Regulation and Selectivity of Exchange Factors for G-proteins of the Ras-family

Milica Popović

ISBN: 978-90-8891-615-1

Cover design by Milica Popović

Printed by Proefschriftmaken.nl || Uitgeverij BOXPress

Published by Uitgeverij BOXPress, ‘s-Hertogenbosch

Financial support for the publication of this thesis from UMC Utrecht and Universiteit Utrecht is gratefully acknowledged.

Regulation and Selectivity of Exchange Factors for G-proteins of the Ras-family

Regulatie en selectiviteit van uitwisselingsfactoren voor G-eiwitten van de Ras-familie

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op donderdag 23 mei 2013 des middags te 12.45 uur

door

Milica Popović
geboren op 11 februari 1981 te Belgrado, Servië

Promotor: Prof.dr. J. L. Bos

Co-promotor: Dr. H. Rehmann

This thesis was accomplished with financial support from TI Pharma.

Mojim roditeljima

Contents

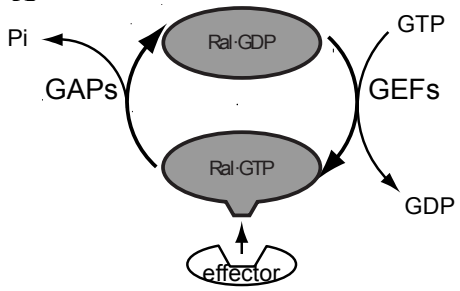
Chapter 1

General introduction

I Small G-proteins – the Ras-superfamily

Small G-proteins are important regulators of cellular signaling pathways. Up to date, over 150 small G-proteins have been identified. They are grouped into five major families on the basis of sequence and functional similarities: Ras, Rho, Rab, Arf and Ran[1]. Among them, Ras family members are involved in the control of central cell signaling pathways such as cell differentiation, proliferation, apoptosis, adhesion and cellular transport. They comprise of over 30 members among which are the Ras, Rap, Ral and Rheb proteins.
The most critical property for the function of G-proteins is the switching between GDP-bound and GTP-bound conformations (Figure 1A). Only the GTP-bound form is able to bind and activate downstream effectors. Therefore, the cycling between GTP- and GDP-bound conformations results in switching the signaling on and off. The inactivation of G-proteins occurs upon hydrolysis of the bound GTP to GDP. G-proteins hydrolyze bound GTP very slowly, and this process is accelerated by the action of GTPase activating proteins (GAPs). Activation of G-proteins requires dissociation of the bound GDP and subsequent binding of GTP. As with GTP hydrolysis, the release of the bound nucleotide is intrinsically slow and inefficient. Guanine nucleotide exchange factors (GEFs) increase the rates of nucleotide exchange manifold. The activation of G-proteins occurs only upon appropriate biological signals, and therefore the activity of GEFs needs to be tightly regulated.

A



B

	G1	G2	G3	G4	G5
consensus sequence	**GxxxxGK**$^{\mathbf{S}}_{\mathbf{T}}$	**T**	**DxxG**	**NKxD**	$^{\mathbf{C}}_{\mathbf{S}}$**A**
H-Ras	**G**AGGV**GKS**	YDP**T**IED	IL**D**TA**G**QE	VG**NK**C**D**L	YIET**SA**K
Rap1B	**G**SGGV**GKS**	YDP**T**IED	IL**D**TA**G**TE	VG**NK**C**D**L	FLES**SA**K
CDC42	**G**DGAV**GKT**	YVP**T**VFD	LF**D**TA**G**QE	VGTQI**D**L	YVEC**SA**L
Ran	**G**DGGT**GKT**	YVA**T**LGV	VW**D**TA**G**QE	CG**NK**V**D**I	YYDI**SA**K
GNAI1	**G**AGES**GKS**	RVK**T**TGI	MF**D**VG**G**QR	FL**NK**K**D**L	THFT**CA**T
EF- TU	**G**HVDH**GKT**	RGI**T**INA	HT**D**CP**G**HA	YV**NK**A**D**A	VIVG**SA**L

Figure 1A. The cycling of the G-protein between inactive and active conformation. Binding of effectors to the active form of G-proteins is indicated with an arrow. **B. Conserved nucleotide-binding motifs G1 to G5 of selected G-proteins.** Adapted from ref[3].

1.1 G-domain structure

All small G-proteins are composed of a single G-domain, consisting of six β-strands that are surrounded by five helices. These G-domains are additionally found in the heterotrimeric G-proteins, where the G_α-subunit adopts the G-domain fold, and in protein factors involved in protein synthesis. There are five conserved sequence elements located in five loops that are referred to as G1 to G5 (Figure 1B)[2,3]. As these loops are involved in the binding of the guanine nucleotide and a Mg^{2+} ion that forms a complex with nucleotide, there is high sequence conservation within the G-loops in G-proteins. G1 is the phosphate-binding loop (P-loop), with the consensus sequence GXXXGKS/T. The glycine residues bind the α- and β-phosphate of the nucleotide, the lysine bridges the β- and γ-phosphates while the serine or threonine contributes to the co-ordination of the Mg^{2+}. G2 contains an invariant threonine involved in Mg^{2+} coordination. G3 is characterized by the DXXG motif where the aspartic acid residue interacts with Mg^{2+} and the glycine with the γ-phosphate. The N/TKXD motif of G4 binds to the guanine ring. Lastly, the conserved alanine residue from the G5 loop binds to the oxygen from the guanine base. This interaction has an important role in determining preference of the guanine base over the adenine base, as the corresponding amino group in adenine, would sterically clash with the alanine residue.

1.2 Switch mechanism

The switch from the active to the inactive conformation results primarily in the structural reorganization of two loops in the G-protein, while the rest of the protein remains mainly unperturbed. Therefore, these two loops are named switch I and switch II. Switch I contains the G2 motif, and switch II the G3 motif. The disruption or the formation of the interactions between the residues from switch I and switch II and the γ-phosphate of GTP trigger the conformational switch. This basic mechanism is the same for all G-proteins. The γ-phosphate is released after the hydrolysis of GTP to GDP to switch off the signaling of the G-protein. This mechanism is described as a "loaded spring", where cleavage of phosphate allows switch I and II to relax into GDP-bound conformation (Figure 2)[2].

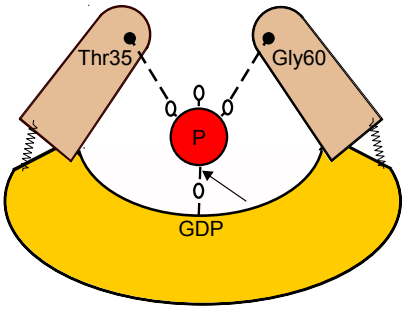


Figure 2. Loaded spring mechanism – change from the GTP-bound to GDP-bound conformation of the G-protein. Adapted from ref[2]. The cleavage of the γ-phosphate is indicated by an arrow. Upon cleavage of the γ-phosphate, hydrogen bonds with Thr35 and Gly60 (H-Ras) are broken, which allows relaxation of switch I and switch II into different conformation, as indicated by the springs.

1.3 Regulation of small G-protein activity

As well as being regulated by their activation and inactivation at defined time points, G-proteins are localized and confined to discreet cellular compartments. Moreover, proteins that regulate G-proteins, such as GEFs, GAPs, lipidating and delipidating enzymes and guaninine nucleotide dissociation inhibitors (GDIs), are all highly regulated themselves, adding more complexity to the process of G-protein regulation.

1.3.1 Activation of small G-proteins by GEFs (Mechanism of GEF catalyzed nucleotide exchange)

The activation of G-protein consists of the GDP dissociation and binding of GTP instead. The affinity of G-protein for GTP and GDP is similar, and is in the nanomolar to picomolar range. Since the affinity for a nucleotide is high, most of the G-proteins exchange their bound nucleotide slowly, with half-life times of several hours. GEFs efficiently accelerate dissociation of G-protein-bound nucleotides. In most cases, there is no GEF induced preference towards rebinding of GTP or GDP. Since the cellular concentration of GTP is around ten times higher than the concentration of GDP, the action of GEFs on G-proteins results mostly in binding of GTP and thereby G-protein activation. One exception is the DOCK family of Rho GEFs that are suggested to interact differently with GDP- and GTP-bound G-proteins[4].
While all G-proteins share the same fold, different GEF families have structurally unrelated catalytic domains. These domains interact with their G-protein substrates in different ways approaching them under various angles[5]. The principle of the exchange mechanism is, however, the same – the GEF interacts with residues from the nucleotide-binding surface of G-proteins, distorts the nucleotide-binding site, and thus decreases the affinity for nucleotide binding. After the release of the bound nucleotide, the G-protein remains in a complex with the GEF. The structures of ternary GEF/G-protein/nucleotide complexes demonstrate that primarily the phosphate binding site is perturbed by the action of GEFs[5]. The structure of the ArfGRF/Arf/GDP complex shows that a negatively charged glutamic acid finger from ArfGRF points into the phosphate binding site. This results in the destabilization of phosphate binding. Another ternary complex that has been crystallized is the plant G-protein Rop in a complex with GDP and RopGEF. Similarly, the interaction with the GEF destabilizes the phosphate binding site. It is, thus, a common element of GEF action to interact with the P-loop in the G-protein and to distort the phosphate-binding region. This reaction sequence is reversible. The incoming nucleotide approaches the G-protein/GEF complex with its base first. The subsequent establishment of the interaction between the phosphates and the G-protein destabilizes the interaction with the GEF and completes the last step of exchange reaction, where the GEF leaves the complex (Figure 3).

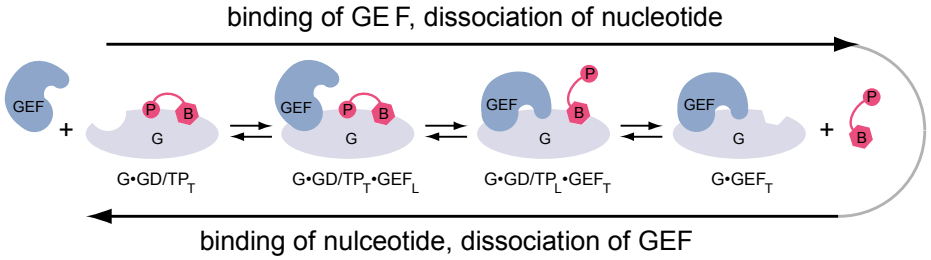


Figure 3. Mechanism of GEF-catalyzed G-protein nucleotide exchange. Adapted from ref[5]. Nucleotide exchange occurs via series of reversible steps. G – G-protein, P –phosphate moiety, of the nucleotide B –base of the nucleotide. In the trimeric complex, the GEF and the nucleotide compete for binding to the G-protein. Tight (subscript T) and loose (subscript L) binding to the G-protein are indicated. The approaching GEF at first interacts loosely with the G-protein, while the nucleotide is tightly bound. The GEF then competes out the phosphate moiety of the nucleotide, leaving the nucleotide loosely bound to the G-protein.

1.3.2 Inactivation of small G-proteins by GAPs

GAPs accelerate the cleavage of the γ-phosphate of GTP by several orders of magnitude. The mechanism of GAP-catalyzed GTP hydrolysis was unraveled in great detail by structures of various G-protein complexes with GAPs, non-hydrolysable GTP analogs and the transition state mimic, GDP•AlF_4^-. In the structure of Ras in complex with the GTP analog GppNHp, a water molecule is positioned in a way compatible with the nucleophilic attack on the γ-phosphate. AlF_3 and AlF_4^- were originally shown to activate G_α subunits. The aluminium fluorides occupy the place of the γ-phosphate in the GDP bound protein[3]. The structure of AlF^{4-} bound to Gα▪GDP revealed that AlF_4^- is complexed not only with residues that are known to bind to γ-phosphate, but also the residues that are important for the catalytic mechanism of GTP hydrolysis[6]. AlF_4^-, therefore, imitates the γ-phosphate in the pentavalent intermediate state of GTP hydrolysis when the γ-phosphate is nucleophilically attacked by a water molecule. This structure demonstrates how a catalytic glutamine residue of the G-protein brings the water molecule into position for the attack. Ras/RasGAP structure shows that RasGAP stabilizes the position of Gln61 that coordinates the nucleophilic water molecule in Ras. In addition, an arginine residue supplied by the GAP inserts into the nucleotide-binding site of Ras and neutralizes the negative charge of the leaving phosphate. The arginine residue, thus, stabilizes the transition state. These residues are not entirely conserved in small G-proteins and GAPs. In Rab GAP catalyzed inactivation of Rab, the Gln residue that orients the water molecule is supplied by the GAP. In RapGAP, an asparagine orients the water molecule. The details of the catalytic mechanism differ between different GAP/G-protein pairs, but the general steps in the GAP catalyzed activation are always the stabilization of the catalytic residues of G-proteins and the contribution of additional catalytic residues in trans[5].

1.3.3 Spatial regulation of small G-proteins

G-proteins are not regulated solely on the level of their activation state. Their cellular localization is regulated by post-translational modifications and protein transporters. The common modification of many Ras superfamily proteins is the covalent attachment of lipids that directs the protein to membrane compartments. The majority of Ras and Rho proteins have a C-terminal CAAX sequence, where X is any amino acid and A any aliphatic amino acid residue[1]. The cysteine residue is the target for prenylation: farnesyl transferases and geraniylgeranyl transferases covalently attach farnesyl- and geranylgeranyl-isoprenoid lipids to the cysteine. This process is irreversible, as the AAX peptide is cleaved off from the protein during processing. In some G-proteins, cysteine residues that are N-terminal to CAAX motif are modified by palmitoylation. These residues are present in H- and N-Ras, and are, together with CAAX sequence, responsible for localization of H- and N-Ras at the Golgi and the plasma membrane[7,8]. The palmitoylation reaction is reversible; acyl protein thiosterase (APT1) has been shown to remove the palmitoyl moieties from Ras proteins[9,10]. K-Ras contains a lysine-rich polybasic region N-terminal to the CAAX motif, which stabilizes the membrane localization by electrostatic interaction with negatively charged lipid head groups. Similarly to K-Ras, Rap1A and Rap1B contain a poly-lysine stretch, while the Rap2 isoforms, Rap2A, Rap2B and Rap2C, contain additional cysteine residues that are target of prenylation. Rab family members can be mono- or doubly-prenylated as well. Instead of a CAAX motif, they possess a CC or CXC C-terminal motif[11]. Arf family members are modified by N-terminal myristoylation, which confines them to cellular membranes[12,13].
Lipidated G-proteins do not always diffuse to the target cell compartment freely. Guanine nucleotide dissociation inhibitors, GDIs, bind to prenylated Rho and Rab proteins masking the prenyl group by placing it in their hydrophobic pocket. In that way, otherwise insoluble lipidated proteins are shuffled through the cytoplasm. GDIs bind GDP- but not GTP-bound Rho and Rab proteins. Thereby, GDIs deliver G-proteins to and extract G-proteins from membranes[14,15]. The GDI-like factor PDEδ has a role in transporting Ras, Rap and Rheb proteins. PDEδ binds to farnesylated and depalmitoylated small G-proteins independent of their nucleotide status and enhances their diffusion in the cytoplasm[16–18].

II Ras family of small G-proteins

The Ras family is comprised of 36 members, among which are Ras, Rap, Ral, Rit, R-Ras and Rheb. One of the most studied ones are the founding members of the family, H-, K- and N-Ras that are commonly mutated in cancer[19]. G-proteins from the Ras family have a central function in the control of proliferation, transcription, migration, exo- and endo-cytosis and cell-polarity.

2.1 Ras

The classical Ras group consists of H-Ras, K-Ras and N-Ras, which control transcription, proliferation, survival and differentiation. Mutations in these genes can, therefore, result in significant perturbations of cell signaling. It is estimated that activating point mutations in Ras are found in 20 to 30% of all human metastatic cancers[20–22]. Furthermore, mutations in genes which are part of the Ras-signaling pathway are common in certain types of cancer: 50% of melanomas contain Raf mutations, while 20-30% of non-small cell lung carcinoma contain mutated K-Ras[23,24].

There are three major Ras controlled signaling pathways: activation of the Raf and MAP kinase cascade, activation of phosphatidylinositol 3-kinase (PI3K) and activation of Ral-GEFs form the RalGDS-family. Raf, PI3K and Ral-GEFs contain an ubiquitin-like fold, which is responsible for the interaction with Ras•GTP. The ubiquitin-like folds are either classified as a Raf- and PI3K-like Ras-Binding Domains (RBDs) or RalGDS-like Ras Association (RA) domains.

The MAPK pathway is initiated by binding of growth factors, hormones and chemokines to receptor tyrosine kinases (RTKs). Ligand-bound RTKs dimerize and undergo auto-phosphorylation in trans. The phosphorylated tyrosine residues serve as the recognition sites for the binding of SH2 domain-containing proteins, such as the adaptor protein Grb2. By binding to RTKs, Grb2 recruits Sos, a GEF for Ras, to the plasma membrane, where Ras is localized. Active Ras then directly binds to Raf, which results in Raf activation by the release of auto-inhibition and translocation to the cell membrane[25]. Activated Raf phosphorylates its substrate MEK1/2, which in turn phosphorylates ERK1/2. Finally, phosphorylated ERK translocates to the nucleus where it phosphorylates and activates several transcription factors[26].

The Ras effector PI3K phosphorylates the 3´-hydroxyl group of phosphoinositides. Class I PI3Ks are heterodimers that consist of various combinations of regulatory and catalytic subunits. The regulatory subunit blocks the catalytic subunit. This autoinhibition is relieved upon binding of the catalytic subunit to Ras•GTP and of the regulatory subunit to RTKs or adaptor proteins. The RBD from the catalytic subunit makes interactions with both switch I and switch II loops in Ras. In addition, direct contact is made between Ras and catalytic domain of catlytic subunit[27]. It is not entirely clear whether the interaction of the catalytic subunit with Ras only relieves the interaction with the regulatory subunit and affects the localization of PI3K, or if there is allosteric activation of kinase activity as well[28]. Once PI3K is activated, it phosphorylates its substrate PIP_2 to PIP_3, which binds to several PH domain-containing proteins and functions to localize them to plasma membrane. The PI3K pathway is negatively regulated by PTEN that dephosphorylates PIP_3 to PIP_2. PKB kinase, that binds to and is activated by PIP_3, functions as an important mediator of PI3K signaling and plays an important role in the regulation of cellular survival. It phosphorylates, and thereby deactivates, several substrates: the pro-apoptotic protein Bad, forkhead transcription factors that are involved in expression of pro-apoptotic genes, and TSC2, a component of the Rheb GAP complex, resulting in the increased protein synthesis. Phosphorylation of Mdm2 results in its activation and antagonizes p53-

mediated apoptosis. PI3K also phosphorylates IKKα, the inhibitor of the transcription factor NF-κB. This phosphorylation results in the activation of NFκB and increased trancription of anti-apoptotic genes[26].

2.2 Ral

The third major group of Ras effectors, RalGDS-like GEFs, connects Ras and Ral signaling. RalGDS-like GEFs bind to Ras•GTP via their RA domain. This interaction brings RalGDS-like GEFs to the membrane where their substrates RalA and RalB are localized (see below). Ral signaling is involved in the regulation of cell proliferation, exocytosis, endocytosis, motility and cancer development[29]. Several effectors of Ral are identified so far, like RLIP76, which acts as a GAP for Rac, the transcription factor, ZONAB, and the components of the exocyst complex, Sec5 and Exo84. Unlike Ras effectors that share the same protein fold, Ral effectors have structurally unrelated domains which mediate binding to Ral•GTP. The RalA/Sec5 structure reveals that Sec5 interacts with switch I of RalA via an immunoglobulin-like domain forming intermolecular β-sheet. Exo84 binds to switch I and switch II via a PH domain, while RLIP76 binds to switch I and switch II by a coiled-coil domain[30].

The two isoforms of Ral, RalA and RalB, are suggested to have different roles in cell signaling. Even though both RalA and RalB are able to bind to components of the exocyst complex, only RalA was found to influence trafficking of E-cadherin in vivo. This functional difference is attributed to differences in cellular localizations of RalA and RalB. Only RalA localizes to perinuclear recycling endosomes where vesicle sorting is regulated. The finding that RalA mutants that do not localize to perinuclear endosomes fail to promote basolateral delivery of E-cadherin, supports the importance of localization in mediating the differential functions of RalA and RalB[31].

The initial studies of the role of RalGEFs in cancer development attributed a minor role for Ral in tumor progression. Subsequent studies, however, demonstrated a crucial role of RalGEFs in the promotion of Ras-mediated tumorigenesis and cancer growth[32]. The transforming role of RalGEFs has been related to their catalytic activity: while Rlf is able to induce transformation, its catalytically dead construct fails to do so[33]. In accordance with the different roles of RalA and RalB in cell signaling, they were also demonstrated to perform different functions in cancer. Depletion of RalA leads to the impairment of anchorage-independent proliferation of cancer cells, while depletion of RalA in adherent cells has no inhibitory effect on proliferation. Depletion of RalB is shown to be necessary for the survival of several tested tumor-derived cell lines, but not for the survival of non-cancerous cells. Interestingly, siRNA-induced depletion RalA reversed the apoptotic phenotype observed upon loss of RalB alone, suggesting antagonistic effects of RalA on RalB induced apoptosis in tumor cell signaling[34]. On the other hand, RalA and RalB were found to function redundantly in tumorigenesis of K-Ras-driven non-small cell lung carcinoma in null or conditional knockout mouse models. Deletion of either RalA or RalB did not interfere with the initiation of the tumor growth, while deletion of both Ral genes

blocked tumor formation. Interestingly, the loss of one Ral isoform was partially compensated by the increase in the activation of the other isoform without an increase in protein level. In accordance with the redundant role in tumor formation, RalA and RalB were found to be mostly co-localized[35]. It may be that RalA and RalB have redundant role during tumor initiation while they could have different functions in established tumors.

2.3 Rap

The group of Rap proteins contains five members: Rap1A, Rap1B, Rap2A, Rap2B and Rap2C. They show high sequence homology, with the exception of their C-termini. Rap is a close homologue of Ras that has initially been identified as reverting transforming effect of Ras[36]. Since Rap shows high affinity for some of the Ras effectors[37], this reversion of Ras-induced transformation was at first explained by sequestration of Ras effectors by Rap in unproductive complexes[38]. However, the reverting effect of Rap on Ras-induced transformation was demonstrated not to include negative regulation of Ras effectors[39]. Instead Rap functions via its own effectors, some of which are the RA domain-containing adaptor proteins Riam, RapL, Tiam, Krit, AF6 Arap1 and Arap3. Some effectors of Ras, such as Raf and PI3K are reported to be activated by Rap as well, but these findings remain controversial[40]. Similarly to Ras, Rap effectors contain ubiquitin-like fold domains that interact with Rap.

Rap functions in the control of cell-cell adhesion, cell-matrix adhesion and regulation of spreading, polarity and cell migration[41,42].

Rap1 has a central role in the control of the cell-cell contacts via regulation of adherens junctions. Cadherin proteins interact with each other via their extracellular regions forming adherens junctions, while their intracellular tails bind to the actin cytoskeleton via a set of junctional proteins. The role of Rap1 in the control of cell-cell junctions was firstly demonstrated in Drosophila melanogaster epithelial wing cells mutant for Rap1. These cells lose their cohesion and disperse into surrounding tissue due to defects in their adherens junctions[43]. This effect comes from the involvement of Rap in recruitment and stabilization of the junctional protein E-cadherin at the membrane[42].

Adhesion of cells to the extracellular matrix is mediated by integrins and Rap has been implicated in the control of this process. For example, Riam has been suggested to be a Rap effector that enhances adhesion in T-cells by influencing actin organization and integrin affinity[44], and RapL mediates lymphocyte adhesion[45].

Rap is involved in the regulation of cell migration as well. Activation of endogenous Rap in epithelial cells inhibits growth factor-induced migration by stabilization of focal adhesions. It is suggested that Rap affects focal adhesion dynamics by modulating connections between integrins and actin[46]. Rap mutants show impairment in cell migration during morphogenesis in the drosophila embryo[47]. C3G, a GEF for Rap, has also been implicated in migration of neuronal cells during development of mouse embryos[48].

The link between Rap and the control of polarity was firstly established in yeast. The Rap homologue, Bud1, was found to determine the position of the bud site in yeast[49]. Rap also

controls the determination of axon position in neurites, polarization of lymphocytes and brush border formation in intestinal epithelial cells[50–54].

2.4 Rheb

The first identified function of Rheb was the regulation of the uptake of the amino acids arginine and lysine[55] and the induction of growth arrest in response to limiting nutrients[56]. Drosophila studies had revealed that these effects arise from the involvement of Rheb in signaling of mTor, and that Rheb most likely signals downstream of the TSC1/TSC2 complex[57]. mTor is a kinase found in two distinct complexes, TORC1 and TORC2. TORC1 regulates growth-related cellular processes, such as protein synthesis, nutrient uptake and ribosome biogenesis. Studies in mammalian cells demonstrated that the TSC1/TSC2 complex functions as GAP for Rheb[58], and confirmed the role of Rheb in mTor signaling[59,60]. The TSC1/TSC2 complex itself is negatively regulated by phosphorylation by PKB, that is triggered in response to insulin stimulation[61].
The mechanism by which Rheb affects mTor activity is not entirely clear. A direct interaction between mTor and Rheb has been demonstrated, but this interaction seemed to be weak and was not confirmed between endogenous proteins[62]. Furthermore, this interaction did not depend on GTP loading, so mTor does not seem to behave as a classical G-protein effector. However, GTP loading of Rheb is shown to be critical for the activation of mTor[63]. Several different proteins have been suggested to function as a GEF for Rheb, but an unambiguous Rheb GEF has still not been identified[64].

III Structure and regulation of Ras family guanine nucleotide exchange factors

Not all G-proteins have assigned GEFs but they are known for the most well characterized G-proteins. Classical Ras proteins are activated by Sos1 and 2, RasGRF1 and 2 and RasGRP1 and 3. The diverse group of GEFs for Rap consists of C3G, Epac1 and 2, PDZGEF1 and 2, Repac (MR-GEF), RasGRP2 and 3, RasGEF1A and B and PLCε. Several GEFs catalyze the activation of RalA and RalB: The RalGDS group of Ral-GEFs consists of RalGDS, Rgl1, Rgl2/Rlf and Rgl3, while RalGPS 1 and 2 belong to the RalGPS group. In this way, various signaling pathways can converge to activate the same G-proteins via acting through different GEFs.

3.1 Structural organization of Ras family GEFs

GEFs for the Ras family of small G-proteins are multidomain proteins (Figure 4). The domains fulfill various functions: catalysis of nucleotide exchange, structural stabilization, regulation of GEF activity by inhibition of the catalytic domain, and regulation of GEF

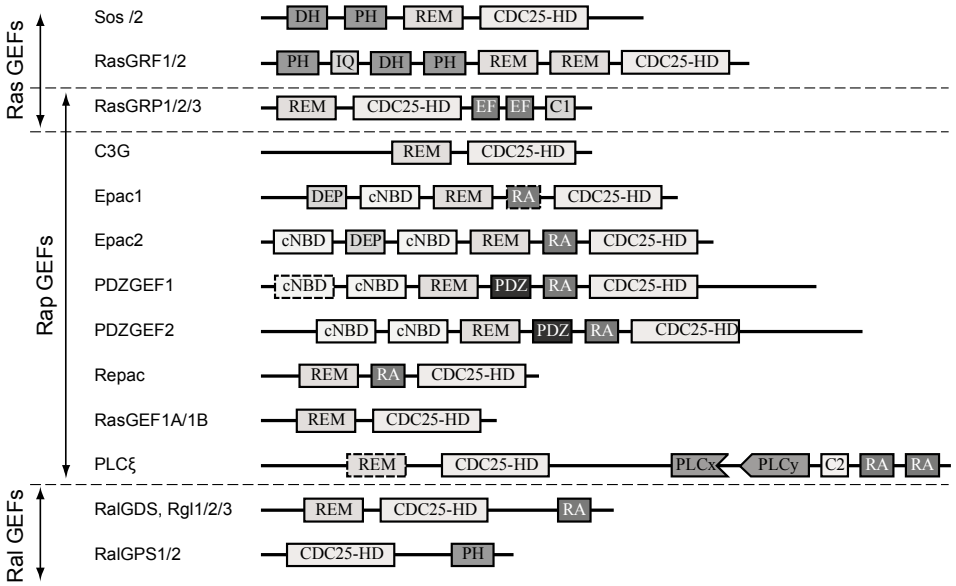


Figure 4. Domain organization of Ras-family GEFs. Dashed lines indicate a putative cNBD in PDZGEF1, a putative RA domain in Epac1 and a putative REM domain in PLCε. REM - Ras Exchange Motif, CDC25-HD – CDC25-Homology Domain, PH – Pleckstrin Homology, DH-PH – Dbl Homology-Pleckstrin Homology, IQ - IQ motif (involved in calmodulin binding), EF - EF hands (involved in Ca2+ binding), C1 – protein kinase C conserved region 1, DEP – Dishevelled, Egl-10, Pleckstrin, cNBD – Cyclic Nucleotide Binding Domain, RA – Ras Association, PDZ – PSD95, DlgA, Zo-1, PLCx and PLCy (catalytic lipase domain of PLC), C2 – protein kinase C conserved region 2 (involved in Ca2+ binding).

localization by binding to lipids and proteins. The catalytic domain of Ras family GEFs is CDC25-homology domain (CDC25-HD). This domain is in basically all GEFs preceded by a Ras-exchange motif (REM) domain that structurally stabilizes the accompanying CDC25-HD. Regulatory domains are involved in the regulation of the activity of the CDC25-HD and in the regulation of the cellular localization of the GEF. For example, the cyclic-nucleotide-binding domains of the Rap-GEF, Epac, interact with the CDC25-HD occluding the Rap-binding site, and, thus, repress Epac activity. Cellular localization of GEFs is regulated in various ways in order to bring GEFs in cellular compartments where their substrates are localized or to keep them separated from G-proteins. As described, RA domains bind to Ras•GTP or Rap•GTP and bring GEFs for other G-proteins to the membrane. PH domains recruit GEFs to membrane-bound lipids or other sites where PH domain-interacting proteins are localized[65]. Some GEFs, such as Sos and C3G, contain proline-rich sequences that are recognized by SH3 domain-containing proteins, and it is via these proteins that Sos and C3G are targeted to particular cellular regions[66,67]. On the other hand, binding of RanBP2 to the CDC25-HD of Epac1 confines Epac to the nuclear envelope and blocks its catalytic activity[68]. Some of these regulatory domains and sequences are discussed below in more detail.

3.1.1 Catalytic region of Ras family GEFs

Every G-protein family has its own type of GEF: DH-PH tandem domains and DOCK DHR-2 domains catalyze the nucleotide exchange of the Rho family, the β-propeller fold in RCC1 catalyzes Ran activation, Sec7 domains are found in Arf GEFs while activation of Rab family G-proteins is catalyzed by Vps-9, Sec2, DENN and Mss4 domains[5,69,70]. These domains are structurally unrelated.
The catalytic domain of GEFs for the Ras family, the CDC25-HD, co-occur with a REM domain, with the exception of RalGPS1 and RalGPS2, GEFs for Ral, and possibly PLCε, a GEF for Rap. All other Ras family GEFs contain a REM domain, whose main role is to stabilize the structure of the CDC25-HD.
The interaction of CDC25-HD with G-proteins has been investigated by mutational and structural studies. The structure of the Ras/Sos complex reveals an extensive binding interface, with over 30 interacting residues. Contacts are established with the P-loop (residues 10-17), switch I (residues 35-40), switch II (residues 57-75) and the α3 helix (residues 95-105) of Ras. Interactions that are the most critical for the GEF catalytic exchange mechanism are established between the helical hairpin from CDC25-HD and switch I in Ras. The helical hairpin displaces switch I and, therefore, destabilizes binding of the phosphate moiety of the nucleotide, which is the first step in the GEF catalyzed dissociation of the nucleotide. The structure also reveals positioning of the REM domain in the proximity of the helical hairpin of Sos. Two conserved hydrophobic side chains from the helical hairpin, Ile956 and Phe958, insert into a hydrophobic groove of the REM domain. This interaction stabilizes the conformation of the helical hairpin, supporting the suggestion that the role of REM domain is primarily the structural stabilization of the CDC25-HD[71].
The general structure and orientation of the CDC25-HD and the REM domain of Epac2 are very similar[72,73]. The mechanism of destabilization of nucleotide binding is the same: the CDC25-HD of Epac2 interacts with switch I of Rap. As in Sos, the helical hairpin of Epac2 makes interactions with the REM domain via the insertion of hydrophobic residues into the hydrophobic pocket of the REM domain.

3.1.2 Regulatory domains of Ras family GEFs

Activation of G-proteins needs to occur only upon the proper biological stimuli, to ensure a tight regulation of GEF activity. There are several different ways by which this regulation is accomplished. Some GEFs are autoinhibited: for example C3G is negatively regulated by its N-terminal part (chapter 2 of this thesis) while Epac2 is negatively regulated by its cyclic nucleotide-binding domains (cNBDs). The release from the inhibition can occur as a consequence of the binding of small molecules or other proteins, by posttranslational modifications or by combination of these. The activity of GEFs is often regulated by translocation to compartments where the substrate G-protein is found. Another possibility for the regulation of GEFs is allosteric regulation by binding of other

proteins or molecules. Sos and Epac are discussed below as examples of the complex regulation mechanisms.

3.1.2.1 RA domain

RA domains bind to Ras•GTP and Rap•GTP and thereby classify as effector proteins. However, RA domains are also present in several GEFs for G-proteins of the Ras family: Members of the RalGDS group, Epac1 and Epac2, PDZGEF1 and PDZGEF2 and Repac contain one RA domain each and PLCε contains two RA domains. These GEFs can, therefore, interlink different G-proteins directly.
The major regulatory function of RA domains in CDC25-HD-containing GEFs seems to be spatial control of GEF activity. The classical example is the recruitment of GEFs from the RalGDS group by Ras to cause activation of Ral signaling. Overexpressed RalGDS is localized in the cytoplasm, but, when co-expressed with an active mutant of Ras, it is membrane localized, and cells show increased GTP loading of Ral[74]. The effect of Ras on RalGDS is demonstrated to be the consequence of the translocation, as artificially membrane localized RalGDS causes Ras independent activation of Ral. Similarly, PLCε translocates to membrane by binding to H-Ras and possibly also to Rap1 via its RA domain[75]. The N-terminal RA domain of PLCε does not bind to H-Ras, however, the second RA domain binds to H-Ras with 10-15 fold higher affinity than to Rap1, Rap2 and R-Ras. The function of Ras binding to PLCε seems not to be the spatial control only. It has been shown that binding of Ras to PLCε increases the lipase activity *in vitro*. Since the removal of both RA-domains leads to the increase in PLCε lipase activity as well, binding of H-Ras has been suggested to relieve RA domain mediated autoinhibition[76].
Epac2 binds to H-, K- and N-Ras via its RA domain[77]. This interaction has been suggested to mediate the translocation of Epac2 to the plasma membrane. Interestingly, the RA domain of the Rap-GEF, PDZGEF1, binds to active Rap and M-Ras[78,79]. The interaction with active Rap might accomplish a positive feedback in Rap signaling. The RA domain of PDZGEF2 has been shown to bind to active M-Ras[80,81]. The interaction of PDZGEF2 and M-Ras plays a role in TNF-α signaling: TNF-α induced activation of M-Ras causes the translocation of PDZGEF2 to plasma membrane where it activates Rap1[81]. Finally, the RA domain of Repac is suggested to bind to M-Ras, resulting in inhibition of Repac[78] albeit the mechanistic basis for this inhibition in not entirely clear.

3.1.2.2 PH domain

PH domains are found in RasGRF1 and RasGRF2, Sos1 and Sos2 and RalGPS1 and RalGPS2. RasGRF1 and RasGRF2 contain two PH domains, one N-terminal PH domain and one PH domain that belongs to a DH-PH module. Sos1 and Sos2 each contain a DH-PH module, and RalGPS contains a single PH domain.
PH domains are shown to mediate interactions with lipids and proteins. Around 10% to 20% of all characterized PH domains are thought to bind to membrane-bound

phospholipids[82]. The PH domain of Sos is reported to bind PIP_2 and phosphatidic acid (PA)[66,83,84]. It has been demonstrated that PA is synthesized in response to growth factor signaling, and that a PH domain mediated interaction with PA causes the translocation of Sos to cellular membranes[83]. This is thus an alternative mechanism to the aforementioned Grb2 mediated translocation. The N-terminal PH domain of RasGRF binds to phosphatidylinositol and contributes to the localization of RasGRF. In addition, the PH domain has been suggested to negatively regulate the catalytic activity of RasGRF[65], but, at the moment, the putative regulatory function of the PH domain remains controversial.
DH-PH modules act as GEFs for the Rho family of small G-proteins, whereby the DH domain is the actual catalytic domain. The general importance of PH domains for the function of DH domains is not entirely clear[85]. The PH domain increases the catalytic efficacy of the DH domain in some cases. In Vav, the PH domain seems even to be required for catalytic activity. Structural studies on Vav point to a possible role of PH domain in the stabilization of Vav DH domain[86]. However, in other cases an inhibitory function of the PH domain has been suggested.

3.2 Regulation of SOS

The human Sos proteins are about 150 kDa in size. The two homologues, Sos1 and Sos2, that show 70% homology are ubiquitously expressed. While the knock-out of Sos1 causes embryonic death in mice, knock-down of Sos2 does not result in any observable developmental defect[87,88]. Interestingly, $Sos1^{-/-}$ cells are resistant to transformation by v-Src and do not respond to EGF stimulation even though the remaining Sos2 is able to activate Ras and to interact as Sos1 via Grb2 with the EGFR.
The N-terminus of Sos proteins contain a histone-fold followed by a DH-PH module and a C-terminal REM domain and CDC25-HD. Sos, thus, functions as a GEF for both Rho and Ras family members via its PH-DH module and its CDC25-HD, respectively. The PH-DH module acts on Rac[89], while CDC25-HD acts on H-, K- and N-Ras proteins[90,91]. The C-terminal part of Sos1 contains multiple proline-rich sequences that bind to SH3 domains of adaptor proteins, such as Grb2, Crk and Nck[92,93].
One of the first insights in the regulation of Sos was that translocation of Sos to the plasma membrane drastically increases levels of Ras•GTP [94]. The adaptor protein Grb2 was found to bind to proline-rich sequence in the C-terminal part of SOS translocating it to plasma membrane where it binds to phosphotyrosine sites of EGF receptor via its SH2 domain[92,95,96]. This brings Sos in the proximity of its substrate, Ras, resulting in the increase in Ras GTP loading[97]. In this way, activation of Ras is coupled to the changes in tyrosine phosphorylation.
The translocation of Sos is, however, not sufficient for the full activation of Ras. Membrane-targeted C-terminally truncated Sos possesses higher catalytic activity than the non-truncated one[94]. Furthermore, a direct negative effect of N-terminal parts on the catalytic activity towards Ras have been shown[98,99]. In addition, Sos was found to be phosphorylated in its C-terminus by MAP kinase[100]. This phosphorylation results in the dissociation of Sos/Grb2 complex[101–103]. Finally, a positive regulation of Sos by an

allosteric effect of Ras•GTP binding adds an additional level of complexity to Sos regulation[104]. The allosteric regulation is mediated by a Ras•GTP molecule that binds to interfaces of the REM domain and CDC25-HD. Sos is, therefore, subjected to positive feedback regulation. Structural analysis of a Sos construct that encompasses DH-PH module next to the REM domain and CDC25-HD has shown that the DH-PH module blocks the allosteric Ras-binding[105]. PIP_2 binding to the PH domain has been suggested to release the block of the allosteric Ras-binding site. A crucial function of the PH- and REM-domains in the regulation of Sos is supported by the observation that mutations in this regions are correlated with Sos gain-of-function in patients with Noonan Syndrome[106].

3.3 Regulation of Epac

Epac1 and Epac2 consist of an N-terminal regulatory and a C-terminal catalytic part. The regulatory part of Epac1 contains a DEP domain and a cNBD, while Epac2 has an additional cNBD domain on its N-terminus. The catalytic part of Epac consists of a REM domain and a CDC25-HD domain, with a RA domain inserted between.
In the autoinhibited conformation, the regulatory part of Epac makes contacts with the CDC25-HD and prevents Rap from binding[72]. Binding of cAMP breaks the inhibitory interaction between the cNBD and CDC25-HD by rearranging the relative position of the two domains and leaving the CDC25-HD accessible to Rap (Figure 5). Interestingly, although Epac2 has two cNBD, the first cNBD is not necessary for the autoinhibition since the deletion of this domain does not change the regulation of Epac2 by cAMP. Furthermore, binding of cAMP to the first cNBD is not necessary for the activation of Epac2 either[73].

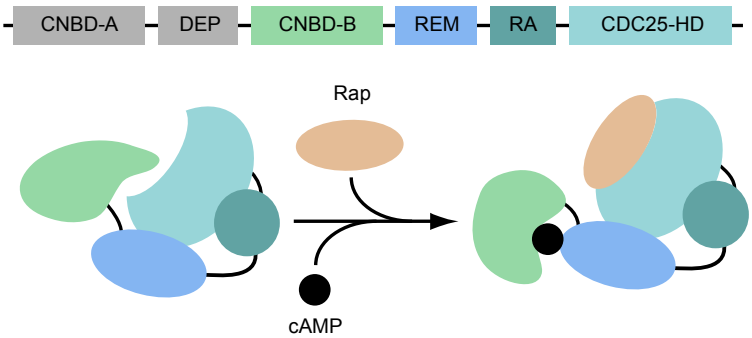


Figure 5. cAMP regulation of the activity of Epac2. cAMP brings Epac2 into the open conformation whereby access of Rap to CDC25-HD becomes free. The crystallized construct of active Epac2 does not include the first cNBD and DEP domains (indicated in grey)[73].

In order to release the autoinhibition of Epac2, the second cNBD needs to swing away from the CDC25-HD. In the absence of cAMP, this movement is sterically blocked by

residues form the phosphate-binding cassette (PBC). The PBC is highly conserved in cNBDs and interacts with the phosphate-sugar moiety of the cyclic nucleotide. Upon cAMP binding, the PBC tightens and releases the steric block of the movement. The active conformation is further stabilized by newly formed interactions of the cNBD with the REM domain.
Another consequence of cAMP binding to Epac1 is the translocation of Epac1 to the plasma membrane[107]. In that way, cAMP exhibits a dual control of Epac1activity by both release of autoinhibition and by translocaion to the membrane. The translocation is mediated by the DEP domain. Independently of this process the ERM family members ezrin, radixin and moesin recruit Epac1 the membrane by binding to the far N-terminus of Epac[108]. Finally, Epac1 localizes to nuclear pore by the interaction with nuclear pore protein RanBP2. RanBP2 binds to CDC25-HD of Epac and keeps it inactive. Upon phosphorylation of RanBP2, Epac1 is released from the nuclear envelope[68].

3.4 Regulation of Ral GEFs

The RalGDS group of Ral-GEFs consists of four proteins that share the same domain organization. These are RalGDS, Rgl1, Rgl2/Rlf and Rgl3. They contain an N-terminal REM domain and CDC25-HD followed by a C-terminal RA domain. The RA domain binds to Ras•GTP and Rap•GTP with an affinity of 1 µM and 0.08 µM, respectively[79,109]. This interaction mediates recruitment of the RalGEFs to membranes where Ral is localized as well (Figure 6)[110]. Artificial targeting of RalGDS or Rlf to the membrane by a CAAX fusion, leads to an increase in Ral•GTP levels[74,111].

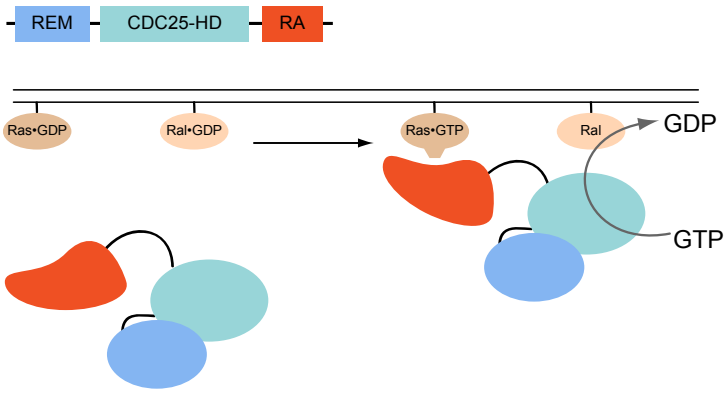


Figure 6. Regulation of the Ral-GEF Rlf by localization. When Ras is inactive, Rlf is localized in the cytoplasm. Upon activation of Ras, the RA domain of Rlf binds to Ras and brings Rlf to plasma membrane. Rlf is thus localized in proximity of Ral and catalyzes activation of Ral.

In addition to spatial regulation, RalGDS is suggested to be inhibited by its N-terminal region[112,113]. PKC-induced phosphorylation of N-terminal residues in RalGDS prevents

the dissociation of the N-terminus from the CDC25HD and keeps RalGDS in the inactive conformation[112]. On the other hand, binding of PDK1 to the N-terminus of RalGDS results in induction of nucleotide exchange activity[113].

A different manner of regulation by phosphorylation is reported for Rgl2. PKA phosphorylates the C-terminal region of Rgl2 and thereby reduces the ability of Rgl2 to interact with Ras and Rap. This prevents Ras induced membrane translocation and results in lower Ral•GTP levels[114].

Next to the interaction with Ras, additional protein interactions are reported to influence localization of RalGEFs. RalGDS interacts with β-arrestin. This interaction keeps RalGDS in the cytoplasm. Upon activation of the fMLP receptor the RalGDS•β-arrestin complex dissociates, and RalGDS can translocate to the membrane[115].

Ral-GEFs are localized differentially to control the activity of RalA and RalB. RalGEFs, as well as RalA and RalB, change their localizations during the progression of cytokinesis. For example, during the early phases of cytokinesis, RalGDS, but not Rgl1, colocalizes with RalA and Sec5 to the midbody ring. In a later phase, Rgl1 and RalB become enriched at the midbody[116].

3.5 Regulation of C3G

The Rap-GEF C3G has been identified as a protein that interacts with SH3 domains of Grb2 and Crk[117]. It has been shown to function as a GEF for Rap1-[118], but not Rap2-isoforms[119]. C3G was also suggested to act as a GEF for R-Ras and TC21 (R-Ras2) and for the member of the Rho family TC10[120–123]. The activation of TC10 does not seem to be direct, since C3G is not able to catalyze the exchange of nucleotide on TC10 *in vitro*[124].

In its C-terminus, C3G contains a REM domain and a CDC25-HD. The N-terminal part of the protein (residues 1 to 690) contains no predicted domains. Several proline-rich sequences that bind to SH3 domain-containing proteins are found in this region. Four highly similar proline motifs are shown to bind to the adaptor protein Crk[125]. Some of the proline motifs are also reported to bind to the SH3 domain of the kinases Hck and Abl[126,127]. A distinct proline motif has been shown to bind to the SH3 domain of adaptor protein p130Cas[128]. Furthermore, the region from 144 to 230 binds E-cadherin, resulting in Rap1 activation to regulate the recruitment of E-cadherin to cell-cell junctions[129].

Regulation of C3G activation has been studied in detail by the group of Matsuda[67,130]: Overexpression of the catalytic domain of C3G resulted in higher Rap1•GTP levels than overexpression of full length C3G, suggesting that the N-terminal part of the protein negatively regulates C3G activity[67].

In cells, overexpression of Crk I, Crk II and CrkL remarkably enhanced Rap1•GTP levels in a C3G-dependent manner. Both SH2 and SH3 domains of Crk were shown to be necessary for this effect. *In vitro* studies found that Crk had no effect on the activity of C3G, suggesting that Crk functions in the spatial control of C3G only. Indeed, in cell fractionation studies, Crk overexpression enriched C3G in the membrane fraction and artificially membrane targeted C3G increased Rap•GTP levels[67].

In addition, overexpression of Crk was found to induce tyrosine phosphorylation of C3G. The phosphorylation was mapped to Tyr504. C3G Y504F was less efficient in inducing Rap•GTP levels than the wild type protein. Therefore, a putative allosteric regulation of C3G by phosphorylation was suggested[130]. The tyrosine kinases Src[131], Lyn[132], Hck[126] and Abl [133,134] and additional phosphorylation sites have been implicated to regulate localization and activity of C3G. In addition, C3G has been shown to be negatively regulated by the phosphatase TC48[135].
Crk recruits C3G to various protein complexes. Upon mechanical stretching of cells[136] or B-cell receptor stimulation[137], C3G is recruited to p130Cas/Crk/C3G complex, while stimulation of integrins have been shown to result in formation of both p130Cas/Crk/C3G[138] and Cbl/Crk/C3G[139] complexes. The interaction between C3G and Crk is subjected to regulation as well: insulin stimulation leads to the dissociation of Crk/C3G complex, perhaps by inducing inhibitory phosphorylation of Crk[140]. The C3G/CrkL complex has been demonstrated to be negatively regulated by c-Cbl-induced ubiquitination of Crk in T-cells[141].

3.6 SH3 domain-containing proteins involved in the regulation of C3G

Around 300 SH3 domains are encoded by the human genome. SH3 domain containing proteins function in many cellular pathways. SH3 domains are present in adaptor proteins, such as Crk, Grb2 and Nck, and are also found in proteins which enzymatic activity such as the protein kinases Src and Abl, RasGAP, PI3K, and phospholipases.

3.6.1 SH3 domain

The SH3 domain was one of the first domains recognized for the establishment of protein-protein interactions. The founding members of the family were the SH3 domains in adaptor proteins like Crk, the protein kinase Src and phospholipase Cγ (PLCγ)[142]. Early work identified proline-rich sequences, with the minimal consensus being PxxP, as the binding sites recognized by SH3 domains[143]. Structural studies on the interaction between SH3 domains and short model peptides have shown that the peptides can bind to the SH3 domain in both orientations. Peptides were, therefore, classified as class I with the consensus R/KxxPxxP and class II with the consensus xPxxPxR/K. The affinity of the model peptides is typically in the micromolar range. The peptides adapt the conformation of a polyproline type II helix, where the amino-acid chain makes a left-turn helix with three residues per turn and where all the proline residues are in the trans conformation[144]. This helical structure has a triangular cross-section with the base of a triangle laying on the SH3 domain surface (Figure 7). The peptide-binding groove of SH3 domains consists of three relatively shallow pockets defined by conserved hydrophobic residues. Two of these pockets are responsible for binding of the proline residues, while the third, the so called “specificity pocket”, interacts with the arginine or lysine side chain.

Apart from the canonical interaction of SH3 domains, atypical interactions with SH3 domains are described. In these cases, the peptides do not adapt the conformation of a polyproline-2 helix and/or do not require additional determinants in place of basic residues for binding to the specificity pocket. For example the SH3 domain of Abl kinase binds hydrophobic rather than basic residues in its specificity pocket[145].

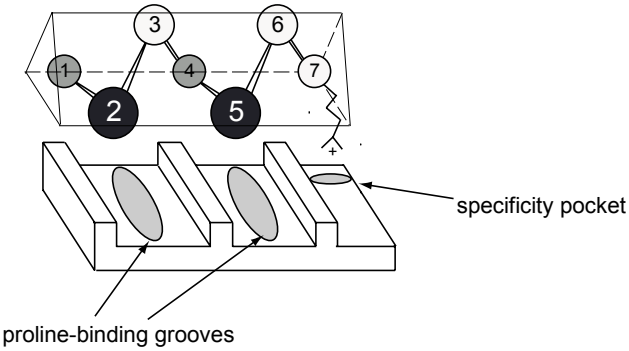


Figure 7. Binding of xPxxPxR/K peptides to SH3 domains. Adapted from ref[181]. The peptide adopts the conformation of a polyproline type II helix, where three amino acid residues form one turn. The proline residues number 2 and 5 bind to the proline grooves of the SH3 domain. The charged amino acid residue binds to the specificity pocket.

3.6.2 Crk

Crk was identified as a transforming factor of avian sarcoma retrovirus, CT10. Crk was shown to selectively increase tyrosine phosphorylation, even though it did not contain any region homologous to the catalytic domains of known protein kinases. Crk was, therefore, postulated to act by modulating the activity of endogenous tyrosine kinase, hence its name CT10 regulator of kinase (Crk)[146]. Soon, two cellular homologues of viral Crk were identified[147]. The two forms, Crk I and Crk II, arise from a single gene by alternative splicing[148]. Crk I contains as the viral version one SH2 and one SH3 domain, while Crk II contains an additional C-terminal SH3 domain (Figure 8). The SH3 domain in CrkI and the corresponding SH3 domain in Crk II are, therefore, referred to as $SH3^N$, while the additional SH3 domain in Crk II is termed $SH3^C$. Another Crk-family protein, CrkL, was identified and shown to have the same domain organization as Crk II[149].
The SH2 domain and the $SH3^N$ domain are involved in the interactions with other proteins and serve to assemble protein complexes. The SH2 domain recognizes pY-X-X-P sequence (where pY stands for phosphorylated tyrosine residue) and mediates interactions with, amongst others, $p130^{Cas}$, Cbl, paxillin and Gab1. The $SH3^N$ domain associates with proteins that contain a PxLPxR/K sequence, such as Abl, Sos, C3G and DOCK180[150]. The $SH3^C$ domain seems not to recognize proline rich sequences[150]. The apparent lack of affinity for proline-rich peptides is in agreement with the absence of conserved aromatic residues which are involved in peptide binding in $SH3^N$ [151].

Crk family members function as adaptor proteins since they serve to link various signaling proteins. By doing so, Crk proteins are involved in the control of transcription, cell proliferation, phagocytosis, endocytosis, cellular adhesion and migration and apoptosis[152]. Increased expression of Crk proteins has been found in several tumor types[153]. CrkL is necessary for the tumorigenic function of Bcr-Abl, an oncogenic kinase causing chronic myelogenous leukemia[154]. Interestingly, the tumorigenic potential seems to differ between Crk isoforms. Crk I but not Crk II expressing-fibroblasts proliferate in soft agar and induce massive tumors in nude mice. This suggests a role of Crk I in anchorage-independent growth[148]. Crk I, therefore, seems to have higher biological activity than Crk II. On the other hand, high overexpression of Crk II does induce growth of NIH 3T3 and several other fibroblast-derived cell lines[155]. Further, Ogawa *et al.* report that overexpression of Crk I has a stronger effect on cellular transformation than Crk II[156]. Taken together, these results support a putative regulatory function of $SH3^C$, which balances the oncogenic properties of Crk.

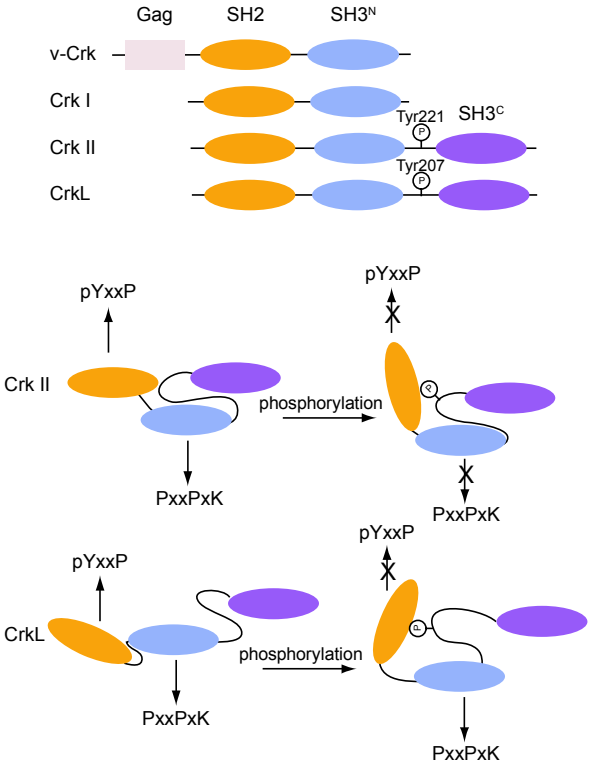


Figure 8. Domain organization of Crk proteins and their regulation. Adapted from ref[161]. Crk II and CrkL are regulated by inhibitory phosphorylations of the indicated tyrosine residues. The SH2 domains of Crk II and CrkL bind to the pYXXP motif. Upon phosphorylation of tyrosine residues within the $SH3^N$-$SH3^C$ linker, the SH2 domain becomes involved in an intramolecular interaction. The N-terminal SH3 domain of phosphorylated CrkL is additionally blocked by $SH3^C$.

In agreement with this, Feller *et al.*[157] found that c-Abl binds to the $SH3^N$ domain and then phosphorylates Y221, in the linker between the two SH3 domains, which is absent in Crk I. The SH2 domain binds to pY221 and is thus prevented from interaction with other proteins (Figure 8). Furthermore, the interaction of the SH2 domain with the pY221 renders the binding site the $SH3^N$ domain occluded[158]. Similarly to Crk II, CrkL was shown to be phosphorylated on Tyr207[159]. Phosphorylation of Tyr207 is demonstrated to negatively regulate the protein[160].

Despite the similarities in the sequence and in the regulation by the establishment of intramolecular interaction between the SH2 domain and a phosphorylated tyrosine residue, CrkL and Crk II display remarkable different features[161]. While the structure of Crk II is compact and the SH2 domain freely accessible[158], the structure of CrkL is elongated and the binding site of the SH2 domain in CrkL seems to be partially occluded[162]. The accessibility of the $SH3^N$ domain of CrkL and Crk II differs as well: While the $SH3^N$ domain in CrkL is freely accessible, $SH3^N$ in Crk II is partially occluded[158,163]. Upon phosphorylation, the SH2 domains of both CrkL and Crk II are inaccessible, but while the $SH3^N$ domain of phosphorylated Crk II is inaccessible as well, the $SH3^N$ domain in phosphorylated CrkL remains exposed (Figure 8).

Crk II is shown to be additionally regulated by a proline switch. The NMR structure of Crk II demonstrates that Pro238, in the linker between the SH3 domains, can interchange between the cis- and trans- conformation. The trans conformation renders Crk II in the open conformation, while the switch into the cis-form results in the association of the $SH3^N$ domain with the linker and the $SH3^C$ domain[163,164].

3.6.3 Src kinase family

Src is the founding member of the Src kinase family of non-receptor tyrosine kinases, that consists of eight members: Src, Fyn, Lyn, Hck, Yes, Blk, Fgr and Lck. Src was discovered as the oncogene responsible for tumorigenic property of Rous sarcoma virus[165]. After sequencing of viral Src (v-Src), the existence of a homolog gene in vertebrates, termed c-Src, was found. Unlike v-Src, c-Src has low transforming activity in mammalian and avian cell lines[166] and soon Src was identified as the first tyrosine kinase[167,168,169].

Src kinases are involved in the transmission of signals from outside to the inside of the cell. They mainly function downstream of RTKs and integrins. Src kinases, in complex with FAK, signal upstream of Ras and PI3K, paxillin, p130Cas, Rho family small G-proteins, regulating processes such as gene expression, cytoskeletal reorganization, adhesion and survival. Src family members are found overexpressed, but rarely mutated, in various types of cancer. It seems that their role is in the maintenance and progression of tumors, rather than in oncogenesis[170].

All members of the family have a similar domain organization (Figure 9). An N-terminal unique domain is followed by one SH2 and one SH3 domain and the kinase domain itself. The N-terminal tail of Src kinases is myristoilated to mediate membrane association and, in addition, some Src kinases can be palmitoylated.

The SH2 domain of Src has a dual role. It mediates interactions with other proteins and is involved in the establishment of the autoinhibited state of Src. Tyr530 (numbering according to the human sequence) in the C-terminal tail can be phosphorylated by, among others, Csk. In consequence, it binds to the SH2 domain and renders the kinase in an inactive[171]. Similarly, the SH3 domain contributes to the maintenance of the inhibited state by an internal interaction[172]. Deletion and mutations of the SH3 domain, as well as the addition of peptides with high affinity for the SH3 domain, result in increased kinase activity[173].

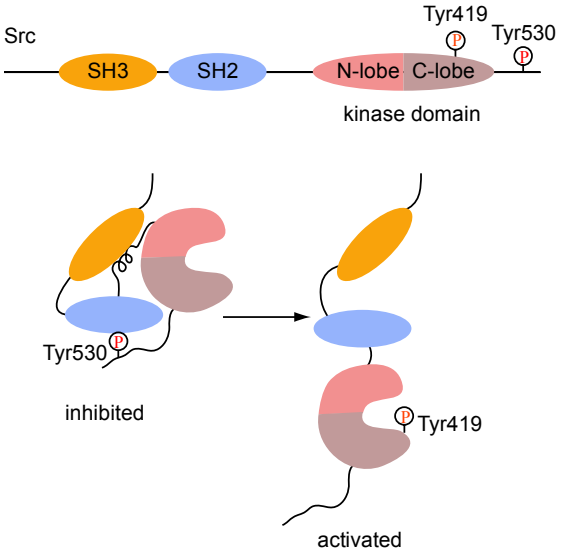


Figure 9. Domain organization and regulation of Src. In the inhibited conformation the SH2 and SH3 domains of Src are packed in the back of Src kinase domain. The SH2 and SH3 domains are engaged in intramolecular interactions. In the active conformation, Tyr419 in the activation loop of Src is phosphorylated and the activation loop adapts a conformation that allows the access of substrates.

It was, at first, speculated that the SH2 domain might inhibit Src kinase by directly blocking the active site. Structures of Src and Hck in their inactive conformations demonstrated that this was not the case[174,175]. Instead, the SH2 domain and the SH3 domain are packed to the back of the kinase domain. Thereby, the two domains influence the conformation of the activation loop. The activation loop is located within the C-terminal lobe of the kinase domain, and, in its inactive conformation, it adapts helical conformation with Tyr419 pointing inwards. In this conformation, the access to the protein substrates is blocked. Upon release of the inhibitory interaction of the SH2- and SH3-domain, the activation loop is pulled and its helical conformation is disturbed. This exposes Tyr419 and makes it accessible for phosphorylation. Subsequent phosphorylation of Tyr419 further stabilizes the active conformation of the activation loop, and that allows the access of the substrate to the catalytic site[176]. This phosphorylation is catalyzed in trans by Src itself [173].

Src itself is rarely mutated in cancer, still, elevated Src activity is correlated with higher malignancy in colon cancer. It seems, therefore, likely that Src activation is achieved by deregulation of regulators of Src activity. For example, PTP1B, the phosphatase that dephosphorylates the inhibitory Tyr530, is found to be overexpressed in colon cancer[177].

IV Selectivity in Ras family signaling

Ras family members are grouped together based on their sequence similarities. The sequences of members of different Ras family subgroups differ enough to provide selective recognition by interacting proteins. For example, Ral-GEFs act only on RalA and RaB and not on any other Ras family member[91] and Sos1 and Sos2 activate H-, K- and N-Ras which exhibit partial functional redundancy due to the ability to interact with the same set of effectors[37]. While the regions in G-proteins that establish interactions with their regulators and effectors are largely identical, their C-termini are very variable. C-terminal variability within the same subgroup of G-proteins is responsible for differences in post-translational processing and for different cellular localizations. The ability to target G-protein isoforms to different cellular compartments is demonstrated to be in part responsible for the differences observed in the outcome of cellular signaling. For example, H-Ras is confined to lipid rafts in the plasma membrane, and fails to activate Raf and PI3K pathways[178]. Another mechanism to control the signaling outcome of different isoforms is via different expression levels[179].

Ras- and Rap-proteins, even though being members of different groups, share high sequence similarity in the region responsible for interaction with regulators and effectors. Therefore, certain GEFs and effectors interact with Rap- and Ras-proteins, in addition to those that are either Rap- or Ras-specific. The RBD of Raf is able to interact with Ras-proteins as well as with Rap1A and Rap1B, and, in a similar way, the RA domains of the proteins RalGDS, affadin-6, Krit1, PDZGEF1 or RIN1 may also interact with both Ras and Rap proteins[79]. Affinity constants for interactions of these proteins with Ras and Rap were measured and found to differ in average around 10 fold. They can be relatively low, as is the case for the interaction of PDZGEF1 with H-Ras (33 μM). On the other hand, the affinity of both H-Ras and Rap1B for Raf is very high (0.67 μM and 0.08 μM, respectively)[79]. High affinity is a good predictor for the biological relevance of the interaction in some cases: a clear correlation between the abilities of various Ras family members to interact with Raf in cells and their abilities to stimulate the kinase activity of Raf was found[37]. The drawback of studies that assess the occurrence of G-protein/effector interactions *in vivo* by simple overexpression of G-proteins is that they detect interactions that occur under resting conditions and often out of the context of normal cellular localization. This explains, in part, why the relevance of certain interactions between G-proteins and their putative effectors is still debated. For example, Raf is reported to act both downstream of Ras and Rap in some studies, while others could not find a direct link between Raf and Rap[40]. Similarly, PI3K is reported to be activated by Rap1 as well. On the other hand, dual activation of Ras and Rap by GEFs is well established. RasGRP3 has been demonstrated to activate both Ras- and Rap-proteins *in vivo* and *in vitro*[180].

V The scope of the thesis

The role of GEFs to activate small G-proteins places them in the center of some of the most critical cellular processes. They need to function timely, in confined cellular regions and to act selectively on particular small G-proteins. Therefore, their activity and localization is highly regulated. In this thesis, we describe our studies of the regulation of several GEFs for G-proteins of the Ras-family.
In Chapter 2, we investigated the regulation of the Rap-GEF C3G. We show that C3G activation is a multistep process that requires both interaction of C3G with the adapter protein, Crk, and phosphorylation of C3G tyrosine residues by Src kinase.
In Chapter 3, we present the structure of catalytic region of the Ral-GEF, Rlf. This structure contains the novel β-sheet element that precedes the REM domain. An exposed proline-rich loop is the binding site for SH3 domain-containing interaction partners of Rlf.
The basis for the selectivity of GEFs of the Ras-family with a focus on Ras- and Rap-GEFs is studied in Chapter 4. Most GEFs show preference for either Ras- or Rap-proteins, with the exception of RasGRP3 that acts on both. We identify key residues in Rap that determine the selectivity of the GEFs and set up rules for the prediction of the selectivity.
A summary and discussion of our results is presented in Chapter 5.

References

1. Wennerberg, K. The Ras superfamily at a glance. *Journal of Cell Science* **118,** 843–846 (2005).
2. Vetter, I. R. The Guanine Nucleotide-Binding Switch in Three Dimensions. *Science* **294,** 1299–1304 (2001).
3. Sprang, S. R. Emerging roles for RGS proteins in cell signalling. *Trends in pharmacological sciences* **20,** 376–382 (1999).
4. Yang, J., Zhang, Z., Roe, S. M., Marshall, C. J. & Barford, D. Activation of Rho GTPases by DOCK exchange factors is mediated by a nucleotide sensor. *Science* **325,** 1398–1402 (2009).
5. Bos, J. L., Rehmann, H. & Wittinghofer, A. GEFs and GAPs: Critical Elements in the Control of Small G Proteins. *Cell* **129,** 865–877 (2007).
6. Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E. & Sigler, P. B. GTPase mechanism of Gproteins from the 1.7-A crystal structure of transducin alpha-GDP-AIF-4. *Nature* **372,** 276–279 (1994).
7. Rocks, O. An Acylation Cycle Regulates Localization and Activity of Palmitoylated Ras Isoforms. *Science* **307,** 1746–1752 (2005).
8. Rocks, O., Peyker, A. & Bastiaens, P. I. Spatio-temporal segregation of Ras signals: one ship, three anchors, many harbors. *Current Opinion in Cell Biology* **18,** 351–357 (2006).
9. Magee, A. I., Gutierrez, L., McKay, I. A., Marshall, C. J. & Hall, A. Dynamic fatty acylation of p21N-ras. *The EMBO journal* **6,** 3353 (1987).
10. Dekker, F. J. *et al.* Small-molecule inhibition of APT1 affects Ras localization and signaling. *Nature Chemical Biology* **6,** 449–456 (2010).
11. Triola, G., Waldmann, H. & Hedberg, C. Chemical Biology of Lipidated Proteins. *ACS Chemical Biology* **7,** 87–99 (2012).
12. Antonny, B., Beraud-Dufour, S., Chardin, P. & Chabre, M. N-terminal hydrophobic residues of the G-protein ADP-ribosylation factor-1 insert into membrane phospholipids upon GDP to GTP exchange. *Biochemistry* **36,** 4675–4684 (1997).
13. Matheson, L. A., Suri, S. S., Hanton, S. L., Chatre, L. & Brandizzi, F. Correct Targeting of Plant ARF GTPases Relies on Distinct Protein Domains. *Traffic* **9,** 103–120 (2008).
14. Seabra, M. C. & Wasmeier, C. Controlling the location and activation of Rab GTPases. *Current Opinion in Cell Biology* **16,** 451–457 (2004).
15. Goody, R. S., Rak, A. & Alexandrov, K. The structural and mechanistic basis for recycling of Rab proteins between membrane compartments. *Cellular and Molecular Life Sciences* **62,** 1657–1670 (2005).
16. Ismail, S. A. *et al.* Arl2-GTP and Arl3-GTP regulate a GDI-like transport system for farnesylated cargo. *Nature Chemical Biology* **7,** 942–949 (2011).
17. Hanzal-Bayer, M., Renault, L., Roversi, P., Wittinghofer, A. & Hillig, R. C. The complex of Arl2-GTP and PDEdelta: from structure to function. *The EMBO journal* **21,** 2095–2106 (2002).
18. Chandra, A. *et al.* The GDI-like solubilizing factor PDEδ sustains the spatial organization and signalling of Ras family proteins. *Nature Cell Biology* **14,** 148–158 (2011).
19. Baines, A. T., Xu, D. & Der, C. J. Inhibition of Ras for cancer treatment: the search continues. *Future Medicinal Chemistry* **3,** 1787–1808 (2011).
20. Karnoub, A. E. & Weinberg, R. A. Ras oncogenes: split personalities. *Nat. Rev. Mol. Cell Biol.* **9,** 517–531 (2008).
21. Santarpia, L., Lippman, S. M. & El-Naggar, A. K. Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy. *Expert opinion on therapeutic targets* **16,** 103–119 (2012).
22. Ward, A. F., Braun, B. S. & Shannon, K. M. Targeting oncogenic Ras signaling in hematologic malignancies. *Blood* **120,** 3397–3406 (2012).

23. Tsao, H., Chin, L., Garraway, L. A. & Fisher, D. E. Melanoma: from mutations to medicine. *Genes Dev.* **26,** 1131–1155 (2012).
24. Califano, R., Landi, L. & Cappuzzo, F. Prognostic and predictive value of K-RAS mutations in non-small cell lung cancer. *Drugs* **72 Suppl 1,** 28–36 (2012).
25. Cutler, R. E., Stephens, R. M., Saracino, M. R. & Morrison, D. K. Autoregulation of the Raf-1 serine/threonine kinase. *Proceedings of the National Academy of Sciences* **95,** 9214–9219 (1998).
26. De Luca, A., Maiello, M. R., D'Alessio, A., Pergameno, M. & Normanno, N. The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches. *Expert Opin. Ther. Targets* **16 Suppl 2,** S17–27 (2012).
27. Pacold, M. E. *et al.* Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase γ. *Cell* **103,** 931–944 (2000).
28. Castellano, E. & Downward, J. RAS Interaction with PI3K More Than Just Another Effector Pathway. *Genes & cancer* **2,** 261–274 (2011).
29. Camonis, J. H. & White, M. A. Ral GTPases: corrupting the exocyst in cancer cells. *Trends in cell biology* **15,** 327–332 (2005).
30. Fenwick, R. B. *et al.* The RalB-RLIP76 complex reveals a novel mode of ral-effector interaction. *Structure* **18,** 985–995 (2010).
31. Shipitsin, M. & Feig, L. A. RalA but not RalB enhances polarized delivery of membrane proteins to the basolateral surface of epithelial cells. *Molecular and cellular biology* **24,** 5746–5756 (2004).
32. Ferro, E. & Trabalzini, L. RalGDS family members couple Ras to Ral signalling and that's not all. *Cellular Signalling* **22,** 1804–1810 (2010).
33. Lim, K. H. *et al.* Activation of RalA is critical for Ras-induced tumorigenesis of human cells. *Cancer cell* **7,** 533–545 (2005).
34. Chien, Y. & White, M. A. RAL GTPases are linchpin modulators of human tumour-cell proliferation and survival. *EMBO reports* **4,** 800–806 (2003).
35. Peschard, P. *et al.* Genetic deletion of RALA and RALB small GTPases reveals redundant functions in development and tumorigenesis. *Curr. Biol.* **22,** 2063–2068 (2012).
36. Kitayama, H., Sugimoto, Y., Matsuzaki, T., Ikawa, Y. & Noda, M. A ras-related gene with transformation suppressor activity. *Cell* **56,** 77–84 (1989).
37. Rodriguez-Viciana, P., Sabatier, C. & McCormick, F. Signaling specificity by Ras family GTPases is determined by the full spectrum of effectors they regulate. *Molecular and cellular biology* **24,** 4943–4954 (2004).
38. Reuther, G. W. & Der, C. J. The Ras branch of small GTPases: Ras family members don't fall far from the tree. *Current opinion in cell biology* **12,** 157–165 (2000).
39. Zwartkruis, F. J. T., Wolthuis, R. M. F., Nabben, N. M. J. M., Franke, B. & Bos, J. L. Extracellular signal-regulated activation of Rap1 fails to interfere in Ras effector signalling. *The EMBO journal* **17,** 5905–5912 (1998).
40. Raaijmakers, J. H. & Bos, J. L. Specificity in Ras and Rap Signaling. *Journal of Biological Chemistry* **284,** 10995–10999 (2008).
41. Kooistra, M. R. H., Dubé, N. & Bos, J. L. Rap1: a key regulator in cell-cell junction formation. *Journal of cell science* **120,** 17–22 (2007).
42. Pannekoek, W. J., Kooistra, M. R. H., Zwartkruis, F. J. T. & Bos, J. L. Cell–cell junction formation: the role of Rap1 and Rap1 guanine nucleotide exchange factors. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1788,** 790–796 (2009).

43. Knox, A. L. & Brown, N. H. Rap1 GTPase regulation of adherens junction positioning and cell adhesion. *Science* **295,** 1285–1288 (2002).
44. Lafuente, E. M. *et al.* RIAM, an Ena/VASP and Profilin ligand, interacts with Rap1-GTP and mediates Rap1-induced adhesion. *Developmental cell* **7,** 585–595 (2004).
45. Katagiri, K., Maeda, A., Shimonaka, M. & Kinashi, T. RAPL, a Rap1-binding molecule that mediates Rap1-induced adhesion through spatial regulation of LFA-1. *Nat. Immunol.* **4,** 741–748 (2003).
46. Lyle, K. S., Raaijmakers, J. H., Bruinsma, W., Bos, J. L. & De Rooij, J. cAMP-induced Epac-Rap activation inhibits epithelial cell migration by modulating focal adhesion and leading edge dynamics. *Cellular signalling* **20,** 1104–1116 (2008).
47. Asha, H., De Ruiter, N. D., Wang, M. G. & Hariharan, I. K. The Rap1 GTPase functions as a regulator of morphogenesis in vivo. *The EMBO journal* **18,** 605–615 (1999).
48. Yip, Y. P., Thomas, T., Voss, A. K. & Yip, J. W. Migration of sympathetic preganglionic neurons in the spinal cord of a C3G-deficient mouse suggests that C3G acts in the reelin signaling pathway. *The Journal of comparative neurology* **520,** 3194–3202 (2012).
49. Park, H. O., Chant, J. & Herskowitz, I. BUD2 encodes a GTPase-activating protein for Bud1/Rsr1 necessary for proper bud-site selection in yeast. *Nature* **365,** 269–274 (1993).
50. Schwamborn, J. C. & Püschel, A. W. The sequential activity of the GTPases Rap1B and Cdc42 determines neuronal polarity. *Nature neuroscience* **7,** 923–929 (2004).
51. Shimonaka, M. *et al.* Rap1 translates chemokine signals to integrin activation, cell polarization, and motility across vascular endothelium under flow. *J. Cell Biol.* **161,** 417–427 (2003).
52. Gérard, A., Mertens, A. E. E., Van Der Kammen, R. A. & Collard, J. G. The Par polarity complex regulates Rap1-and chemokine-induced T cell polarization. *The Journal of cell biology* **176,** 863–875 (2007).
53. Lorenowicz, M. J., Van Gils, J., De Boer, M., Hordijk, P. L. & Fernandez-Borja, M. Epac1-Rap1 signaling regulates monocyte adhesion and chemotaxis. *Journal of leukocyte biology* **80,** 1542–1552 (2006).
54. Gloerich, M. *et al.* Rap2A links intestinal cell polarity to brush border formation. *Nat. Cell Biol.* **14,** 793–801 (2012).
55. Urano, J., Tabancay, A. P., Yang, W. & Tamanoi, F. The Saccharomyces cerevisiae Rheb G-protein is involved in regulating canavanine resistance and arginine uptake. *Journal of Biological Chemistry* **275,** 11198–11206 (2000).
56. Mach, K. E., Furge, K. A. & Albright, C. F. Loss of Rhb1, a Rheb-related GTPase in fission yeast, causes growth arrest with a terminal phenotype similar to that caused by nitrogen starvation. *Genetics* **155,** 611–622 (2000).
57. Stocker, H. *et al.* Rheb is an essential regulator of S6K in controlling cell growth in Drosophila. *Nature cell biology* **5,** 559–566 (2003).
58. Garami, A. *et al.* Insulin activation of Rheb, a mediator of mTOR/S6K/4E-BP signaling, is inhibited by TSC1 and 2. *Molecular cell* **11,** 1457–1466 (2003).
59. Inoki, K., Li, Y., Xu, T. & Guan, K. L. Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. *Genes & development* **17,** 1829–1834 (2003).
60. Tee, A. R., Manning, B. D., Roux, P. P., Cantley, L. C. & Blenis, J. Tuberous sclerosis complex gene products, Tuberin and Hamartin, control mTOR signaling by acting as a GTPase-activating protein complex toward Rheb. *Current biology* **13,** 1259–1268 (2003).
61. Wullschleger, S., Loewith, R. & Hall, M. N. TOR signaling in growth and metabolism. *Cell* **124,** 471–484 (2006).
62. Yonezawa, K. & Avruch, J. Rheb binds and regulates the mTOR kinase. *Current Biology* **15,** 702–713 (2005).

63. Sancak, Y. *et al.* PRAS40 is an insulin-regulated inhibitor of the mTORC1 protein kinase. *Molecular cell* **25,** 903–915 (2007).
64. Durán, R. V. & Hall, M. N. Regulation of TOR by small GTPases. *EMBO Rep.* **13,** 121–128 (2012).
65. Fernández-Medarde, A. & Santos, E. The RasGrf family of mammalian guanine nucleotide exchange factors. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **1815,** 170–188 (2011).
66. Pierre, S. & Coumoul, X. Understanding SOS (Son of Sevenless). *Biochemical pharmacology* **82,** 1049–1056 (2011).
67. Ichiba, T. *et al.* Enhancement of guanine-nucleotide exchange activity of C3G for Rap1 by the expression of Crk, CrkL, and Grb2. *J. Biol. Chem.* **272,** 22215–22220 (1997).
68. Gloerich, M. *et al.* The nucleoporin RanBP2 tethers the cAMP effector Epac1 and inhibits its catalytic activity. *Science Signalling* **193,** 1009 (2011).
69. Barr, F. & Lambright, D. G. Rab GEFs and GAPs. *Current opinion in cell biology* **22,** 461–470 (2010).
70. Wu, X., Ramachandran, S., Lin, M.-C. J., Cerione, R. A. & Erickson, J. W. A minimal Rac activation domain in the unconventional guanine nucleotide exchange factor Dock180. *Biochemistry* **50,** 1070–1080 (2011).
71. Boriack-Sjodin, P. A., Margarit, S. M., Bar-Sagi, D. & Kuriyan, J. The structural basis of the activation of Ras by Sos. *Nature* **394,** 337–343 (1998).
72. Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A. & Bos, J. L. Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state. *Nature* **439,** 625–628 (2006).
73. Rehmann, H. *et al.* Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B. *Nature* **455,** 124–127 (2008).
74. Matsubara, K. *et al.* Plasma membrane recruitment of RalGDS is critical for Ras-dependent Ral activation. *Oncogene* **18,** 1303–1312 (1999).
75. Song, C. *et al.* Regulation of a novel human phospholipase C, PLCepsilon, through membrane targeting by Ras. *Science Signalling* **276,** 2752 (2001).
76. Bunney, T. D. *et al.* Structural and mechanistic insights into ras association domains of phospholipase C epsilon. *Mol. Cell* **21,** 495–507 (2006).
77. Li, Y. *et al.* The RAP1 guanine nucleotide exchange factor Epac2 couples cyclic AMP and Ras signals at the plasma membrane. *Journal of Biological Chemistry* **281,** 2506–2514 (2006).
78. Rebhun, J. F., Castro, A. F. & Quilliam, L. A. Identification of guanine nucleotide exchange factors (GEFs) for the Rap1 GTPase. *Journal of Biological Chemistry* **275,** 34901–34908 (2000).
79. Wohlgemuth, S. *et al.* Recognizing and Defining True Ras Binding Domains I: Biochemical Analysis. *Journal of Molecular Biology* **348,** 741–758 (2005).
80. Gao, X. *et al.* Identification and characterization of RA-GEF-2, a Rap guanine nucleotide exchange factor that serves as a downstream target of M-Ras. *Journal of Biological Chemistry* **276,** 42219–42225 (2001).
81. Yoshikawa, Y. *et al.* The M-Ras-RA-GEF-2-Rap1 Pathway Mediates Tumor Necrosis Factor-alpha–dependent Regulation of Integrin Activation in Splenocytes. *Molecular biology of the cell* **18,** 2949–2959 (2007).
82. Dinitto, J. & Lambright, D. Membrane and juxtamembrane targeting by PH and PTB domains. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1761,** 850–867 (2006).

83. Zhao, C., Du, G., Skowronek, K., Frohman, M. A. & Bar-Sagi, D. Phospholipase D2-generated phosphatidic acid couples EGFR stimulation to Ras activation by Sos. *Nat. Cell Biol.* **9,** 706–712 (2007).
84. Findlay, G. M. & Pawson, T. How is SOS activated? Let us count the ways. *Nat. Struct. Mol. Biol.* **15,** 538–540 (2008).
85. Viaud, J., Gaits-Iacovoni, F. & Payrastre, B. Regulation of the DH–PH tandem of guanine nucleotide exchange factor for Rho GTPases by phosphoinositides. *Advances in Biological Regulation* **52,** 303–314 (2012).
86. Rapley, J., Tybulewicz, V. L. J. & Rittinger, K. Crucial structural role for the PH and C1 domains of the Vav1 exchange factor. *EMBO Rep.* **9,** 655–661 (2008).
87. Qian, X. *et al.* The Sos1 and Sos2 Ras-specific exchange factors: differences in placental expression and signaling properties. *The EMBO journal* **19,** 642–654 (2000).
88. Esteban, L. M. *et al.* Ras-guanine nucleotide exchange factor sos2 is dispensable for mouse growth and development. *Molecular and cellular biology* **20,** 6410–6413 (2000).
89. Nimnual, A. S., Yatsula,, B. A. & Bar-Sagi, D. Coupling of Ras and Rac Guanosine Triphosphatases Through the Ras Exchanger Sos. *Science* **279,** 560–563 (1998).
90. Bonfini, L., Karlovich, C. A., Dasgupta, C. & Banerjee, U. The Son of sevenless gene product: a putative activator of Ras. *Science (New York, NY)* **255,** 603 (1992).
91. Quilliam, L. A., Rebhun, J. F. & Castro, A. F. A growing family of guanine nucleotide exchange factors is responsible for activation of Ras-family GTPases. *Prog. Nucleic Acid Res. Mol. Biol.* **71,** 391–444 (2002).
92. Buday, L. & Downward, J. Epidermal growth factor regulates the exchange rate of guanine nucleotides on p21ras in fibroblasts. *Mol. Cell. Biol.* **13,** 1903–1910 (1993).
93. Hall, B. E., Yang, S. S. & Bar-Sagi, D. Autoinhibition of Sos by intramolecular interactions. *Front Biosci* **7,** 288–94 (2002).
94. Aronheim, A. *et al.* Membrane targeting of the nucleotide exchange factor Sos is sufficient for activating the Ras signaling pathway. *Cell* **78,** 949–961 (1994).
95. Lowenstein, E. J. *et al.* The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling. *Cell* **70,** 431–442 (1992).
96. Egan, S. E. *et al.* Association of Sos Ras exchange protein with Grb2 is implicated in tyrosine kinase signal transduction and transformation. *Nature* **363,** 45–51 (1993).
97. Quilliam, L. A. *et al.* Membrane-targeting potentiates guanine nucleotide exchange factor CDC25 and SOS1 activation of Ras transforming activity. *Proceedings of the National Academy of Sciences* **91,** 8512–8516 (1994).
98. Chen, R. H., Corbalan-Garcia, S. & Bar-Sagi, D. The role of the PH domain in the signal-dependent membrane targeting of Sos. *The EMBO journal* **16,** 1351–1359 (1997).
99. Corbalan-Garcia, S., Margarit, S. M., Galron, D., Yang, S. & Bar-Sagi, D. Regulation of Sos activity by intramolecular interactions. *Molecular and cellular biology* **18,** 880–886 (1998).
100. Corbalan-Garcia, S., Yang, S. S., Degenhardt, K. R. & Bar-Sagi, D. Identification of the mitogen-activated protein kinase phosphorylation sites on human Sos1 that regulate interaction with Grb2. *Molecular and cellular biology* **16,** 5674–5682 (1996).
101. Cherniack, A. D., Klarlund, J. K., Conway, B. R. & Czech, M. P. Disassembly of Son-of-sevenless proteins from Grb2 during p21ras desensitization by insulin. *J. Biol. Chem.* **270,** 1485–1488 (1995).
102. Rozakis-Adcock, M., Van der Geer, P., Mbamalu, G. & Pawson, T. MAP kinase phosphorylation of mSos1 promotes dissociation of mSos1-Shc and mSos1-EGF receptor complexes. *Oncogene* **11,** 1417–1426 (1995).

103. Waters, S. B. *et al.* Desensitization of Ras activation by a feedback disassociation of the SOS-Grb2 complex. *Journal of Biological Chemistry* **270,** 20883–20886 (1995).
104. Margarit, S. *et al.* Structural evidence for feedback activation by Ras· GTP of the Ras-specific nucleotide exchange factor SOS. *Cell* **112,** 685–695 (2003).
105. Sondermann, H. *et al.* Structural analysis of autoinhibition in the Ras activator Son of sevenless. *Cell* **119,** 393–405 (2004).
106. Tartaglia, M. *et al.* Gain-of-function SOS1 mutations cause a distinctive form of Noonan syndrome. *Nature Genetics* **39,** 75–79 (2006).
107. Ponsioen, B. *et al.* Direct spatial control of Epac1 by cyclic AMP. *Science Signalling* **29,** 2521 (2009).
108. Gloerich, M. *et al.* Spatial regulation of cyclic AMP-Epac1 signaling in cell adhesion by ERM proteins. *Molecular and cellular biology* **30,** 5421–5431 (2010).
109. Spaargaren, M. & Bischoff, J. R. Identification of the guanine nucleotide dissociation stimulator for Ral as a putative effector molecule of R-ras, H-ras, K-ras, and Rap. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 12609–12613 (1994).
110. Kishida, S. *et al.* Colocalization of Ras and Ral on the membrane is required for Ras-dependent Ral activation through Ral GDP dissociation stimulator. *Oncogene* **15,** 2899–2907 (1997).
111. Wolthuis, R. M. F., De Ruiter, N. D., Cool, R. H. & Bos, J. L. Stimulation of gene induction and cell growth by the Ras effector Rlf. *The EMBO journal* **16,** 6748–6761 (1997).
112. Rusanescu, G., Gotoh, T., Tian, X. & Feig, L. A. Regulation of Ras signaling specificity by protein kinase C. *Mol. Cell. Biol.* **21,** 2650–2658 (2001).
113. Tian, X., Rusanescu, G., Hou, W., Schaffhausen, B. & Feig, L. A. PDK1 mediates growth factor-induced Ral-GEF activation by a kinase-independent mechanism. *EMBO J.* **21,** 1327–1338 (2002).
114. Ferro, E. *et al.* G-protein binding features and regulation of the RalGDS family member, RGL2. *Biochem. J.* **415,** 145–154 (2008).
115. Bhattacharya, M. *et al.* Beta-arrestins regulate a Ral-GDS Ral effector pathway that mediates cytoskeletal reorganization. *Nat. Cell Biol.* **4,** 547–555 (2002).
116. Cascone, I. *et al.* Distinct roles of RalA and RalB in the progression of cytokinesis are supported by distinct RalGEFs. *EMBO J.* **27,** 2375–2387 (2008).
117. Tanaka, S. *et al.* C3G, a guanine nucleotide-releasing protein expressed ubiquitously, binds to the Src homology 3 domains of CRK and GRB2/ASH proteins. *Proceedings of the National Academy of Sciences* **91,** 3443–3447 (1994).
118. Gotoh, T. *et al.* Identification of Rap1 as a target for the Crk SH3 domain-binding guanine nucleotide-releasing factor C3G. *Molecular and cellular biology* **15,** 6746–6753 (1995).
119. Van den Berghe, N., Cool, R. H., Horn, G. & Wittinghofer, A. Biochemical characterization of C3G: an exchange factor that discriminates between Rap1 and Rap2 and is not inhibited by Rap1A(S17N). *Oncogene* **15,** 845–850 (1997).
120. Gotoh, T. *et al.* Activation of R-Ras by Ras-guanine nucleotide-releasing factor. *Journal of Biological Chemistry* **272,** 18602–18607 (1997).
121. Mochizuki, N. *et al.* Crk activation of JNK via C3G and R-Ras. *Journal of Biological Chemistry* **275,** 12667–12671 (2000).
122. Ohba, Y. Regulatory Proteins of R-Ras, TC21/R-Ras2, and M-Ras/R-Ras3. *Journal of Biological Chemistry* **275,** 20020–20026 (2000).
123. Chiang, S. H. *et al.* Insulin-stimulated GLUT4 translocation requires the CAP-dependent activation of TC10. *Nature* **410,** 944–948 (2001).
124. Pannekoek, W. J. *New insights in Rap-mediated control of cell adhesion*. (2012).

125. Knudsen, B. S., Feller, S. M. & Hanafusa, H. Four proline-rich sequences of the guanine-nucleotide exchange factor C3G bind with unique specificity to the first Src homology 3 domain of Crk. *Journal of Biological Chemistry* **269,** 32781–32787 (1994).
126. Shivakrupa, R., Radha, V., Sudhakar, C. & Swarup, G. Physical and functional interaction between Hck tyrosine kinase and guanine nucleotide exchange factor C3G results in apoptosis, which is independent of C3G catalytic domain. *Journal of Biological Chemistry* **278,** 52188–52194 (2003).
127. Radha, V., Rajanna, A., Mitra, A., Rangaraj, N. & Swarup, G. C3G is required for c-Abl-induced filopodia and its overexpression promotes filopodia formation. *Exp. Cell Res.* **313,** 2476–2492 (2007).
128. Kirsch, K. H., Georgescu, M. M. & Hanafusa, H. Direct binding of p130Cas to the guanine nucleotide exchange factor C3G. *Journal of Biological Chemistry* **273,** 25673–25679 (1998).
129. Hogan, C. *et al.* Rap1 regulates the formation of E-cadherin-based cell-cell contacts. *Molecular and cellular biology* **24,** 6690–6700 (2004).
130. Ichiba, T. *et al.* Activation of C3G guanine nucleotide exchange factor for Rap1 by phosphorylation of tyrosine 504. *J. Biol. Chem.* **274,** 14376–14381 (1999).
131. Fukuyama, T. *et al.* Involvement of the c-Src-Crk-C3G-Rap1 signaling in the nectin-induced activation of Cdc42 and formation of adherens junctions. *J. Biol. Chem.* **280,** 815–825 (2005).
132. He, Y. *et al.* The non-receptor tyrosine kinase Lyn controls neutrophil adhesion by recruiting the CrkL-C3G complex and activating Rap1 at the leading edge. *J. Cell. Sci.* **124,** 2153–2164 (2011).
133. Nolz, J. C. *et al.* The WAVE2 complex regulates T cell receptor signaling to integrins via Abl- and CrkL-C3G-mediated activation of Rap1. *J. Cell Biol.* **182,** 1231–1244 (2008).
134. Mitra, A. & Radha, V. F-actin-binding domain of c-Abl regulates localized phosphorylation of C3G: role of C3G in c-Abl-mediated cell death. *Oncogene* **29,** 4528–4542 (2010).
135. Mitra, A., Kalayarasan, S., Gupta, V. & Radha, V. TC-PTP dephosphorylates the guanine nucleotide exchange factor C3G (RapGEF1) and negatively regulates differentiation of human neuroblastoma cells. *PLoS ONE* **6,** e23681 (2011).
136. Tamada, M., Sheetz, M. P. & Sawada, Y. Activation of a signaling cascade by cytoskeleton stretch. *Dev. Cell* **7,** 709–718 (2004).
137. Ingham, R. J. *et al.* B cell antigen receptor signaling induces the formation of complexes containing the Crk adapter proteins. *Journal of Biological Chemistry* **271,** 32306–32314 (1996).
138. Vuori, K., Hirai, H., Aizawa, S. & Ruoslahti, E. Introduction of p130cas signaling complex formation upon integrin-mediated cell adhesion: a role for Src family kinases. *Mol. Cell. Biol.* **16,** 2606–2613 (1996).
139. Uemura, N. & Griffin, J. D. The adapter protein Crkl links Cbl to C3G after integrin ligation and enhances cell migration. *J. Biol. Chem.* **274,** 37525–37532 (1999).
140. Okada, S., Matsuda, M., Anafi, M., Pawson, T. & Pessin, J. E. Insulin regulates the dynamic balance between Ras and Rap1 signaling by coordinating the assembly states of the Grb2-SOS and CrkII-C3G complexes. *EMBO J.* **17,** 2554–2565 (1998).
141. Zhang, W. Negative Regulation of T Cell Antigen Receptor-mediated Crk-L-C3G Signaling and Cell Adhesion by Cbl-b. *Journal of Biological Chemistry* **278,** 23978–23983 (2003).
142. Mayer, B. J. SH3 domains: complexity in moderation. *Journal of Cell Science* **114,** 1253–1263 (2001).
143. Kay, B. K., Williamson, M. P. & Sudol, M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14,** 231–241 (2000).

144. Moradi, M., Babin, V., Roland, C., Darden, T. A. & Sagui, C. Conformations and free energy landscapes of polyproline peptides. *Proceedings of the National Academy of Sciences* **106,** 20746–20751 (2009).
145. Saksela, K. & Permi, P. SH3 domain ligand binding: What's the consensus and where's the specificity? *FEBS letters* **586**, 2609– 2614 (2012).
146. Mayer, B. J., Hamaguchi, M. & Hanafusa, H. A novel viral oncogene with structural similarity to phospholipase C. *Nature* **332,** 272–275 (1988).
147. Reichman, C. T., Mayer, B. J., Keshav, S. & Hanafusa, H. The product of the cellular crk gene consists primarily of SH2 and SH3 regions. *Cell Growth Differ.* **3,** 451–460 (1992).
148. Matsuda, M. *et al.* Two species of human CRK cDNA encode proteins with distinct biological activities. *Mol. Cell. Biol.* **12,** 3482–3489 (1992).
149. Ten Hoeve, J., Morris, C., Heisterkamp, N. & Groffen, J. Isolation and chromosomal localization of CRKL, a human crk-like gene. *Oncogene* **8,** 2469–2474 (1993).
150. Feller, S. M. Crk family adaptors-signalling complex formation and biological roles. *Oncogene* **20,** 6348–6371 (2001).
151. Sriram, G. & Birge, R. B. Commentary: The carboxyl-terminal Crk SH3 domain: Regulatory strategies and new perspectives. *FEBS Lett.* **586,** 2615–2618 (2012).
152. Birge, R. B., Kalodimos, C., Inagaki, F. & Tanaka, S. Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. *Cell Commun Signal* **7,** 13 (2009).
153. Sriram, G. & Birge, R. B. Emerging roles for crk in human cancer. *Genes & cancer* **1,** 1132–1139 (2010).
154. Seo, J.-H. *et al.* A specific need for CRKL in p210BCR-ABL-induced transformation of mouse hematopoietic progenitors. *Cancer Res.* **70,** 7325–7335 (2010).
155. Iwahara, T., Akagi, T., Shishido, T. & Hanafusa, H. CrkII induces serum response factor activation and cellular transformation through its function in Rho activation. *Oncogene* **22,** 5946–5957 (2003).
156. Ogawa, S. *et al.* The C-terminal SH3 domain of the mouse c-Crk protein negatively regulates tyrosine-phosphorylation of Crk associated p130 in rat 3Y1 cells. *Oncogene* **9,** 1669–1678 (1994).
157. Feller, S. M., Knudsen, B. & Hanafusa, H. c-Abl kinase regulates the protein binding activity of c-Crk. *EMBO J.* **13,** 2341–2351 (1994).
158. Kobashigawa, Y. *et al.* Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK. *Nat. Struct. Mol. Biol.* **14,** 503–510 (2007).
159. De Jong, R., Ten Hoeve, J., Heisterkamp, N. & Groffen, J. Tyrosine 207 in CRKL is the BCR/ABL phosphorylation site. *Oncogene* **14,** 507–513 (1997).
160. Senechal, K., Heaney, C., Druker, B. & Sawyers, C. L. Structural requirements for function of the Crkl adapter protein in fibroblasts and hematopoietic cells. *Mol. Cell. Biol.* **18,** 5082–5090 (1998).
161. Bell, E. S. & Park, M. Models of crk adaptor proteins in cancer. *Genes Cancer* **3,** 341–352 (2012).
162. Jankowski, W. *et al.* Domain organization differences explain Bcr-Abl's preference for CrkL over CrkII. *Nat. Chem. Biol.* **8,** 590–596 (2012).
163. Sarkar, P., Saleh, T., Tzeng, S.-R., Birge, R. B. & Kalodimos, C. G. Structural basis for regulation of the Crk signaling protein by a proline switch. *Nat. Chem. Biol.* **7,** 51–57 (2011).
164. Sarkar, P., Reichman, C., Saleh, T., Birge, R. B. & Kalodimos, C. G. Proline cis- trans Isomerization Controls Autoinhibition of a Signaling Protein. *Molecular cell* **25,** 413–426 (2007).
165. Martin, G. S. The hunting of the Src. *Nat. Rev. Mol. Cell Biol.* **2,** 467–475 (2001).

166. Parker, R. C., Varmus, H. E. & Bishop, J. M. Expression of v-src and chicken c-src in rat cells demonstrates qualitative differences between pp60v-src and pp60c-src. *Cell* **37,** 131–139 (1984).
167. Collett, M. S. & Erikson, R. L. Protein kinase activity associated with the avian sarcoma virus src gene product. *Proc. Natl. Acad. Sci. U.S.A.* **75,** 2021–2024 (1978).
168. Levinson, A. D., Oppermann, H., Levintow, L., Varmus, H. E. & Bishop, J. M. Evidence that the transforming gene of avian sarcoma virus encodes a protein kinase associated with a phosphoprotein. *Cell* **15,** 561–572 (1978).
169. Eckhart, W., Hutchinson, M. A. & Hunter, T. An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. *Cell* **18,** 925–933 (1979).
170. Summy, J. M. & Gallick, G. E. Src family kinases in tumor progression and metastasis. *Cancer and metastasis reviews* **22,** 337–358 (2003).
171. Liu, X. *et al.* Regulation of c-Src tyrosine kinase activity by the Src SH2 domain. *Oncogene* **8,** 1119–1126 (1993).
172. Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S. A. & Draetta, G. Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. *EMBO J.* **12,** 2625–2634 (1993).
173. Ingley, E. Src family kinases: regulation of their activities, levels and identification of new pathways. *Biochim. Biophys. Acta* **1784,** 56–65 (2008).
174. Sicheri, F., Moarefi, I. & Kuriyan, J. Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385,** 602–609 (1997).
175. Xu, W., Harrison, S. C. & Eck, M. J. Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385,** 595–602 (1997).
176. Xu, W., Doshi, A., Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol Cell* **3,** 629–638 (1999).
177. Guarino, M. Src signaling in cancer invasion. *Journal of cellular physiology* **223,** 14–26 (2010).
178. Jaumot, M., Yan, J., Clyde-Smith, J., Sluimer, J. & Hancock, J. F. The linker domain of the Ha-Ras hypervariable region regulates interactions with exchange factors, Raf-1 and phosphoinositide 3-kinase. *J. Biol. Chem.* **277,** 272–278 (2002).
179. Lampson, B. L. *et al.* Rare Codons Regulate KRas Oncogenesis. *Current Biology* (2012). doi:10.1016/j.cub.2012.11.031
180. Yamashita, S. CalDAG-GEFIII Activation of Ras, R-Ras, and Rap1. *Journal of Biological Chemistry* **275,** 25488–25493 (2000).
181. Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann, R. N. & Lim, W. A. Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors. *Science* **282,** 2088–2092 (1998).

Chapter 2

Regulation of Guanine Nucleotide Exchange Factor C3G

Milica Popovic, Marije Rensen-De Leeuw and Holger Rehmann

Molecular Cancer Research, Centre of Biomedical Genetics and Cancer Genomics Centre, University Medical Center Utrecht, Utrecht, The Netherlands

Abstract

C3G is a Guanine Nucleotide Exchange Factor (GEF) for the small G-protein Rap1. C3G is involved in embryonic development and the control of adherens junction formation, migration and cell adhesion. It contains four highly similar proline-rich sequences that can bind to the N-terminal SH3 domain of Crk. Crk is adaptor protein that functions in tyrosine kinase signal cascades. The regulation of C3G was suggested to include Crk-mediated translocation to cellular membrane and Crk-dependent phoshorylation on Tyr504. Here we investigate the regulation mechanism of C3G *in vitro*. We demonstrate that C3G is autoinhibited. We find that Crk has a direct role in C3G activation as its binding partially activates C3G. Different Crk-binding sites are found to have distinct contributions to the activation of C3G. We further demonstrate that Src kinase phosphorylates C3G and that tyrosine phosphorylation increases the activity of C3G. We propose a mechanism of C3G activation, in which Src-catalyzed phosphorylation and interaction with Crk function together to induce full activation of C3G.

Introduction

C3G (Crk SH3 domain binding GEF) was identified as a binding partner of Crk (CT10 regulator of kinase)[1]. The presence of a CDC25-homology domain (CDC25-HD) indicated that C3G functions as a GEF for members of the Ras-family of small G-proteins. Indeed, C3G was demonstrated to be a Rap1 selective GEF, as it lacks catalytic activity towards Rap2[2,3]. In addition, it is reported to act as a GEF for R-Ras, TC10 and TC21[4–7]. However, *in vitro* data do not support that C3G acts as a GEF for R-Ras (Chapter 4, this thesis) or TC10[8].

The function of C3G is mostly related to its Rap1-activating function, although it has been found to have some Rap-independent roles[9–11]. Studies of knockout and low expressed C3G in mice established its role in development, cell adhesion and migration[12,13]. C3G has also been implicated in integrin signalling[14], cytoskeletal reorganization[15], trafficking[7,16] and cell junction integrity[17].

The domain organization of C3G is fairly simple; apart from the Ras family catalytic region that consists of a REM domain and a CDC25-HD, the N-terminal half is predicted to be mainly unstructured (Supplementary Figure 1). This region contains four proline-rich sequences that bind to Crk (referred to as site I, II, II and IV throughout the thesis)[18]. In addition, p130Cas binds to an additional fifth proline-rich motif[19]. Other binding partners of C3G are Hck[11], c-Abl[15] and E-cadherin[16], all of them interacting with the N-terminal region.

Crk is a well-established interacting partner of C3G that is suggested to exist in a constitutive complex with C3G[20,21]. It was discovered as a cancer-promoting protein in avian CT10 virus-induced tumors. Transformation with Crk induces massive increase in cellular tyrosine phosphorylation, while Crk itself was found not to contain kinase activity. It was therefore suggested to function by activating tyrosine kinase[22].

There are three mammalian Crk proteins: Crk I and Crk II, splice variants of a single gene, and CrkL that is encoded by another gene. Crk I contains one SH2 and one SH3 domain, while Crk II and CrkL contain an additional C-terminal SH3 domain. The SH2 domain of Crk binds to sequences that contain a pYxxP motif (pY stands for phosphorylated tyrosine). This domain mediates interactions of Crk with p130Cas, paxillin, Cbl, IRS and Gab. The N-terminal SH3 domain of Crk ($SH3^N$) preferentialy binds PxxPxK motifs. It interacts with C3G, Sos, DOCK180 and Abl kinase[21,23]. The C-terminal SH3 domain of Crk II and Crk L ($SH3^C$) is classified as an atypical SH3 domain as it has very low affinity for proline-rich sequences. It is involved in the negative regulation of $SH3^N$ in Crk II, by partially occluding its proline-sequence binding cleft[24].

The mechanism by which Crk induces tyrosine kinase activity is not fully understood. Abl and Src are identified as major kinase candidates for Crk-induced activation. Abl directly interacts with Crk by binding to a proline-rich sequence in the SH2 domain of Crk via its SH3 domain. Interaction between Crk and Abl is suggested to increase the activity of Abl kinase[25]. Crk is suggested to positively regulate the activity of Src as well, by displacing Csk, a kinase that negatively regulates Src, from the Src-signaling complexes[26].

Apart from its involvement in the regulation of activity of tyrosine kinases, Crk functions

as an important adaptor protein. One of his major interacting partners is p130Cas, an adaptor protein that is phosphorylated by Src and FAK. p130Cas contains 15 YxxP motifs in its central region, that upon phosphorylation of tyrosine serve as docking sites for the SH2 domain of Crk. By binding to p130Cas on one side, and C3G, Sos, DOCK180 and Abl on another, Crk is involved in the assembly of large signaling complexes[23]. Crk-interacting proteins are thus translocated in the response to tyrosine kinase signaling. The translocation promotes their function, by bringing them in the contact with their substrates as it is suggested for Sos[27], or in addition by direct activation as in the case of Abl[25].
Regulation of C3G has been previously studied by Matsuda et al. They have demonstrated that overexpression of Crk results in a C3G-dependent manner in increased Rap•GTP levels *in vivo*. Crk was shown not to affect the activity of C3G *in vitro*, and is thus assumed to act indirectly by enriching C3G at the plasma membrane. Indeed, targeting of C3G to the plasma membrane by the addition of the membrane targeting sequence also leads to an increase in GTP loading of Rap1. Since Rap1 is localized on membrane, translocation of C3G brings it in contact with its substrate. Apart from translocation of C3G to the membrane, Crk is suggested to have a role in tyrosine phosphorylation of C3G, as overexpression of Crk results in increased C3G phosphorylation. Tyr504 is suggested to be the main site of tyrosine phosphorylation in C3G. It is proposed that Crk-induced phosphorylation of C3G results in the repression of a cis-acting negative regulatory domain outside of the catalytic region[28,29]. Several tyrosine kinases, Src, Lyn, Hck and Abl, have so far been implicated in the regulation of C3G[11,30–33].
In this study we investigate the activation of C3G *in vitro*. We demonstrate that C3G is autoinhibited. N-terminally truncated C3G construct that lacks first 544 residues remains autoinhibited. We further demonstrate that Crk directly increases the activity of C3G. This activation is only partial, as Crk does not induce full C3G activation. Similar as Crk, tyrosine phosphorylation of C3G by Src results in a partial activation of C3G. Together, phosphorylation of C3G by Src and interaction with Crk induce full activation of C3G.

Materials and Methods

Protein preparation

$C3G^{cat}$ (aa 830-1077) was expressed from pET15b and $C3G^{fl}$ (aa 1-1077), $C3G^{388-1077}$, $C3G^{454-1077}$, $C3G^{544-1077}$, $C3G^{fl}$Y504A, $C3G^{fl}$Y504E, $C3G^{fl}$Y554F, $C3G^{fl}$Y561F, $C3G^{fl}$Y570F, $C3G^{fl}$Y579F and $C3G^{fl}$Y590F from a modified version of pET20b which contains an additional 5xHis-tag in BL21 Star Rosetta2 LysS strain.
C3G peptides (aa 273-290, aa 446-464, aa 535-550 and aa 598-620), $C3G^{230-1077}$, $C3G^{273-1077}$ and $C3G^{535-1077}$ were expressed as GST fusion proteins from pGEX vectors in the bacterial strain CK600K.
Protein production in bacteria was induced by 100 nM IPTG (isopropyl β-D-1-thiogalactopyranoside). After 20 hrs bacteria were collected by centrifugation, resuspended in the first affinity chromatography buffer supplemented with PMSF and lysed by sonication. The lysate was cleared by centrifugation and applied to the respective

affinity column.
His-tagged C3G proteins were eluted in an immidazol gradient from 5 to 250 mM immidazol in buffer containing 50 mM Tris pH 8.0, 500 mM NaCl and 1 mM β-mercaptoethanol.
GST-tagged C3G peptides and proteins were purified as GST-fusion proteins according to standard protocols. For GST-tagged C3G proteins the concentration of NaCl was kept at 400 mM NaCl throughout all steps.
Catalytically active Src (251-533) in the vector pET28a was a kind gift from J. Kuriyan. It was purified in the same way as His-tagged C3G proteins, with an additional Superdex 75 gel chromatography step. The buffer used for gel chromatography consisted of 50mM TrisHCl pH 8.0, 500 mM NaCl and 5 mM β-mercaptoethanol.
Crk I, Crk II, CrkL, $SH3^N$ (CrkII 130-192), Hck (56-116) and Src (85-144) were cloned in the pGEX4T3 vector, purified and expressed as GST fusion proteins according to standard protocols whereby the GST tag was removed by thrombin cleavage. Rap1B (residues 1 – 167) Rap1B (residues 1-167) was expressed from the ptac vector as described previously[34].

In vitro GEF assay
GEF assays were performed as described previously[35] in buffer containing 50mM Tris▪HCl pH7.5, 10% glycerol, 200mM NaCl and 5mM dithiothreitol. For the Crk titration experiment (Figure 6) a buffer containing 50 mM NaCl was used. Rap1 is loaded with the fluorescently labeled GDP analog mGDP. The fluorescence intensity of G-protein bound mGDP is approximalty twice as high as of mGDP free in solution. The nucleotide exchange of mGDP against an excess of unlabeled nucleotide can therefore be monitored as decay in fluorescence intensity. Fluorescence decay is fitted exponentially to obtain C3G reaction rates.

Limited proteolysis of C3G
Trypsin was typically used in a ratio 1 ng trypsin for 1 µg C3G. The reaction mixture was incubated for 5 min at room temperature and the reaction was stopped by the addition of PMSF.

Src phosphorylation assay
Src 251-533, containing the kinase domain only, was used for phosphorylation of C3G. A typical phosphorylation reaction was performed by mixing 20 volumes of C3G (180 µM C3G in buffer containing 50 mM Tris pH 8.0, 500 mM NaCl and 5 mM β-mercaptoethanol), 4 volumes of Src reaction buffer (100 mM TrisHCl pH 7.6. 125 mM $MgCl_2$ and 25 mM $MnCl_2$) and 1 volume of 2.5 mM ATP in Src reaction buffer. Src was added to this reaction mixture to the final concentration of 6 µM. The reaction was incubated for 1 hr on room temperature before determination of GEF activity.

Isothermal titration calorimetry (ITC) measurements
Titrations were performed in a NANO ITC 2G (TA Instruments) equipped with a 1 ml cell at 20°C. In a typical experiment C3G was titrated with a solution containing the Crk

construct in steps of 3.7 µl. All proteins were brought into buffer of the same preparation either by gel filtration or by extensive dialysis. Data were analysed by the manufactures software.

Results

C3G is autoinhibited by its N-terminal region

In order to investigate a putative allosteric regulation of C3G, we compared the nucleotide exchange rates of full length C3G ($C3G^{fl}$) and a construct of the CDC25-HD only ($C3G^{cat}$). $C3G^{fl}$ exhibits significantly lower activity towards Rap1B than $C3G^{cat}$ (Figure 1A).

To exclude the possibility that the low activity of $C3G^{fl}$ occurs as a result of improper folding, we subjected $C3G^{fl}$ to limited proteolysis by trypsin. We analyzed the progress of proteolysis by SDS-page and observed that a proteolysis product with an approximated size of 30 kDa is more resistant to digestion than the rest of the protein (Figure 1B). After prolonged exposure to trypsin, this product is completely degraded as well. We have stopped the proteolysis at the time point in which the biggest amount of the 30 kDa product is obtained. The GEF activity of this sample was measured. A significant increase in the reaction rate of the proteolytically degraded $C3G^{fl}$ as compared to the non-treated $C3G^{fl}$ was observed. The catalytic rate of the trypsin-degraded sample is comparable to the activity of $C3G^{cat}$ (Figure 1A).

As the N-terminal part of the protein is expected to be rather unstructured, as judged from the absence of any recognized domains and low content of predicted secondary structure, it is therefore likely to be degraded early during proteolysis (Supplementary Figure 1). A fragment of about 30 kDa would be the minimal size required in order to contain the intact CDC25-HD (about 250 amino acids). Both the size of the product and its GEF activity indicate that the 30 kDa proteolytic product contains the CDC25-HD. Purified recombinant $C3G^{fl}$ is thus in principal catalytically competent but autoinhibited. The autoinhibition can be relieved by proteolytic digestion of the N-terminal part of C3G.

A series of N-terminal truncation constructs of C3G starting at position 230, 273, 388, 454, 535 and 544 was generated (Figure 1C). All constructs are characterized by a low GEF activity similar to $C3G^{fl}$ (Figure 1D and data not shown), and exhibit high catalytic activity after trypsin treatment (data not shown). Autoinhibition is thus preserved in the absence of residues 1 to 544. As the REM domain starts around residue 690 (as predicted by the domain prediction tool SMART), the region between 544 and 690 is referred to as inhibitory region.

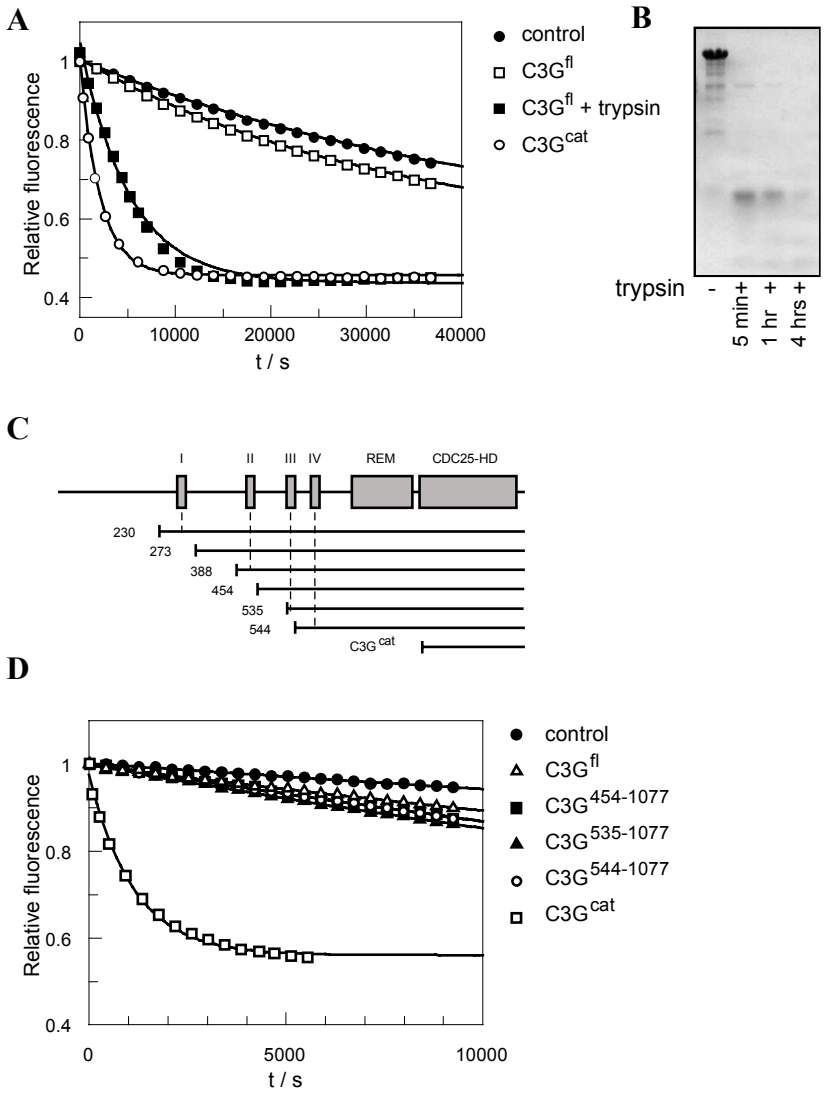


Figure 1. C3G is autoinhibited. A, Nucleotide exchange activity of 150 nM $C3G^{fl}$, trypsin digested $C3G^{fl}$ and $C3G^{cat}$ was measured towards 200nM Rap1B. **B**, SDS-PAGE analysis of the time dependent proteolysis of $C3G^{fl}$ by trypsin. **C**, Schematic representation of the domain organization of C3G. The different N-terminal truncated constructs are indicated. Proline-rich Crk binding sites are labeled as I, II, III and IV. REM, Ras Exchange Motif; CDC45-HD, CDC25-homology domain **D**, Nucleotide exchange activity of $C3G^{fl}$, $C3G^{454-1077}$, $C3G^{535-1077}$, $C3G^{544-1077}$ and $C3G^{cat}$ towards Rap1B.

Binding of Crk activates C3G

Addition of Crk I, Crk II, CrkL or the isolated N-terminal SH3 domain ($SH3^N$) enhances C3G activity (Figure 2A, 2B and data not shown). All N-terminally truncated constructs,

with the exception of $C3G^{544-1077}$, are still activated similarly to $C3G^{fl}$ (data not shown). $C3G^{535-1077}$ contains sites III and IV. As site III in $C3G^{544-1077}$ is lacking, the presence of site III is required for the activation of C3G by Crk.

We used isothermal titration calorimetry (ITC) to investigate the interaction between C3G and Crk in more detail. We firstly determined binding affinities of the four isolated Crk-binding peptides. All four peptides bind to Crk I and $SH3^N$ with affinities of about 1 μM (Table 1). The similar binding affinities of Crk to the C3G peptides are in agreement with their high sequence identity and with previous reports in literature[18]. The apparent binding stoichiometry is in the range of 0.3 to 0.4.

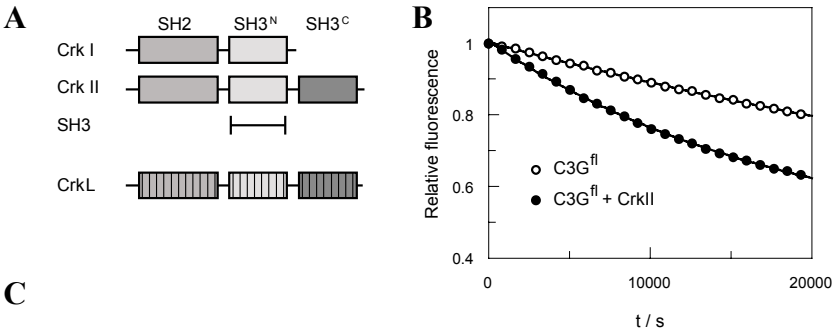
As $SH3^N$ resembles the binding properties of Crk I (Figure 2D and Table1), it was used to study the interaction further. The presence of four Crk-binding sites in C3G complicates the analysis of the titration. On the other hand, all four sites are highly similar in sequence and the isolated peptides showed almost identical binding characteristics. The titrations of $SH3^N$ into solutions of $C3G^{535-1077}$ or $C3G^{273-1077}$ were therefore analyzed by assuming four identical and independent binding sites. The SH3 domain binds to $C3G^{273-1077}$ and to $C3G^{535-1077}$ with overall affinities of 6.7 μM and 20.9 μM and apparent stoichiometry constants of 1.2 and 0.8, respectively. The apparent stoichiometric constant of $C3G^{535-1077}$ is a multiple of 0.4, whereby the factor 2 corresponds to the number of binding sites in the constructs. In case of $C3G^{273-1077}$, the stoichiometry constant of 1.2 corresponds to 3 and 4 multiples of 0.4 and 0.3, respectively. This suggests that all four sites can be occupied simultaneously. It is not entirely clear why the expected 1:1 binding ratio for a single site was not obtained. Likely this is the consequence of a systematic underestimation of the Crk concentration.

For a further analysis $C3G^{535-1077}$, which contains only two binding sites, but which is still activated by Crk, was used. To investigate the contribution of the two Crk-binding sites the conserved leucine of the proline-rich sequence was mutated to glutamic acid. This mutation is expected to be incompatible with Crk-binding as the binding pocket in Crk is likely not able to adapt to the sterically more demanding and charged residue. Indeed, no binding of $SH3^N$ to $C3G^{535-1077}$L543E,L611E was observed (Figure 2E). Binding to $C3G^{535-1077}$L543E occurs with an apparent stoichiometric constant of 0.4, which would correspond to one functional binding site. $C3G^{535-1077}$L611E shows a more complicated titration pattern, with reasonable heat release in the first two-three titration steps and hardly any additional heat release in the following steps.

When analyzed for its exchange activity, $C3G^{535-1077}$L543E shows no activation upon Crk-$SH3^N$ addition, whereas $C3G^{535-1077}$L611E is activated (Figure 2F). This confirms that binding of Crk to the site III is necessary for the activation of C3G. These results are in agreement with the observation that $C3G^{544-1077}$ (that lacks site III) can not be activated upon Crk addition. $C3G^{535-1077}$L543E,L611E shows already high activity in the absence of Crk and is not activated any further upon addition of Crk, indicating that the mutation itself already influences the conformation of C3G.

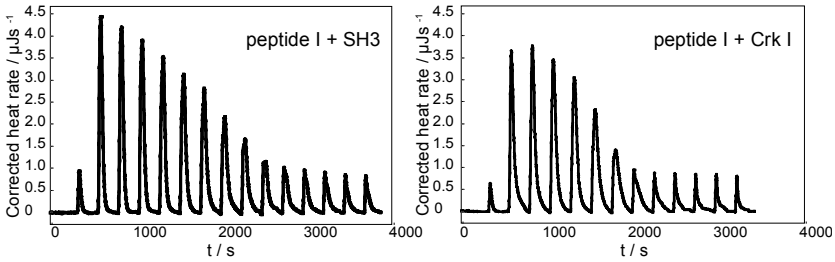
	$SH3^N$		Crk I		Crk II		SH3 (Src)		SH3 (Hck)	
	n	**Kd**	**n**	**Kd**	**n**	**Kd**	**n**	**Kd**	**n**	**Kd**
peptide I	0.33	0.9 μM	0.40	0.8 μM			0.33	47.3 μM		
peptide II	0.31	1.2 μM	0.41	0.5 μM			0.34	51.2 μM		
peptide III	0.33	1.9 μM	0.49	2.0 μM			0.33	62.7 μM	0.36	24.2 μM
peptide IV	0.30	1.8 μM	0.34	1.4 μM			0.41	47.2 μM		
$C3G^{535-1077}$	0.79	20.9 μM	0.44	31.9 μM	1.00	9.5μM				
$C3G^{535-1077}$L543E	0.40	7.8 μM								
$C3G^{535-1077}$L611E	0.014	170 μM								
$C3G^{535-1077}$L543E,L611E	no binding									
$C3G^{273-1077}$	1.17	6.7 μM								

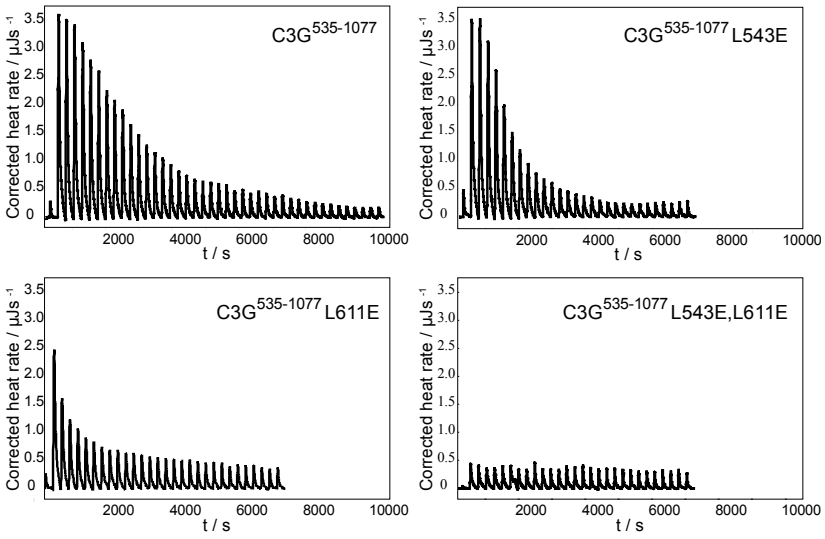
Table1. Summary of ITC data. n – stoichiometric constant, Kd – constant of dissociation.



C

I	GST-273-290	-PGIRVVDNSPPALPPKK
II	GST-446- 464	---PGQQTDTPPALPEKKRRSA
III	GST-535-550	-----GDPEKPPPLPEKKNKH
IV	GST-598-620	GVDSVQELAPPPALPPKQRQLEP





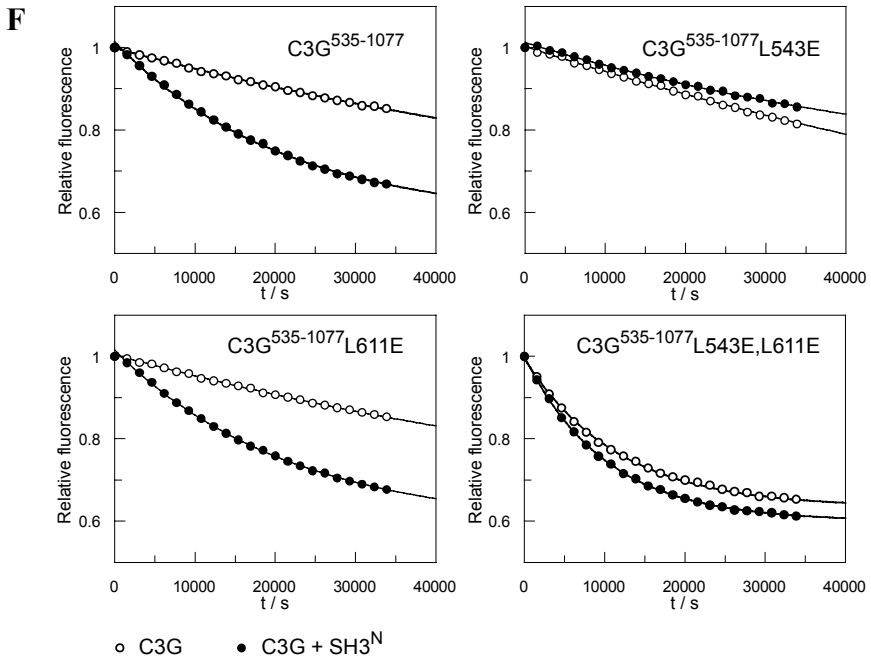


Figure 2. Crk binding to C3G. A, Domain organization CrkI and CrkII, which are splice variants of the same gene, and CrkL. The borders of $SH3^N$ are indicated. **B,** Effect of Crk II on the exchange activity of $C3G^{fl}$, $C3G^{535-1077}$ and $C3G^{544-1077}$. Crk II is used at a concentration of 5μM. **C**, Alignment of the peptides corresponding to the four Crk-binding site in C3G. **D**, Heat traces of the titration of GST-C3G-peptide I with $SH3^N$ and Crk I as monitored by ITC. **E**, Heat traces of the titration of $C3G^{535-1077}$, $C3G^{535-1077}$L543E, $C3G^{535-1077}$L611E and $C3G^{535-1077}$L543E,L611E with $SH3^N$. **F**, Nucleotide exchange activity of $C3G^{535-1077}$, $C3G^{535-1077}$L543E, $C3G^{535-1077}$L611E and $C3G^{535-1077}$L543E,L611E towards Rap1B in the presence or absence of 220 μM $SH3^N$.

Tyrosine phosphorylation of C3G by Src kinase increases the activity of C3G

C3G is reported to be phosphorylated by several tyrosine kinases, most of them being members of the Src kinase family. C3G was therefore subjected to Src-catalyzed phosphorylation *in vitro*. To estimate the amount of phosphorylated C3G, we analyzed the mobility of phosphorylated and non-phosphorylated C3G by SDS-PAGE. The phosphorylation induces a complete shift of the C3G band relative to non-phosphorylated C3G (Figure 3A). As judged from the gel shift, more than 90% of C3G is phosphorylated. Phosphorylation of C3G by Src results in an increase in GEF activity (Figure 3B). N-terminally truncated mutants were subjected to phosphorylation by Src, and were found to be activated by phosphorylation including $C3G^{544-1077}$ that contains only site IV and is not activated by Crk any more (Figure 3B).

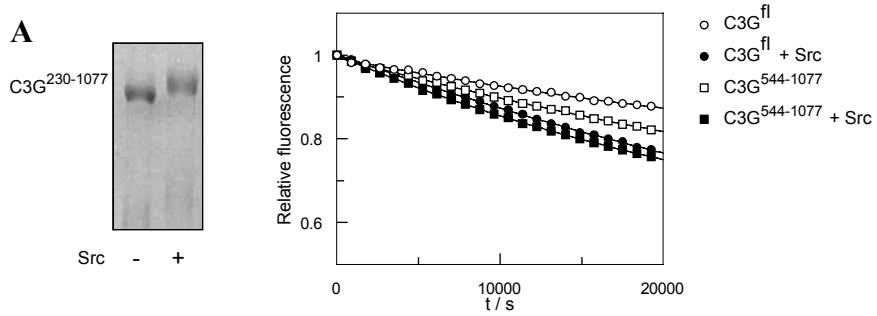


Figure 3. Src-catalyzed phosphorylation of C3G A, SDS-PAGE analysis of the mobility of phosphorylated and non-phosphorylated $C3G^{230-1077}$ **B**, Nucleotide exchange activity of non-phosphorylated and phosphorylated $C3G^{fl}$ and $C3G^{544-1077}$ towards Rap1B.

Crk binding and tyrosine phosphoporylation by Src co-operate to activate C3G

Combination of phosphorylation and Crk binding induces much higher levels of C3G activity than the separated events (Figure 4A). The obtained activity is close to the activity of $C3G^{cat}$ and thus close to the maximal possible activation. The shortest construct that still shows the behavior of $C3G^{fl}$ is $C3G^{535-1077}$ (Figure 4B). $C3G^{544-1077}$ still responds to phosphorylation by Src but no effect of Crk is obtained anymore.

It has been previously suggested that phosphorylation of Y504 contributes to the regulation of C3G *in vivo*. We tested if this residue is necessary for the activation of C3G *in vitro*. To that end, C3GY504E and C3GY504A were generated. The mutation of tyrsosine to glutamic acid introduces negative charge and therefore might mimic phosphorylation to some extent. However, $C3G^{fl}$Y504E and $C3G^{fl}$Y504A respond to Src treatment and Crk binding as $C3G^{fl}$ (Figure 5A). Since tyrosine phosphorylation still leads to the activation of $C3G^{fl}$Y504E and $C3G^{fl}$Y504A, other tyrosine(s) have to be involved in the activation of C3G.

The inhibitory region (544 – 690) contains five tyrosine residues, which are good candidates to mediate Src-induced activation. However, single mutations of the tyrosine residues at position 554, 561, 570, 579 or 590 to phenylalanine do not abolish Src induced activation of C3G (Figure 5B and data not shown).

SH3 domains of Src and Hck are less effective C3G activators than Crk

The Src kinase family member Hck has been reported not only to phosphorylate C3G but also to directly bind to C3G via its SH3 domain. This raises the interesting possibility that the SH3 domain of Src kinases may mimic the effect of Crk binding.

To test if Src and Hck can bind to the Crk binding sites in C3G, we measured the binding of the SH3 domains of Src and Hck to the C3G peptides using ITC. The SH3 domains of Src and Hck bind with affinities of around 60 μM and 25 μM, respectively, to peptide driven from the third proline rich motif compared to 1.9 μM for the binding of Crk-$SH3^N$. In addition, we tested if the SH3 domains of Src and Hck have an effect on the activity of C3G. Upon addition of the SH3 domains to phosphorylated C3G only a small effect on the activity is observed (data not shown). However, due to the relative low binding

affinity the SH3 domains were tested at non-saturating conditions. Thus it seems that in principle binding of the SH3 domains of Src and Hck mimic Crk binding albeit the effect is limited by the low affinity.

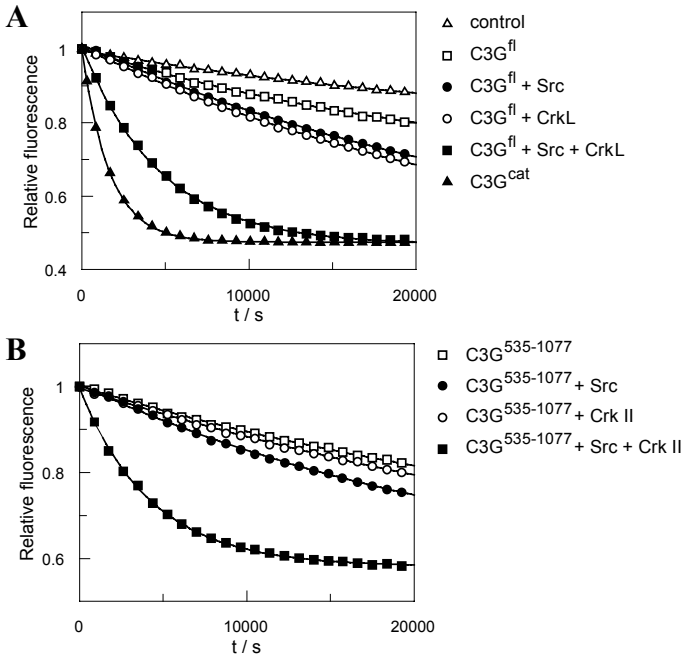


Figure 4. Combined effect of Src-catalyzed phosphorylation and CrkL binding on C3G. A, B, Nucleotide exchange activity of $C3G^{fl}$ **(A)** or $C3G^{535-1077}$ **(B)** towards Rap1B was measured with or without pre-treatment with Src kinase in the presence or absence of 2 μM CrkL **(A)** and 5μM Crk II **(B)**.

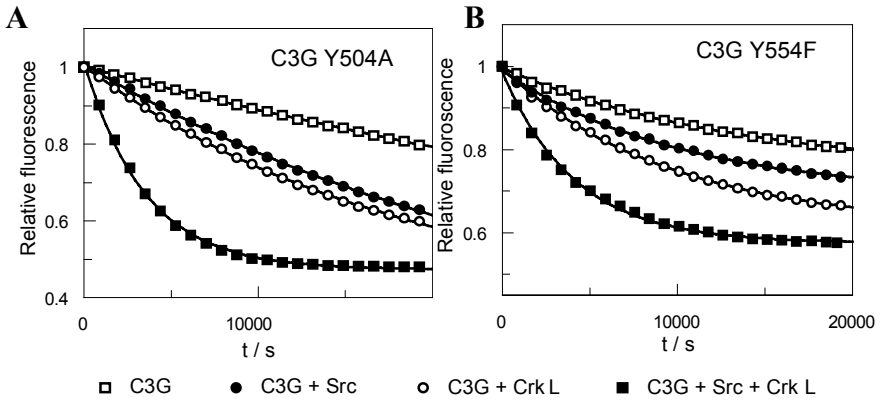


Figure 5. Contribution of tyrosine residues from the inhibitory region to the regulation of C3G. A, B, Nucleotide exchange activity $C3G^{fl}$Y504A **(A)** or $C3G^{fl}$Y554F **(B)** towards Rap1B was measured with or without pre-treatment with Src kinase in the presence or absence of 2 μM CrkL.

Different Crk proteins activate C3G to various extents

We compared the effect of different Crk proteins on C3G activation, by titration of both phosphorylated and non-phosphorylated $C3G^{535-1077}$ with increasing amounts of $SH3^N$, Crk I and Crk II (Figure 6). In case of non-phosphorylated C3G, higher concentrations of $SH3^N$ are necessary to induce the same level of C3G activation as of Crk I or Crk II, but upon saturation the same level of activity is reached. The same trend in affinity is observed for phosphorylated C3G. However, upon saturation different levels of maximal activity are reached. The level reached by $SH3^N$ is lowest and that by Crk II highest.

To confirm the differential effect of $SH3^N$ and Crk II on C3G activity, we performed a competition experiment. While Crk II induces higher C3G activity than $SH3^N$, addition of $SH3^N$ to Crk II decreases Crk II- induced C3G activity (data not shown). Clearly, the isolated SH3 domain is able to compete Crk II out of the complex with C3G, but is not able to activate the bound C3G to the same extent as Crk II.

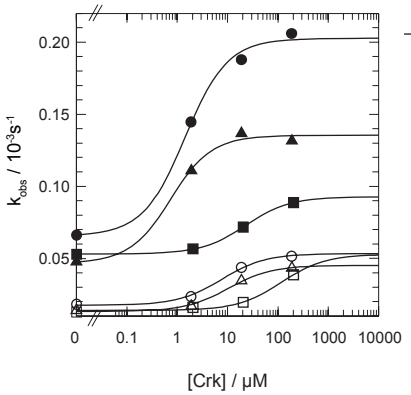


Figure 6. Effect of $SH3^N$, Crk I and CrkII on the activity of phosphorylated and non-phosphorylated C3G. Nucleotide exchange activity of non-phosphorylated and phosphorylated C3G towards Rap1B was measured in the presence of various concentrations of $SH3^N$, Crk I and Crk II. The obtained decay in fluorescence was fitted as a single exponential decay to obtain the rate constants (k_{obs}), which were plotted against the Crk concentration.

Discussion

Autoinhibition is a common way of regulation of GEFs. Among GEFs for G-proteins of the Ras family, Sos[36], Epac[35,37] and PDZGEF[38] are reported to be autoinhibited. The regulation mechanism of the Rap-GEF Epac2 has been studied in detail. Structural studies reveal that the autoinhibition of Epac2 is a result of a direct interaction of the second cyclic nucleotide binding domain (cNBD) with the CDC25-HD that prevents Rap from approaching the catalytic. Upon binding of cAMP to the cNBD, the cNBD swings away and thereby allow Rap to access the CDC25-HD freely[37,39].

Analogous to the regulation of Epac, we propose that the inhibitory region of C3G establishes direct contacts with the CDC25-HD and thereby prevents the access of Rap. The low intrinsic activity of $C3G^{fl}$ would therefore be the result of stochastic breaking of interactions between the CDC25-HD and the inhibitory region of C3G, bringing a small percentage of C3G into active conformation. The cooperative effect of Crk-binding and

Src-catalyzed phosphorylation suggests the existence of two components of the autoinhibitory surface: one is established between sequences that include one or more Crk-binding sites while the other involves tyrosine residue(s)-containing regions (Figure 7).

Activation of C3G by binding of Crk is at variance with the observations by Matsuda et al, who found that Crk had no direct effect on the catalytic activity of C3G *in vitro*[28]. This discrepancy might be due to the use of C3G which was obtained by immunoprecipitation from mammalian cells in the work of Matsuda et al, as cellular Crk might have been co-precipitated. Alternatively, our experimental setup might have higher sensitivity.

All four binding sites in C3G can be occupied by Crk simultaneously as suggested by the ITC experiments. Not all of them are necessarily involved in the release of autoinhibition. We have demonstrated a crucial function of site III in auto-inhibition, as Crk does not increase the activity of $C3G^{535-1077}$L543E, in which the third binding site is mutated, or the activity of $C3G^{544-1077}$ in which the third binding site is removed. Binding of Crk to site III is thus necessary for the activation, but is it also sufficient? $C3G^{535-1077}$L611E (in which site IV is mutated) is activated by Crk to as similar extent as $C3G^{535-1077}$ suggesting that this site is not necessary for activation. On the other hand, $C3G^{535-1077}$L543E,L611E shows activity higher than Crk-activated $C3G^{535-1077}$. This would suggest that site IV has a role in keeping the inhibited conformation of C3G. However, at the current state it is difficult to distinguish whether the leucine to glutamic acid mutations only abolish interaction with Crk, whether in addition these mutations mimic binding of Crk by interfering with the interaction between the inhibitory region and the CDC25-HD or whether the mutation simply disturbs local protein folding.

The ITC data distinguish site III as well. $SH3^N$ interacts with the isolated peptide of site III with very similar affinity as with the peptides of the other sites. When binding to $C3G^{535-1077}$L611E, in which site IV is mutated but site III left intact, was subjected to ITC a complicated titration pattern was obtained in which strong heat release in the very first injection was followed by multiple injections with minor and slowly decaying heat release. Reasonable heat release in only the very few first titration steps is indicative for a low concentration of binding sites. Minor but slowly decaying heat release over many injections is indicative for low binding affinity. This could be explained by a direct interaction of site III with the CDC25-HD: If Crk has to compete for binding with the catalytic domain, parts of the free enthalpy of Crk binding should be “consumed” by disrupting the intramolecular interaction in C3G. Thus the affinity for binding to the protein would be lower than the affinity for the isolated peptide and probably the binding enthalpy would be lower too. Indeed the affinity for the free peptide of $C3G^{273-1077}$ (site I, II, III and IV), $C3G^{535-1077}$ (site III and IV) and $C3G^{535-1077}$L611E (site III) were 1.9 μM, 6.7 μM and 20.9 μM, respectively. Thus when $C3G^{273-1077}$ and $C3G^{535-1077}$ were analyzed under the assumption of four independent and identical binding sites, the affinity decreased with increasing contribution of site III. The first two injection steps might then reflect binding of Crk with high affinity to the fraction of C3G which is in the open conformation anyway. Furthermore, the binding sites might not be independent and binding of Crk to site IV might prime binding site III. This would explain why $C3G^{535-1077}$

behaves as containing two binding sites whereas both sites are impaired upon mutation of the fourth site.

The importance of sites I and II is not established. $C3G^{535-1077}$, that lacks first two sites, is inhibited and addition of Crk and phosphorylation by Src activates it in the same manner as $C3G^{fl}$. Still, it is possible that these sites contribute to the establishment of the autoinhibition in the context of full length C3G.

$SH3^{N}$, CrkI and CrkII activate phosphorylated C3G under saturating conditions to different extents but non-phosphorylated C3G to the same extent. A possible explanation for the observed differences might be the involvement of the SH2 domain in activation of tyrosine-phosphorylated C3G. On the other hand, a role of the SH2 domain can not account for the difference between Crk I and Crk II. Another possibility might be that the differences arise due to the increasing size of the protein: The larger proteins might be more demanding in space and therefore weaken the interaction between the inhibitory region and the CDC25-HD more efficiently. However, in the simplest scenario, this effect should be observable with non-phosphorylated C3G as well.

Tyr504 has been identified as the main site of tyrosine phosphorylation in C3G and Tyr504 phosphorylation has been correlated with increased Rap•GTP levels *in vivo*[29]. We find that Tyr504 is not necessary for the activation of C3G by Crk and Src. Tyr504 is thus likely not required or at least not sufficient for the allosteric regulation of C3G. Tyr504 might however contribute to the regulation of C3G "indirectly" as for example by the control of cellular localization.

$C3G^{535-1077}$ is still activated by Src in the same way as $C3G^{fl}$. Therefore tyrosine residues C-terminal to position 535 must be involved in the regulation. We demonstrate that mutation of single tyrosine residues from the inhibitory region is not sufficient to prevent activation by Src. Comparison of the sequences of human and drosophila C3G, reveals the high conservation of all four Crk-binding sites (Supplementary Figure 2). This suggests that the regulation mechanism of C3G may be conserved as well. The REM domains and CDC25-HDs display high sequence homology, with one and six conserved tyrosine residues, respectively. The inhibitory region contains one conserved tyrosine residue at position Tyr554, but $C3G^{fl}$Y554F is still activated by Src. In principle the conserved residues within the REM domain and the CDC25-HD could be mutated as well. However, mutations in the CDC25-HD domain might influence the catalytic competency directly. A direct approach for the identification of phosphorylated residues, as for example an analysis by mass spectroscopy, would therefore be more feasible.

For the phosphorylation of C3G by Src a construct of Src that contains only the kinase domain was used. Src kinase family members do however contain SH3 domains. Indeed, Hck has been reported to directly interact with the Crk-binding sites II and IV[11]. It would therefore be intriguing, if binding of Src by its SH3 domain to C3G would result in activation of C3G. In such a scenario, Src could then fully activate C3G by subsequent phosphorylation.

However, the measured affinity of the isolated SH3 domain of Crk for binding the site III peptide is 10 and 30 fold higher than the affinities of the isolated SH3 domains of Hck and Src, respectively. However, we can not exclude the possibility that SH3 domains of Hck

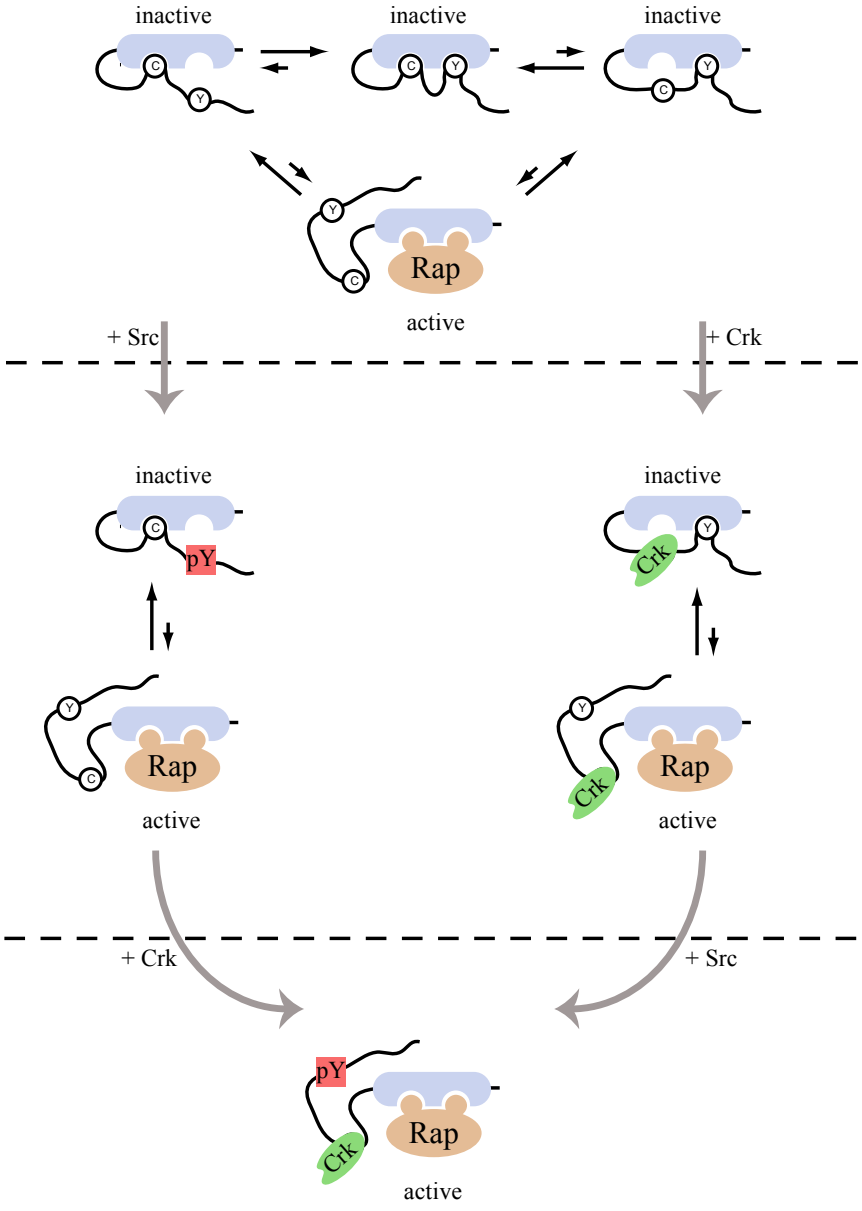


Figure 7. Proposed model of C3G activation. The inhibitory region of C3G establishes two contacts with catalytic CDC25-HD. The first contact site is established by Crk-binding site(s), and labeled as "C". The second site is established by tyrosine residue(s) containing regions and is labeled as "Y". The tyrosine residues might either be part of the inhibitory region or of the CDC25-HD. Full activation of C3G is achieved through series of steps, from the mainly inactive basal state (**top**), through partially activated Src-catalyzed phosphorylated or Crk-bound forms (**middle**) to the full active state of phosphorylated and Crk bound C3G (**bottom**). "pY" indicates tyrosine phoshorylation and "Crk" binding of Crk. The model considers only one binding site for Crk as

relevant for the activation and a simple two-state conformational transition of "C" site: Crk-free (C site interacts with CDC25-HD) and Crk-bound (C-site does not interact with CDC25-HD). For putative extension of the model see discussion. The equilibria between two conformations are indicated by two arrows.

and Src, in the context of full length proteins, can achieve the physiologically relevant levels of C3G activation. It might be possible that binding of C3G by the SH3 domain of Src would direct the phosphorylation reaction towards tyrosine residues in the direct proximity of the kinase domain. While bound to C3G Src could "wait" for the tyrosine residue to become accessible as the tyrosine residue might be masked in the interaction between the inhibitory region and the CDC25-HD. On the other hand, the relative low affinity would prevent Src from "staying" with one and the same C3G molecule as this would not be compatible with the function of Src as a catalyst of C3G activation.

In the simplest model of C3G activation, auto inhibition of C3G occurs by binding of the N-terminal inhibitory region to the CDC25 homology domain. The intramolecular interaction surface could be considered as composed of two contact regions (Figure 7). Activation of C3G would require the release of the interaction from both contact regions, a process that could be described as the transition from a closed to an open conformation.

One contact region includes one or more Crk binding sites while the other contact region contains tyrosine residues which are phosphorylated by Src. For a small percentage of C3G, both contacts are simultaneously broken as result of a stochastic process. This results in an "open" conformation in which the access to CDC25-HD is free for Rap binding. As seen in Figure 7, several conformations of C3G are suggested to be catalytically inactive: The conformation in which both contact sites are established and the conformations in which only one of the sites is broken. Only the conformation in which both interactions are broken is active.

Binding of Crk to C3G induces breaking of the interactions in the first contact region, while the second contact region is unaffected. Tyrosine phosphorylation induces interruption of the second contact region while keeping the first contact point intact. In each case the overall inhibitory interaction would be weakened and thus a higher proportion of C3G would adapt the active conformation. Finally, when both phosphorylation and interaction with Crk are combined, both contact sites are interrupted and C3G is in the active conformation.

The depicted model is simplified but can be adapted to a more complex behavior of C3G. The two contact points do not necessarily have to be independent and for example phosphorylation might ease binding of Crk. Furthermore, the Crk contact point could be subdivided in two Crk binding sites. Furthermore, binding of Crk does not necessarily have to be fully incompatible with an interaction between the Crk contact point and the CDC25-HD. Instead, different versions of Crk might interrupt the interaction within the Crk point with different efficiency (as suggested from titration of phosphorylated C3G with different Crk proteins). Structural analysis of C3G is required to fully understand the regulation mechanism.

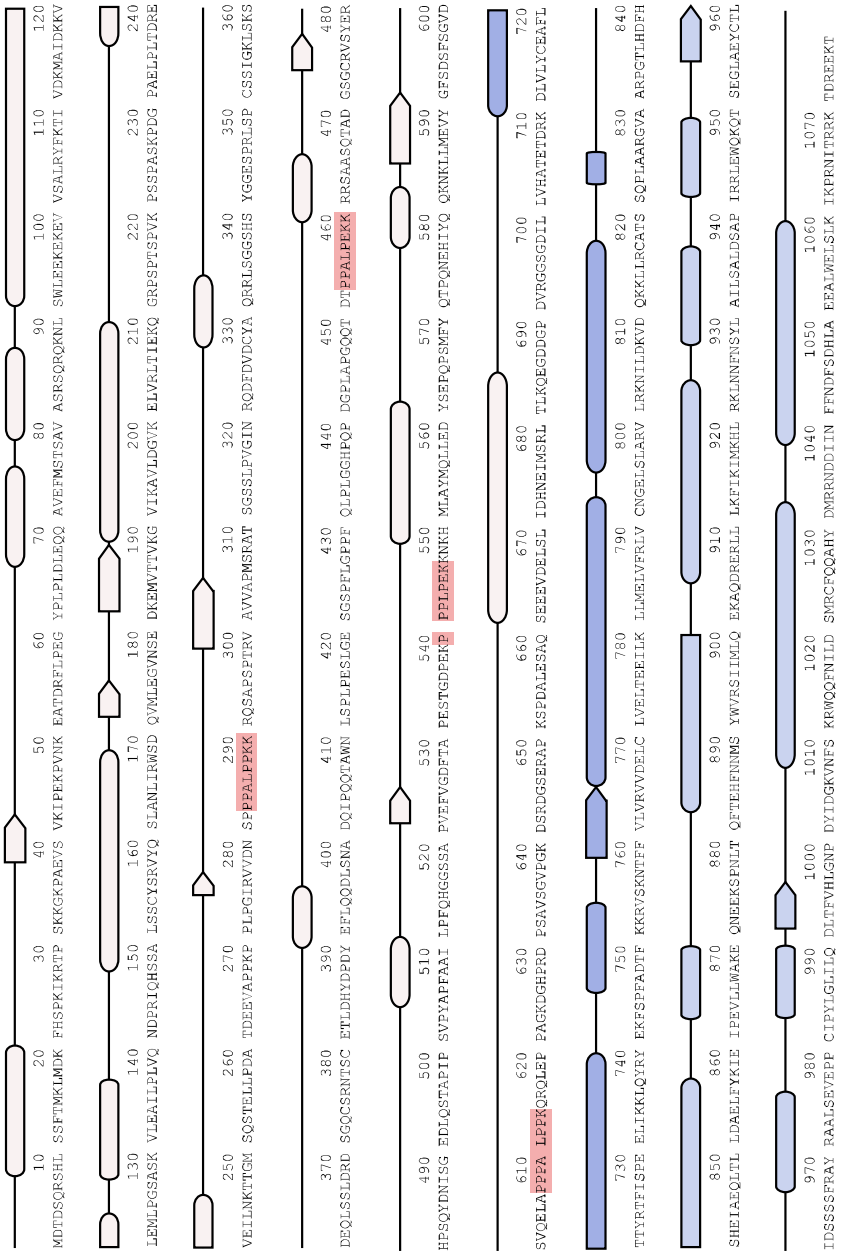
Acknowledgements
The authors are thankful to J. Kuriyan for providing DNA constructs. We thank members of the J. L. Bos lab for discussions and critical reading of the manuscript. MP and MRL were supported by the TI Pharma Project T3-106 to J.L. Bos, and HR by the Chemical Sciences of the Netherlands Organization for Scientific Research (NWO).

References

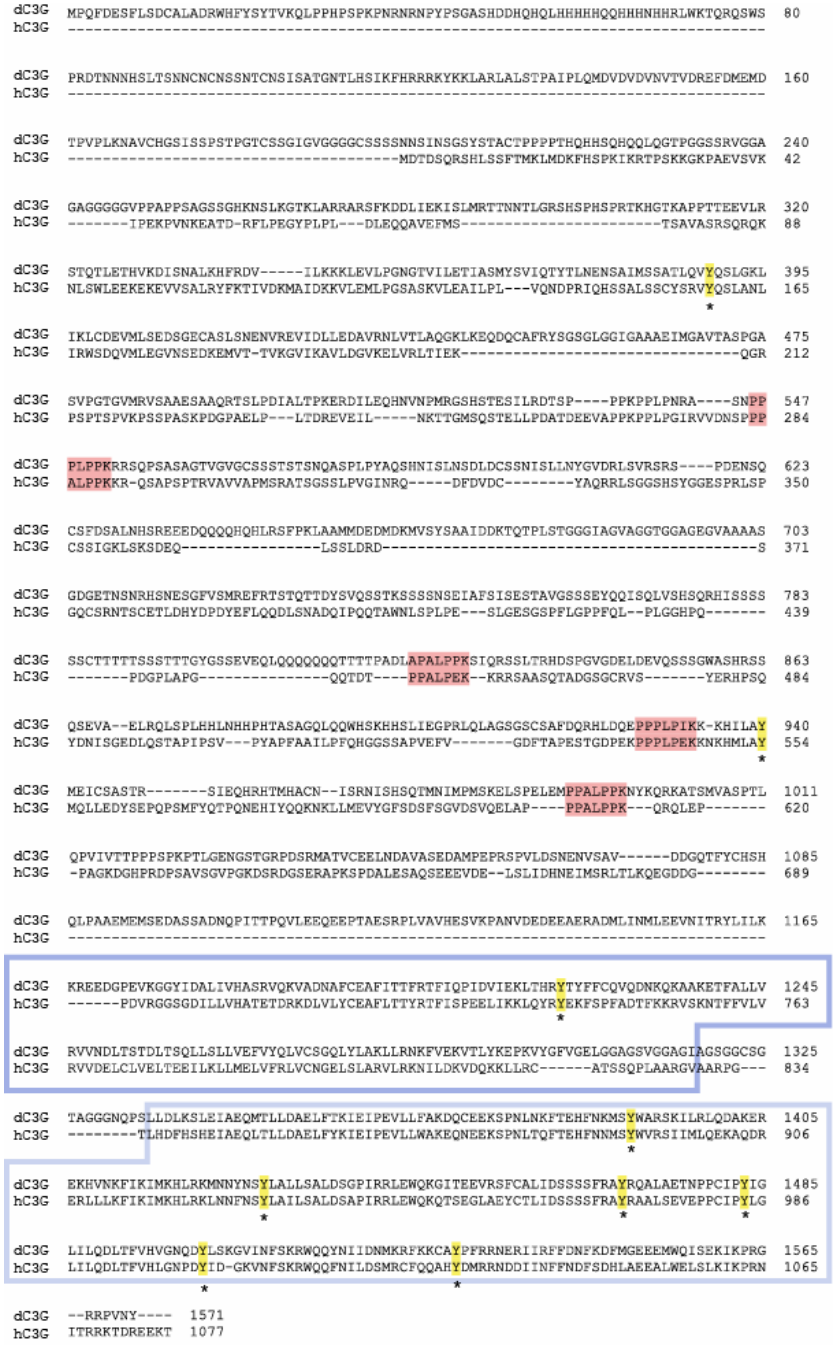
1. Tanaka, S. *et al.* C3G, a guanine nucleotide-releasing protein expressed ubiquitously, binds to the Src homology 3 domains of CRK and GRB2/ASH proteins. *Proceedings of the National Academy of Sciences* **91,** 3443–3447 (1994).
2. Gotoh, T. *et al.* Identification of Rap1 as a target for the Crk SH3 domain-binding guanine nucleotide-releasing factor C3G. *Molecular and cellular biology* **15,** 6746–6753 (1995).
3. Van den Berghe, N., Cool, R. H., Horn, G. & Wittinghofer, A. Biochemical characterization of C3G: an exchange factor that discriminates between Rap1 and Rap2 and is not inhibited by Rap1A(S17N). *Oncogene* **15,** 845–850 (1997).
4. Gotoh, T. *et al.* Activation of R-Ras by Ras-guanine nucleotide-releasing factor. *Journal of Biological Chemistry* **272,** 18602–18607 (1997).
5. Mochizuki, N. *et al.* Crk activation of JNK via C3G and R-Ras. *Journal of Biological Chemistry* **275,** 12667–12671 (2000).
6. Ohba, Y. Regulatory Proteins of R-Ras, TC21/R-Ras2, and M-Ras/R-Ras3. *Journal of Biological Chemistry* **275,** 20020–20026 (2000).
7. Chiang, S. H. *et al.* Insulin-stimulated GLUT4 translocation requires the CAP-dependent activation of TC10. *Nature* **410,** 944–948 (2001).
8. Pannekoek, W. J. *New insights in Rap-mediated control of cell adhesion*. (2012).
9. Guerrero, C. *et al.* Transformation suppressor activity of C3G is independent of its CDC25-homology domain. *Oncogene* **16,** 613–624 (1998).
10. Gutiérrez-Uzquiza, A. *et al.* C3G down-regulates p38 MAPK activity in response to stress by Rap-1 independent mechanisms: involvement in cell death. *Cell. Signal.* **22,** 533–542 (2010).
11. Shivakrupa, R., Radha, V., Sudhakar, C. & Swarup, G. Physical and functional interaction between Hck tyrosine kinase and guanine nucleotide exchange factor C3G results in apoptosis, which is independent of C3G catalytic domain. *Journal of Biological Chemistry* **278,** 52188–52194 (2003).
12. Ohba, Y. *et al.* Requirement for C3G-dependent Rap1 activation for cell adhesion and embryogenesis. *EMBO J.* **20,** 3333–3341 (2001).
13. Voss, A. K., Gruss, P. & Thomas, T. The guanine nucleotide exchange factor C3G is necessary for the formation of focal adhesions and vascular maturation. *Development* **130,** 355–367 (2003).
14. Sakakibara, A., Ohba, Y., Kurokawa, K., Matsuda, M. & Hattori, S. Novel function of Chat in controlling cell adhesion via Cas-Crk-C3G-pathway-mediated Rap1 activation. *J. Cell. Sci.* **115,** 4915–4924 (2002).
15. Radha, V., Rajanna, A., Mitra, A., Rangaraj, N. & Swarup, G. C3G is required for c-Abl-induced filopodia and its overexpression promotes filopodia formation. *Exp. Cell Res.* **313,** 2476–2492 (2007).
16. Hogan, C. *et al.* Rap1 regulates the formation of E-cadherin-based cell-cell contacts. *Molecular and cellular biology* **24,** 6690–6700 (2004).
17. Asuri, S., Yan, J., Paranavitana, N. C. & Quilliam, L. A. E-cadherin dis-engagement activates the Rap1 GTPase. *J. Cell. Biochem.* **105,** 1027–1037 (2008).
18. Knudsen, B. S., Feller, S. M. & Hanafusa, H. Four proline-rich sequences of the guanine-nucleotide exchange factor C3G bind with unique specificity to the first Src homology 3 domain of Crk. *Journal of Biological Chemistry* **269,** 32781–32787 (1994).
19. Kirsch, K. H., Georgescu, M. M. & Hanafusa, H. Direct binding of p130Cas to the guanine nucleotide exchange factor C3G. *Journal of Biological Chemistry* **273,** 25673–25679 (1998).

20. Matsuda, M. *et al.* CRK protein binds to two guanine nucleotide-releasing proteins for the Ras family and modulates nerve growth factor-induced activation of Ras in PC12 cells. *Mol. Cell. Biol.* **14,** 5495–5500 (1994).
21. Feller, S. M. Crk family adaptors-signalling complex formation and biological roles. *Oncogene* **20,** 6348–6371 (2001).
22. Mayer, B. J., Hamaguchi, M. & Hanafusa, H. A novel viral oncogene with structural similarity to phospholipase C. *Nature* **332,** 272–275 (1988).
23. Birge, R. B., Kalodimos, C., Inagaki, F. & Tanaka, S. Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. *Cell Commun Signal* **7,** 13 (2009).
24. Sriram, G. & Birge, R. B. Commentary: The carboxyl-terminal Crk SH3 domain: Regulatory strategies and new perspectives. *FEBS Lett.* **586,** 2615–2618 (2012).
25. Shishido, T. *et al.* Crk family adaptor proteins trans-activate c-Abl kinase. *Genes Cells* **6,** 431–440 (2001).
26. Bell, E. S. & Park, M. Models of crk adaptor proteins in cancer. *Genes Cancer* **3,** 341–352 (2012).
27. Teng, K. K. *et al.* v-Crk modulation of growth factor-induced PC12 cell differentiation involves the Src homology 2 domain of v-Crk and sustained activation of the Ras/mitogen-activated protein kinase pathway. *J. Biol. Chem.* **270,** 20677–20685 (1995).
28. Ichiba, T. *et al.* Enhancement of guanine-nucleotide exchange activity of C3G for Rap1 by the expression of Crk, CrkL, and Grb2. *J. Biol. Chem.* **272,** 22215–22220 (1997).
29. Ichiba, T. *et al.* Activation of C3G guanine nucleotide exchange factor for Rap1 by phosphorylation of tyrosine 504. *J. Biol. Chem.* **274,** 14376–14381 (1999).
30. Fukuyama, T. *et al.* Involvement of the c-Src-Crk-C3G-Rap1 signaling in the nectin-induced activation of Cdc42 and formation of adherens junctions. *J. Biol. Chem.* **280,** 815–825 (2005).
31. Nolz, J. C. *et al.* The WAVE2 complex regulates T cell receptor signaling to integrins via Abl- and CrkL-C3G-mediated activation of Rap1. *J. Cell Biol.* **182,** 1231–1244 (2008).
32. Mitra, A. & Radha, V. F-actin-binding domain of c-Abl regulates localized phosphorylation of C3G: role of C3G in c-Abl-mediated cell death. *Oncogene* **29,** 4528–4542 (2010).
33. He, Y. *et al.* The non-receptor tyrosine kinase Lyn controls neutrophil adhesion by recruiting the CrkL-C3G complex and activating Rap1 at the leading edge. *J. Cell. Sci.* **124,** 2153–2164 (2011).
34. Herrmann, C., Horn, G., Spaargaren, M. & Wittinghofer, A. Differential interaction of the ras family GTP-binding proteins H-Ras, Rap1A, and R-Ras with the putative effector molecules Raf kinase and Ral-guanine nucleotide exchange factor. *Journal of Biological Chemistry* **271,** 6794–6800 (1996).
35. De Rooij, J. *et al.* Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature* **396,** 474–477 (1998).
36. Hall, B. E., Yang, S. S. & Bar-Sagi, D. Autoinhibition of Sos by intramolecular interactions. *Front Biosci* **7,** 288–94 (2002).
37. Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A. & Bos, J. L. Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state. *Nature* **439,** 625–628 (2006).
38. De Rooij, J. *et al.* PDZ-GEF1, a guanine nucleotide exchange factor specific for Rap1 and Rap2. *J. Biol. Chem.* **274,** 38125–38130 (1999).
39. Rehmann, H. *et al.* Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B. *Nature* **455,** 124–127 (2008).

Supplementary Figures



Supplementary Figure 1. Secondary structure prediction of C3G. α-helical sequences are represented as rounded rectangles, and β-strands as arrows. Regions predicted to be a part of REM domain and the CDC25-HD are shown in dark blue and light blue, respectively. The proline-rich Crk-binding sites are highlighted in pink.



Supplementary Figure 2. Sequence alignment of human and drosophila C3G. The conserved Crk-binding sites are highlighted in pink, tyrosine residues in yellow, and the predicted REM and CDC25-HD in dark blue and light blue, respectively. Conserved tyrosine residues are indicated by asterisks.

Chapter 3

The conformation of a docking site for SH3 domains is pre-selected in the Guanine Nucleotide Exchange Factor Rlf

Milica Popovic[1], Arjen Jakobi[2], Marije Rensen-de Leeuw[1] and Holger Rehmann[1]

[1]Molecular Cancer Research, Centre of Biomedical Genetics and Cancer Genomics Centre, University Medical Center Utrecht, Utrecht, The Netherlands
[2]Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, The Netherlands

Summary

Rlf is a Guanine Nucleotide Exchange Factor for the small G-proteins RalA and RalB and couples Ras signalling to Ral signalling. Here the crystal structure of the catalytic module of Rlf consisting of a REM- and a CDC25-homology domain is determined. The structure is distinguished by an extended three stranded β-sheet that precedes the REM domain. The β-sheet element is conserved in the RalGDS family of Guanine Nucleotide Exchange Factors and stabilises the orientation of the REM domain relative to the CDC25-homology domain. A proline-rich sequence at its top is unique to Rlf. Several proteins, which interact with the β-sheet by SH3 domains, are identified. This interaction is characterised thermodynamically. Conformational pre-selection results in a gain of affinity and contributes to the establishment of SH3 domain selectivity.

Introduction

The members of the Ras-family of small G-proteins are involved in the regulation of such diverse effects as proliferation, differentiation, adhesion and exocytosis. They do so by acting as molecular switches cycling between an off- and an on-state, whereby the latter mediate downstream signalling. The off-state corresponds to the GDP and the on-state to the GTP bound version of the G-protein. Transition to the on-state is caused by the exchange of GDP for GTP. Nucleotide exchange is catalysed by Guanine Nucleotide Exchange Factors (GEFs). Switching off is achieved by the hydrolysis of GTP to GDP. The intrinsically low GTPase activity of small G-proteins is enhanced by GTPase activating proteins (GAPs), which terminate signalling. Effector proteins selectively interact with the GTP bound conformation of the G-protein and are responsible for transmitting the signal[1-3].

The small G-protein Ral is involved in the control of exocytosis[4], endocytosis[5], gene regulation[6-8] and cellular transformation[9,10]. The RalGDS family of GEFs connects signalling of the small G-proteins Ras and Ral and consists of four members in human: Rgl1, Rlf/Rgl2, Rgl3, and RalGDS. In all family members a N-terminal REM domain is followed by a CDC25 homology domain (CDC25-HD) and a C-terminal RA-domain. CDC25-HDs are the catalytic domains found in GEFs for members of the Ras-family of small G-proteins[11]. The CDC25-HDs of the RalGDS family are selective for the small G-proteins RalA and RalB[12-14]. REM domains always co-occur with CDC25-HDs and mainly stabilise the fold of the CDC25-HD[11]. RA-domains bind to the GTP bound form of Ras- and Rap-proteins[15]. Thereby GEFs of the RalGDS family act as effectors mainly of Ras-signalling.

The well-established Ras to Ral signalling occurs at cell membranes. Members of the Ras-family are membrane localised by a C-terminal lipid modification. Many GEFs are intrinsically active, but are kept away from their substrates in the cytosol. GEFs of the RalGDS-family are recruited to the membrane by binding to Ras•GTP[16,17]. Formation of Ras•GTP itself is induced by several GEFs, one of which is SOS1[18]. SOS1 occurs in complex with the adaptor protein Grb2. Grb2 binds to phosphorylated transmembrane receptors via its SH2 domain and thereby co-recruits SOS1 to the plasma membrane[18].

Grb2 is an example of the important contribution of adaptor proteins to the spatial regulation of signal transduction. Grb2 consists of SH2 and SH3 domains. SH3 domains interact with proline-rich sequences with the minimal consensus being PxxP[19] and binding of Grb2 to SOS1 is mediated by such an interaction. About 300 SH3 domains are found in the human genome. This number highlights the evolutionary success in using small domains as modules, which can be shuffled and integrated into many genes, as a tool to create protein interactions. This process is further eased by the need of only a short stretch of about 7 residues containing a PxxP motif as the requirement for SH3 domain binding in the target protein. On the other hand this raises the question whether elements other than the primary structure surrounding the proline residues contribute to the establishment of selectivity.

Here the crystal structure of the catalytic module of Rlf is determined. From the structure

an exposed SH3 domain binding site is predicted and indeed several SH3 domain containing proteins are identified in a yeast two-hybrid screen and confirmed to interact with the exposed binding site in Rlf.

Results

Several constructs of Rlf were subjected to crystallisation trials of which Rlf^{50-514} resulted in crystals suitable for structure determination (Table 1). Rlf^{50-514} contains the REM and the CDC25 homology domain but lacks the far N-terminus, which is expected to be unstructured, and the C-terminal RA-domain. Rlf^{50-514} is fully functional in terms of its catalytic activity towards Ral (Figure 1) in agreement with earlier studies on Rgl1[20]. Furthermore, addition of Rap loaded with a none-hydrolysable GTP analogue has no effect on the activity of full length Rlf (Rlf^{fl}) (Figure 1). Rap was chosen instead of Ras in this experiment, as Rap•GTP has a higher affinity for Rlf than Ras•GTP[21,22]. This confirms the classification of Rlf as a constitutively active GEF, whereby the interaction of Ras•GTP with the RA domain results in translocation of Rlf rather than in an allosteric regulation of the GEF activity itself.

Rlf^{50-514} crystallised into space group $P2_1$ with two molecules per asymmetric unit. The two molecules are related approximately by a two-fold rotation axis through a crystallographic dimer interface within the CDC25-homology domain (Figure 2). Differences between the two molecules are mainly restricted to the N-terminal β-sheet (Figure S1).

The CDC25-homology of Rlf distinguishes itself by two insertions spanning residue 307 to 322 and 402 to 427. Both insertions are at solvent exposed regions. Whereas residues 403 to 426 are unstructured, residues 307 to 322 form a sheet-loop-sheet structure, which is inserted between two α-helices (Figure 2B). The insertions are distant to the catalytic

Data collection		Refinement	
space group	$P2_1$	resolution range (Å)	31-2.3
cell dimensions (Å)	a=70.9 b=75.6 c=101.3	R_{cryst} (%)	25.3
cell angels (°)	α=90 β=98.1 γ=90	R_{free} (%)[b]	28.7
wavelength (Å)	0.97626	average B factor (Å^2)	42.2
resolution range (Å)[a]	31-2.3 (2.4-2.3)	rmsd from ideal values:	
number of reflections	220886	bond lengths (Å)	0.007
number unique reflections	46676	bond angles (°)	0.998
completeness (%)[a]	98.6 (98.2)		
I/σ[a]	24.3 (3.4)		
R_{meas} (%)[a]	4.6 (47.0)		

Table 1. Crystallographic data

[a] Values in parenthesis correspond to highest resolution shell.

[b] The Free-R factor was calculated with 5% of the data omitted from structure refinement.

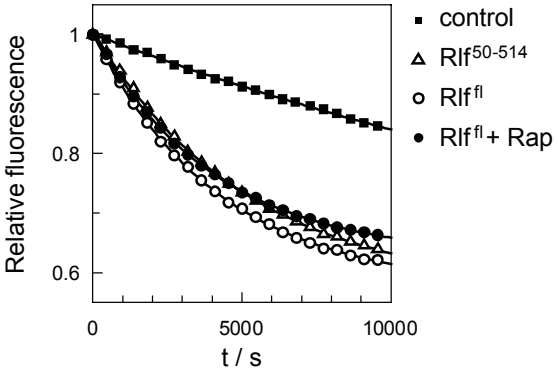


Fig. 1 The activity of Rlf is independent of its RA domain. Nucleotide exchange activity of 50 nM full length Rlf (Rlf^{fl}) in the presence or absence of 10 µM of Rap•GppNHp, or Rlf^{50-514} towards RalB was measured as decay in the fluorescence signal caused by the exchange of a fluorescent GDP analogue bound to Ral (see Material and Methods).

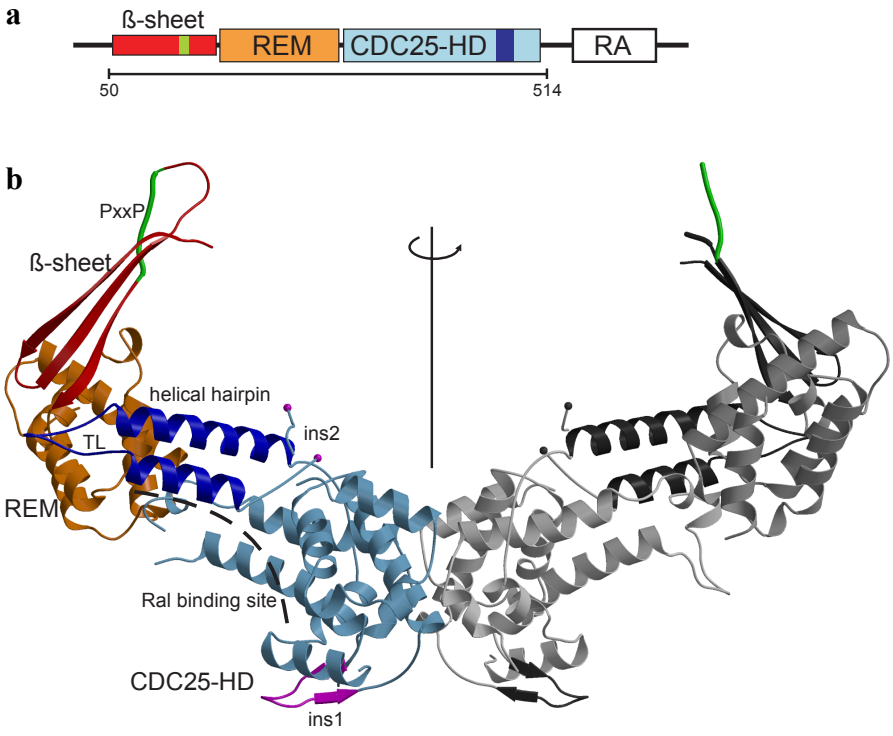


Fig. 2 Crystal structure of Rlf^{50-514}. (a) Domain organisation of Rlf, with start and stop of Rlf^{50-514} indicated. The colour code is used throughout the figures. REM, Ras exchange motif; CDC25-HD, CDC25-homology domain; RA, Ras association domain. (b) Ribbon diagram of the crystallographic dimer of Rlf^{50-514}, with molecule A coloured as coded in (a) and molecule B in shades of gray. The two-fold rotation axis is indicated. The predicted binding site for Ral in the CDC25-HD is indicated by dotted line. Rlf specific insertions in the CDC25-HD are highlighted in magenta and are labelled ins1 (aa 307 to 322) and ins2 (aa 403 to 426, start and end position indicated; the residues itself are unstructured). PxxP, proline-rich region (aa 82 to 87) are represented by thick coil in green; TL tip loop within the helical hairpin.

site and a function is not directly evident. The catalytic site of the CDC25-homology domain can be assigned based on the structures of SOS1[23] and Epac2[24] in complex with the small G-proteins H-Ras and Rap1B, respectively. The catalytic site is freely accessible for binding to Ral (Figure 2B and S2). In analogy to the structures of SOS1 and Epac2, the α-helix formed by residues 432 to 444 would be inserted into the nucleotide binding cleft of Ral (Figure S2). This α-helix is the N-terminal α-helix of the so-called helical hairpin, a helix-loop-helix structure. As the helical hairpin protrudes out of the CDC25-homology domain, the loop between the two helices is referred to as "the tip-loop".
An important function of the REM domain is the stabilisation of the helical hairpin. The REM domain shields the hydrophobic residues Phe457 and I455 in the C-terminal helix of the helical hairpin (Figure 3). Both residues and thereby the mode of the interaction between the REM-domain and the CDC25-homology domain are conserved. In variance to SOS1 and Epac2, the interaction is stabilised by Phe464 and Tyr454 additionally, which function like cantles at both ends of a saddle (Figure 3). These residues are conserved in the RalGDS family.

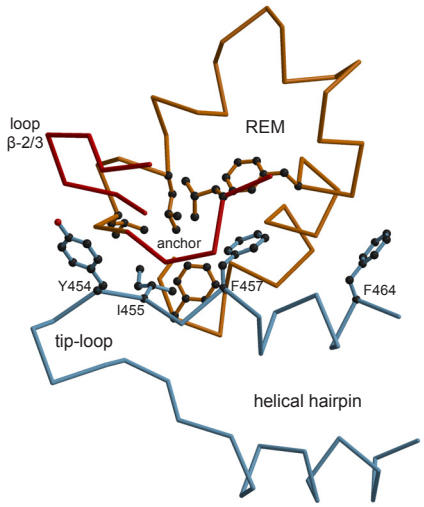


Fig. 3 Hydrophobic interaction between the REM domain and the CDC25-HD. Parts of the REM domain (orange) which are shielding Ile455 and Phe457 are shown. Phe464 and Tyr454 are stabilising the position of the REM domain. Loop β2/β3 refers to the loop between β-strand 2 and 3 of the β-sheet element.

The first forty residues of Rlf$^{50\text{-}514}$ form a striking three stranded β-sheet of some 10 residues per strand, which protrudes out of a compact unit formed by the REM domain and the CDC25-HD (Figure 2 and 4A). The β-sheet element is C-terminally directly connected to the first helix of the REM domain (Figure 4A). The last residues of the third β-strand anchor the REM-domain to the CDC25 homology domain and stabilise the relative orientation between these domains. They are therefore referred to as "the anchor". The anchor forms a β-sheet-like interaction with the tip-loop of the helical hairpin in the CDC25-homology domain. The tip-loop expands the β-sheet element by two short "pseudo"-β-strands (Figure 4A).
The same anchoring mechanism is found in SOS1 and Epac2 (Figure 4B and C). The tip-loop of the helical hairpin in SOS1 is longer by some residues than in Epac2 or Rlf and forms two clear, albeit short β-strands. The anchor in SOS1 is a short β-strand on its own and not the end of a long β-strand as in Rlf (Figure 4C). In SOS1 the β-sheet formed by

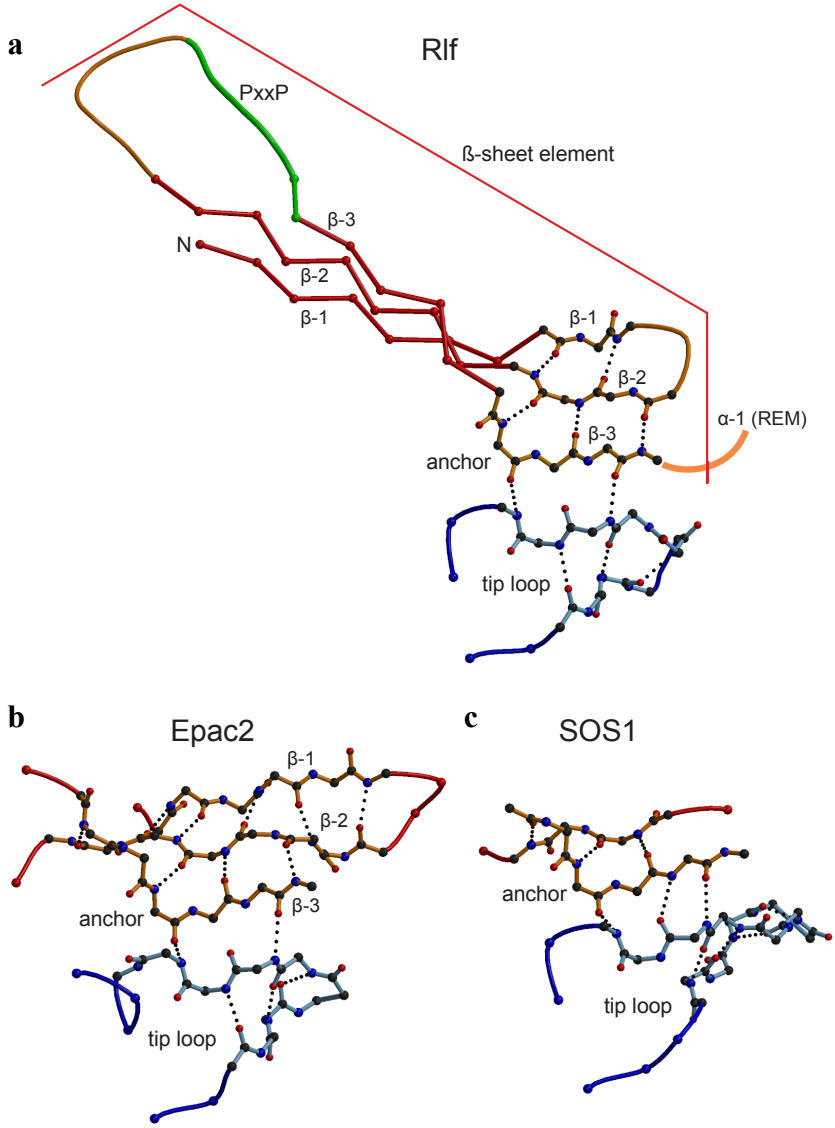


Fig. 4 The β-sheet element and its relation to the CDC25-HD. (a) The three stranded β-sheet formed by residues 53 to 95 is shown with the individual strands labelled β1, β2 and β3. The last residues of β3, labelled as anchor, form a β-sheet like interaction with the tip-loop (blue) of the helical hairpin and are connected to the first α-helix of the REM domain as graphically indicated. Residues, which are involved in this interaction, are shown in ball-and-stick representations with the side chains omitted for clarity. Dotted lines indicate hydrogen bonds. PxxP, proline-rich region (residues 82 to 87) are represented by thick coil in green. N, N-terminus. (b,c) The structural equivalent elements in Epac2 (b) and SOS1 (c) are shown. Note the equivalent to β1 is missing in SOS1. All three panels are at the same scale and in the same orientation based on superposition of the CDC25-HDs.

the tip-loop and the anchor is extended by a β-sheet like interaction to a short sequence element N-terminal to the anchor. Structurally this element mimics the N-terminus of the second β-strand of the β-sheet element (Figure 4C). Epac2 contains a structural element that could be considered as a short version of the β-sheet element. As in Rlf a three stranded β-sheet is extended by a β-sheet like interaction with the tip-loop of the helical hairpin (Figure 4B).

The loop between the second and the third β-strand of the β-sheet element in Rlf contains a proline-rich sequence (Figure 2 and 4A). At the top of the β-sheet element, the loop is highly exposed and accessible and could thus function as a binding site for SH3 domains. Indeed, a yeast two-hybrid screen with Rlf had identified with high confidence the SH3 domain containing proteins PLCγ-1, Myosin-1E and Intersectin-2 as Rlf interacting proteins. Whereas PLCγ-1 and Myosin-1E contain one SH3 domain each, Intersectin-2 contains five SH3 domains. Sequence analysis indicates highest homology of the SH3 domains of PLCγ-1 and Myosin-1E to the fifth SH3 domain of Intersectin-2 (Figure S3). These SH3 domains display the sequence features for binding to PxPPxR motives (Figure S3). Such a motive is found in the β-sheet element in the sequence P_{76}LDPLA**P**L**PP**P**R**$_{87}$. The interaction between Rlf and the hits of the yeast two-hybrid screen were confirmed by the ability of the isolated SH3 domains to co-precipitate $Rlf^{50\text{-}514}$ (Figure 5A).

Next to the proline-rich sequence in the β-sheet element, the region from residue 217 to 236 in the unresolved linker region between the REM- and the CDC25-homology domain contains several prolines and several additional PxxP motives are found distributed over the protein. Isothermal titration calorimetry (ITC) in combination with a mutagenesis study was used to assign the SH3 domain binding site in Rlf unambiguously. The SH3 domain of PLCγ-1 (PLCγ-1^{SH3}) was selected as a model for more detailed analysis. PLCγ-1^{SH3} binds to $Rlf^{50\text{-}514}$ and to $Rlf^{75\text{-}91}$ with 0.9 μM and 16 μM affinity, respectively, indicating that the proline-rich sequence of the β-sheet element is sufficient for binding (Figure 5B). No heat-traces were observed upon titration of PLCγ-1^{SH3} into solutions of $Rlf^{50\text{-}514}$ containing the mutations P82E, P84E or R87E (Figure 5C). Pro82, Pro84 and Arg87 would point into the hydrophobic cleft of the bound SH3 domain, if a similar binding mode is assumed as found for the SH3 domain of Myosin I from *Acanthamoeba castellanii* in complex with the peptide RPK**PV****PPPR**G (compare to **P**$_{82}$L**PPPR**$_{87}$ in Rlf) (Figure 6A and B) (pdb entry 2DRK). A glutamic acid at positions 82, 84 or 87 would be incompatible with binding which is in agreement with the obtained data. In precipitation experiment neither PLCγ-1^{SH3} nor the SH3 domain of Myosin1E or the fifth SH3 domain of Intersectin2 is able to interact with $Rlf^{50\text{-}514}$R87E, indicating that the β-sheet element is the docking site for all three proteins.

The peptide in the SH3 domain of Myosin I from *Acanthamoeba castellanii* can be superimposed onto the mapped binding site in Rlf (Figure 6A and B). This modelling study demonstrates that the PLPPPR motif in Rlf adopts a conformation ready to accept binding of an SH3 domain (Figure 6B). Furthermore, the engagement of the PLPPPR motif in the end of the first β-stand and the following loop is basically compatible with an SH3 domain interaction. Binding in this docking pose would only require a slight adaptation in the loop region following the PLPPPR motif as judged from the loop conformation in molecule A of the asymmetric unit (Figure 6A). In molecule B this region

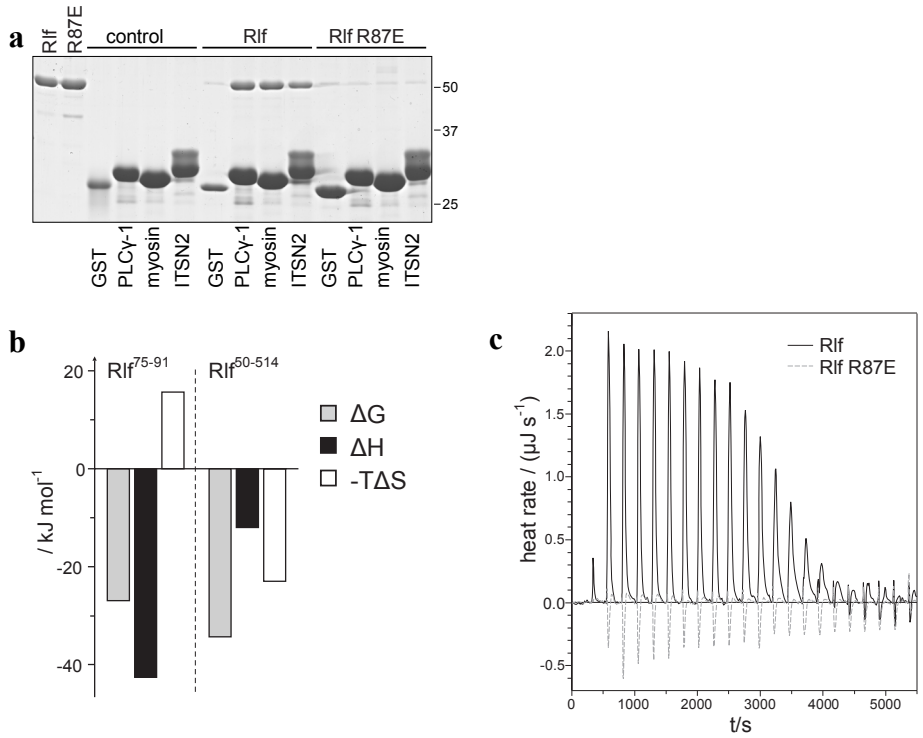


Fig. 5 Interaction between Rlf and SH3 domains. (a) GST-fusion proteins of the SH3 domains of PLCγ-1, myosin1E and intersectin2 (ITSN2) were immobilised on glutathione beads and used to precipitate either Rlf^{50-514} (Rlf) or Rlf^{50-514}R87E (Rlf R87E). Rlf^{50-514} and Rlf^{50-514}R87E were loaded to the first two lanes as reference. Shown is a Coomassie Briliant Blue stain of a SDS-PAGE. (b) GST-Rlf^{75-91} or Rlf^{50-514} were titrated with PLCγ-1^{SH3} in an isothermal titration calorimeter to determine the affinity and the enthalpy of the interaction. From this the free enthalpy and the binding entropy were calculated. (c) Heat trace from the titration of Rlf^{50-514} and Rlf^{50-514}R87E with PLCγ-1^{SH3}.

is unresolved. The remaining parts of the β-sheet elements are highly similar in both molecules, except a slight difference in the angle by which the β-sheet elements protrude (Figure 6C and S1). This is indicative for an intrinsic conformational flexibility within the loop region, which would likely allow adaption for SH3 domain binding in the proposed docking pose.

The docking pose suggests the absence of additional contacts between Rlf and a bound SH3 domain except the interaction with the PLPPPR motif as the β-sheet element keeps the SH3 domain distant to the remaining molecule (Figure 6A). In agreement with this is the catalytic activity of Rlf^{50-514} not influenced by binding of PLCγ-1^{SH3} (Figure 7).

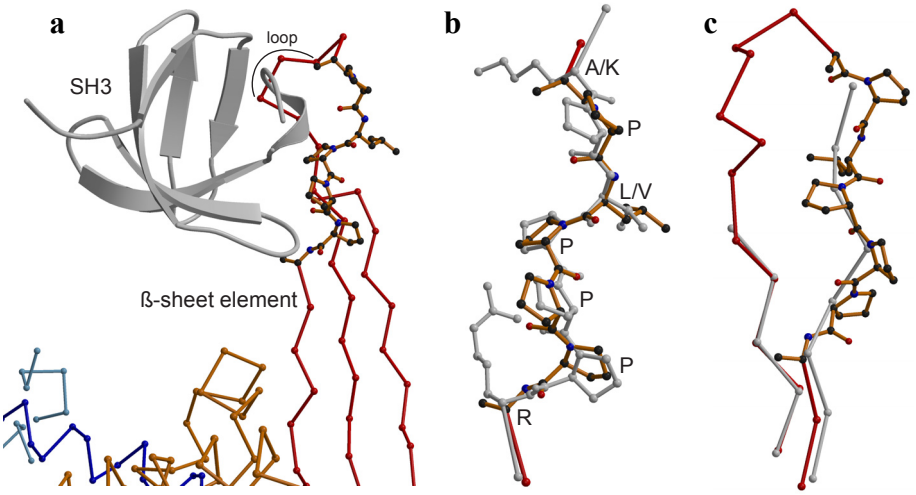


Fig. 6 Model of the Rlf-SH3 domain interaction. (a) The SH3 domain of Myosin I from *acanthamoeba castellanii* (pdb entry 2DRK) shown in grey was positioned into to structure of Rlf (molecule A) based on the superposition of the bound proline-rich peptide onto residues 81 to 88 in Rlf, which are shown in ball-and-stick representation. The residues in the labelled loop would have to undergo slight conformation rearrangements to accommodate the docking pose of the SH3 domain. (b) Residues 80 to 89 of Rlf are shown with the superimposed peptide (grey). The side chain of Arg87 is not resolved in the electron density in Rlf (molecule A). (c) The loop between β1 and β2 in molecule A (red) and molecule B (grey) is shown after superposition of the β-sheet elements onto each other. The SH3 domain binding residues of molecule A are shown in ball-and-stick representation.

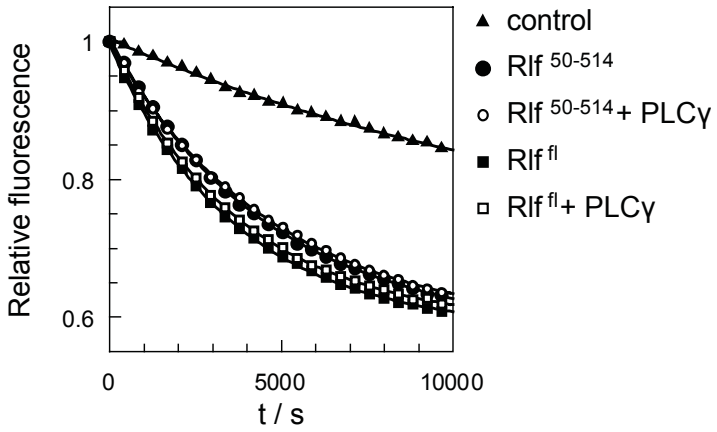


Fig. 7 The SH3 domain interaction does not influence the catalytic activity of Rlf. Nucleotide exchange activity of 50 nM full length Rlf (Rlf^{fl}) or Rlf^{50-514} towards RalB was measured as decay in the fluorescence signal caused by the exchange of a fluorescent GDP analogue bound to Ral in the presence or absence of 10 µM PLCγ-1^{SH3} (see Material and Methods).

Discussion

The N-terminal β-sheet is a distinctive element in the structure of Rlf. Sequence analysis suggests that the β-sheet is a common element in the RalGDS family (Figure S4). Even though sequence homology is only moderate in this region, secondary structure prediction supports the existents of the β-sheet. The proline-rich sequence in the loop between the second and the third β-strand is, however, unique in Rlf/Rgl2 (Figure S4).

The β-sheet element can be considered as an independent structural component N-terminal to the REM domain. In agreement with this, sequence based automated domain assignments as the SMART-algorithms define the N-terminus of the first helix in the REM-domain as the start. On the other hand, the anchor is crucial in stabilising the relative orientation between the REM domain and the CDC25-HD. REM domains always co-occur with CDC25-HD[11] and the anchor is found in all structures of this module known so far[23,25]. For that reason, we had previously proposed to consider the anchor containing β-strand as part of the REM domain[25].

This proposal is further justified by the functionality of the Rap-GEF Epac2. In Epac2 the REM domain is preceded by a cyclic nucleotide binding domain. cAMP interacts with the core of the cyclic nucleotide binding domain as well as with the first and second β-strand of the β-sheet element-like structure in Epac2. Consequently these two β-strands were considered as being part of the cyclic nucleotide binding domain[24]. However, the β-strands are not conserved in other cyclic nucleotide binding domains. Still, common to all cyclic nucleotide binding domains are conserved interactions of the cyclic nucleotide with the domain core and divergent interactions with structurally unrelated C-terminal elements[26,27].

Thus REM- and cyclic nucleotide binding domains are examples in which the folding units do not correspond to the functional unit. The classical definition of the REM domain is clearly a folding unit, but for its functionality it requires the anchor. The classical definition of the cyclic nucleotide binding domain combines the domain core, corresponding to a folding unit, with C-terminal elements that are structurally unrelated to each other (ranging form β-strands to α-helices). The β-sheet element-like structure in Epac2 could be considered as a folding unit serving two other folding units to acquire functionality.

The PDB database contains about 350 structures of SH3 domains, many of which in complex with a short proline-rich peptide. Only about five structures of different proteins display the interaction of an SH3 domain with more than a short proline-rich peptide and in most of these cases, the interaction is intra-molecular. It is thus of interest to analyse the interaction of SH3 domains with proline-rich motifs in the context of a full protein. On one hand the β-sheet element in Rlf mimics the situation of a free and short peptide, by separating the binding motif from the rest of protein. Furthermore, the β-sheet element fixes the six residues of the binding motif into the stretched conformation of a proline type II helix, which is required for docking of the SH3 domain, without blocking the accessibility of the residues. This is likely an important principle in gaining selectivity in signalling: identical motifs are excluded from interaction with an SH3 domain as soon as they adopt a less stretched and less accessible conformation, which might even be more

common in folded proteins. Such a mechanism would put a limitation to motif predictions purely based on primary structure.
On the other hand, binding of PLCγ-1^{SH3} to the Rlf$^{75-91}$ or Rlf$^{50-515}$ is distinct thermodynamically. The lower affinity of the peptide, 16 μM instead of 0.9 μM, might be explained by additional interactions of the SH3 domain with Rlf outside the proline motif. However, the performed docking study does not support such interactions (Figure 5). Furthermore, the thermodynamic characteristics are more fundamentally different. Binding to the peptide is accompanied by a negative change in entropy; binding is purely driven by enthalpy. Binding to Rlf$^{50-515}$ is favoured both by enthalpy and entropy, but ΔH is only -12 kJ/mol instead of the -42.5 kJ/mol observed for the peptide (Figure 5B). This binding characteristic is in agreement with the following considerations: The isolated peptide displays a high degree of conformational freedom. The stretched conformation required for binding is only one of many possible conformations. The SH3 domain has to select the stretched conformation and to reduce conformational freedom, resulting in a decrease in entropy. If the same residues are part of the β-sheet element, the conformational freedom is restricted. The β-sheet element has fixed the residues and thereby preselected a binding conformation. The entropic price had been paid during protein folding. The pre-selection has entropic advantages but disadvantages regarding binding enthalpy. If bound, the isolated peptide can adopt the most optimal conformation, as for example for the formation of hydrogen bonds. The β-sheet element, however constrains the residues in a preferential but not in the most optimal conformation, which reduces the binding enthalpy.
The physiological relevance of the β-sheet element interaction remains to be determined. The data presented here have proven a direct allosteric regulation of Rlf to be unlikely, as binding of the isolated SH3 domain does not influence the catalytic activity of Rlf *in vitro*. It is, however, intriguing to speculate that β-sheet element-mediated interactions contribute to the cellular localisation of Rlf. The β-sheet element-mediated interaction might open ways for Ras-independent translocations. PLCγ-1 is recruited to tyrosine phosphorylated transmembrane receptors via interactions of its SH2 domains[28]. Rlf might thus co-translocate with PLCγ-1. Alternatively, Ras•GTP might control the localisation of PLCγ-1 by using Rlf as an adaptor protein. Interestingly, PLCγ was recently shown to be recruited to focal adhesions based on its interaction with the guanine nucleotide exchange factors Ras-GRF1 and Ras-GRF2[29]. The interaction with Intersectin-2, a guanine nucleotide exchange factor for the small G-protein CDC42[30], might orchestrate Ras and Ral with CDC42 mediated effects. Finally, myosin1E associated generation of Ral•GTP might play a role in Ral controlled vesicle transport.

Material and Methods

Constructs

The SH3 domains of PLCγ-1 (PLCG1, *homo sapiens*, aa 790 to 848) referred to as PLCγ-1^{SH3}, myosin-1E (MYO1E, homo sapiens, aa 1054 to 1108) and intersectin-2 (ITSN2, *homo sapiens*, aa 1117 to 1185) were cloned into the pGEX4T3 expression vector. Amino acid residues 1 to 778, 1 to 735, 50 to 735 or 50 -514 of Rlf (Rgl2, *mus musculus*) were cloned into the pGEX6P3 expression vector.

cDNAs for PLCγ-1 (IRAKp961M0798Q) and myosin-1E (IRATp970C10100D) were obtained from Source BioScience and cDNAs for intersectin-2 (Addgene plasmid 25174) from Addgene.

Protein purification

Ral and the SH3 domains of PLCγ-1, myosin-1E and intersectin-2 were expressed and purified as GST fusion proteins according to standard procedures, if required, the GST-tag were removed by thrombin cleavage. H-Ras (*homo sapiens*, aa 1-166) was expressed from the ptac plasmid and loaded with the hydrolysis resistant GTP analogue GppNHp as described [21].

Rlf proteins were expressed as GST fusion proteins in the bacterial strain CK600K upon induction with 100 µM IPTG at 25°C over night in Standard I medium (Merck). The bacteria were harvested by centrifugation, re-suspended in buffer containing 50 mM Tris HCl pH 7.5, 50 mM NaCl, 5 mM EDTA, 5% glycerol, 5 mM β-mercaptoethanol and 0.1 mM PMSF and lysed by sonication. The lysate was cleared by centrifugation at 50,000g and loaded to a GSH column (Pharmacia) equilibrated with the pervious buffer. The column was washed with 6 column volumes of buffer containing 50 mM Tris HCl pH 7.5 400 mM NaCl, 5% glycerol and 5 mM β-mercaptoethanol, with 20 column volumes of buffer containing 50 mM Tris HCl pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 5% glycerol, 5 mM β-mercaptoethanol and 0.25 mM ATP at low flow rate, with two column volumes of 50 mM Tris HCl pH 7.5, 50 mM NaCl 2,5% glycerol and 5 mM β-mercaptoethanol (buffer A) and eluted with buffer A containing 20 mM glutathione. The protein containing fractions were pooled, concentrated, supplemented with 1 mg homemade GST-PreScission Protease per 100 mg protein, dialysed against buffer A and re-apply to a GST column equilibrated with buffer A, concentrated and further purified by gel filtration on a Superdex75 26/60 column (GE-Healthcare) in buffer A.

Crystallography

Rlf^{fl}, Rlf$^{1-735}$, Rlf$^{50-735}$ and Rlf$^{50-514}$ were subjected to crystallisation trials. Only Rlf$^{50-514}$ resulted in crystals. The initial hit was optimised and final crystals were grown at 289 K in sitting drops using a reservoir solution containing 0.1 M Bis-Tris propane buffer, pH 6.5, 0.2 M $NaNO_3$ and 12% polyethylene glycol 3350. Crystals were cryoprotected in a solution containing the mother liquor supplemented with 20% glycerol and flash cooled in liquid nitrogen. Data were collected at 100 K at beamline ID23eh1 of ESRF and processed with XDS[31]. Molecular replacement was carried out in MOLREP[32] using the CDC25 homology domain of RalGPS1 (residues 23-289) and RasGRF1 (1028-1262) as

poly-alanine search model. The program O[33] was used to build the model into 2Fo-Fc and Fo-Fc maps and refinement was carried out with REFMAC[34]. The Ramachandran plot depicts 94.9 % of main chain torsion angels in the most favoured and 5.1% in additional allowed regions with 0 residues in disallowed regions.
Figures were generated using the programs Molscript[35], Bragi[36] and Raster3D[37].

Determination of GEF activity
RalX was loaded with the fluorescent GDP analogue 2'-/3'-O-(N-methylanthraniloyl)-guanosine diphosphate (mGDP) as described previously for Rap1B[38]. The fluorescence intensity of Ral bound mGDP is approximately twice as intense as of free mGDP and thus nucleotide exchange can be observed as decay in fluorescence upon addition of an excess unlabelled GDP. All reactions were performed in buffer containing 50 mM TrisHCl pH7.5, 50 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 5% glycerol and 20µM GDP at 20°C, with 200 nM Ral•mGDP and x nM Rlf as described[38]. Ras•GppNHp and PLCγ-1^{SH3} were added at concentration as indicated in the figure legends.

Yeast two-hybrid
A human placenta cDNA library was screened for Rlf interacting proteins by Hybrigenics, SA, Paris, France.

Isothermal Titration Calorimetry
Titrations were performed in a NANO ITC 2G (TA Instruments) equipped with a 1 ml cell at 20°C. In a typical experiment a solution of 190 µM Rlf^{50-514} or GST-Rlf^{79-91} was titrated with a solution of 1.5 mM PLCγ-1^{SH3} in steps of 6 µl. Experiments were performed in buffer containing 50 mM TrisHCl pH 7.5, 50 mM NaCl, 2.5 % glycerol and 0.5 mM Tris(2-carboxyethyl)phosphine hydrochloride. All proteins were brought into buffer of the same preparation either by gel filtration or by extensive dialysis. Data were analysed by the manufacturer's software.

Precipitation Assays
GST or GST-fusion proteins were coupled to glutathione beads from lysates of bacteria over-expressing the respective proteins (see protein purification) and washed three times with buffer containing 50 mM TrisHCl pH 7.5, 50 mM NaCl, 2 % glycerol and 5 mM β-mercaptoethanol. 20 µg of Rlf^{50-514} or Rlf^{50-514}R87E, or buffer control, were added to the pre-coupled beads and incubated for 20 minutes at room temperature under gentle tumbling. The beads were washed three times with buffer as before and eluted by application of 18µl of Laemmli loading buffer and heating to 95 °C for three minutes. The samples were subjected to analysis by SDS-PAGE on 12.5% gels and protein bands were stained by Coomassie Brilliant Blue.

Acknowledgements
We thank Piet Gros for access to crystallisation robots, the European Synchrotron Radiation Facility for providing synchrotron facilities, the scientists at ID23eh1 for help with data collection, Cheryl Arrowsmith for providing cDNA and the members of the Bos lab for discussions and critical reading the manuscript. MP and MRL were supported by the TI Pharma Project T3-106 to J.L. Bos, and HR by the Chemical Sciences of the Netherlands Organization for Scientific Research (NWO).

Reference List

1. I. R. Vetter and A. Wittinghofer, The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299 (2001).
2. J. L. Bos, H. Rehmann, and A. Wittinghofer. GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865 (2007).
3. K. Wennerberg, K. L. Rossman, and C. J. Der. The Ras superfamily at a glance. *J. Cell Sci.* **118**, 843 (2005).
4. S. Moskalenko, *et al*. The exocyst is a Ral effector complex. *Nat. Cell Biol.* **4**, 66 (2002).
5. V. Jullien-Flores, *et al*. RLIP76, an effector of the GTPase Ral, interacts with the AP2 complex: involvement of the Ral pathway in receptor endocytosis. *J. Cell Sci.* **113**, 2837 (2000).
6. N. D. de Ruiter, et al. Ras-dependent regulation of c-Jun phosphorylation is mediated by the Ral guanine nucleotide exchange factor-Ral pathway. *Mol. Cell Biol.* **20**, 8480 (2000).
7. N. D. de Ruiter, B. M. Burgering, and J. L. Bos. Regulation of the Forkhead transcription factor AFX by Ral-dependent phosphorylation of threonines 447 and 451. *Mol. Cell Biol.* **21**, 8225 (2001).
8. D. M. Ouwens, *et al.* Growth factors can activate ATF2 via a two-step mechanism: phosphorylation of Thr71 through the Ras-MEK-ERK pathway and of Thr69 through RalGDS-Src-p38. *EMBO J*. **21**, 3782 (2002).
9. T. Urano, R. Emkey, and L. A. Feig. Ral-GTPases mediate a distinct downstream signaling pathway from Ras that facilitates cellular transformation. *EMBO J.* **15**, 810 (1996).
10. M. A. White, *et al*. A role for the Ral guanine nucleotide dissociation stimulator in mediating Ras-induced transformation. *J. Biol. Chem.* **271**, 16439 (1996).
11. L. A. Quilliam, J. F. Rebhun, and A. F. Castro. A growing family of guanine nucleotide exchange factors is responsible for activation of Ras-family GTPases. *Prog. Nucleic Acid Res Mol Biol* **71**, 391 (2002).
12. C. F. Albright, *et al*. Characterization of a guanine nucleotide dissociation stimulator for a ras-related GTPase. *EMBO J*. **12**, 339 (1993).
13. R. M. Wolthuis, *et al*. Stimulation of gene induction and cell growth by the Ras effector Rlf. *EMBO J* **16**, 6748 (1997).
14. E. Ferro, *et al*. G-protein binding features and regulation of the RalGDS family member, RGL2. *Biochem. J*. **415**, 145 (2008).
15. C. Herrmann. Ras-effector interactions: after one decade. *Curr. Opin. Struct. Biol.* **3**, 122 (2003).
16. K. Matsubara, *et al*. Plasma membrane recruitment of RalGDS is critical for Ras-dependent Ral activation. **18**, 1303 (1999).
17. S. Kishida, *et al*. Colocalization of Ras and Ral on the membrane is required for Ras-dependent Ral activation through Ral GDP dissociation stimulator. *Oncogene* **15**, 2899 (1997).
18. P. Chardin, *et al*. Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2. **260**, 1338 (1993).
19. B. K. Kay, M. P. Williamson, and M. Sudol. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14**, 231 (2000).
20. H. Murai, *et al*. 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 (1997).

21. C. Herrmann, *et al*. Differential interaction of the ras family GTP-binding proteins H-Ras, Rap1A, and R-Ras with the putative effector molecules Raf kinase and Ral-guanine nucleotide exchange factor. *J. Biol. Chem.* **271**, 6794 (1996).
22. S. Wohlgemuth, et al. Recognizing and defining true Ras binding domains I: biochemical analysis. *J. Mol. Biol.* **348**, 741 (2005).
23. P. A. Boriack-Sjodin, *et al*. The structural basis of the activation of Ras by Sos. *Nature* **394**, 337 (1998).
24. H. Rehmann, *et al*. Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B. *Nature* **455**, 124 (2008).
25. H. Rehmann, *et al*. Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state. *Nature* **439**, 625 (2006).
26. H. Rehmann, A. Wittinghofer, and J. L. Bos. Capturing cyclic nucleotides in action: snapshots from crystallographic studies. *Nat. Rev. Mol. Cell Biol.* **8**, 63 (2007).
27. J. M. Canaves and S. S. Taylor. Classification and Phylogenetic Analysis of the cAMP-Dependent Protein Kinase Regulatory Subunit Family. *J Mol Evol* **54**, 17 (2002).
28. P. G. Suh, *et al*. Multiple roles of phosphoinositide-specific phospholipase C isozymes. *BMB. Rep.* **41**, 415 (2008).
29. Q. Wang, *et al*. Ras-guanine-nucleotide-releasing factors 1 and 2 interact with PLCgamma at focal adhesions to enable IL-1-induced Ca2+ signalling, ERK activation and MMP-3 expression. *Biochem. J.* **449**, 771 (2013).
30. K. L. Rossman, C. J. Der, and J. Sondek. GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors. *Nat. Rev. Mol. Cell Biol*. **6**, 167 (2005).
31. W Kabsch. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Cryst* **26**, 795 (1993).
32. A. Vagin and A. Teplyakov. An approach to multi-copy search in molecular replacement. *Acta Crystallogr. D. Biol. Crystallogr.* **56**, 1622 (2000).
33. T. A Jones, *et al*. Improved methods for building protein models in electron- density maps and the location of errors in these models. *Acta Crystallogr A* **47**, 110 (1991).
34. G. N. Murshudov, A. A. Vagin, and E. J. Dodson. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr*. **53**, 240 (1997).
35. P. J Kraulis. Molscript - a program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* **24**, 946 (1991).
36. D. Schomburg and J. Reichelt. Bragi - A Comprehensive Protein Modeling Program System. *Journal of Molecular Graphics* **6**, 161 (1988).
37. E. A Merritt and M. E. P Murphy. Raster3D version-2.0 - A program for photorealistic molecular graphics. *Acta Crystallogr D Biol Crystallogr* **50**, 869 (1994).
38. H. Rehmann. Characterization of the activation of the Rap-specific exchange factor Epac by cyclic nucleotides. *Methods Enzymol* **407**, 159 (2006).

Supplementary Figures

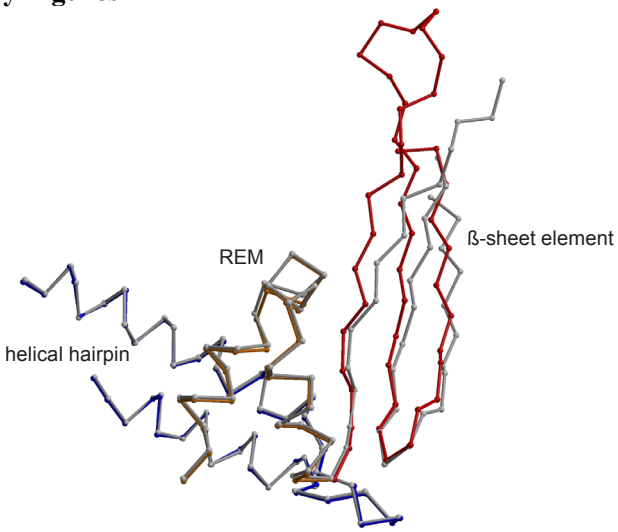


Fig. S1 Orientation of the β-sheet element. The two molecules of the asymmetric unit were superimposed onto each other by excluding the β-sheet element from the fitting procedure. Residues of molecule A are colour coded as in figure 1. Molecule B is depicted in grey. The β-sheet element, the helical hairpin and parts of the REM domain are shown.

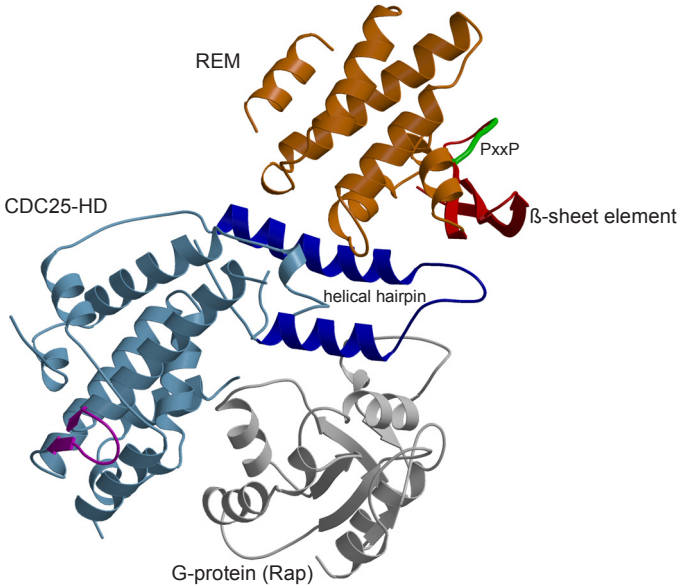


Fig. S2 Identification of the Ral binding site. The CDC25-HD of the Rap-GEF Epac2 (pdb entry 3CF6) in complex with its substrate Rap was superimposed onto the CDC25-HD of $Rlf^{50\text{-}514}$ (molecule A). Based on the superimposition Rap was placed into the structure of $Rlf^{50\text{-}514}$. $Rlf^{50\text{-}514}$ is colour coded as in figure 1 and Rap as a model for Ral is depicted in grey.

Fig. S3 Sequence alignment of SH3 domains. The SH3 domains of PLCγ-1 (aa 798 to 829) and myosin1E (aa 1093 to 1008) and the fifth SH3 domain of intersectin2 (aa 1133 to 1180) were aligned using ClustalW. Residues identical in all three proteins are highlighted by bold letter type. Sequence motifs associated with the recognition of the proline-rich motif are marked by grey boxes.

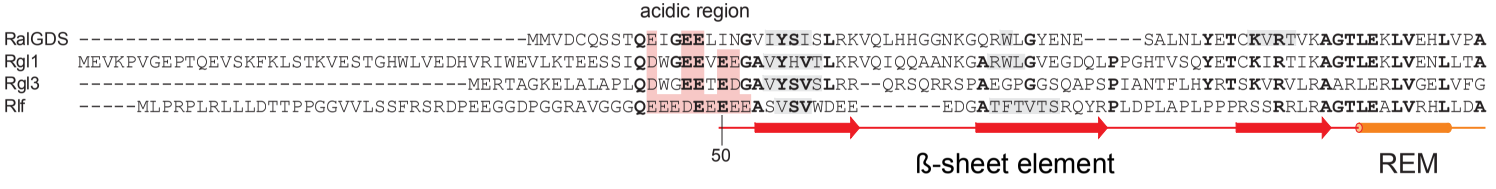


Fig. S4 The β-sheet element is conserved in members of the Ral-GDS family. The sequence alignment shows the N-terminal part of the RalGDS-family members RalGDS (homo sapiens), Rgl1 (homo sapiens), Rlf/Rgl2 (mus musculus) and Rgl3 (homo sapiens) including the β-sheet element and the first helix of the REM domain. The β-sheet element is preceded by an acidic region in which the acidic residues are on red background. The secondary structure as observed in the structure of Rlf^{50-514} is indicated underneath the alignment. Residues, which are predicted to be part of a β-strand based on consensus secondary prediction (http://xtal.nki.nl/ccd) are on grey background.

Chapter 4

Selectivity of CDC25 homology domain-containing Guanine Nucleotide Exchange Factors

Milica Popovic, Marije Rensen-De Leeuw and Holger Rehmann

Molecular Cancer Research, Centre of Biomedical Genetics and Cancer Genomics Centre, University Medical Center Utrecht, Utrecht, The Netherlands

Abstract

The Ras family of small G-proteins has a central function in the regulation of a variety of signal transmitting processes ranging from cell cycle control to regulation of exocytose. Signalling by Ras G-proteins is initiated by CDC25-homology domain (CDC25-HD) containing Nucleotide Exchange Factors (GEFs). Each GEF is characterised by a specific selectivity profile towards G-proteins and commonly acts on a small subset of the Ras family members only. Thereby GEFs pivotally determine which downstream signalling routes are activated. The structural basis for the establishment of selectivity in the GEF-G-protein interaction is only partially understood. The selectivity of GEFs is frequently discussed controversially in the literature. Here a systematic approach is undertaken to determine the selectivity of CDC25-HD towards members of the Ras family. A dataset of 126 pairings is generated in a standardised *in vitro* approach with purified recombinant proteins. Furthermore, the basis of the selectivity is analysed in a comprehensive mutational study. Together these data show distinct selectivity profiles of the GEFs and allow clear predictions for untested pairings.

Introduction

Small G-proteins process biological signals in various cellular signalling pathways by acting as molecular switches[4,32]. GDP and GTP-bound G-proteins differ in conformation, and only the GTP-bound form can interact with effector proteins and convey the biological signal. The nucleotide state of G-proteins is tightly regulated. Guanine nucleotide exchange factors (GEFs) catalyse the exchange of GDP with GTP thereby inducing the signalling active state of G-proteins. GTPase activating proteins (GAPs) catalyse the hydrolysis of a bound GTP to GDP by accelerating the intrinsic hydrolysis activity of G-proteins leading to termination of signalling.
The human Ras super-family of small G-proteins consist of 154 members, which are grouped into the Arf, Rab, Ran, Ras and Rho family[34]. Each family has its own type of GEFs, which cannot act on G-proteins from other families. For example, almost all known GEFs for G-proteins of the Ras family contain a CDC25-homology domain as the catalytic domain[23]. The CDC25-homology domain is structurally unrelated to GEFs for other G-protein families[4]. Structurally unrelated GEFs have in a number of cases claimed for given G-proteins. GEFs of the DOCK type were suggested to act on Ras and Rho family members[35], but careful biochemical and structural studies could confirm only a direct interaction with Rho family members[37]. Similarly, the ability of TCTP, which adapts the fold of canonical Rab-GEFs[29], to act as a GEF for Rheb[11] was questioned in independent studies[26,33].
Whereas conservation of the domains allows a straight forward assignment of GEFs to G-protein families, it is much more difficult to predict the specificity profile of an individual GEF within its cognate G-protein family. Although several partial studies have been published often in course of the identification of a novel GEF, a clear overall picture is still lacking. The present study therefore sought for an approach to systematically and comprehensively analyse the selectivity of CDC25 homology domain containing GEFs for G-proteins of the Ras family.

Results and Discussion

Experimental Approach

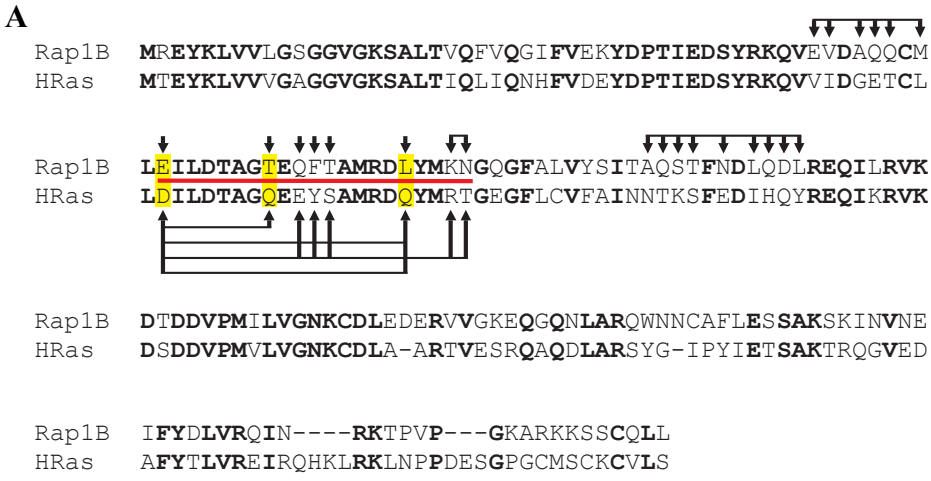
Mechanistically the GEF reaction is a multi-step process[4]. GEFs compete with the nucleotide for binding to the G-protein by partially overlapping interaction surfaces. In course of the exchange reaction a transient trimeric complex of GEF, G-protein and nucleotide is formed, in which either the binding interface of the GEF or the nucleotide is only partially established. The establishment of the full interaction between the G-protein and the GEF, or the G-protein and the nucleotide forces the respective other binding partner out of the trimeric complex. As the cellular GTP concentration is higher than the GDP concentration, this mechanism results overall in the exchange of GDP for GTP.
The thermodynamic and kinetic constants can be determined for all steps of this process and selected pairs of GEF and G-protein were analysed in biophysical detail[16]. Such a

characterisation requires high amounts of proteins, as the affinities of the trimeric complexes are low, and very high time resolution to monitor exchange rates at saturating concentrations of G-proteins. This makes a detailed analysis of many pairs unfeasible. Furthermore, such a characterisation is of mechanistic interest, but might be considered of limited practical use, as GEFs, G-proteins and nucleotides occur in a restricted concentration range under physiological conditions.

If a comprehensive biophysical characterisation is replaced by the determination of the exchange rate at fixed concentrations of GEF and G-protein, many combinations can be tested. The concentrations can be chosen far from saturation, which might also well reflect the physiological situation. The obtained exchange rates can then function as a measure of the GEF activity, even tough they are not absolute microscopic constants. By using identical reaction conditions, the obtained selectivity profiles of different G-proteins can be compared to identify preferences of the GEFs.

Here, this approach is used for Ras-family G-proteins in combination with CDC25-homology domain containing GEFs. The Ras family is constituted by 36 members (Fig. 1), some of which are extensively characterised: “The Ras-proteins” (H-Ras, K-Ras, N-Ras, M-Ras and R-Ras), which are involved in the control of cell growth and differentiation, which are established oncogenes and which are found frequently mutated in tumours[13]; “the Rap-proteins” (Rap1A, Rap1B, Rap2A, Rap2B and Rap2C), which were originally identified due to their anti-proliferative (Ras-phenotype reverting) activity and which regulate cell adhesion and cell junction formation[3]; “the Ral-proteins” (RalA and RalB), which are involved in the control of exocytose and transport of vesicles[12,22]; and finally the G-protein Rheb, which controls cell size by regulating the activity of the mTOR complex[1]. Much less is known of the remaining family members.

Representative CDC25-homology domain containing GEFs were selected based on the possibility to purify recombinant proteins and the following considerations. Rgl1, Rgl2, Rgl3 and RalGDS form a group of evolutionary conserved CDC25-homology domain-containing GEFs for RalA and RalB without any reported activity to other G-proteins and



B

Fig. 1 Sequence homology within the Ras-family. (a) Sequence alignment of Rap1B and H-Ras. Identical residues are highlighted in bold. Regions, which constitute the interaction surface for binding of the GEF are marked by a red line. Positions 54, 61 and 70 are marked by yellow background. Arrows indicate mutations which swap residues between Ras and Rap or vice versa; single arrow, single point mutation; connected arrows, multiple mutations. (b) Sequence alignment of all Ras family members covering the region from position 36 to 78 (Rap1A numbering). Residues matching the consensus sequence of the alignment are highlighted in bold. Residues identical or highly similar to Rap1b at position 54, 61 or 70 are marked by yellow and grey background, respectively. X_n, insertion of n residues.

are therefore excluded form the analysis[23,30]. Ras and Rap proteins are more homologous to each other than to Ral proteins (Fig. 1) and thus the analysis was focused on Ras- and Rap-GEFs. CDC25 homology domains can be highly related in sequence to each other, as PDZ-GEF1 and PDZ-GEF2 but can also be rather distant in evolution[30]. Proteins were selected to cover the different CDC25 homology domains as comprehensive as possible.
First, selectivity profiles of SOS1, Epac1, Epac2, C3G, PDZ-GEF1, PDZ-GEF2, RasGRP1, RasGRP2 and RasGRP3 were determined for H-Ras, K-Ras M-Ras, N-Ras, R-Ras, Rheb, Rap1A, Rap1B, Rap2A, Rap2B, Rap2C, Di-Ras1, Di-Ras2 and Rit (Fig. 2). Second, as structural information on GEF-G-protein complexes is only available for SOS1 and H-Ras and for Epac2 and Rap1B (Fig. 3), H-Ras and Rap1B were subjected to a mutational study to analyse the basis of discrimination between Ras and Rap proteins (Fig. 4).

Redundancy in Selectivity

PDZ-GEF1 and PDZ-GEF2 and to a large extent Epac1 and Epac2 have identical selectivity profiles. Similarly, H-Ras, K-Ras, N-Ras and M-Ras, Rap1A and Rap1B, and Rap2A, Rap2B and Rap2C respond to all tested GEFs almost identically. Regarding the sole enzyme-substrate interactions these groups are redundant GEFs or G-proteins. This is in agreement with the high sequence homology in the regions responsible for the interaction within these groups. This does, however, not mean that these proteins are redundant in the biological context. The catalytic activity of the CDC25-homology domain is combined with different (regulatory) properties. For example, Epac1 contains one and Epac2 two cyclic nucleotide binding domains and PDZ-GEF1 and PDZ-GEF2 show major differences in their C-terminal third. The most obvious differences within the G-proteins lie in the extreme C-termini, which mediate the cellular localisation of the protein. Therefore, signalling by a combination of one of the redundant GEFs and one of the redundant G-proteins becomes likely a unique pathway by the specific regulation of the GEF and the specific localisation of the G-protein in many cases.

Discrimination between Ras and Rap

With the exception of RasGRP3 all tested GEFs discriminate between Rap and Ras proteins and can thus be classified as Rap- and Ras-GEFs respectively (Fig. 2). Epac1, Epac2, C3G, PDZ-GEF1, PDZ-GEF2 and RasGRP2 act as Rap-GEFs, whereby some additionally discriminate between Rap1 and Rap2 (see below). Ras-GRP1 acts as Ras-GEF. RasGRP3 acts as Ras- and Rap-GEF, though with preference for Ras. The Ras-GEFs do not act on R-Ras (see below) but do not discriminate any further between H-Ras, K-Ras, N-Ras and M-Ras.
The major differences between Rap and Ras within the core interactions surface are found at positions 54 (Glu in Rap and Asp in Ras), 61 (Thr in Rap and Gln in Ras) and 70 (Leu in Rap and Gln in Ras) (Figs, 1, 3). The surrounding residues, in particular those that point towards the interaction surface are identical. The residues at position 54, 61 and 70 are thus expected to be the major determinants of selectivity. Indeed, in Rap the corresponding residue of Ras and vice versa in Ras the corresponding residue of Rap would be incompatible with the local environment found in the interaction surface and

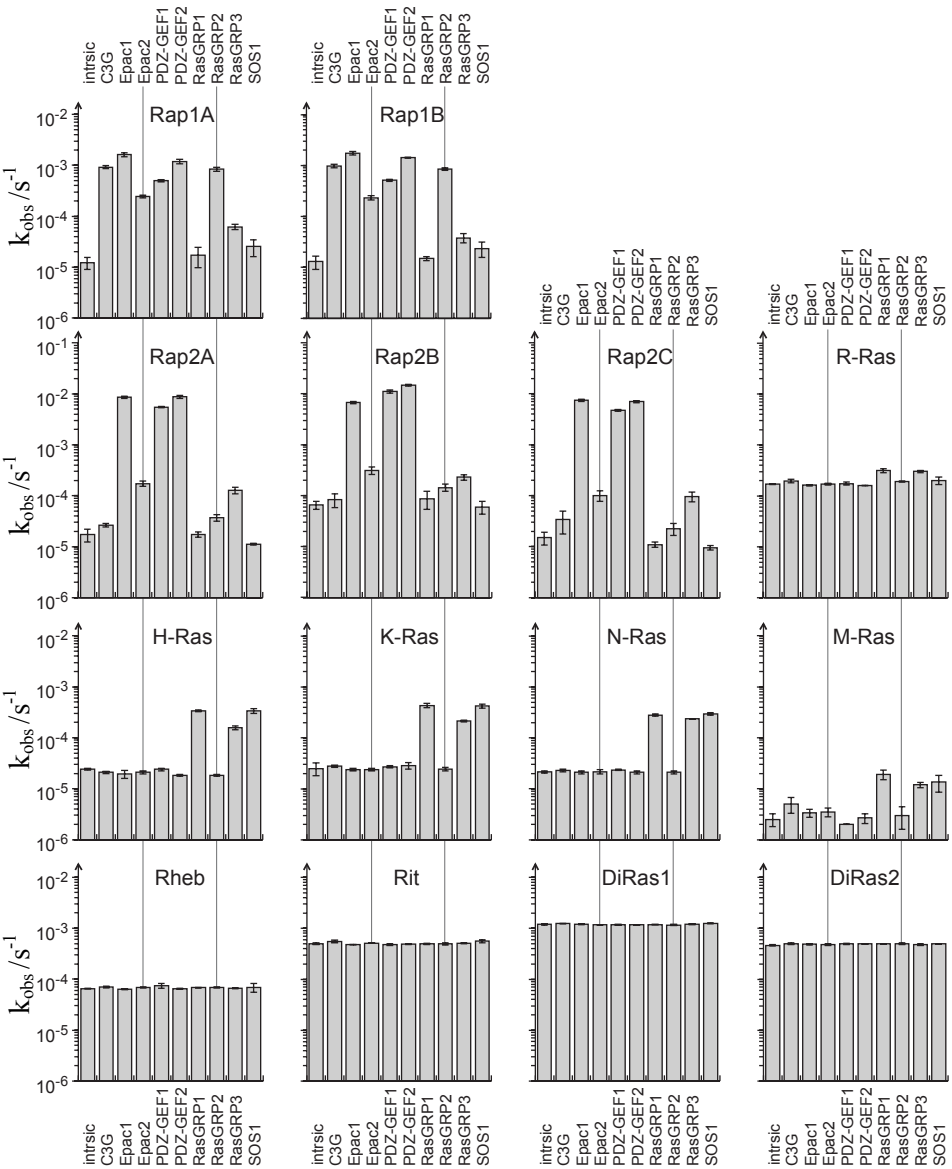


Fig. 2 Selectivity profile of CDC25-homology domain containing GEFs. To measure GEF activity G-proteins were loaded with the fluorescent GDP analogue 2'-/3'-O-(N-methylanthraniloyl)-guanosine diphosphate (mGDP) as described previously for Rap1B.[24] The fluorescence intensity of G-protein bound mGDP is approximately twice as intense as of free mGDP and thus nucleotide exchange can be observed as decay in fluorescence upon addition of an excess unlabelled GDP. The time dependency of the fluorescence signal can be fitted as single exponential decay with offset, whereby the obtained rate constant k_{obs} is a measure for GEF activity[24]. Each combination of GEF and G-protein were measured in triplicates and the obtained k_{obs} are presented in bar diagrams with logarithmic scale. The sole purpose of the light gray background lines is to guide the eye trough the

figure. All reactions were performed in 50 mM TrisHCl pH7.5, 100 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 5% glycerol and 20µM GDP at 20°C, with 200 nM G-protein and 150 nM GEF as described[24]. As an exception, the activity of SOS1 was recorded at 1500 nM, as the activity obtained with SOS1 is rather low towards all G-proteins tested under standard conditions. To allow more reliable conclusions regarding the selectivity profile of SOS1, exchange activity was measured at a 10 times higher concentration of SOS1. The catalytic activity of SOS1 is enhanced by binding of a second Ras molecule to a distant second binding site[20,28]. This effect might be missed with the truncated construct under the experimental conditions used here. Activity of Epac1 and Epac2, which are auto-inhibited and activated upon binding of cAMP[5,14], were determined in the presence of 100 µM cAMP. All proteins were expressed recombinantly in bacteria and purified as follow: H-Ras (aa 1-166, *homo sapiens*), Rap1A (aa 1-167, *homo sapiens*) and Rap1B (aa 1-167, *homo sapiens*) were expressed from the ptac plasmid as described[10]. K-Ras (aa 1-170, *homo sapiens*), M-Ras (aa 1-183, *homo sapiens*), N-Ras (aa 1-171, *homo sapiens*), R-Ras (aa 2-218 *homo sapiens*), Rheb (aa 1-184, *rattus norvegicus*), Rap2A (aa 1-183, *homo sapiens*), Rap2B (aa 1-166, *homo sapiens*), Rap2C (aa 1-183, *homo sapiens*), Di-Ras1 (aa 1-198, *homo sapiens*), Di-Ras2 (aa 1-199, *homo sapiens*) and Rit (aa 1-184 *homo sapiens*) were expressed as GST-fusion proteins, whereby the GST-tag was removed by thrombin cleavage as describe[24]. Epac1 (aa 149-881, *homo sapiens*) and Epac2 (aa 280-993, *mus musculus*) were expressed as GST fusion proteins and purified as described[24]. SOS1 (aa 564-1049, *homo sapiens*), C3G (aa 830-1078, *homo sapiens*), PDZ-GEF1 (aa 250-990, *homo sapiens*), PDZ-GEF2 (aa 410-1137, *homo sapiens*), RasGRP1 (aa 1-461, *rattus norvegicus*), RasGRP2 (aa 1-417, *homo sapiens*) and RasGRP3 (aa 1-383, *homo sapiens*) were expressed as His-tagged proteins by the used of the pET vector system and purified by nickel affinity chromatography in an imidazole gradient and gelfiltration according to standard procedures. cDNAs for Di-M-Ras (IRATp970E1274D) Ras1 (IRATp970D0628D) and Di-Ras2 (HU4_p940E09182D) were obtained from Source BioScience and cDNAs for K-Ras[38] (Addgene plasmid 31200) and N-Ras[6] (Addgene plasmid 14723) from Addgene.

would often induce serious steric clashes (Fig. 3).
To confirm the relevance of these residues, the Rap1B to H-Ras mutants $Rap1B^{E54D}$, $Rap1B^{T61Q}$ and $Rap1B^{L70Q}$ were generated. The exchange activity of Epac1 and Epac2 is drastically reduced towards all three mutants. PDZGEF1 and PDZGEF2 tolerate the T61Q mutation relative well, but are severely affected by the other two mutations. C3G on the other hand seems to act even slightly more efficient on $Rap1B^{T61Q}$ (see also [31]) and tolerates the E54D mutation.
If the selectivity would only be determined by these key residues, the corresponding mutations in Ras should result in the gain of Rap-GEF activity. However, neither Epac nor PDZGEF is acting efficiently on $H\text{-}Ras^{D54E,Q70L}$ or on $H\text{-}Ras^{D54E,Q61T,Q70L}$. This indicates that additional residues contribute to the selectivity of the interaction between RapGEFs and G-proteins.

Additional determinants of selectivity
To identify additional determinants of selectivity, residues in direct proximity to the core interaction surface in Rap1B were mutated to the corresponding amino acid in H-Ras. The activities of all RapGEFs are reduced towards $Rap1B^{F64Y}$ and to a lesser extend towards $Rap1B^{T65S}$, whereas almost no reduction is observed towards $Rap1B^{K73R,N74T}$. Only very limited effects are observed when Rap1B is swapped to H-Ras in the regions 45 to 52, 86

to 96 or 101 to 106, suggesting that these regions are not significantly involved in interaction with GEFs.
Indeed a significant gain of RapGEF activity is observed towards $Ras^{D54E,Q61T,E63Q,Y64F,S65T,Q70L}$, which almost resembles the activity towards Rap1B. Hardly any further gain is obtained upon addition of the mutations R73K and T74N.

Discrimination between Rap1 and Rap2
Rap1 and Rap2 proteins are highly similar in sequence and most of the residues in the Rap-Epac interaction interface are identical. Indeed RasGRP3, PDZGEF1, PDZGEF2, Epac1 and Epac2 do not discriminate between Rap1 and Rap2, even though Epac2 has a slight preference for Rap1 over Rap2. Interestingly, C3G and RasGRP2 clearly discriminate against Rap2. C3G is an established Rap1 selective RapGEF and selectivity has be shown to depend on residues 65 and 66 as well as by the region from residue 105 to 108 in Rap1[31].

The selectivity determinants are interconnected
RasGRP3 acts as Ras- and Rap-GEF. In the simplest scenario it should therefore be indifferent towards swapping mutations of the main selectivity determinates at position 54, 61 and 70. However, its activity towards $Rap1B^{T61Q}$ and $Rap1B^{L70Q}$ is almost completely abolished (Fig. 4). Based on the crystal structures of Epac2•Rap1B[25] and Sos1•H-Ras[2], this effect can be explained by the influence of the mutation of one of these residues on the respective two others.
Leu70 of Rap1B points into a hydrophobic environment formed amongst others by Met882 and Met850 of Epac2 (Fig. 3a). This pocket can not host the more space filling Gln70 of H-Ras. In SOS1 the methionine residues are replaced by serine residues, one of which forms a hydrogen bond with Gln70. In addition the Cα-atom of Gln70 in H-Ras is shifted by 2.2 Å relative to Leu70 in Rap1B. This causes changes within the whole helix formed by the neighbouring residues, which are thus, even though conserved, reoriented. These changes are propagated trough the protein: The orientation of Leu56 in H-Ras is flipped as compared to Leu56 in Rap1B such that Tyr71 can be accommodated (Fig. 3a). Tyr71 is performing different interactions in the Epac2•Rap1B and the SOS1•H-Ras complex (Fig. 3b). These interactions are incompatible with the respective other conformation of Tyr71. The conformation of Tyr71 in H-Ras is in addition favoured by a hydrogen bond with Gln61, which is in turn positioned by Thr935 of SOS1 (Fig. 3c). In contrast, the corresponding Lys911 in Epac2 favours the conformation of Tyr71 in Rap1B by a direct hydrogen bond to the hydroxyl group. In the Epac2•Rap1B complex, Thr61 is coordinating a water molecule together with the carboxy oxygen of the peptide bonds of Asp57 and Ala59. This stabilises the loop between Asp57 and Ala59 in the Epac2•Rap1B complex in a different conformation than in the SOS1•H-Ras complex. The positioning of Thr58 in H-Ras would be incompatible with the orientation of Tyr71 in Rap1B. Position 70 and 61 are thus coupled via Tyr71.
Similarly, Tyr71 is coupling via Leu56 to position 54 (Fig. 3d). The conformation of Leu56 in the Epac2•Rap1B complex would be sterically incompatible with the position of His911 in SOS1, which forms a hydrogen bond with Asp54 in H-Ras. Thus the three

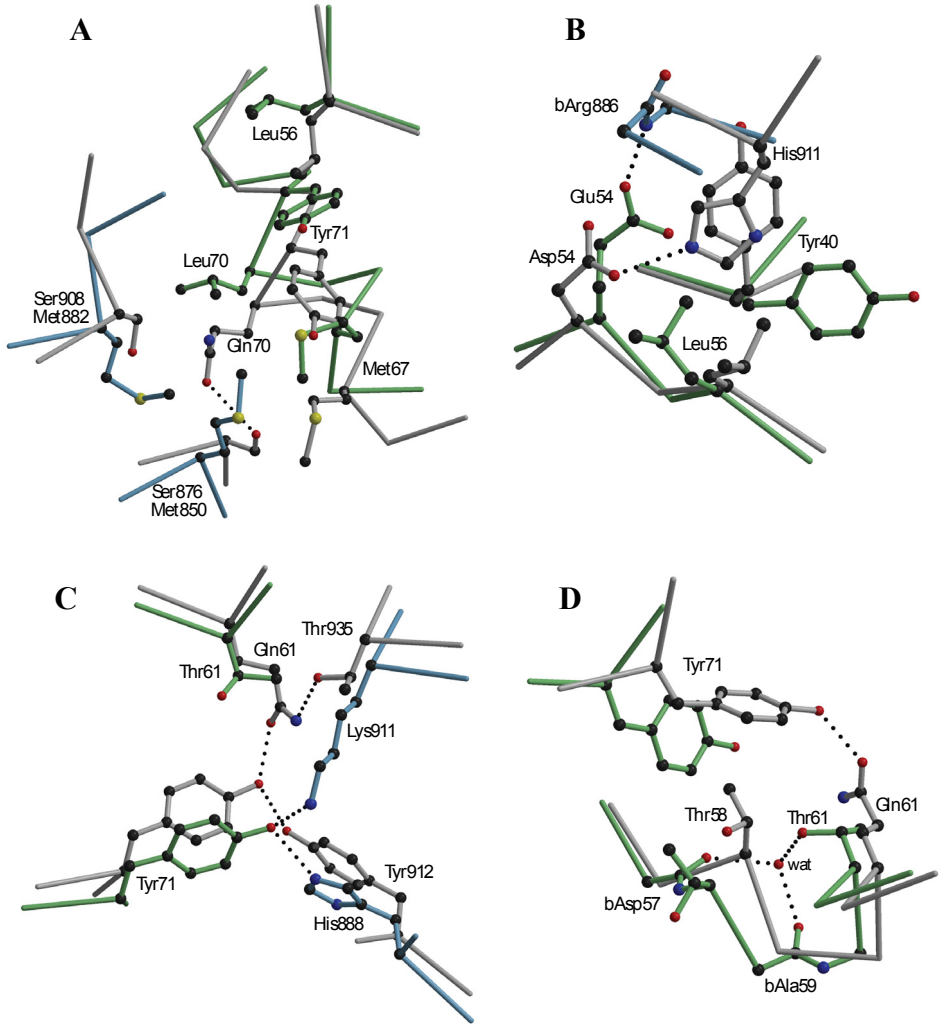


Fig. 3 Comparison of the GEF-G-protein interaction in the Epac2•Rap1B and the SOS1•H-Ras complex. The crystal structures of the Epac2•Rap1B complex (pdb entry 3CF6)[25] and the SOS1•H-Ras complex (pdb entry 1BKD)[2] were used to analyse the interaction between the G-proteins and CDC25-homology domains. The G-proteins were superimposed onto each other by the algorithms of Bragi[27]. Epac2 is shown in light blue, Rap in light green and SOS1 and Ras in light grey. Hydrogen bonds are indicated by dotted lines. Amino acid residues are labelled by the three letter code; lower case b indicates that only the backbone is depicted. (a) Surrounding of Leu70 in Rap and Gln70 in Ras. (b), (c) Surrounding of Thr61 in Rap and Gln61 in Ras; wat, water molecule. (d) Surrounding of Glu 54 in Rap and Asp54 in Ras. The figures were generated using the programs MolScript and Raster3D[17,21].

positions, which are the major determinant of selectivity are interconnected with each other. Mutants, in which not all three residues are swapped simultaneously, cannot be

considered as a linear combination of Ras and Rap properties. Instead, they have to adapt their conformation to accommodate for an otherwise incompatible combination of residues and thereby acquire novel properties, which are not intrinsic to either H-Ras or Rap1B. This is nicely illustrated by C3G, which first gains ($Ras^{D54E,Q70L}$) than loses ($Ras^{D54E,Q61T,Q70L}$) and finally gains again ($Ras^{D54E,Q61T,E63Q,Y64F,S65T,Q70L}$) activity towards Ras mutants with increasing Rap character.

G-proteins without a GEF

None of the tested GEFs works efficiently on R-Ras, Rheb, Rit, Di-Ras1 or Di-Ras2. All four G-proteins are characterised by a relative high intrinsic exchange rate, which is about 10 times higher than the averaged intrinsic rate of most other G-proteins of the Ras-family tested. It might be that these G-proteins are not regulated by GEFs. Modulation of the nucleotide status might occur via the regulated action of GTPase activating proteins. In such a scenario regulation would impinge rather on "off-switching" than on "on-switching" of signalling. Regulation of Rheb by its GAP TSC1/2 is well established[39] and RapGAP1 is reported to act on Di-Ras[7]. On the other hand, nucleotide exchange of Rap2B, which has a relative high intrinsic exchange rate as well, is found to be efficiently catalysed by several GEFs (Fig. 2).

The inability of CDC25-homology domains to act as a GEF for Rheb is in agreement which the general view that a Rheb-GEF is not discovered yet. Here, CDC25-homology domains were tested for the first time systematically with Rheb. Hsu et al. have proposed the non CDC25-homolgy domain containing TCTP protein to act as a GEF for Rheb[11], but these findings were questions by subsequent structural and functional studies[26,33].

Epac[19], C3G[8] and some RasGRPs[15,36] have been suggested to act on R-Ras. In consequence R-Ras is often considered to be more "Rap-like" than "Ras-like". This assignment was supported by the biological function of R-Ras in cell-adhesion processes[9,18] in which Rap-proteins are involved as well[3]. Weak catalytic activity of RasGRP1 and RasGRP3 was found in this study. Likely this activity has prompted earlier work to classify RasGRP proteins as GEFs for R-Ras. Similarly, the intrinsic exchange rate of R-Ras is enhanced about 1.2 fold but that of Rap1B about 70 fold by C3G. The mutagenesis study on Rap1B and H-Ras has identified positions 54, 61 and 70 as key discriminators between Ras and Rap selective GEFs. All three positions in R-Ras are Ras-like and not Rap-like (Fig. 1). Formally it cannot be excluded that very weak activities towards R-Ras are sufficient to initiate signalling *in vivo*, where proteins are organised in a cellular micro-environment. However, signalling of different G-proteins is often interconnected and GEFs for one G-protein functions as effectors of others. Thus, GEFs like Epac or C3G might affect R-Ras GTP levels indirectly, which would explain the observed effect of these GEFs on R-Ras in biological systems[19].

Prediction of unknown GEFs

Not all G-proteins of the Ras-family nor all CDC25-homology domain containing GEFs were included in this study. However, the tested CDC25 homology domain containing proteins were selected as representatives of their family. No fundamental differences were observed for highly related CDC25 homology domains in Epac1 and Epac2 or in PDZ-

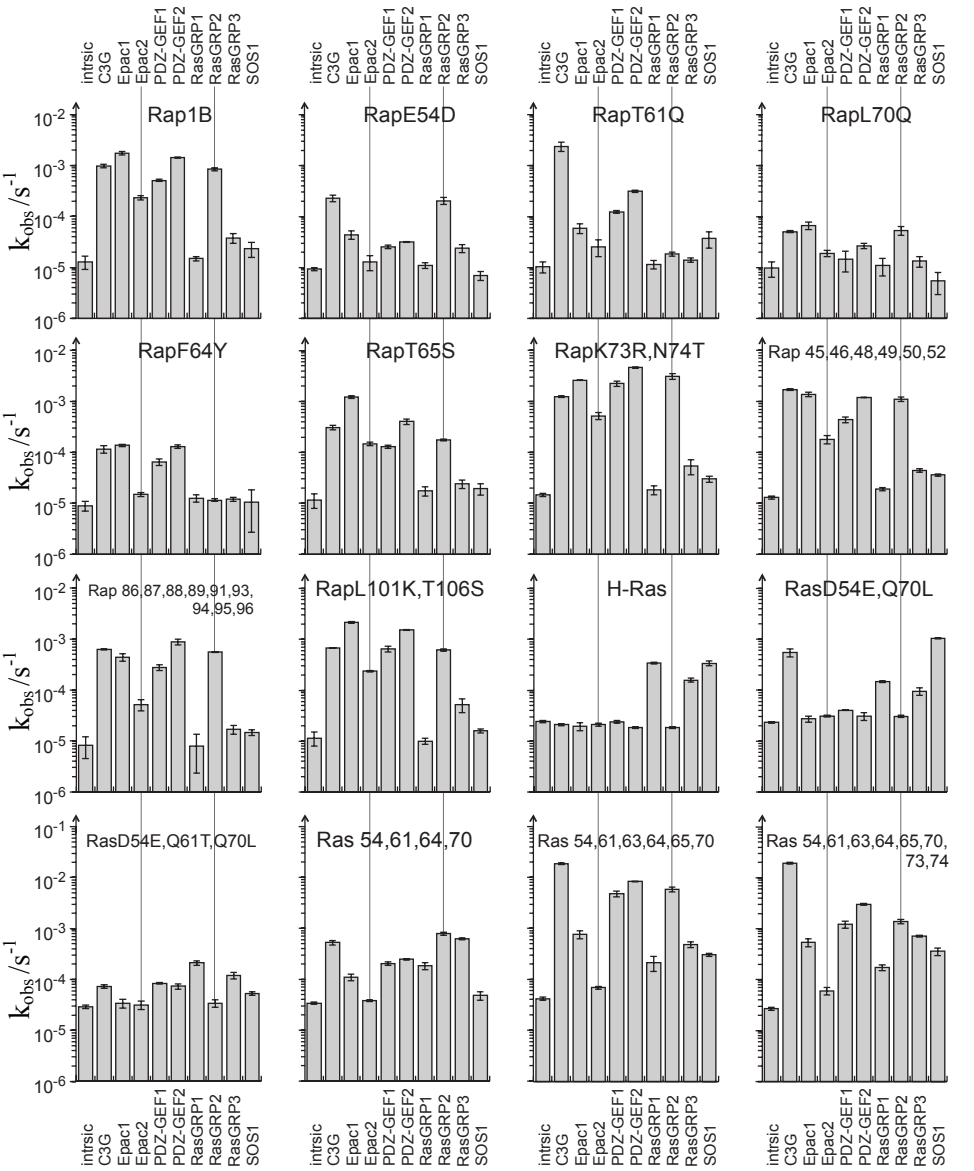


Fig. 4 Mutational study of H-Ras and Rap1B. Rap1B and H-Ras were subjected to mutagenesis, whereby mutations in Rap1B exchange amino acids to the corresponding residue in H-Ras and vice versa. For clarity of comparison the data of the wild type proteins are copied from Fig. 2. GEF activity was determined as described in the legend of Fig. 2.

GEF1 and PDZ-GEF2. Thus, it seems reasonable to consider for example SOS1 as a representative for SOS2 as well. Furthermore it is unlikely that the investigated GEFs could act on other members of the Ras-family, which are more distinct in sequence. The combined conservation of glutamic acid, threonine and leucine at positions 54, 61 and 70

is a unique feature of the Rap proteins. Mutation of any of these residues to the corresponding residue in Ras drastically reduces the catalytic activity of the Rap selective GEFs, in particular of Epac. This makes it unlikely that Epac is able to act as a GEF for any other member of the Ras family. Di-Ras1 and Di-Ras2 might be closest related to Rap in respect to positions 54, 61 and 70. Still no activity of Epac or any other tested GEF was observed towards these G-proteins. Similarly, R-Ras, which is highly related to H-Ras, K-Ras, N-Ras and M-Ras is not responding to the tested Ras-GEFs. The action of CDC25 homology domain containing GEFs seems therefore be restricted to Ras-, Rap- and Ral-proteins and GEFs for the other Ras-family members seem to be missing.

Validity of the approach

In this study selectivity profiles for a variety of G-proteins from the Ras family were generated by relative simple measurements, which made the analysis of many GEF and G-protein pairs possible. This approach takes into consideration that the intrinsic exchange rate varies from G-protein to G-protein. The comparison of absolute rates of one GEF for different G-proteins might therefore be misleading and it is therefore preferable to group the data as selectivity profiles of individual G-protein. Formally it would be possible to express the GEF-catalysed exchange rates relative to the intrinsic exchange rate. However, the intrinsic exchange rate is by orders of magnitude smaller than the catalysed exchange rate for an efficient GEF. Furthermore slow intrinsic exchange rates might be determined incorrectly due to issues of protein stability. Normalisation to relatively small and putatively incorrect values should be avoided. Instead, the visual comparison of specificity profiles of different G-proteins as presented in Figs 2 and 4 accounts for differences in the intrinsic exchange rates.

As mentioned the measured exchange rates are not absolute. A meaningful conclusion for a GEF is only possible if the GEF shows significant exchange activity for at least one G-protein. Otherwise the GEF might be inactive as for example due to a lacking co-factor or simply due to improper protein folding. In case of an overall weak activity of a particular GEF, this GEF can be analysed at higher concentration, as it was done here for SOS1.

Conclusion

To the knowledge of the authors, this study presents the most comprehensive data set on the selectivity of CDC25 homology domain containing GEFs and thereby on the selectivity of the initiation of signalling in the Ras family of small G-proteins. The data set helps to clarify the assignment of GEFs to G-proteins in cases which are controversial in literature. For example, the proposed action of several GEFs on R-Ras is put into perspective by the generated activity profile of the GEFs. Signalling of individual G-proteins is often directly coupled trough the regulation of GEFs and GTPase activating proteins by G-proteins. The presented data set should allow a good prediction of whether a by biological means discovered link between a GEF and a G-protein, is direct or indirect and can thereby help to direct subsequent functional studies. Furthermore this study described an experimental approach to compare under standardised conditions many pairs of GEFs and G-proteins and to deal with the kinetically complex GEF reaction. Structural information is available only for two complexes of GEF and G-protein. However, the

combination of this information with a mutational study as presented here, allows the identification of selectivity determinants and a clear prediction of the ability of a GEF to act on a certain G-protein.

Acknowledgements
Adrienne Cox, David Root and Alfred Wittinghofer are acknowledged for providing cDNA constructs. We thank Jelena Linnemann and Nayara Cavalcante Silva for contributions to cloning and protein purifications and member of the Bos lab for discussions and critical reading of the manuscript. MP and MRL were supported by the TI Pharma Project T3-106 to J.L. Bos, and HR by the Chemical Sciences of the Netherlands Organization for Scientific Research (NWO).

References

1. J. Avruch, *et al*. Insulin and amino-acid regulation of mTOR signaling and kinase activity through the Rheb GTPase. *Oncogene* **25**, 6361 (2006).
2. P. A. Boriack-Sjodin, *et al*. The structural basis of the activation of Ras by Sos. *Nature* **394**, 337 (1998).
3. J. L. Bos. Linking Rap to cell adhesion. *Curr. Opin. Cell Biol.* **17**, 123 (2005).
4. J. L. Bos, H. Rehmann, and A. Wittinghofer. GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865 (2007).
5. J. de Rooij, *et al*. Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature* **396**, 474 (1998).
6. J. J. Fiordalisi, *et al*. Mammalian expression vectors for Ras family proteins: generation and use of expression constructs to analyze Ras family function. *Methods Enzymol.* **332**, 3 (2001).
7. R. Gasper, B. Sot, and A. Wittinghofer. GTPase activity of Di-Ras proteins is stimulated by Rap1GAP proteins. *Small Gtpases* **1**, 133 (2010).
8. T. Gotoh, *et al.* Activation of R-Ras by Ras-guanine nucleotide-releasing factor. *J. Biol. Chem.* **272**, 18602 (1997).
9. G. S. Griffiths, *et al*. R-Ras interacts with filamin a to maintain endothelial barrier function. *J. Cell Physiol* **226**, 2287 (2011).
10. C. Herrmann, *et al*. Differential interaction of the ras family GTP-binding proteins H-Ras, Rap1A, and R-Ras with the putative effector molecules Raf kinase and Ral-guanine nucleotide exchange factor. *J. Biol. Chem*. **271**, 6794 (1996).
11. Y. C. Hsu, *et al.* Drosophila TCTP is essential for growth and proliferation through regulation of dRheb GTPase. *Nature* **445**, 785 (2007).
12. V. Jullien-Flores, *et al*. RLIP76, an effector of the GTPase Ral, interacts with the AP2 complex: involvement of the Ral pathway in receptor endocytosis. *J. Cell Sci.* 113, 2837 (2000).
13. A. E. Karnoub and R. A. Weinberg. Ras oncogenes: split personalities. *Nat. Rev. Mol. Cell Biol.* **9**, 517 (2008).
14. H. Kawasaki, *et al*. A family of cAMP-binding proteins that directly activate Rap1. *Science* **282**, 2275 (1998).
15. H. Kawasaki, *et al*. A Rap guanine nucleotide exchange factor enriched highly in the basal ganglia. *Proc Natl Acad Sci* **95**, 13278 (1998).
16. C. Klebe, *et al*. The kinetic mechanism of Ran--nucleotide exchange catalyzed by RCC1. *Biochemistry* **34**, 12543 (1995).
17. P. J Kraulis. Molscript - a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst*. **24**, 946 (1991).
18. M. Lehto, *et al*. The R-Ras interaction partner ORP3 regulates cell adhesion. *J. Cell Sci*. **121**, 695 (2008).
19. Jesus M. Lopez De, *et al*. Cyclic AMP-dependent and Epac-mediated activation of R-Ras by G protein-coupled receptors leads to phospholipase D stimulation. *J. Biol. Chem*. **281**, 21837 (2006).
20. S. M. Margarit, et al. Structural evidence for feedback activation by Ras.GTP of the Ras-specific nucleotide exchange factor SOS. Cell **112**, 685 (2003).
21. E. A Merritt and M. E. P Murphy. Raster3D version-2.0 - A program for photorealistic molecular graphics. Acta Crystallogr D Biol crystallogr **50**, 869 (1994).
22. S. Moskalenko, *et al*. The exocyst is a Ral effector complex. *Nat. Cell Biol*. **4**, 66 (2002).

23. L. A. Quilliam, J. F. Rebhun, and A. F. Castro. A growing family of guanine nucleotide exchange factors is responsible for activation of Ras-family GTPases. *Prog. Nucleic Acid Res Mol Biol* **71**, 391 (2002).
24. H. Rehmann. Characterization of the activation of the Rap-specific exchange factor Epac by cyclic nucleotides. *Methods Enzymol* **407**, 159 (2006).
25. H. Rehmann, *et al*. Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B. *Nature* **455**, 124 (2008).
26. H. Rehmann, *et al*. Biochemical characterisation of TCTP questions its function as a guanine nucleotide exchange factor for Rheb. *FEBS Lett*. **582**, 3005 (2008).
27. D. Schomburg and J. Reichelt. Bragi - A Comprehensive Protein Modeling Program System. *Journal of Molecular Graphics* **6**, 161 (1988).
28. H. Sondermann, *et al*. Structural analysis of autoinhibition in the Ras activator Son of sevenless. *Cell* **119**, 393 (2004)..
29. P. Thaw, *et al*. Structure of TCTP reveals unexpected relationship with guanine nucleotide-free chaperones. *Nat. Struct. Biol.* **8**, 701 (2001).
30. T. J. van Dam, *et al.* Phylogeny of the CDC25 homology domain reveals rapid differentiation of Ras pathways between early animals and fungi. *Cell Signal*. **21**, 1579 (2009).
31. N. van den Berghe, R. H. Cool, and A. Wittinghofer. Discriminatory residues in Ras and Rap for guanine nucleotide exchange factor recognition. *J Biol Chem* **274**, 11078 (1999).
32. I. R. Vetter and A. Wittinghofer. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299 (2001).
33. X. Wang, *et al*. Re-evaluating the roles of proposed modulators of mammalian target of rapamycin complex 1 (mTORC1) signaling. *J. Biol. Chem.* **283**, 30482 (2008).
34. K. Wennerberg, K. L. Rossman, and C. J. Der. The Ras superfamily at a glance. *J. Cell Sci.* **118**, 843 (2005).
35. V. Yajnik, *et al*. DOCK4, a GTPase activator, is disrupted during tumorigenesis. Cell **112**, 673 (2003).
36. S. Yamashita, *et al*. CalDAG-GEFIII activation of Ras, R-ras, and Rap1. *J Biol Chem* **275**, 25488 (2000).
37. J. Yang, *et al*. Activation of Rho GTPases by DOCK exchange factors is mediated by a nucleotide sensor. Science **325**, 1398 (2009).
38. X. Yang, *et al*. A public genome-scale lentiviral expression library of human ORFs. *Nat. Methods* **8**, 659 (2011).
39. Y. Zhang, *et al*. Rheb is a direct target of the tuberous sclerosis tumour suppressor proteins. *Nat Cell Biol* **5**, 578 (2003).

Chapter 5

General discussion and summary

Activation mechanism of C3G

In Chapter 2 we have presented the analysis of the activation mechanism of C3G. We have demonstrated that C3G is autoinhibited by its N-terminal part. Autoinhibition is a common mechanism of regulation of GEFs. Structures of several GEFs regulated by the autonhibition have been solved[1,2]. Autoinhibition is established by the interaction of the catalytic region with other parts or domains of the GEF. This interaction prevents the access of the G-protein to the catalytic site. We have therefore suggested that C3G is autoinhibited by the interaction of its N-terminal region with the catalytic CDC25-HD.

We have found that autoinhibition of C3G is partially relieved by binding of Crk. The individual Crk-binding sites have different roles in the activation process. Binding to site III has been demonstrated to be required for the release of autoinhibition. When binding of Crk to this site is abolished by mutations activation by Crk is lost. Binding of Crk to site IV seems not to be necessary as mutations in this site do not prevent activation by Crk. On the other hand, binding of Crk to site IV might prime site III for binding as ITC data suggest interplay between these sites. Sites I and II are not required to maintain the autoinhibited state as N-terminal truncated C3G which lacks these sites is still autoinhibited. It therefore seems that site III plays the key role in Crk-mediated C3G activation and that site IV might have a supportive function.

We have further shown that tyrosine phosphorylation of C3G by Src leads to increased activity as well. Mapping studies have suggested that phosphorylation of tyrosine residues C-terminal to Pro544 are required for the activation. We failed, however, to prevent Src-mediated activation by mutations of single tyrosine residue in the region between Pro544 and the start of the catalytic region. It could be that several tyrosine residues participate in inhibitory interactions with the CDC25-HD, but that phosphorylation of only one residue is sufficient to break the interaction. Alternatively, a tyrosine residue in the CDC25-HD might be the target of the kinase.

Binding of Crk or phosphorylation by Src causes only a partial activation of C3G. When the two are combined, a substantial increase in C3G activity is obtained. Based on these findings we have proposed that the inhibitory interaction between the N-terminal part and the CDC25-HD consists of two components, one that responds to Crk and one that responds to tyrosine phosphorylation. Binding of Crk or phosphorylation by Src each relieve only one inhibitory component, while the other one remains mainly unaffected. By simultaneous tyrosine phosphorylation and interaction with Crk, the interactions of both inhibitory components are broken leaving C3G fully active (Figure 7, Chapter 2). As discussed in Chapter 2, this model is somewhat simplified, but can be adapted to explain more complicated activation characteristics.

A striking feature is the different maximal activation levels of phosphorylated C3G induced by Crk I and Crk II. Crk II differs from Crk I by an additional C-terminal SH3 domain. The C-terminal SH3 domain is expected to be incapable of binding to proline-rich sequence[3]. Thus the two SH3 domains of Crk II can not simply bind to two proline-site simultaneously. In fact the C-terminal SH3 domain has been implicated in negatively influencing the N-terminal SH3 domain of Crk II[4]. Structural studies have demonstrated that the module composed of the linker between the SH3 domains and the C-terminal SH3

domain blocks the binding grove of the N-terminal SH3 domain[5]. In agreement with this, Crk I binds proline rich peptides with higher affinities than Crk II[4]. In the biological context Crk I is considered as being more potent than Crk II, as it has higher transforming activity[6]. In contradiction to this, we have found in initial ITC experiments that CrkII binds with higher affinity to C3G than Crk I and that Crk II induces higher maximal activity of C3G independently of the affinity difference. Thus Crk II is a more potent C3G activator than Crk I.

What are the implications of our findings for the function of C3G in cellular context? Crk was considered to act solely as an adaptor protein that recruits C3G to cellular membranes. The complex of C3G and Crk is mainly considered to be constitutive. Regulation of C3G was assumed to occur at the level of localization. Furthermore, phosphorylation of C3G upon stimulation of cells was observed and considered to correlate with C3G activity *in vivo*. We have demonstrated that phosphorylation of C3G results in a direct activation of C3G. Furthermore, this activation is strongly enhanced by binding of Crk. The ability of the $SH3^N$ domain of Crk II to interact with proline-rich motifs is regulated. For example, Abl acts as an inhibitor of Crk II-mediated signaling. Abl induces phosphorylation of Tyr221 in the linker region between the N- and the C-terminal SH3 domain. Tyr221 binds to the SH2 domain of CrkII in cis and adapts thereby a position which blocks the accessibility of the binding grove in the N-terminal SH3 domain[7–9].

Upon activation of tyrosine kinases, binding sites for Crk are created in proteins such as p130Cas[10], paxillin[11], Cbl[12], Dab1[13,14] and IRS[15,16]. So far, C3G is reported to be present in Crk-induced complex with p130Cas[17,18], Cbl[18,19] and Dab1[20].

p130Cas is an adaptor protein identified as the major phosphorylated protein upon vCrk-induced transformation. $p130Cas^{-/-}$ fibroblasts have short disorganized actin filaments, defects in actin bundling and poorly organized focal adhesions[21]. Its central region contains 15 YxxP motifs that are the target of phosphorylation and that function as SH2-domain docking sites. Apart from Crk, this motifs are reported to bind Nck and SHIP2[22]. The formation of the C3G/Crk/p130Cas complex is well established and mediates Rap activation during cell stretching[23]. Phosphorylation of the various YxxP motifs in p130Cas occurs at random order, suggesting that the repetitiveness of the YxxP motifs does not serve to organize a temporal signaling pattern. Likely the multiple motifs serve to amplify the Crk signaling pathways (Figure 1). Apart from amplification, the repetitive YxxP motifs could have a role in bringing together various Crk-complexes, such as Crk/C3G, Crk/Abl, Crk/DOCK180 or Crk/JNK and/or Nck- and SHIP2-mediated complexes[24]. These complexes would then be localized closely and could therefore influence the signaling of each other. For example, Abl has been reported to phosphorylate C3G[25]. On the other hand, Abl-mediated phosphorylation of Crk II would result in the release of C3G from the complex and thus contribute to termination of C3G-mediated signaling.

As Src is bound to p130Cas, C3G is upon translocation to p130Cas by Crk immediately localized in the proximity of Src. This leads to a simple model of activation of C3G in the C3G/Crk/p130Cas/Src complex. The activation sequence would be started by activation of tyrosine kinases, which then phosphorylate p130Cas to create binding sites for the SH2 domain of Crk. The Crk/C3G complex would translocate to p130Cas where Src is already

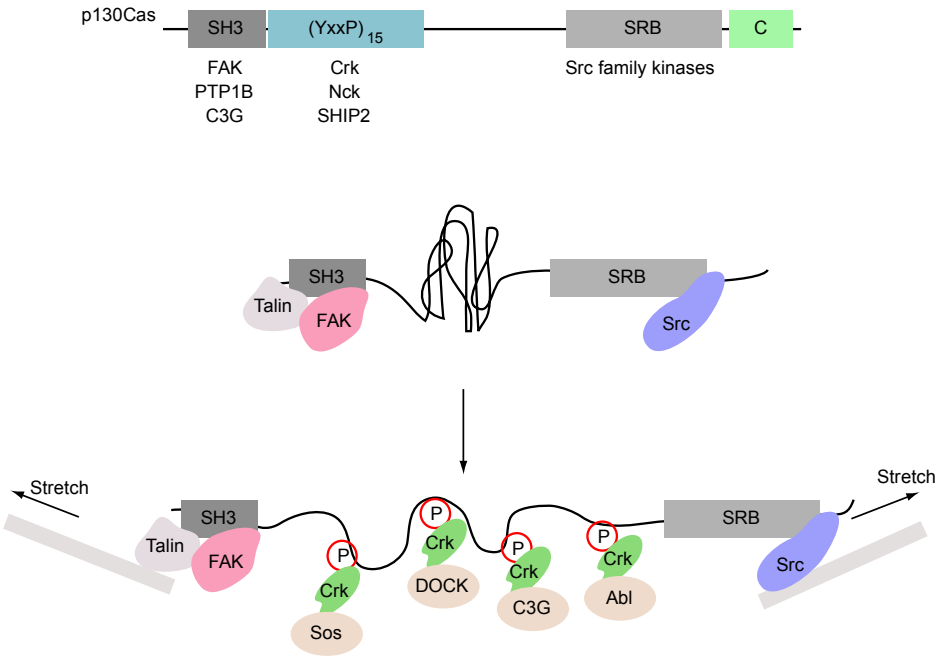


Figure 1. Domain organization of p130Cas and interactions assembly of Crk-mediated protein complexes.
Adapted from [24]. Some of the p130Cas-interacting proteins are indicated bellow p130 Cas SH3 – SH3 domain, (YxxP)15 – repeats of the YXXP motif in the substrate binding region, SRB – Src Binding region, C – C-terminal region. The substrate binding region is shown as a compressed structure when the tyrosine residues from the YXXP motifs are non-phosphorylated, and as an extended structure upon tyrosine phosphorylation. Mechanical force leads to stretching of the substrate binding region (indicated by arrows). The Crk-mediated complexes with C3G and other Crk-interacting proteins may be spatially integrated.

present to directly act on C3G.
Similarly to activation in p130Cas complexes, C3G is reported to be phosphorylated in the reelin pathway. The secreted glycoprotein reelin initiates signaling upon binding to its cellular receptors, that leads to phosphorylation of the adaptor protein Dab1[26,27]. Crk/C3G then binds to Dab1 where C3G becomes phosphorylated and causes activation of Rap1 downstream of reelin[20]. The kinase that phosphorylates C3G upon stimulation by reelin is not identified, but Dab1 is a substrate of Src and Fyn[28,29]. This would indicate that Src is localized in proximity of the Dab1/C3G complex and that all components for the activation of C3G are present in the Dab1-complex.
Another reported complex of C3G is C3G/Crk/Cbl. Interestingly, Src is bound to Cbl. This interaction is not mediated by the SH2 domain of Src as in the case of p130Cas but by the SH3 domain of Src. Furthermore Cbl is suggested to act as Src inhibitor by binding to its catalytically important Tyr419 residue[30]. Inhibition of Src would be incompatible

with the activation of C3G in this complex. Experimental data on the activity of C3G in the C3G/Crk/Cbl complex are conflicting. While some indeed report that Cbl has a negative effect on C3G activity and Rap signaling[31,32], others report a positive effect of Cbl on Rap activation[33,34].

Could there be a role for Src in C3G activation besides tyrosine phosphorylation of C3G? We have demonstrated that the SH3 domains of Src and Hck bind to C3G, although with significantly weaker affinity than Crk. However, it may be possible that Src would establish interaction with C3G in the conditions where their local concentrations would be high, for example in p130Cas complexes where both C3G and Src are enriched. The SH3 domain of p130Cas directly interacts with a proline rich sequence distinct to the Crk-binding sites[35]. It may be possible that C3G would be activated via a Crk-independent route. In this case Src that is also bound to p130Cas may interact with the Crk-binding sequences of C3G and phosphorylate C3G. Another possibility would be that the localization of C3G via the direct interaction with p130Cas would not result in C3G activation, but rather in promotion of Rap-independent roles of C3G.

Considering the activation sequence of C3G, turning off C3G-signaling could be achieved by several mechanisms: dephosphorylation of activatory tyrosine residues of C3G, breaking the interaction with Crk (or other proteins that would bind to the proline-rich sites involved in the activation of C3G), delocalization of active C3G by e.g. dephosphorylation of adaptor proteins that are involved in assembling Crk/C3G complexes (such as p130Cas) and by combination of these.

C3G has been reported to be dephosphorylated by TC48 phosphatase[36]. Phosphorylation of Crk and dissociation of the C3G/Crk complex has been reported to occur downstream of EGF- and insulin receptor-signaling[37,38]. Furthermore, ubiquitination of Crk by Cbl results in the dissociation of C3G/Crk complex [32]. PTP1B phosphatase has been reported to dephosphorylate CrkII, thereby positively regulating C3G signaling[37]. However, PTP1B could also have an overall negative effect on the activity of C3G, as it has been shown to dephosphorylate p130Cas too.

Insights in the structure and regulation of Rlf

In Chapter 3 we have investigated functional and structural properties of the Ral-GEF Rlf. Rlf belongs to the RalGDS-family of Ral-GEFs, that share the same domain organization. Apart from the catalytic module constituted by the REM-domain and the CDC25-HD, they contain a regulatory RA domain. The RA domain controls the cellular localization of RalGEFs. In several cases RA domains are implicated in the allosteric regulation of enzymatic activity. For example, the RA-domain of PLCε is proposed to influence lipase activity[39] and the RA domain in ARAP3 is proposed to influence GAP activity[40]. We have therefore investigated if binding of active Rap or Ras could influence the activity of Rlf (Chapter 3). The activity of Rlf is high and found to be in the range of the activity typically observed for constitutive active GEFs (Chapter 4), suggesting that Rlf is not autoinhibited. Furthermore, the activity of Rlf remains unchanged upon binding to active Ras and Rap (Chapter 3).

The structure of Rlf^{50-514} identified a prominent three stranded β-sheet preceding the REM domain. The β-sheet protrudes away from the compactly folded REM-CDC25-HD module. Among the so far crystallized Ras-family GEFs only Rlf contains such a pronounced β-sheet structure. The Ras-GEF Sos has only two short β-strands preceding the REM domain. The Rap-GEF Epac2 contains three β-strands that are N-terminal to its REM domain, but they are much shorter than the β-strands present in Rlf and do not form an extensive rod-like structure. The β-sheet element stabilizes the structure of Rlf, as its deletion leaves Rlf insoluble.

The β-sheet element is a conserved feature in GEFs of the Ral-GDS family. A unique feature of the β-sheet in Rlf is a proline-rich sequence present in the tip of a loop connecting the second and the third β-strand. As it is located in a most protruding part of the Rlf structure, it is very accessible and thus ideally located for the establishment of protein-protein interactions. Indeed, a yeast two-hybrid screen for Rlf-interacting proteins identified the SH3 domain-containing proteins, myosin1E, PLCγ and intersectin2. We have confirmed that these interactions are mediated by binding of the SH3 domain to the proline rich sequence in the β-sheet. The conformation adapted by the β-sheet is almost identical to that of a proline-rich peptide bound to the SH3 domain of myosin I from *A. castellanii*. The thermodynamic analysis of the interaction of an SH3 domain with the isolated Rlf-peptide and with Rlf^{50-514} suggest that the β-sheet fixes the conformation of the proline-rich stretch in a way that is optimized for binding to SH3 domains.

Although other GEFs of the RalGDS-family are expected to have a similar β-sheet structure, the proline rich sequence is unique to Rlf. Therefore the other family members are not expected to interact with SH3 domains. However, it may still be that in these GEFs the β-sheet mediates other types of protein-protein interactions. Alternatively, the function of the β-sheet could be reduced to its stabilising effect on the orientation of the REM domain relative to the CDC25-HD in theses GEFs.

The physiological relevance of the interactions between Rlf and myosin1E, PLCγ-1 and intersectin2 is not established. Interestingly, PLCγ-1 seems to be related to Ral-signaling as it also interact with RalGPS2, a Ral-GEF not belonging to the Ral-GDS family[41].

Selectivity of GEFs for the Ras family of G-proteins

In biological systems it is often difficult to distinguish whether observed effects of a GEF on a G-protein are direct or indirect. For example, active Ras causes the activation of Ral by interacting with Ral-GEFs. Thereby Ras-GEFs are placed in the signaling pathways upstream of Ral. The most direct approach to assess the specificity of GEFs is measuring their activity *in vitro* using purified proteins. In Chapter 4 we have investigated the selectivity of Ras-family GEFs. We set to determine selectivity profile of representative CDC25-HD domain containing GEFs towards G-proteins of the Ras family with a focus on Rap- and Ras-proteins.

Even though C3G has been suggested to directly activate R-Ras[42], this is not supported by our measurements. Therefore reported effects of C3G on R-Ras in biological systems are most likely indirect. This effect is not necessarily Rap-dependent as Rap-independent functions of C3G have been reported[43–45]. Similarly, Epac has been suggested to act on R-Ras[46] but we were not able to detect any activity of Epac1 and Epac2 towards R-Ras.

In fact, we found no GEFs which were able to act on R-Ras. This classifies R-Ras together with Rit, Rheb, Di-Ras1 and Di-Ras2 as G-proteins with unknown GEFs. The intrinsic nucleotide exchange rates of these G-proteins are in average 10 fold higher compared to other G-proteins. It could be that these G-proteins are not regulated by GEFs. However, it is intriguing to speculate that so far unrecognized GEFs still exist. Canonical GEFs for G-proteins of the Ras-family contain CDC25-HDs. We have not tested all CDC25-HD containing GEFs, but we have covered CDC25-HDs rather comprehensively by analysis representatives of highly homolog CDC25-HDs. Therefore all known CDC25-HD can be excluded with high certainty. It is unlikely that additional so far unrecognized CDC25-HD containing proteins exist. Genomes were extensively searched for CDC25-HD containing proteins without success in identifying novel proteins[47]. One might thus expect structurally unrelated unknown GEFs.

We performed a mutational analysis to determine the residues that are most critical for the establishment of selectivity of Rap-selective GEFs to G-proteins. Ras is highly homologous to Rap, but not activated by Rap GEFs (with a single exception of a dual Ras- and Rap-selective GEF RasGRP3). We therefore investigated the importance of certain residues by creating Rap to Ras mutants and vice versa. This approach has been used previously[48], but is employed here with much larger number of mutants and tested G-protein/GEF pairs. We identified residues at positions 54, 61, and 70 to have a major effect on the recognition of G-protein by Rap-selective GEFs. The identity of the amino acids at these positions can be used to predict the ability of Rap-GEFs to act on non characterized G-proteins.

Of all other G-proteins Di-Ras1 and Di-Ras2 are closest to but not identical to Rap based on our Rap-consensus defined by position 54, 61 and 70 (Figure 1B, Chapter 4). In agreement with our prediction, Di-Ras1 and Di-Ras2 are not activated by Rap-GEFs. Interestingly, different selections rules apply for the action of GAPs. Rap1GAP acts on Di-Ras1 and Di-Ras2 although not as efficiently as on Rap1[49]. Similarly, on the effector site many RA domains interact with Ras- and with Rap-proteins. Thus activators, inhibitors and effectors of Ras-family G-proteins used different mechanisms to establish

selectivity. This would allow evolution to impinge to a certain extent independently on the interaction with the three main classes of G-protein interacting proteins.

References

1. Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A. & Bos, J. L. Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state. *Nature* **439,** 625–628 (2006).
2. Sondermann, H. *et al.* Structural analysis of autoinhibition in the Ras activator Son of sevenless. *Cell* **119,** 393–405 (2004).
3. Muralidharan, V. *et al.* Solution structure and folding characteristics of the C-terminal SH3 domain of c-Crk-II. *Biochemistry* **45,** 8874–8884 (2006).
4. Kobashigawa, Y. *et al.* Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK. *Nat. Struct. Mol. Biol.* **14,** 503–510 (2007).
5. Sarkar, P., Saleh, T., Tzeng, S.-R., Birge, R. B. & Kalodimos, C. G. Structural basis for regulation of the Crk signaling protein by a proline switch. *Nat. Chem. Biol.* **7,** 51–57 (2011).
6. Matsuda, M. *et al.* Two species of human CRK cDNA encode proteins with distinct biological activities. *Mol. Cell. Biol.* **12,** 3482–3489 (1992).
7. Rosen, M. K. *et al.* Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein. *Nature* **374,** 477–479 (1995).
8. De Jong, R., Ten Hoeve, J., Heisterkamp, N. & Groffen, J. Tyrosine 207 in CRKL is the BCR/ABL phosphorylation site. *Oncogene* **14,** 507–513 (1997).
9. Jankowski, W. *et al.* Domain organization differences explain Bcr-Abl's preference for CrkL over CrkII. *Nat. Chem. Biol.* **8,** 590–596 (2012).
10. Sakai, R. *et al.* A novel signaling molecule, p130, forms stable complexes in vivo with v-Crk and v-Src in a tyrosine phosphorylation-dependent manner. *EMBO J.* **13,** 3748–3756 (1994).
11. Birge, R. B. *et al.* Identification and characterization of a high-affinity interaction between v-Crk and tyrosine-phosphorylated paxillin in CT10-transformed fibroblasts. *Mol. Cell. Biol.* **13,** 4648–4656 (1993).
12. Ribon, V., Hubbell, S., Herrera, R. & Saltiel, A. R. The product of the cbl oncogene forms stable complexes in vivo with endogenous Crk in a tyrosine phosphorylation-dependent manner. *Mol. Cell. Biol.* **16,** 45–52 (1996).
13. Chen, K. *et al.* Interaction between Dab1 and CrkII is promoted by Reelin signaling. *J. Cell. Sci.* **117,** 4527–4536 (2004).
14. Huang, Y. *et al.* Tyrosine phosphorylated Disabled 1 recruits Crk family adapter proteins. *Biochem. Biophys. Res. Commun.* **318,** 204–212 (2004).
15. Beitner-Johnson, D. *et al.* The proto-oncogene product c-Crk associates with insulin receptor substrate-1 and 4PS. Modulation by insulin growth factor-I (IGF) and enhanced IGF-I signaling. *J. Biol. Chem.* **271,** 9287–9290 (1996).
16. Sorokin, A., Reed, E., Nnkemere, N., Dulin, N. O. & Schlessinger, J. Crk protein binds to PDGF receptor and insulin receptor substrate-1 with different modulating effects on PDGF- and insulin-dependent signaling pathways. *Oncogene* **16,** 2425–2434 (1998).
17. Vuori, K., Hirai, H., Aizawa, S. & Ruoslahti, E. Introduction of p130cas signaling complex formation upon integrin-mediated cell adhesion: a role for Src family kinases. *Mol. Cell. Biol.* **16,** 2606–2613 (1996).
18. Ingham, R. J. *et al.* B cell antigen receptor signaling induces the formation of complexes containing the Crk adapter proteins. *Journal of Biological Chemistry* **271,** 32306–32314 (1996).
19. Reedquist, K. A. *et al.* Stimulation through the T cell receptor induces Cbl association with Crk proteins and the guanine nucleotide exchange protein C3G. *J. Biol. Chem.* **271,** 8435–8442 (1996).
20. Ballif, B. A. *et al.* Activation of a Dab1/CrkL/C3G/Rap1 pathway in Reelin-stimulated neurons. *Curr. Biol.* **14,** 606–610 (2004).

21. Honda, H. *et al.* Cardiovascular anomaly, impaired actin bundling and resistance to Src-induced transformation in mice lacking p130Cas. *Nat. Genet.* **19,** 361–365 (1998).
22. Morimoto, C. & Tachibana, K. Beta 1 integrin-mediated signaling in human T cells. *Hum. Cell* **9,** 163–168 (1996).
23. Tamada, M., Sheetz, M. P. & Sawada, Y. Activation of a signaling cascade by cytoskeleton stretch. *Dev. Cell* **7,** 709–718 (2004).
24. Birge, R. B., Kalodimos, C., Inagaki, F. & Tanaka, S. Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. *Cell Commun Signal* **7,** 13 (2009).
25. Mitra, A. & Radha, V. F-actin-binding domain of c-Abl regulates localized phosphorylation of C3G: role of C3G in c-Abl-mediated cell death. *Oncogene* **29,** 4528–4542 (2010).
26. Sheldon, M. *et al.* Scrambler and yotari disrupt the disabled gene and produce a reeler-like phenotype in mice. *Nature* **389,** 730–733 (1997).
27. Howell, B. W., Herrick, T. M. & Cooper, J. A. Reelin-induced tyrosine [corrected] phosphorylation of disabled 1 during neuronal positioning. *Genes Dev.* **13,** 643–648 (1999).
28. Bock, H. H. & Herz, J. Reelin activates SRC family tyrosine kinases in neurons. *Curr. Biol.* **13,** 18–26 (2003).
29. Kuo, G., Arnaud, L., Kronstad-O'Brien, P. & Cooper, J. A. Absence of Fyn and Src causes a reeler-like phenotype. *J. Neurosci.* **25,** 8578–8586 (2005).
30. Sanjay, A. *et al.* Cbl Associates with Pyk2 and Src to Regulate Src Kinase Activity, v 3 Integrin-Mediated Signaling, Cell Adhesion, and Osteoclast Motility. *The Journal of Cell Biology* **152,** 181–196 (2001).
31. Shao, Y., Elly, C. & Liu, Y.-C. Negative regulation of Rap1 activation by the Cbl E3 ubiquitin ligase. *EMBO Rep.* **4,** 425–431 (2003).
32. Zhang, W. Negative Regulation of T Cell Antigen Receptor-mediated Crk-L-C3G Signaling and Cell Adhesion by Cbl-b. *Journal of Biological Chemistry* **278,** 23978–23983 (2003).
33. Boussiotis, V. A., Freeman, G. J., Berezovskaya, A., Barber, D. L. & Nadler, L. M. Maintenance of human T cell anergy: blocking of IL-2 gene transcription by activated Rap1. *Science* **278,** 124–128 (1997).
34. Lee, H., Gaughan, J. P. & Tsygankov, A. Y. c-Cbl facilitates cytoskeletal effects in v-Abl transformed fibroblast through Rac1- and Rap1-mediated signaling. *Int. J. Biochem. Cell Biol.* **40,** 1930–1943 (2008).
35. Kirsch, K. H., Georgescu, M. M. & Hanafusa, H. Direct binding of p130Cas to the guanine nucleotide exchange factor C3G. *Journal of Biological Chemistry* **273,** 25673–25679 (1998).
36. Mitra, A., Kalayarasan, S., Gupta, V. & Radha, V. TC-PTP dephosphorylates the guanine nucleotide exchange factor C3G (RapGEF1) and negatively regulates differentiation of human neuroblastoma cells. *PLoS ONE* **6,** e23681 (2011).
37. Okada, S. & Pessin, J. E. Insulin and epidermal growth factor stimulate a conformational change in Rap1 and dissociation of the CrkII-C3G complex. *J. Biol. Chem.* **272,** 28179–28182 (1997).
38. Okada, S., Matsuda, M., Anafi, M., Pawson, T. & Pessin, J. E. Insulin regulates the dynamic balance between Ras and Rap1 signaling by coordinating the assembly states of the Grb2-SOS and CrkII-C3G complexes. *EMBO J.* **17,** 2554–2565 (1998).
39. Song, C. *et al.* Regulation of a novel human phospholipase C, PLCepsilon, through membrane targeting by Ras. *Science Signalling* **276,** 2752 (2001).
40. Krugmann, S., Williams, R., Stephens, L. & Hawkins, P. T. ARAP3 is a PI3K- and rap-regulated GAP for RhoA. *Curr. Biol.* **14,** 1380–1384 (2004).
41. Ceriani, M. *et al.* Functional analysis of RalGPS2, a murine guanine nucleotide exchange factor for RalA GTPase. *Exp. Cell Res.* **313,** 2293–2307 (2007).

42. Mochizuki, N. *et al.* Crk activation of JNK via C3G and R-Ras. *Journal of Biological Chemistry* **275,** 12667–12671 (2000).
43. Guerrero, C. *et al.* Transformation suppressor activity of C3G is independent of its CDC25-homology domain. *Oncogene* **16,** 613–624 (1998).
44. Shivakrupa, R., Radha, V., Sudhakar, C. & Swarup, G. Physical and functional interaction between Hck tyrosine kinase and guanine nucleotide exchange factor C3G results in apoptosis, which is independent of C3G catalytic domain. *Journal of Biological Chemistry* **278,** 52188–52194 (2003).
45. Gutiérrez-Uzquiza, A. *et al.* C3G down-regulates p38 MAPK activity in response to stress by Rap-1 independent mechanisms: involvement in cell death. *Cell. Signal.* **22,** 533–542 (2010).
46. De Jesus, M. L. Cyclic AMP-dependent and Epac-mediated Activation of R-Ras by G Protein-coupled Receptors Leads to Phospholipase D Stimulation. *Journal of Biological Chemistry* **281,** 21837–21847 (2006).
47. Van Dam, T. J. P., Rehmann, H., Bos, J. L. & Snel, B. Phylogeny of the CDC25 homology domain reveals rapid differentiation of Ras pathways between early animals and fungi. *Cell. Signal.* **21,** 1579–1585 (2009).
48. Van den Berghe, N., Cool, R. H. & Wittinghofer, A. Discriminatory residues in Ras and Rap for guanine nucleotide exchange factor recognition. *J. Biol. Chem.* **274,** 11078–11085 (1999).
49. Gasper, R., Sot, B. & Wittinghofer, A. GTPase activity of Di-Ras proteins is stimulated by Rap1GAP proteins. *Small GTPases* **1,** 133–141 (2010).

Nederlandse samenvatting

In een gezond organisme is een optimale regulatie van de cellulaire processen van kritisch belang. Cellen bevatten veel verschillende eiwitten, die verantwoordelijk zijn voor de meeste cellulaire processen. Bepaalde types eiwitten kunnen verkeren in een actieve of inactieve staat, wat vaak weer gereguleerd wordt door andere eiwitten.

In dit proefschrift hebben wij de regulatie onderzocht van een specifieke groep eiwitten, namelijk GTPases behorende tot de Ras familie. Leden van deze familie hebben een centrale, regulerende functie in veel verschillende cellulaire activiteiten. De meest bestudeerde leden van de Ras familie zijn Ras, Rap en Ral. Ras is betrokken in de regulatie van celgroei en celdeling; Rap reguleert de adhesie tussen cellen en de extracellulaire matrix en de onderlinge adhesie tussen naburige cellen en Ral beïnvloedt het transport van moleculen in de cel. Aangezien de eiwitten uit de Ras familie diverse, belangrijke cellulaire processen reguleren, dient hun activiteit nauwkeurig gecontroleerd te worden. GTPases kunnen verkeren in twee verschillende staten, actief en inactief. Het kan echter voorkomen dat een cel zijn vermogen verliest om de activiteit van een GTPase te reguleren, waardoor bijvoorbeeld het eiwit Ras permanent in een actieve staat verkeert. Dit kan verscheidene problemen veroorzaken in het functioneren van de cel en kan uiteindelijk leiden tot ziektes zoals kanker. De overgang van de inactieve staat naar de actieve staat wordt gereguleerd door ‘guanine nucleotide exchange factors’ (GEFs). GEFs bestaan uit verschillende domeinen, waarbij een specifiek domein in de GEF bindt aan de GTPase en de GTPase hierdoor activeert. Andere domeinen van de GEF zijn betrokken in onder andere de regulatie van de GEF.

Enerzijds reguleren GEFs de activiteit van GTPases, anderzijds wordt hun eigen activiteit ook gereguleerd. Sommige GEFs worden gereguleerd door hun locatie in de cel. Wanneer een GTPase zich bijvoorbeeld op het celmembraan bevindt en de GEF zich in de cel bevindt, zijn ze fysiek gescheiden en kunnen ze hierdoor geen interactie met elkaar aangaan. Om een GTPase te kunnen activeren, moet de GEF naar dezelfde locatie getransporteerd kunnen worden. De locatie en transport van een GEF wordt vaak beïnvloedt door andere moleculen.

Een andere manier om de activiteit van een GEF te reguleren is auto-inhibitie. Hierbij verhindert een specifiek domein van de GEF de interactie tussen GEF en GTPase. Auto-inhibitie kan op verschillende manieren opgeheven worden. Soms kan de interactie van kleine moleculen of eiwitten met het inhiberende domein van de GEF de vorm van de GEF zodanig veranderen, dat de GEF in een actieve staat verkeert. Een andere manier is de chemische modificatie van de GEF. Bepaalde eiwitten kunnen additionele atomen aan de GEF plaatsen, wat de vorm van de GEF omzet van inactief naar actief. Een voorbeeld hiervan zijn kinases, die fosfaatgroepen aan eiwitten toevoegen.

In Hoofdstuk 1 hebben wij een overzicht gegeven van de functie en regulatie van de verschillende GTPases uit de Ras-familie. Ook bespreken wij hier de structuur en regulatie van de bijbehorende GEFs voor deze GTPases.

In Hoofdstuk 2 hebben wij de regulatie van het eiwit C3G onderzocht. C3G is een GEF voor de GTPase Rap. Het domein in C3G dat met Rap kan interacteren, bevindt zich in de

tweede helft van het eiwit. Wij laten zien dat C3G gereguleerd wordt door middel van auto-inhibitie: de eerste helft van C3G voorkomt dat de tweede helft zijn GEF functie uitoefent. Naar analogie met andere GEFs stellen wij voor dat het eerste gedeelte van C3G bindt aan het tweede gedeelte, waardoor het bindingsdomein voor Rap in het tweede deel niet langer bereikbaar is. In dit hoofdstuk hebben wij onderzocht hoe de auto-inhibitie van C3G opgeheven kan worden. Interactie van C3G met het adapter eiwit Crk zorgt voor een gedeeltelijke activatie van C3G. Crk is een reeds bekende bindingspartner van C3G, maar werd verondersteld de lokalisatie van C3G in de cel te reguleren in tegenstelling tot de activiteit. Verder tonen wij dat de kinase Src C3G ook gedeeltelijk activeert door één of meerdere fosfaatgroepen aan C3G te plaatsen. De combinatie van Crk binding en fosforylatie door Src leidt tot de volledige activatie van C3G. Gebaseerd op deze observaties stellen wij voor dat de auto-inhibitie van C3G gemedieerd wordt door twee contactpunten. Deze twee contactpunten worden verbroken door respectievelijk de binding met Crk en de fosforylatie door Src. Alleen wanneer beide contactpunten verbroken zijn, ligt het tweede deel van C3G volledig vrij en kan het zijn GEF functie ongehinderd uitvoeren.

In Hoofdstuk 3 hebben wij het eiwit Rlf, een GEF voor de GTPase Ral, onderzocht. Eén van de domeinen van Rlf is het RA domein. Van gelijkende domeinen in andere eiwitten is aangetoond dat deze de activiteit van het eiwit moduleert. Echter, uit onze proeven blijkt dat het RA domein geen invloed heeft op de activiteit van Rlf. Rlf kan op zichzelf in een actieve staat verkeren; dit in tegenstelling tot C3G dat zich in een geïnhibeerde staat bevindt en andere eiwitten nodig heeft om geactiveerd te worden.

Verder hebben wij in dit hoofdstuk de structuur van Rlf geanalyseerd. Hierbij wordt bepaald hoe de atomen in het eiwit Rlf zijn gearrangeerd. Het deel van Rlf dat verantwoordelijk is voor zijn GEF activiteit, vertoont een grote gelijkenis met corresponderende domeinen in de GEFs Sos1 en Epac2. Een belangrijk verschil in Rlf is een naar buiten uitstekende β-sheet, een unieke staaf-achtige structuur. De top van deze β-sheet faciliteert de binding met eiwitten met een SH3 domein. Uit onze experimenten blijkt dat de β-sheet in Rlf bindt aan de eiwitten PLCγ-1, myosine I en intersectin 2. Deze eiwitten bevatten alle drie een SH3 domein, dat deze interactie faciliteert. De interactie tussen de β-sheet en deze eiwitten lijkt geen invloed te hebben op de activiteit van Rlf.

In Hoofdstuk 4 hebben wij weer een ander aspect van GEFs onderzocht. Elke cel bevat meerdere GEFs en GTPases, die niet allemaal compatibel zijn. In plaats daarvan is een specifieke GEF alleen in staat een selectief aantal GTPases te activeren. Om een uitgebreid overzicht te creëren van de selectiviteitsprofielen van de individuele GEFs voor de GTPases uit de Ras familie, hebben wij een serie metingen op deze GEFs en GTPases uitgevoerd onder gelijke condities. Hierdoor hebben wij vergelijkbare selectiviteitsprofielen gecreëerd, die van nut kunnen zijn om de cellulaire functie van de GEFs en GTPases te analyseren. Ook hebben wij onderzocht welke kenmerken van de GTPases herkend worden door de desbetreffende GEF, zogenoemde selectiviteitsdeterminanten. Gebaseerd op deze informatie geven wij een voorspelling van de selectiviteit van de individuele GEFs voor de GTPases. Deze selectiviteit hebben wij niet experimenteel bevestigd.

In dit proefschrift hebben wij een aantal aspecten in de functie van GEFs onderzocht. De algemene samenvatting en de discussie van onze resultaten staan beschreven in Hoofdstuk 5.

Summary

Optimal regulation of cellular processes is of critical importance for a healthy organism. Cells contain many different proteins that perform most of the cellular functions. Some of these proteins can be active or inactive, and other proteins regulate their activity.

In this thesis we have investigated the regulation of small G-proteins from the Ras-family. Members of the Ras-family have a central function in the control of many cellular activities. The most well studied members of the Ras family are Ras, Rap and Ral. Ras is involved in control of cell growth and division; Rap controls adhesion of cells to neighboring cells or the extracellular matrix and Ral controls the transport of material within cells. As Ras-family members control various important cellular processes, their activity has to be tightly regulated. Ras-family proteins exist in two different states, active and inactive. Sometimes cell loses the ability to regulate Ras-family proteins. For example, it can happen that Ras is always in the active state. This causes various problems in the proper functioning of cells and may lead to diseases such as cancer. The transition from the inactive to the active state of Ras-family proteins is controlled by guanine nucleotide exchange factors (GEFs). GEFs consist of different modules, whereby a specific module of the GEF binds to the small G-protein in order to activate it. Other modules of the GEF might be involved in the regulation of the GEF or might be involved in different functions.

While GEFs are able to regulate the activity of small G-proteins, their activity is regulated as well. Some GEFs are regulated by the control of their cellular localization. For example, if a G-protein is located at the membrane that surrounds the cell and GEF is located in the inside of the cell, they are physically separated and therefore can not interact. In order for a GEF to activate the G-protein the GEF must come to the same location. Other molecules can influence the location of GEF and bring it to the cellular membrane.

Another way to control the activity of GEFs is autoinhibition. That means that one module of the GEF prevents the GEF from interaction with the G-protein. Autoinhibition can be relieved in various ways. Sometimes small molecules or proteins bind to the inhibitory module of the GEF and change its shape, thereby rendering the GEF active. Another way is to chemically modify the GEF. Other proteins can attach additional atoms to the GEF which then change the shape of GEF from inactive to active. Kinases are such proteins, as they attach phosphate groups to proteins.

In Chapter 1 we gave the overview of the function and regulation of Ras-family members. We have further discussed the structure and regulation of GEFs for G-proteins in the Ras-family.

In Chapter 2 we have investigated the regulation of C3G. C3G is a GEF for the small G-protein Rap. The binding place for Rap is located in the second half of C3G. We have demonstrated that C3G is autoinhibited: The first half of C3G prevents the second half from acting as a GEF. We propose in analogy to other GEFs, that that first half of C3G binds to the second half. Thereby the binding place for Rap in the second half of C3G is blocked. We have investigated how the autonihibition of C3G is relieved. We have found

that interaction of C3G with the adaptor protein Crk partially activates C3G. Crk is a known binding partner of C3G, but it was thought that Ck controls the localization of C3G within the cell and not the activity of C3G. We further found that the kinase Src partially activates C3G as well. To do so, Src attaches one or more phosphate groups to C3G. The combination of Crk-binding and treatment with Src results in full activation of C3G. Based on these observations, we proposed that the interaction between the first and the second half of C3G is mediated by two contact points. One contact point can be disrupted by binding of Crk and the other contact point by the action of Src. Only if both contact points are disrupted the second half of C3G is fully liberated and can act as a GEF without being perturbed by the first half.

In Chapter 3 we have investigated the GEF Rlf, which activates the small G-protein Ral. One module of Rlf is referred to as RA domain. Similar parts in other proteins have been shown to modulate the activity of these proteins. We found that the RA domain has no influence on the activity of Rlf. Whereas C3G was autoinhibited and needed other proteins to be activated, Rlf is in the active state on its own.

We have further analyzed the structure Rlf. This means we have determined how the atoms of Rlf are arranged. The part of Rlf, that is responsible for the GEF activity looks largely the same as the corresponding parts the GEFs Sos1 and Epac2. However, one part of Rlf adopts a unique rod-like structure that protrudes outwards, named β-sheet. At its top the β-sheet contains a specific structure that allows proteins, which contain a SH3 domain, to bind to it. Indeed, we have found that the β-sheet binds to the proteins PLCγ-1, myosin I and intersectin2. All three proteins contain SH3 domains, which are responsible for this interaction. The interactions between the β-sheet and these proteins seem not to be involved in the regulation of the activity of Rlf.

Another aspect of GEFs was investigated in Chapter 4. Each cell contains several GEFs and several G-proteins. However, GEFs and G-proteins are not promiscuous. Instead, each GEF can only activate a specific selection of G-proteins. In order to create an extensive overview of the selectivity profiles of GEFs for G-proteins from the Ras-family, we have performed a series of measurements on GEFs and G-proteins under the same conditions. Thereby we have generated comparable selectivity profiles that can be of use when analyzing the function of these GEFs and G-protein in the cell. In addition, we analyzed which features of the G-proteins are recognized by the GEF, the so-called selectivity determinants. Based on this information we predict the selectivity of GEFs for G-proteins which have not been tested experimentally.

We have therefore investigated various aspects of the function of GEFs. The general summary and the discussion of our results are given in Chapter 5.

Curriculum Vitae

Milica Popović was born on February 11th, 1981, in Belgrade, Serbia. From October 2000 she studied biochemistry at the Faculty of Chemistry, University of Belgrade. She received her diploma in December 2005. After graduating, she worked in the laboratory of Prof Tanja Ćirković Veličković until December 2006. From January 2007 to March 2008 she worked at the Faculty of Chemistry, University of Basel.
From May 2008 she has been working as a PhD student under supervision of Prof. Dr. Johannes L. Bos and Prof. Dr. Holger Rehmann at the University Medical Center, Utrecht. The results of her work are described in this thesis.

List of Publications

Stojkovic, M., Uda, N.R., Brodmann, P., **Popovic, M.**, and Hauser, P.C. (2012). Determination of PCR products by CE with contactless conductivity detection. *J Sep Sci* **35**, 3509–3513.

Woller, B., Luiskandl, S., **Popovic, M.**, Prieler, B.E.M., Ikonge, G., Mutzl, M., Rehmann, H., and Herbst, R. (2011). Rin-like, a novel regulator of endocytosis, acts as guanine nucleotide exchange factor for Rab5a and Rab22. *Biochim. Biophys. Acta* **1813**, 1198–1210.

Shirinian, M., **Popovic, M.**, Grabbe, C., Varshney, G., Hugosson, F., Bos, H., Rehmann, H., and Palmer, R.H. (2010). The Rap1 guanine nucleotide exchange factor C3G is required for preservation of larval muscle integrity in Drosophila melanogaster. *PLoS ONE* **5**.

Stanic, D., Radosavljevic, J., Polovic, N., Jadranin, M., **Popovic, M.**, Vuckovic, O., Burazer, L., Jankov, R., and Velickovic, T.C. (2009). Removal of N-terminal peptides from β-lactoglobulin by proteolytic contaminants in a commercial phenol oxidase preparation. *International Dairy Journal* **19**, 746–752.

Blanusa, M., Perovic, I., **Popovic, M.**, Polovic, N., Burazer, L., Milovanovic, M., Gavrovic-Jankulovic, M., Jankov, R., and Cirkovic Velickovic, T. (2007). Quantification of Art v 1 and Act c 1 being major allergens of mugwort pollen and kiwi fruit extracts in mass-units by ion-exchange HPLC-UV method. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **857**, 188–194.

Acknowledgements/ Zahvalnica

I would like to express my gratitude to all people who have helped me to complete my PhD thesis.

Hans, I would like to thank you for the supervision during my PhD research. Since my first visit during my interview I was impressed with the group, and I am very grateful for the opportunity to be a part of it during these five years.

Holger, your deep knowledge, ideas and enthusiasm were always motivating, especially during the first years when none of the GEFs seemed to work. Thank you for your guidance, patience, teaching and great help during every day of my PhD research.

I would further like to thank all present and former members of Bos and Burgering groups for the useful discussions, presentations, group retreats and help in the lab, but also for their friendliness and fun outside of the lab.

On Ying, I have enjoyed our always-improving tennis games and chats. Thank you for the help with my first cell biology steps and for the help with my thesis. Marlous, I liked our trip to Christmas market (even though the car seat could have been warmer). Good luck with your last steps in finishing PhD! Marije, I can not thank you enough for all the times you gave me the hand during these five years!! Thank you for all the help with movings, Dutch translations and of course all the experiments. Nayara, I hope you will settle well in Sao Carlos. I wish you best luck with everything you will do in the future! Anneke, I wish you a lot of success with your proteins. In case you decide to do some of the in vitro experiments, I would be more than glad to help. Sarah C, thank you for the microscopy instructions. I wish you all the best with finishing up your work! Willem-Jan, thank you for always being helpful with my thesis-related and other questions! Marjolein, thank you for your help when I was starting in the lab. Good luck! Fried, thank you for your supportive attitude, and I wish you a lot of success in your future projects! Lucas, the convincing TNT-man, good luck with the brush borders. Rene, Carla, Hugo and Bas, I wish you a lot of success with organoids! Emma, good luck with your thesis! Dear Sarah R, thank you for all the chats and outings. I hope that science in Dundee will go well. Anouk, thank you for the help in the lab as well as outside of the lab. Good luck with all the new things in your life! Rianne, thank you for the gezelligheid and help with purifications and good luck with finishing your studies! Lars and Anna, thank you for your friendliness and help. All the best! Martijn thank you for the help with the cell experiments and good luck with your new projects! Patricia, thank you for your enthusiastic help whenever there was something to organize. Ingrid, it was nice to share the office with you. Best wishes for the future! Jorien, good luck at your new position!

Evi, I am glad to see that your DNA work is paying off. Thank you for all the movies, Greek snacks and jokes and good luck with finishing-up. David, thank you for the

helpfulness with my computer problems and questions. Maaike, I think Rlf is a great protein. Good luck with finishing your thesis! Harmjan, I wish you a lot of success with your peptides and MS! Nive, I enjoyed our time together in Utrecht. Best luck in completing your PhD. Marrit, Hesther, Astrid, Boudewijn, Marieke, Marten, Tobias, Lydia, Cristina, Miranda and Paulien, I wish you all the best in your future!

Dear Wim and IT team, I would like to thank you for always being so helpful with my computer questions.

Furthermore, I would like to thank Prof. Piet Gros for access to the crystallization robot. Michael, Arjen and Arie, thank you very much for your kind help with my crystallization screens.

I am also very grateful to my diploma thesis supervisor, Prof. Tanja Ćirković Veličković, for her help and supervision during and after my undergraduate studies.

Veliku zahvalnost dugujem i svojim prijateljicama Milici, Bojani, Jasni, Bilji, Biljani i Nataši na podršci i prijateljstvu.

Na kraju, želim da se zahvalim Mihajlu, mojim roditeljima Radu i Jelki, mojoj sestri Mariji i bratu Pavlu, na pomoći, ljubavi i podršci.