrophobic groove in the REM domain. Corresponding sites in RasGRF1 are Leu1202 and Phe1204, indicating similar interactions between the helical hairpin and REM domain in RasGRF1. In RalGPS1a, residues at the same positions are hydrophilic, and no REM exists in RalGPS1a.

1 region (Fig. 4A). These favorable interactions fix the substrate switch 1 region away from its GTP/GDP binding site and maintain an open conformation of the nucleotide binding pocket in the substrate GTPase. In our modeled RalGPS1a-RalA complex, however, these polar interactions appear to be absent. Although RalA switch 1 remains primarily negatively charged (e.g. by the residues Glu41, Asp42, and Glu44) similar to Ras switch 1, the positively charged region found in SOS α 12 was missing in RalGPS1a α 12. In particular, the key residues Arg950 and Lys963 of SOS which interact directly with Ras switch 1 are not conserved in the RalGPS1a amino acid sequence (Fig. 1, asterisk). In fact, this region of RalGPS1a α 12 is primarily negatively charged (Fig. 4B), and therefore unfavorable for the formation of salt bridge bonds with RalA switch 1. These variations suggest that other types of interactions, such as hydrogen bonds, between the RalGPS1a helical hairpin and RalA switch 1 might exist. Therefore, determination of the detailed activation mechanism of RalA by the RalGPS1a

helical hairpin would require further structural studies of corresponding complexes.

In addition to the helical hairpin, another major interface in our complex model was located in the bottom area of the bowl shaped Cdc25 domain (Fig. 4C). Switch 2 of the substrate RalA was buried into a central cavity in the RalGPS1a Cdc25 domain. A hydrophobic patch formed by residues Phe94, Phe98, Met142, Tyr184, and Tyr202 of RalGPS1a anchor the RalA switch 2 via stacking of aromatic groups and hydrophobic interactions with the side chains of Tyr75, Ile78, and Tyr82 in RalA. Surrounding this core hydrophobic interaction region is a set of polar interactions formed by switch 2 and α 3 of RalA with residues on helices α 4, α 9, α 14, and the loop between α 2 and α 3 of RalGPS1a. Although the overall architecture of the interface between RalA switch 2 and RalGPS1a in our model is highly similar to that of the SOS-Ras complex, we find a complementary change on both sides of the switch 2 interface in SOS-Ras and RalGPS1a-RalA complex structure. In the SOS-Ras complex structure, Ras

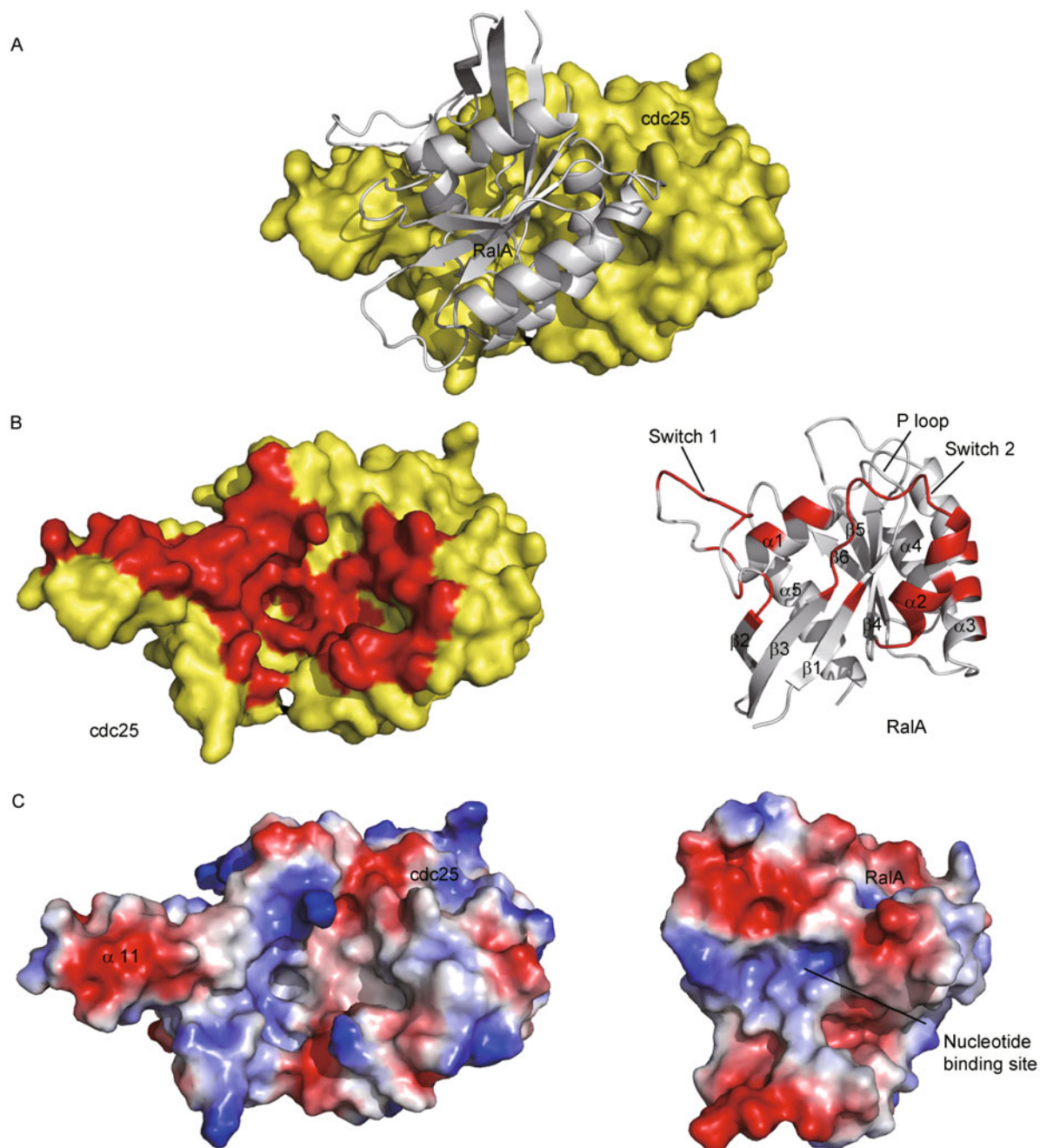


Figure 3. Modeled complex structure of the RalGPS1a Cdc25 domain and RalA. (A) Complex structure modeled by Swiss-Model prediction and superposition. The RalGPS1a Cdc25 domain is shown by a surface representation (yellow), and RalA is shown by a ribbon representation (grey). (B) Interface of the modeled RalGPS1a-RalA complex. The predicted interface involved in complex formation is colored red. The orientation of Cdc25 is similar to that in A, and the orientation of RalA is about 180° from that in A. (C) Electrostatic potential surface of the RalGPS1a Cdc25 domain and RalA. Helix $\alpha 11$ is negatively charged (red), and the RalA nucleotide binding site is positively charged (blue), suggesting that interaction between them is primarily electrostatic.

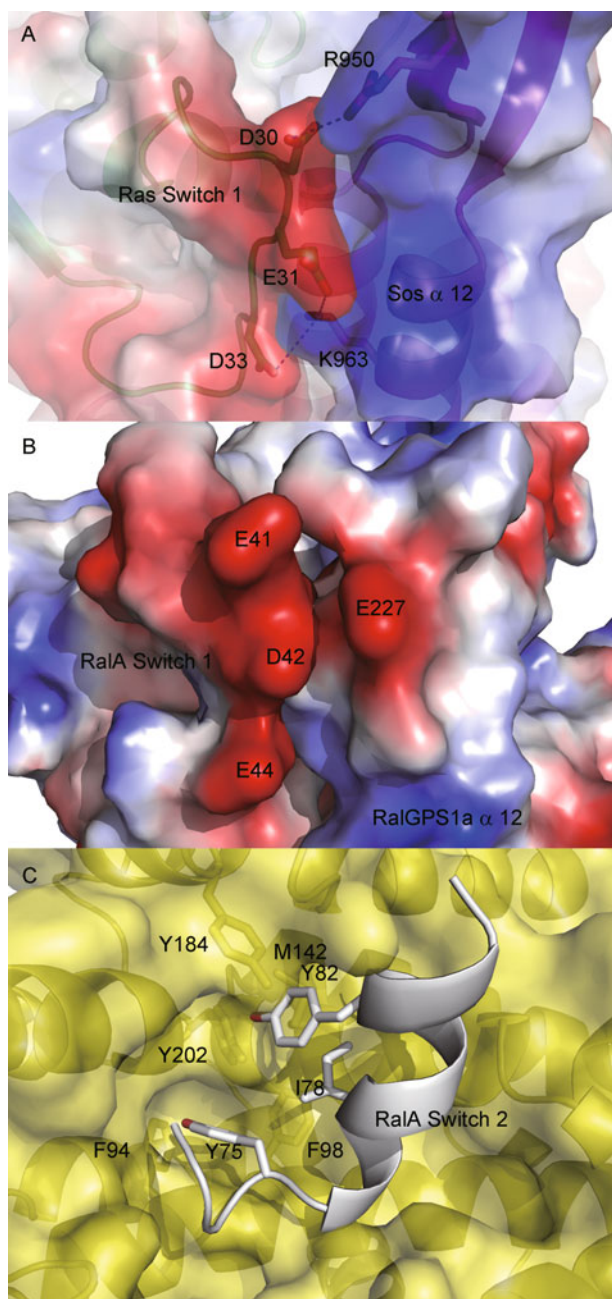


Figure 4. Interactions between the helical hairpin and switch 1. (A) Electrostatic surface potential shows extensive electrostatic interactions between the negatively charged Ras switch 1 and positively charged SOS α 12. Arg950 and Lys963 of SOS form a few critical salt bridge bonds with Asp30, Glu31, and Asp33 of Ras. Side chains of relevant residues are shown in sticks; and nitrogen and oxygen atoms are colored in blue and red, respectively. (B) Electrostatic surface potential representation of the RalGPS1a helical hairpin and RalA switch 1. Selected residues are labeled. (C) Interactions between RalA switch 2 (grey) and the bottom area of the RalGPS1a Cdc25 domain (yellow). The side chains of Tyr75, Ile78, and Tyr82 in RalA form extensive hydrophobic interactions with Phe94, Phe98, Met142, Tyr184, and Tyr202 in RalGPS1a. Side chains of relevant residues are shown in stick representation.

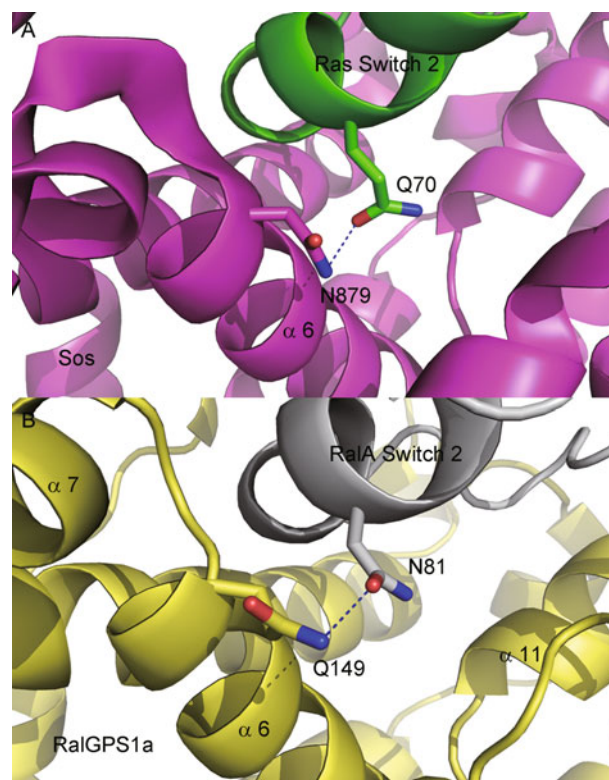


Figure 5. Co-evolution amino acids variation in GEF-switch 2 interface. (A) Hydrogen bond between SOS (magenta) Asn879 and Ras (green) Gln70 in SOS-Ras complex. Side chains of relevant residues are shown in stick representation. (B) Similar hydrogen bond is found between RalGPS1a (yellow) and RalA (grey) at the corresponding site in the modeled RalGPS1a-RalA complex. In RalGPS1a, the related residue is changed from Asn to Gln149, and in RalA, Gln is changed to Asn81. This co-evolution amino acids variation may help to explain the different substrate specificities between SOS and RalGPS1a.

Gln70 interacts in a buried environment with Asn879 of SOS through a hydrogen bond, and the interface would not accommodate two Gln residues (Fig. 5A). This Asn879 residue in SOS *cdc25* is conserved in the other above mentioned, Ras-specific GEF protein, RasGRF1. While in switch 2 of RalA, the residue is changed from Gln (in Ras) to Asn (one carbon bond shorter than Gln) at position 81, its counterpart in RalGPS1a changes from Asn (in SOS and RasGRF1) to Gln149 (one carbon bond longer than Asn) (Fig. 5B); such a co-evolution amino acids variation would partially explain the substrate specificity of different GEFs.

Conformation of the helical hairpin

The helical hairpin is the active site of the Cdc25 domain. We therefore compared the three available structures of Cdc25 domains in order to understand the relationship between the conformation of the helical hairpin and the activity of the

Cdc25 domain. Although the Cdc25 domains of RalGPS1a, SOS, and RasGRF1 share high similarity in their 3D structures, their catalytic activities differ significantly. The SOS Cdc25 domain exerts GEF catalytic activity *in vitro* only when it is expressed together with its N-terminal REM domain and is allosterically stimulated by a GTP bound Ras molecule binding to SOS at a large cleft between the Cdc25 and REM domains (Margarit et al., 2003). An isolated SOS Cdc25 domain was unable to stimulate nucleotide exchange of Ras. However, an isolated RasGRF1 Cdc25 domain has comparable catalytic ability to the activated SOS (Freedman et al., 2006). Earlier studies attributed this variation in catalytic ability to conformational variations of the helical hairpin in the corresponding Cdc25 domains (Freedman et al., 2006). The helical hairpin of the inactive SOS Cdc25 domain adopts an inward-rotated conformation which hinders binding of the Ras substrate. In particular, the binding of Tyr64 from Ras switch 2 to a deep, hydrophobic pocket near the bottom of the helical hairpin is unfavorably disturbed. Such a knot-and-hole interaction has been proven to be essential for SOS-catalyzed nucleotide release, since a single Y64A point mutation at this site in Ras totally abolishes its ability to bind with SOS (Hall et al., 2001). Consistently, the helical hairpins of RasGRF1 (PDB ID: 2IJE) and activated SOS (PDB ID: 1NVW) pivot outwards by about 14° compared with that of inactive SOS (PDB ID: 2II0), resulting in a more open and favorable binding site for substrate Ras.

Like the Cdc25 domain of RasGRF1, the isolated Cdc25 domain from RalGPS1a is able to catalyze nucleotide exchange of Ral *in vitro* (de Bruyn et al., 2000), making it a good candidate to validate the proposed relationship between the helical hairpin conformation and catalytic activity of the corresponding Cdc25 domain. After superimposing the core domain (excluding the helical hairpin) of the three Cdc25 domains (i.e. RalGPS1a, inactive SOS, and RasGRF1), we compared the conformations of their respective helical hairpins and corresponding tyrosine binding sites. RalGPS1a has a helical hairpin which pivots outwards by about 10° relative to that of inactive SOS (Fig. 6A). In addition, in our modeled RalGPS1a-RalA complex structure, Tyr75 in Ral switch 2 (corresponding to Tyr64 in Ras) is buried into the conserved hydrophobic pocket abutting the bottom of the helical hairpin. The side chain of Tyr75 forms extensive hydrophobic interactions with bulky side chains of surrounding residues, including Phe94, Phe98, His139, Tyr202, and Leu203 of RalGPS1a (Fig. 6B and 6C). A similar interaction pattern has also been observed in the SOS-Ras complex, although the corresponding residues are not absolutely conserved between RalGPS1a and SOS. For instance, the side chain of Ras Tyr64 is stabilized by similar hydrophobic interactions with Tyr796, Val799, Ile825, Phe929, Phe930, and Leu934 of SOS (Fig. 6D and 6E). In contrast, in a modeled complex structure of inactive SOS and Ras (Freedman et al., 2006), structural collisions would occur between

Ras Tyr64 and SOS Phe930 and Leu934 because of an inward rotation of the helical hairpin (Fig. 6E), providing a structural explanation for the hindered GEF activity. Taking these results together, we conclude that the activity of the Cdc25 domain depends on an outward rotation of the helical hairpin relative to the structure of inactive SOS.

Stabilization of the helical hairpin

The helical hairpin should be maintained in the appropriate conformation for activation as well as regulation of the Cdc25 domain. For example, in SOS the conformation of the helical hairpin is regulated by the relative position of the REM domain to the main body of the Cdc25 domain. In contrast, because of the lacking of a REM domain, conformation of the helical hairpin in the RalGPS1a Cdc25 domain is stabilized only by its interaction with the main body. A previous study on the RasGRF1 Cdc25 domain identified two structural protrusions, termed flap1 and flap2, extending from the main body and tightly nipping the helical hairpin from both sides (Freedman et al., 2006). Similarly, interactions of the helical hairpin with flap1 (66–88) and flap2 (179–194) were also observed in RalGPS1a. In our structure, flap1 is composed of helices $\alpha 2$, $\alpha 3$, and surrounding loops. Side chains of Phe66, Leu76, and Trp79 of flap1 formed extensive hydrophobic interactions with Leu203, Leu207, Leu210, Leu237, and Ile240 from the helical hairpin. Compared with the extensive interface between flap1 and the helical hairpin, the interface between flap2 and the helical hairpin was smaller, but likely to be critical for maintaining the active conformation of the helical hairpin. Asn183 and Arg188 from flap2 form tight polar interactions with Asp209 and Tyr212 in the helical hairpin (Fig. 6F). The tight clamping of the helical hairpin by flap1 and flap2 helps to maintain the active conformation of the helical hairpin, consistent with the observation that separated Cdc25 domains from both RalGPS1a and RasGRF1 have nucleotide exchange activity *in vitro* (de Bruyn et al., 2000; Freedman et al., 2006). In contrast, flap2 in SOS does not interact with the helical hairpin, allowing conformation of the helical hairpin to switch between active and inactive forms via an allosteric regulation mechanism.

DISCUSSION

The Ras family of small GTPases plays essential roles in coupling cellular signals from cell surface receptors to downstream effector proteins (Colicelli, 2004). Conversion of small GTPase from a GDP bound, inactive form to a GTP bound, active form is usually catalyzed by guanine-nucleotide exchange factors (GEFs) (Boguski and McCormick, 1993). RalGPS1a, a member of the RalGPS family (de Bruyn et al., 2000; Rebhun et al., 2000), specifically activates small GTPase Ral and therefore regulates multiple downstream cellular activities involving cytoskeleton regulation, cell

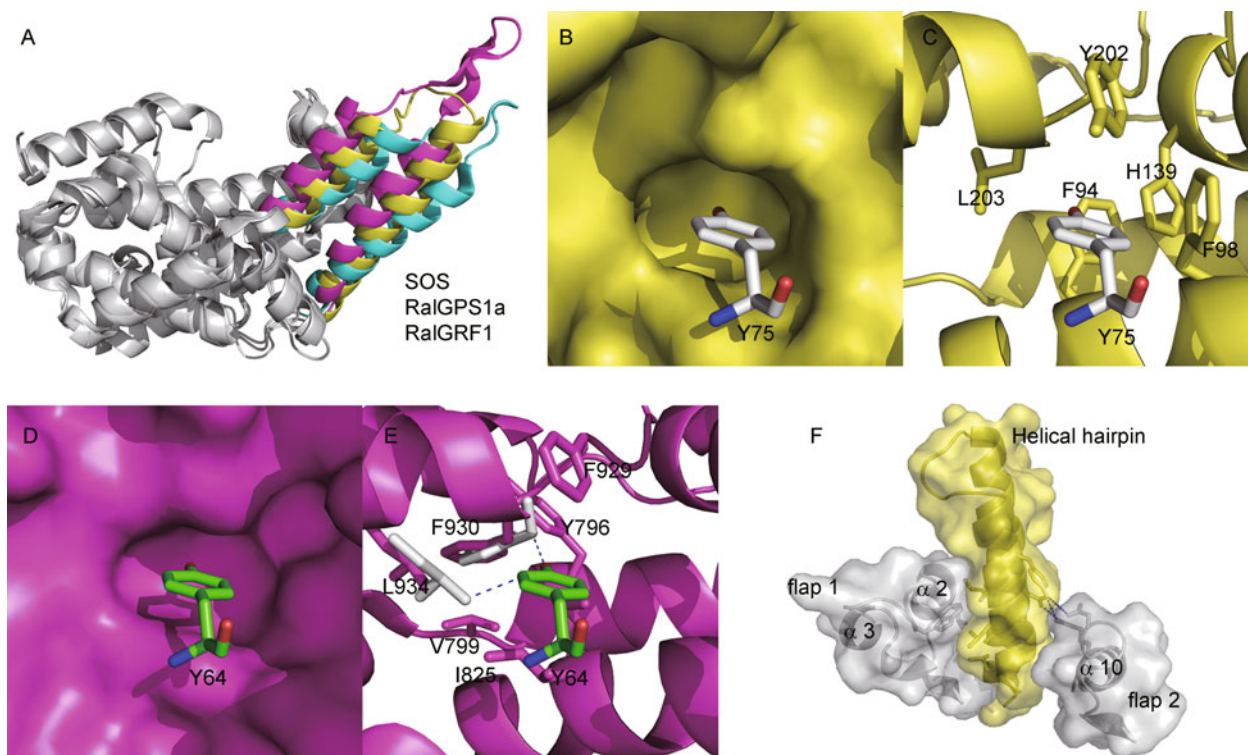


Figure 6. Comparison of helical hairpins and tyrosine binding sites of RalGPS1a, SOS and RasGRF1. (A) Superposition of core domains (i.e. with the helical hairpin omitted) of RalGPS1a (yellow), inactive SOS (magenta), and RasGRF1 (cyan) Cdc25 domains. Helical hairpins of RasGRF1 and RalGPS1a pivot outwards by 14° and 10°, respectively, relative to that of inactive SOS. (B and C) Modeled complex structure shows that Tyr75 of RasA (located between $\beta 3$ and $\alpha 4$) (grey) would be properly buried in a conserved hydrophobic pocket formed by Phe94, Phe98, His139, Tyr202, and Leu203 of RalGPS1a (yellow). Side chains of relevant residues are shown in stick representation. (D and E) Ras Tyr64 (green) is buried in a similar pocket according to the SOS-Ras complex (PDB ID: 1NVW). Relevant residues are shown in stick representation (magenta). The inwardly rotated helical hairpin of inactive SOS changes the position of Leu934 by about 2.5 Å, and Phe930 by 1.5 Å (grey). These movements would result in clashes with Ras Tyr64, thus fully abolishing nucleotide exchange of Ras catalyzed by SOS. (F) Flap1 and flap2 form hydrophobic and hydrophilic interactions, respectively, with the helical hairpin. Tight clamping by flap1 and flap2 from both sides keeps the helical hairpin in the active conformation.

proliferation and tumorigenesis (Ceriani et al., 2007; Bode-mann and White, 2008; Cascone et al., 2008). RalGPS1a harbors an N-terminal Cdc25 domain, which is responsible for activation of the substrate Ral proteins, followed by a PXXP motif and a C-terminal PH domain. Unlike members of the other Ral specific GEF family, RalGDS, RalGPS1a contains neither a Ras exchange motif (REM) nor Ras binding domain (RBD). Previous studies have demonstrated that the REM domain in the Ras specific GEF protein SOS stabilizes the Cdc25 domain by forming an inter-domain β -sheet and a small hydrophobic interface (Boriack-Sjodin et al., 1998), and binding of an activated Ras molecule at the REM-Cdc25 interface stimulates the activity of Cdc25 (Margarit et al., 2003). Similarly, the RBD domain couples signals from activated Ras to the activation of Ral proteins. Lack of both the REM and RBD domains suggests distinct mechanisms in both Cdc25 domain activity and whole RalGPS1a regulation *in vivo*.

In the current study, we determined the crystal structure of RalGPS1a Cdc25 domain at 2.2 Å resolution (Fig. 2A). Consistent with their 30% sequence identity (Fig. 1), the RalGPS1a Cdc25 domain shares a highly similar overall fold with the Cdc25 domains of SOS and RasGRF1 (Fig. 2B). Based on the reported SOS-Ras complex structure (Margarit et al., 2003) and the structural conservation between Ras and RalA, we proposed a complex model of the RalGPS1a Cdc25 domain and its substrate RalA (Fig. 3A). In this modeled complex structure, the interaction between the RalGPS1a Cdc25 domain and switch 2 of the substrate RalA is quite similar to the corresponding part of the SOS-Ras complex. The main differences between the cCdc25 domains of RalGPS1a and SOS occur in their helical hairpin regions. The helical hairpin is an essential part of the Cdc25 active site and effectively disrupts the interaction between nucleotide phosphate groups and the switch 1 region of the substrate small GTPase by replacing the switch 1 position with helix $\alpha 11$

from the GEF cdc25 domain. The other helix of the helical hairpin, $\alpha 12$, helps to stabilize the open conformation of switch 1. SOS $\alpha 12$ holds the opening switch 1 of Ras via a few salt bridge bonds (Fig. 4A). Although the RasGRF1-Ras complex structure is currently unavailable, RasGRF1 features conserved basic residues at several key positions in the helical hairpin and on the Cdc25-Ras interface, implying similar polar interactions with Ras. In contrast, in the RalGPS1a sequence, substitution of a few critical basic residues on SOS $\alpha 12$ by acidic residues essentially eliminates the favorable electrostatic interactions observed in the RalGPS1a-RalA complex (Fig. 4B). These variations in the helical hairpin may help to partially explain the substrate specificity of RalGPS1a and SOS. Further studies on the 3D structure of the RalGPS1a-Ral complex are required for insights into the mechanism of substrate binding specificity of RalGPS1a.

Previous studies have linked the GEF activity of the Cdc25 domain with the conformation of its helical hairpin. Unlike the inward rotation of the helical hairpin of an inactive SOS (Freedman et al., 2006), the helical hairpin of the RalGPS1a Cdc25 domain adopts a relatively open conformation in our crystal structure, favoring substrate binding (Fig. 5A). This structural observation is consistent with the previous observation that the recombinant RalGPS1a Cdc25 domain alone was sufficient to trigger nucleotide exchange of substrate Ral (de Bruyn et al., 2000). Full length RalGPS1a was also proposed to be constitutively active in the cytoplasm (de Bruyn et al., 2000), and the full cellular function of RalGPS1a depends on its intact PH domain which ensures its correct membrane localization. Although the PXXP motif in RalGPS1a can directly interact with an SH3 domain of an upstream adaptor protein Grb2, activity of the full length RalGPS1a was not enhanced by over-expression of Grb2 *in vivo* (Rebhun et al., 2000), implying that conformation of the RalGPS1a Cdc25 domain may not be regulated by the binding of the PXXP motif with Grb2. The need to understand the regulatory mechanisms of RalGPS1a and its interplay with Grb2 warrants further investigations.

MATERIALS AND METHODS

Cloning, expression, and purification

The gene encoding for the Cdc25 domain of human RalGPS1a (i.e. residues 24–289; GenBank ID: NM_014636) was cloned into the pGEX 6p-1 vector (GE Healthcare) using the *Bam*HI/*Sal*I restriction cleavage sites. The N-terminal 23 amino acid residues were predicted to be a signal peptide by the software SIGNALP (Emanuelsson et al., 2007) and were thus omitted from the cloning. The expression plasmid was transformed into the *Escherichia coli* BL21 (DE3) strain (Novagen) to express the GST-target fusion protein. The cell culture was grown in LB medium supplemented with 100 mg/L ampicillin at 37°C until the cell density reached an absorbance of

0.7 at 600 nm. A final concentration of 400 μ mol/L isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to the medium after the temperature was lowered to 25°C, and the cells were cultured for another 12 h for protein production. Cells were harvested by centrifugation at 4000 g for 40 min and then resuspended with ice-cold phosphate buffered saline (PBS, i.e. 137 mmol/L NaCl, 2.7 mmol/L KCl, 4.3 mmol/L Na₂HPO₄, and 1.4 mmol/L KH₂PO₄, pH 7.4). Cell lysate was prepared by sonication and centrifuged at 30,000 g for 1 h. The supernatant was loaded onto a Glutathione Sepharose 4B column (GE Healthcare) pre-equilibrated with PBS. After washing off non-specifically bound proteins, the fusion protein was eluted from the column using a buffer containing 20 mmol/L Tris (pH 7.5), 50 mmol/L NaCl, and 15 mmol/L reduced glutathione. The eluted fusion protein was then digested overnight with PreScission Protease (GE Healthcare) at 4°C. The digested protein sample was then loaded on a Hitrap SP HP (GE Healthcare) ion-exchange column pre-equilibrated with 20 mmol/L Tris (pH 7.5) and 50 mmol/L NaCl. The dissociated GST protein could not bind to the column under this buffer condition and flowed through. The RalGPS1a Cdc25 domain was then purified with a gradient elution of 50–500 mmol/L NaCl in a buffer of 20 mmol/L Tris (pH 7.5), and the protein peak was eluted at about 200 mmol/L NaCl. All purification steps were carried out at a temperature of 16°C or lower.

Crystallization of the RalGPS1a Cdc25 domain

Fractions containing the target protein were pooled and concentrated to a final concentration of 2–5 mg/mL using a centrifugal filter device (Millipore, USA) before crystallization trials. About 700 crystallization conditions were screened using the hanging drop vapor diffusion method (a 0.2 + 0.2 μ L drop over an 80 μ L reservoir) with the Mosquito crystallization robot (TTP LabTech, UK). Conditions No. 23 and No. 24 from the Natrix Kit (Hampton Research) yielded tiny needle shape crystals and were chosen for further optimization. By varying the concentration of precipitants, pH of the buffer and the protein concentration, we obtained needle-shape single crystals which diffracted well at a synchrotron X-ray source. The optimized reservoir solution was composed of 200 mmol/L KCl, 10 mmol/L MgCl₂, 50 mmol/L sodium cacodylate (pH 6.5), and 8.5% (*w/v*) PEG 3350, and the optimal protein concentration was determined to be 2 mg/mL. Crystals appeared one day after setting drops and reached dimensions of 30 μ m \times 30 μ m \times 200 μ m in about a week.

Data collection, processing, and structure determination

Diffraction data were collected on beamline BL17a of the Photon Factory synchrotron facility (KEK, Japan). Freshly grown crystals were soaked in a cryo-protectant buffer (85% (*v/v*) reservoir solution with 15% (*v/v*) glycerol) for 5 min before being flash cooled in a stream of nitrogen gas (100 K). Data were collected from a sweeping range of 120° using a 1° oscillation per frame. Data processing and scaling were carried out with the software package HKL2000 (Otwinowski and Minor, 1997). Initial phases were determined using the molecular replacement method implemented in the program PHASER (McCoy et al., 2007) with the crystal structure of RasGRF1 Cdc25 domain as the search model (PDB ID: 2IJE). The initial model was built using the rebuild-in-place strategy of the program RESOLVE (Terwilliger, 2003) and was then subject to iterative rounds of refinement with the software package PHENIX (Adams et al., 2010) and manual

adjustment with the graphics program COOT (Emsley and Cowtan, 2004). A total of 124 water molecules were included in the final refined model under the guidance of a Fo-Fc difference *Fourier* map. The final model was verified with the program PROCHECK (Laskowski et al., 1993), and statistical data for the final model are summarized in Table 1. Part of the structural analysis was performed with the software EdPDB (Zhang and Matthews, 1995). All structural figures in this study were prepared with the program PyMol (DeLano, 2002).

Deposition of coordinates

Coordinates and structure factors for the crystal structure of the RalGPS1a GEF domain have been deposited to the Research Collaboratory for Structural Bioinformatics Protein Data Bank (<http://www.rcsb.org/pdb/>). The accession ID is 3QXL.

ACKNOWLEDGEMENTS

The authors are grateful to Xiaoxia Yu, Yuanyuan Chen, and other staff members of the Structural Biology Core Facility in the Institute of Biophysics, Chinese Academy of Sciences (CAS), Beijing, for their excellent technical assistance. This work was supported by grants from the State Key Development Program for Basic Research of the Ministry of Science and Technology of China (973 Project) (Grant Nos. 2007CB914304, 2011CB915501, and 2011CB910304), the National High Technology Research and Development Program of China (863 Project) (Grant No. 2006AA02A322), and National Key Technologies R&D Program (Grant No. 2009ZX10603) of the Ministry of Health of China.

ABBREVIATIONS

GEF, guanine-nucleotide exchange factor; PH, pleckstrin homology (domain); RalGPS1a, Ral guanine-nucleotide exchange factor with a PH domain and an SH3 binding motif; RasGRF1, Ras guanine nucleotide-releasing factor 1; RBD, Ras binding domain; REM, Ras exchanger motif; SH3, *src* homology domain 3; SOS, son of sevenless (domain)

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkóczy, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66, 213–221.
- Albright, C.F., Giddings, B.W., Liu, J., Vito, M., and Weinberg, R.A. (1993). Characterization of a guanine nucleotide dissociation stimulator for a ras-related GTPase. *EMBO J* 12, 339–347.
- Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22, 195–201.
- Bodemann, B.O., and White, M.A. (2008). Ral GTPases and cancer: linchpin support of the tumorigenic platform. *Nat Rev Cancer* 8, 133–140.
- Boguski, M.S., and McCormick, F. (1993). Proteins regulating Ras and its relatives. *Nature* 366, 643–654.
- Boriack-Sjodin, P.A., Margarit, S.M., Bar-Sagi, D., and Kuriyan, J. (1998). The structural basis of the activation of Ras by Sos. *Nature* 394, 337–343.
- Broek, D., Toda, T., Michaeli, T., Levin, L., Birchmeier, C., Zoller, M., Powers, S., and Wigler, M. (1987). The *S. cerevisiae* CDC25 gene product regulates the RAS/adenylate cyclase pathway. *Cell* 48, 789–799.
- Cascone, I., Selimoglu, R., Ozdemir, C., Del Nery, E., Yeaman, C., White, M., and Camonis, J. (2008). Distinct roles of RalA and RalB in the progression of cytokinesis are supported by distinct RalGEFs. *EMBO J* 27, 2375–2387.
- Ceriani, M., Scanduzzi, C., Amigoni, L., Tisi, R., Berruti, G., and Martegani, E. (2007). Functional analysis of RalGPS2, a murine guanine nucleotide exchange factor for RalA GTPase. *Exp Cell Res* 313, 2293–2307.
- Chardin, P., and Tavittian, A. (1986). The *ral* gene: a new ras related gene isolated by the use of a synthetic probe. *EMBO J* 5, 2203–2208.
- Chien, Y., and White, M.A. (2003). RAL GTPases are linchpin modulators of human tumour-cell proliferation and survival. *EMBO Rep* 4, 800–806.
- Colicelli, J. (2004). Human RAS superfamily proteins and related GTPases. *Sci STKE* 2004, RE13.
- de Bruyn, K.M., de Rooij, J., Wolthuis, R.M., Rehmann, H., Wesenbeek, J., Cool, R.H., Wittinghofer, A.H., and Bos, J.L. (2000). RalGEF2, a pleckstrin homology domain containing guanine nucleotide exchange factor for Ral. *J Biol Chem* 275, 29761–29766.
- de Rooij, J., Rehmann, H., van Triest, M., Cool, R.H., Wittinghofer, A., and Bos, J.L. (2000). Mechanism of regulation of the Epac family of cAMP-dependent RapGEFs. *J Biol Chem* 275, 20829–20836.
- DeLano, W.L. (2002). The PyMOL Molecular Graphics System. San Carlos, CA: Delano Scientific.
- Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2, 953–971.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126–2132.
- Freedman, T.S., Sondermann, H., Friedland, G.D., Kortemme, T., Bar-Sagi, D., Marqusee, S., and Kuriyan, J. (2006). A Ras-induced conformational switch in the Ras activator Son of sevenless. *Proc Natl Acad Sci U S A* 103, 16692–16697.
- Gouet, P., Courcelle, E., Stuart, D.I., and Métoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, 305–308.
- Hall, B.E., Yang, S.S., Boriack-Sjodin, P.A., Kuriyan, J., and Bar-Sagi, D. (2001). Structure-based mutagenesis reveals distinct functions for Ras switch 1 and switch 2 in Sos-catalyzed guanine nucleotide exchange. *J Biol Chem* 276, 27629–27637.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J Appl Cryst* 26, 283–291.
- Lim, K.H., Baines, A.T., Fiordalisi, J.J., Shipitsin, M., Feig, L.A., Cox, A.D., Der, C.J., and Counter, C.M. (2005). Activation of RalA is critical for Ras-induced tumorigenesis of human cells. *Cancer Cell* 7, 533–545.
- Lo Conte, L., Chothia, C., and Janin, J. (1999). The atomic structure of protein-protein recognition sites. *J Mol Biol* 285, 2177–2198.

- Lyskov, S., and Gray, J.J. (2008). The RosettaDock server for local protein-protein docking. *Nucleic Acids Res* 36, W233–238.
- Margarit, S.M., Sondermann, H., Hall, B.E., Nagar, B., Hoelz, A., Pirruccello, M., Bar-Sagi, D., and Kuriyan, J. (2003). Structural evidence for feedback activation by Ras.GTP of the Ras-specific nucleotide exchange factor SOS. *Cell* 112, 685–695.
- Matthews, B.W. (1968). Solvent content of protein crystals. *J Mol Biol* 33, 491–497.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658–674.
- Moskalenko, S., Henry, D.O., Rosse, C., Mirey, G., Camonis, J.H., and White, M.A. (2002). The exocyst is a Ral effector complex. *Nat Cell Biol* 4, 66–72.
- Murai, H., Ikeda, M., Kishida, S., Ishida, O., Okazaki-Kishida, M., Matsuura, Y., and Kikuchi, A. (1997). Characterization of Ral GDP dissociation stimulator-like (RGL) activities to regulate c-fos promoter and the GDP/GTP exchange of Ral. *J Biol Chem* 272, 10483–10490.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* 276, 307–326.
- Quilliam, L.A. (2006). Specificity and expression of RalGPS as RalGEFs. *Methods Enzymol* 407, 108–114.
- Rameh, L.E., Arvidsson, A., Carraway, K.L. 3rd, Couvillon, A.D., Rathbun, G., Crompton, A., VanRenterghem, B., Czech, M.P., Ravichandran, K.S., Burakoff, S.J., *et al.* (1997). A comparative analysis of the phosphoinositide binding specificity of pleckstrin homology domains. *J Biol Chem* 272, 22059–22066.
- Rangarajan, A., Hong, S.J., Gifford, A., and Weinberg, R.A. (2004). Species- and cell type-specific requirements for cellular transformation. *Cancer Cell* 6, 171–183.
- Rebhun, J.F., Chen, H., and Quilliam, L.A. (2000). Identification and characterization of a new family of guanine nucleotide exchange factors for the ras-related GTPase Ral. *J Biol Chem* 275, 13406–13410.
- Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A., and Bos, J.L. (2006). Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state. *Nature* 439, 625–628.
- Terwilliger, T.C. (2003). Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr D Biol Crystallogr* 59, 38–44.
- Wilbur, W.J., and Lipman, D.J. (1983). Rapid similarity searches of nucleic acid and protein data banks. *Proc Natl Acad Sci U S A* 80, 726–730.
- Wolthuis, R.M., de Ruiter, N.D., Cool, R.H., and Bos, J.L. (1997). Stimulation of gene induction and cell growth by the Ras effector Rlf. *EMBO J* 16, 6748–6761.
- Zhang, X.J., and Matthews, B.W. (1995). Edpdb - a Multifunctional Tool for Protein-Structure Analysis. *J Appl Cryst* 28, 624–630.