

# Mechanisms of autoregulation of C3G, activator of the Rap1 GTPase, and its catalytic deregulation in lymphomas

Arturo Carabias<sup>1\*</sup>, María Gómez-Hernández<sup>1</sup>, Sergio de Cima<sup>1</sup>, Antonio Rodríguez-Blázquez<sup>1</sup>, Alba Morán-Vaquero<sup>1</sup>, Patricia González-Sáenz<sup>1</sup>, Carmen Guerrero<sup>1,2</sup>, José M de Pereda<sup>1†</sup>

<sup>1</sup>Centro de Investigación del Cáncer and Instituto de Biología Molecular y Celular del Cáncer, Consejo Superior de Investigaciones Científicas (CSIC)- Universidad de Salamanca, 37007 Salamanca, Spain.

<sup>2</sup>Departamento de Medicina, Facultad de Medicina, Universidad de Salamanca, Instituto de Investigación Biomédica de Salamanca (IBSAL), 37007 Salamanca, Spain.

\*Present address: Structural Molecular Biology Group, Novo Nordisk Foundation Centre for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3-B, DK-2200 Copenhagen N, Denmark.

† Corresponding author: pereda@usal.es

This is the author's version of the work. It is posted here by permission of the AAAS for personal use, not for redistribution. The definitive version was published in **Science Signaling Vol. 13, Issue 647, eabb7075** on 1 Sep 2020, DOI: [10.1126/scisignal.abb7075](https://doi.org/10.1126/scisignal.abb7075).

**One Sentence Summary:** C3G is autoinhibited by an intramolecular interaction that is released by CrkL during activation and it is disrupted by mutations found in lymphomas leading to aberrant Rap1 signaling.

**Abstract:** C3G is a guanine nucleotide exchange factor (GEF) that regulates cell adhesion and migration by activating Rap1 GTPases. Resting C3G has low GEF activity, which is stimulated by Crk adaptor proteins and tyrosine phosphorylation. Here, we describe the mechanisms of C3G autoinhibition and activation. Two intramolecular interactions regulate the activity of C3G. First, an autoinhibitory region (AIR) within the central domain binds to and blocks the catalytic Cdc25H domain. Secondly, binding of the N-terminal domain to the Ras exchanger motif (REM) is required to stimulate the GEF activity. CrkL activates C3G by displacing the AIR/Cdc25HD interaction. Two missense mutations in the AIR (Y554H and M555K) observed in non-Hodgkin lymphomas disrupt the autoinhibition. Expression of C3G-Y554H or C3G-M555K in Ba/F3 cells caused constitutive activation of Rap1 and integrin LFA-1. Sustained Rap1 activation by deregulated C3G might promote progression of lymphomas, suggesting that C3G could be targeted for treating these malignancies.

## Introduction

Small GTPases act as molecular switches that play key roles in signal transduction. They cycle between an inactive conformation, bound to GDP, and an active GTP-bound state that binds to downstream effectors. Spontaneous exchange of the bound GDP by GTP is very slow. This conversion is accelerated by guanine nucleotide exchange factors (GEFs) that induce the dissociation of GDP from the GTPase and the binding of GTP, which is more abundant than GDP in the cells.

C3G (Crk SH3 domain-binding Guanine nucleotide exchange factor, also known as RapGEF1) is a GEF that directly activates the small GTPases of the Ras-family Rap1a and Rap1b (1, 2) and the GTPase of the Rho-family TC10 (3). C3G also activates moderately R-Ras and TC21 (4). C3G is ubiquitously expressed and is required for mouse embryonic development (5). C3G promotes integrin-mediated adhesion through the activation of Rap1 (6). C3G also regulates other cellular functions including migration (7), actin remodeling (8), apoptosis (9-11), proliferation (12), differentiation (13, 14), and exocytosis (15).

The main splice variant of C3G (isoform a, 1077 residues) has a three-segment structure (16) (Fig. 1A). The N-terminal domain (NTD, residues 1-245) mediates the interaction with E-cadherin (17). The central region (residues 246-670) contains five proline-rich motifs (P0 to P4) that are binding sites for proteins containing SH3 domains; hence, this region is named SH3-binding or SH3b. Crk adaptor proteins bind to motifs P1 to P4 (18, 19), p130Cas binds to the P0 (20), SH3-containing tyrosine kinases Hck (10), c-Abl (8, 21) and the oncoprotein Bcr-Abl (22) also bind to the SH3b. The C-terminal catalytic region consists of a Ras Exchanger Motif (REM) and a Cdc25 homology domain (Cdc25HD), the latter harbors the GEF activity.

In resting C3G, non-catalytic regions block the activity of the Cdc25HD preventing uncontrolled signaling (23-25). Other GEFs of the Cdc25 family are also repressed by autoinhibitory mechanisms (26); yet, there is no similarity between the non-catalytic regions of C3G and other GEFs.

C3G-Rap1 signaling is activated by several stimuli including ligation of B- and T-cell receptors (27, 28), growth factors (29), cytokines (30), and mechanical cues (31). In unstimulated cells C3G is associated to Crk proteins in the cytoplasm (27, 32); these adapter proteins (CrkII and CrkL) contain an SH2 and two SH3 (SH3N and SH3C) domains; only the SH3N binds to Pro-rich motifs, while the non-canonical SH3C does not (33). During activation the Crk-C3G complex is recruited to the plasma membrane through the interaction of the SH2 domain of Crk with tyrosine-phosphorylated proteins (25). C3G is then phosphorylated in Tyr<sup>504</sup> of the SH3b domain and possibly other tyrosine residues (23). C3G is substrate of kinases Src (34), Hck (10), Fyn (28), c-Abl (35), and Bcr-Abl (22). In addition, Crk also activates C3G directly in cells (23) and in vitro (24), though the mechanisms of activation remain unknown.

C3G has a dual role in cancer, acting as tumor suppressor or promoter depending on the cellular context. C3G is down-regulated in cervical squamous cell carcinoma (36). On the other hand, C3G is overexpressed in non-small cell lung cancer (37), head and neck squamous cell carcinoma

(38), and in metastatic epithelial ovarian tumors, in which C3G-Rap1 signaling promotes cell invasion (39). The Crk-C3G-Rap1 pathway is involved in the transformation of papillary thyroid carcinoma (40), and alterations in C3G associate with reduced survival of hepatocellular carcinoma patients (41). The oncoprotein Bcr-Abl, which causes chronic myeloid leukemia (CML), induces proliferation, survival, and migration of CML cells through the activation of Crk-C3G-Rap1 (42). CML cells express abundantly a truncated variant of C3G named p87C3G, which lacks the NTD, and the levels of p87C3G protein diminish upon treatment with the Bcr-Abl inhibitor imatinib (22). CrkL and C3G are phosphorylated and recruited to signaling complexes in a Bcr-Abl dependent manner, suggesting that Rap1 activation in CML cells occurs through CrkL-C3G (43). C3G is constitutively phosphorylated in human lymphoblastoid cell lines derived from B-lymphocytes transformed with Epstein-Barr Virus (EBV) (44) and the *RAPGEF1* gene coding for C3G is highly expressed in EBV-transformed lymphocytes (publicly available GTEX portal data). A high-throughput screening in which the random insertion of the *Sleeping Beauty* transposon was used to test the ability of murine Ba/F3 pro-B-cells to grow independently of interleukin-3 (IL-3), found recurrent transposon insertions in the *Rapgef1* gene, suggesting that increased expression of C3G contributes to oncogenic signaling (45).

Despite the widely held view that the GEF activity of C3G is autoregulated, the details of this mechanism remained unknown. Here, we report the identification of two regulatory intramolecular interactions between non-catalytic and catalytic regions: (i) a contact between the NTD and REM domains stimulates the GEF activity, and (ii) an autoinhibitory region within the SH3b domain interacts with the Cdc25HD and blocks its activity. We also show that CrkL activates C3G by displacing this autoinhibitory interaction. We have mapped residues in the autoinhibitory region that are essential for blocking the activity of C3G. Noteworthy, two point mutations detected in hematological tumors target the autoinhibitory interaction and cause constitutive activation of C3G-Rap1 in cells.

## Results

### The NTD of C3G interacts with the REM domain.

Because the non-catalytic regions of C3G might regulate the GEF activity through interactions with the catalytic region, we searched systematically for potential intramolecular contacts. Initially, we analyzed whether the NTD interacts with other regions of C3G. The NTD with a FLAG-tag (FLAG-NTD) was transiently co-expressed in COS-1 cells with HA-tagged C3G, full length or deletion mutants (Fig. 1A), and the interaction was assessed by co-immunoprecipitation (coIP) (Fig. 1B). Full-length C3G did not interact with the NTD, while a construct that lacks the NTD (C3G- $\Delta$ NTD, 246-1077), did coIP with FLAG-NTD. This suggests that in full-length C3G the NTD-binding site is internally masked by an intramolecular interaction, and the site is exposed in the absence of the NTD.

Four constructs carrying increasing deletions of the central SH3b region (277-1077, 292-1077, 465-1077, and 550-1077) and the catalytic region alone (REM-Cdc25HD, 670-1077) also interacted with FLAG-NTD. On the other hand, the isolated SH3b region did not interact with FLAG-NTD. Within the catalytic region, the REM domain associated with the NTD, while the Cdc25HD did not. Two constructs that contain the REM domain and segments of the SH3b domain, P3-P4-REM (465-814) and P4-REM (550-814), also interacted with FLAG-NTD. Finally, none of the HA-tagged fragments coIP when they were co-transfected with empty pCEFLAG vector, indicating that the interaction was specifically mediated by the NTD.

To analyze if the NTD/REM interaction occurs within the same molecule, we created C3G constructs carrying the fluorescent proteins CFP and Venus at the N- and C-terminus, respectively (Fig. 2A). CFP and Venus form a donor-acceptor pair for Förster resonance energy transfer (FRET), and the FRET signal yields information about their proximity. Upon excitation of CFP in COS-1 cells expressing CFP-C3G-Venus (abbreviated C-C3G-V) a peak appeared at ~527 nm, which corresponds to the emission of Venus due to FRET (Fig. 2, B and C). This suggested proximity between the fluorophores at both ends of C3G. The Venus peak was not detected in cells co-expressing the individually labeled constructs C-C3G and C3G-V. Thus, the FRET signal of C-C3G-V is related to an intramolecular event.

We also analyzed deletion mutants of C-C3G-V in a similar manner. The ratios of the fluorescence intensity at the CFP (475 nm) and Venus (527 nm) maxima were used to assess the FRET signal (Fig. 2, B and C). Mutants that lack either the NTD (C-C3G- $\Delta$ NTD-V) or the complete catalytic region (C-C3G- $\Delta$ REM-Cdc25HD-V) showed lower FRET ratios than the full-length construct, suggesting that they adopt open conformations (Fig. 2D). In contrast, deletion of only the Cdc25HD (C-C3G- $\Delta$ Cdc25HD-V) resulted in the highest FRET ratio. This suggests that CFP at the N-terminus is apparently closer to the REM domain than to the C-terminus of the Cdc25HD. Collectively, coIP and FRET data support the notion of a direct intramolecular interaction between the NTD and REM domains, which does not require the Cdc25HD.

To map the NTD-binding site, we generated point mutants in the REM domain and analyzed their interaction with the NTD. To assist in the design and interpretation of the mutagenesis data, we created a structural homology model of the REM domain of C3G (Fig. 2E) using the structure of the REM domain of SOS as template (46). REM domains consist of five  $\alpha$ -helices ( $\alpha$ 2 to  $\alpha$ 6) that form a helical bundle and an additional short helix ( $\alpha$ 1) that packs on one side of the bundle. To sample multiple areas of the REM domain, we generated the single substitutions L714R, E731R, K756E, R764E, L774R, E784R, and K803E, which are spread along the solvent-exposed surface of the REM domain and are not predicted to be occluded by the Cdc25HD (Fig. 2E). Mutations were introduced in the construct HA-P3-P4-REM, which is expressed in higher levels than the isolated REM, and their interaction with FLAG-NTD was analyzed by coIP (Fig. 2F). Only the mutations E731R and E784R prevented the interaction with the NTD. The double mutant E731R/E784R also showed marginal

interaction with FLAG-NTD. We also analyzed the effect of these point mutations on the conformation of C3G by FRET (Fig. 2, G and H). We used the construct C-C3G- $\Delta$ Cdc25HD-V because it had the highest FRET signal (see above) and because the Cdc25HD is dispensable for the NTD/REM interaction. Mutations L714R, K756E, R764E and K803E did not alter the FRET signal. The mutant L774R had slightly, but not significantly, lower FRET ratio. Mutants E731R and E784R had significantly lower FRET ratios, and when combined (E731R/E784R) caused a further reduction of the FRET signal, suggesting an additive effect of these two substitutions. In summary, Glu<sup>731</sup> and Glu<sup>784</sup>, which are adjacent on the surface of the REM domain, define a putative binding site for the NTD around helices  $\alpha$ 3 and  $\alpha$ 5 (Fig. 2E).

We then analyzed the role of the NTD/REM interaction on the regulation of the GEF activity. We measured the release of the fluorescent nucleotide mant-dGDP from Rap1b in the presence of purified C3G fragments (Fig. 2, I and J, and fig. S1). The spontaneous nucleotide exchange of Rap1b:mant-dGDP is extremely slow in these conditions; the apparent dissociation constant ( $k_{\text{obs}}$ ) was  $\sim 0.02 \times 10^{-3} \text{ s}^{-1}$ . Full-length C3G induced a moderate increase in the exchange rate ( $k_{\text{obs}} 0.55 \times 10^{-3} \text{ s}^{-1}$ ), while the activity of the isolated Cdc25HD ( $k_{\text{obs}} 2.8 \times 10^{-3} \text{ s}^{-1}$ ) was 5-times faster. The truncated construct C3G- $\Delta$ NTD had slightly lower activity ( $k_{\text{obs}} 0.41 \times 10^{-3} \text{ s}^{-1}$ ) and the full-length mutant E731R/E784R was significantly less active ( $k_{\text{obs}} 0.15 \times 10^{-3} \text{ s}^{-1}$ ) than wild type C3G. The role of the NTD/REM interaction on the regulation of the activity of C3G was analyzed later in detail (see below). In summary, the GEF activity of C3G is self-repressed and the NTD is dispensable for the autoinhibition.

### **The SH3b region binds directly to the Cdc25HD and blocks the GEF activity.**

Because C3G- $\Delta$ NTD is autoinhibited, we next explored if the SH3b domain might regulate the GEF activity. HA-tagged constructs of the catalytic region (REM-Cdc25HD) or the individual REM and Cdc25H domains were expressed in HEK293T cells and their interaction with purified GST-SH3b (274-646), or GST alone as a control, was analyzed using a pull-down assay (Fig. 3A). HA-REM-Cdc25HD and HA-Cdc25HD interacted with the SH3b, but no binding of HA-REM was observed. Because HA-REM was expressed in very low levels, we also created constructs of these three fragments fused to mEGFP (Fig. 3B). Both REM-Cdc25HD-mEGFP and Cdc25HD-mEGFP bound specifically to GST-SH3b, whereas no specific interaction was observed between REM-mEGFP and GST-SH3b. Finally, purified Cdc25HD produced in bacteria was also pulled down with GST-SH3b (Fig. 3C). In summary, the SH3b domain binds directly to the Cdc25HD.

We used deletion mapping to identify the segment within the SH3b that binds to the Cdc25HD. GST-SH3b fragments were produced in bacteria and their binding to purified Cdc25HD was assessed by pull-down (Fig. 3, C and D). The Cdc25HD did not bind to fragments of the first half of the SH3b (274-371 and 274-500), but it bound to the second half of the SH3b (501-646), which contains the P3 (539-547) and P4 (607-616) motifs. The segment 501-536, which contains Tyr<sup>504</sup>, was

neither required nor sufficient for the interaction, as the Cdc25HD bound to 537-646 but not to 501-536. The final part of the SH3b (579-646), which contains the P4, did not bind to the Cdc25HD and this region was not required for binding, as observed by the binding of the Cdc25HD to the fragments 274-578, 501-578, 537-588, and 537-578. A construct further shortened at the C-terminus, 537-569, also bound to Cdc25HD. Additional trimming, as in the fragment 537-560, reduced the interaction. Similarly, the segment 545-560 also bound weakly. Other two fragments that lack the P3 motif, 545-646 and 545-569, bound to the Cdc25HD, revealing that the P3 is not required for binding. In summary, the minimal segment of the SH3b that binds to the Cdc25HD corresponds to a 25-residue segment (545-569) located immediately downstream of the P3 motif.

Next, we analyzed whether the SH3b/Cdc25HD interaction affects the GEF activity. We measured the exchange activity over Rap1b *in vitro* of the isolated Cdc25HD, alone and in the presence of the GST-SH3b fragments (40  $\mu$ M) that bind to the Cdc25HD, and compared it to the activity of the autoinhibited full-length C3G (Fig. 3, D, E, and F). The complete SH3b (274-646) reduced the activity of the Cdc25HD to the levels of full-length C3G, revealing that the SH3b/Cdc25HD interaction alone is sufficient for the autoinhibition, since it does not require the presence of the NTD or REM domains. The SH3b fragments 372-646, 501-646, 537-646, and 545-646 also blocked efficiently the activity of the Cdc25HD (80-100 % inhibition). On the other hand, fragments that lack the P4 and the C-terminal segment of the SH3b (274-578, 501-578, 537-588, 537-578, 537-569, and 545-569) did not block the GEF activity or they did only partially (20-45 % inhibition). When the fragment 537-646 was produced without any tag, it blocked the Cdc25HD as efficiently as the equivalent GST-tagged construct; excluding any contribution from the GST moiety. A dose-dependent analysis of the effect of the fragment 537-646 on the exchange activity of the Cdc25HD yielded a half maximal inhibitory concentration (IC<sub>50</sub>) of ~13  $\mu$ M (Fig. 3G).

Collectively, our results revealed an autoinhibitory region within the SH3b domain, hereafter AIR (residues 545-646) that contains two functionally distinct segments (Fig. 3 D and H). A first short segment of the AIR (residues 545-569), which is adjacent to the P3 motif, drives the binding to the Cdc25HD; hence we name it Cdc25HD-binding region or AIR-CBR. A second segment located in the final part of the AIR (residues 570-646) is required to repress the GEF activity of the Cdc25HD; therefore we name it AIR-inhibitory tail or AIR-IT.

### **The AIR/Cdc25HD interaction is the main autoinhibitory mechanism of C3G**

Analysis of the sequences of C3G from 209 species (table S1) revealed that the catalytic region is the most conserved segment (Fig. 4A). The SH3b is globally the less conserved domain; yet, the Pro-rich motifs are highly conserved. The segment 545-583, that is adjacent to the P3 and includes the AIR-CBR, is almost as conserved as the catalytic region; suggesting that the regulatory interaction is preserved in C3G from multiple species.

To identify residues of the AIR that contribute to the interaction with the Cdc25HD, we generated point mutants within and near the AIR-CBR, focusing on conserved positions. Initially, we created six single and eight double mutants that collectively probed 22 residues in the 546-581 segment; acidic and non-charged residues were replaced by Arg, and basic residues were changed to Glu. Mutations were introduced in GST-AIR 537-646 and binding to the Cdc25HD was analyzed by pull-down (Fig. 4B). Only the individual change Y554R and the double mutations H550E/M551R and M555R/Q556R strongly reduced binding to the Cdc25HD. The single substitutions M551R and M555R compromised the interaction; while H550E and Q556R did not affect the binding. Replacing Tyr<sup>554</sup> by Ala also reduced severely the interaction, while mutants M551A and M555A showed partial reduction of the binding (Fig. 4C). The AIR is predicted to be mostly disordered (fig. S2), yet an  $\alpha$ -helix is predicted in the AIR-CBR; with Met<sup>551</sup>, Tyr<sup>554</sup>, and Met<sup>555</sup> being on the same side of that helix (Fig. 4D). Collectively, Tyr<sup>554</sup> makes essential contributions to the binding to the Cdc25HD, and it is flanked by Met<sup>551</sup> and Met<sup>555</sup>, suggesting that these three residues form a continuous site that contacts the Cdc25HD.

To assess the contribution of the AIR/Cdc25HD interaction to the regulation of C3G, the point mutations that disrupt this contact were introduced in the full-length protein and the exchange activities of C3G mutants were analyzed (Fig. 4, E and F). C3G mutants M551R, Y554R, and M555R were constitutively in a high-activity state. C3G-Y554R showed the highest exchange activity ( $k_{\text{obs}}$   $20.5 \times 10^{-3} \text{ s}^{-1}$ ) that was ~50-fold higher than that of C3G-WT. C3G-M551R and M555R showed slightly slower exchange rates ( $k_{\text{obs}}$   $12.9 \times 10^{-3} \text{ s}^{-1}$  and  $k_{\text{obs}}$   $10.3 \times 10^{-3} \text{ s}^{-1}$ ) than Y554R, in concordance with their moderate reduction of the AIR/Cdc25HD binding. In summary, disruption of the AIR/Cdc25HD interaction is sufficient to unleash the GEF activity of C3G and suggests that this contact is the main autoinhibitory mechanism (Fig. 4G).

### **Cancer-related missense mutations in the AIR activate C3G constitutively.**

The Catalogue of Somatic Mutations in Cancer database (COSMIC, version 90) describes ~170 missense single nucleotide variants (SNVs) in the *RAPGEF1* gene; 12 of those are in or near the AIR-CBR (table S2). Two mutations found in patients with non-Hodgkin lymphomas, Y554H and M555K (47, 48), affect residues that are critical for the autoinhibition. This prompted us to analyze their impact on the regulation of C3G. Mutation Y554H abrogated the binding of GST-AIR to the Cdc25HD, and M555K partially reduced the interaction (Fig. 4C). When introduced in the full-length C3G, Y554H induced constitutive high GEF activity (Fig. 4, E and F, and fig. S1). M555K also activated C3G, although to a lesser extent.

Next, we analyzed the effect of C3G-activating mutations on the activation of Rap1 in cells. First, WT and point mutants of C3G-mEGFP, or isolated mEGFP, were transiently expressed in HEK293T cells and the levels of Rap1-GTP were measured (Fig. 5, A and B). As a positive control we expressed C3G-WT-mEGFP carrying the CAAX sequence of K-Ras, which is constitutively

located on the plasma membrane causing the activation of Rap1 (23). Expression of C3G-WT, the mutant C3G-E731R/E784R (with reduced NTD/REM interaction), or mEGFP alone did not result in detectable levels of Rap1-GTP. On the other hand, C3G-M551R and the lymphoma mutants C3G-Y554H and C3G-M555K caused a constitutive increase of Rap1-GTP. Rap1 activation induced by C3G and its mutants in HEK293T cells correlated with their activity *in vitro* (Fig. 5C).

Because mutations Y554H and M555K were described in lymphomas, we also analyzed their impact in the pro-B cell line Ba/F3, in which C3G-mEGFP, WT and mutants, or mEGFP alone, were stably expressed by lentiviral infection. Expression of C3G-Y554H or C3G-M555K resulted in constitutively higher levels of Rap1-GTP than when Ba/F3 cells expressed C3G-WT, C3G-E731R/E784R, isolated mEGFP, or when cells did not express heterologous proteins (Fig. 5D). Upon stimulation with IL-3, which is a known activator of Rap1 in hematopoietic cells (30), Rap1-GTP increased in Ba/F3 cells expressing C3G-WT or C3G-Y554H; Rap1-GTP levels being higher at all time points in the presence of C3G-Y554H (Fig. 5, E). Rap1 was also activated by IL-3 in cells expressing mEGFP, although the levels of Rap1-GTP were lower than in the presence of heterologous expression of C3G; in this case, Rap1 may be activated by endogenous C3G or other GEFs. Finally, we analyzed if C3G mutants affected the activation of integrins. Since Rap1 is a potent activator of the leukocyte function-associated antigen 1 (LFA-1) integrin in Ba/F3 cells (49), we analyzed the impact of mutant C3G on its activation. Active LFA-1 was detected in Ba/F3 cells by flow cell cytometry using the antibody mAb24 that recognizes an active conformation of this integrin (Fig. 5F). Cells expressing C3G-Y554H or C3G-M555K had higher basal levels of active LFA-1 than cells expressing C3G-WT, C3G-E731R/E784R, or mEGFP. In summary, the missense SNVs, Y554H and M555K, found in non-Hodgkin lymphoma patients cause catalytic deregulation of C3G, resulting in increased basal stimulation of Rap1 and integrin LFA-1, and strong activation of Rap1 upon stimulation with IL-3.

### **CrkL activates C3G by releasing the autoinhibitory contact**

Binding of Crk proteins to the SH3b and tyrosine phosphorylation of C3G directly stimulate the GEF activity (23, 24). Given that disruption of the AIR/Cdc25HD interaction by point mutations activates C3G, we analyzed whether Crk or Src-mediated phosphorylation may act on the autoinhibitory interaction. First, we assessed the contribution of these two stimuli to the activation of C3G *in vitro* (Fig. 6, A to C). CrkL induced an ~8-fold increase of the exchange activity of full-length C3G. Phosphorylation of C3G with the kinase domain of Src (SrcKD) caused a ~4-fold increase of the exchange rates with respect to unphosphorylated C3G. Adding CrkL to phospho-C3G (p-C3G) induced the highest activity, ~30-fold increase, suggesting that these two stimuli are independent and additive. Neither CrkL nor phosphorylation altered the activity of the isolated Cdc25HD, supporting the notion that these stimuli act on the autoinhibitory mechanism, not directly on the catalytic domain.



The activity of p-C3G with CrkL was similar to that of the constitutively active mutants C3G-Y554R and Y554H. The activity of C3G-Y554R was not increased by CrkL, phosphorylation, or their combination, and those stimuli only produced a minor increase in the activity of C3G-Y554H (~1.2 to ~1.6-fold) (Fig. 6C and fig. S3, A to D). On the other hand, the partially active mutant C3G-M551R was stimulated by CrkL or phosphorylation to similar levels as C3G-Y554R/H, and the combination of CrkL and phosphorylation did not activate C3G-M551R any further than the individual stimuli. Collectively, our data suggest that CrkL and phosphorylation by Src stimulate C3G by acting on the autoinhibitory mechanism.

We analyzed the impact of CrkL on the AIR/Cdc25HD interaction using independently purified AIR and Cdc25HD. CrkL competed in a dose-dependent manner with the Cdc25HD for binding to the AIR construct 537-646 that contains the P3 and P4 motifs (Fig. 6D). To analyze the independent contribution of CrkL-binding to the P3 and P4, we used two constructs of the AIR each containing only one of these motifs. The P4 was inactivated by introducing simultaneously five point mutations P608A, P609A, L611A, P612A, and K614A (hereafter P4A). CrkL bound to GST-537-646-P4A and displaced the binding to the Cdc25HD with a similar efficiency as for GST-537-646-WT. CrkL also bound to GST-545-646-WT, which only contains the P4; but CrkL binding only caused a minor displacement of the Cdc25HD. Finally, CrkL did not bind to GST-545-646-P4A and consequently did not displace Cdc25HD from binding to this construct.

We also analyzed the effect of the phosphorylation on the AIR/Cdc25HD interaction. The untagged AIR (537-646) was directly phosphorylated by SrcKD (Fig. 6E). The Cdc25HD bound to phospho-GST-AIR (537-646) similarly as to the unphosphorylated protein (Fig. 6F).

Next, we analyzed the effect of CrkL and SrcKD phosphorylation on the ability of the isolated AIR to block the Cdc25HD activity (Fig. 6, G and H and fig. S3E). Binding of CrkL to the AIR 537-646 prevented its inhibition of the Cdc25HD. CrkL also precluded the inhibition of Cdc25HD by the P3-only mutant AIR 537-646-P4A. On the other hand, when CrkL bound only to the P4 in the fragment AIR 545-646, this complex still blocked the activity of the Cdc25HD, although slightly less efficiently than 537-646-P4A alone. Finally, the phosphorylated AIR blocked the activity of the Cdc25HD, albeit slightly less efficiently than unphosphorylated AIR (Fig. 6, G and H).

In summary, CrkL activates C3G by releasing the AIR/Cdc25HD autoinhibitory interaction; while the P3 is not required for the autoinhibition, binding of CrkL to the P3 apparently disrupts the interaction between the adjacent AIR-CBR and the Cdc25HD.

### **The NTD/REM interaction is required for the activation of C3G.**

The isolated Cdc25HD has lower GEF activity than full-length C3G activated by point mutations (for example Y554H) or by the combined stimuli of CrkL and phosphorylation, suggesting that the activity is favored by other regions outside the Cdc25HD. In other GEFs, the REM domain makes close contact with the Cdc25HD, suggesting that in C3G the REM domain could stabilize the Cdc25HD and

favor its activity. Unfortunately, the activity of the region REM-Cdc25HD of C3G could not be analyzed directly because this construct is unstable. Yet, we had observed that autoinhibited full-length C3G has higher residual activity than the C3G-E731R/E784R mutant (see Fig. 2J). This prompted us to analyze in detail the role of the NTD/REM interaction on the regulation of C3G.

We compared the activity of full-length C3G and five constructs in which the NTD/REM interaction was absent or disturbed: C3G- $\Delta$ NTD, C3G-454-1077, C3G-530-1077, and full-length C3G-E731R/E784R (Fig. 7A). We also analyzed C3G-454-1077-E731R/E784R to assess the direct effect of these mutations in the absence of the NTD. Before stimulation, all constructs had low GEF activity; yet, full-length C3G was slightly but significantly more active than the other proteins (Fig. 7B and fig. S4A). Three independent purifications of full-length C3G were used, discarding that the differences could be due to variability between protein batches. CrkL (10  $\mu$ M) activated all these constructs, but those with the disrupted NTD/REM interaction reached slower dissociation rates ( $k_{\text{obs}}$  between 1.1 and 2.2  $\times 10^{-3}$  s $^{-1}$ ) than full-length C3G WT ( $k_{\text{obs}}$  4.3  $\times 10^{-3}$  s $^{-1}$ ) (Fig. 7C and fig. S4B). All the constructs were phosphorylated by the SrcKD to apparently similar levels (fig. S4C); yet, p-C3G full-length WT showed higher activity than the other proteins (Fig. 7D and fig. S4D). Similarly, when CrkL was added to the phosphorylated proteins, C3G-WT was more active than all the other proteins (Fig. 7E and fig. S4E).

We also analyzed if the NTD/REM interaction favors the GEF activity of C3G in cells. mEGFP-tagged C3G-WT, C3G-E731R/E784R, or C3G-WT-CAAX, were expressed in HEK293T cells, either alone or in combination with heterologous expression of CrkL. The latter induces activation of C3G in cells (25). Expression of C3G-WT or C3G-E731R/E784R did not increase the basal levels of Rap1-GTP in HEK293T (Fig. 5A and Fig. 7F) nor in Ba/F3 cells (Fig. 5D). CrkL activated Rap1 in HEK293T cells expressing C3G-WT but not in those with C3G-E731R/E784R. In summary, the NTD/REM interaction indirectly enhances the activity of the Cdc25HD and favors the activation of C3G by CrkL and Src phosphorylation (Fig. 7G).

## Discussion

The direct interaction of the AIR with the Cdc25HD is the main mechanism that represses the GEF activity of unstimulated C3G. This interaction is driven by the AIR-CBR (residues 545-569). Residues Met<sup>551</sup>, Tyr<sup>554</sup>, and Met<sup>555</sup>, are predicted to form a single continuous site that interacts with the Cdc25HD. Because the AIR-CBR alone does not block the GEF activity, most likely it contacts the Cdc25HD outside the catalytic GTPase-binding site. Inhibition requires the AIR-IT (570-646), which is predicted to be disordered. Once that the AIR-CBR binds to the Cdc25HD, the AIR-IT is likely to interact, albeit weakly, with the catalytic binding site, blocking the access of the GTPase. This bipartite interaction of the AIR with the Cdc25HD resembles a lock-lid system in which the AIR-CBR acts as the lock and the AIR-IT as a lid that blocks the catalytic site. The mechanism of autoinhibition of C3G has similarities with that of RasGRP1, a GEF of Ras. In RasGRP1, the EF domain binds to a

surface on the Cdc25HD that is adjacent to the catalytic Ras-binding site, and the linker that connects the Cdc25HD and EF domain blocks the Ras-binding site (50).

The AIR-CBR is a key regulatory site for the activation of C3G. Mutations that disrupt the AIR-CBR/Cdc25HD interaction, such as those described in non-Hodgkin lymphomas, are sufficient to activate C3G to a similar extent as when C3G WT is stimulated by CrkL and Src-mediated phosphorylation. The deregulated active mutants (Y554H and Y554R) are not further stimulated by CrkL and phosphorylation, suggesting that these stimuli operate through the release of the AIR-CBR/Cdc25HD interaction. Binding of CrkL to the P3, which is juxtaposed to the AIR-CBR, displaces the interaction of the AIR with the Cdc25HD. On the other hand, Src-phosphorylation of the isolated AIR did not reduce noticeably its binding to Cdc25HD in pull-down experiments. Yet, phospho-AIR blocked the GEF activity slightly less efficiently than unphosphorylated AIR (Fig. 6G) and phosphorylation of full-length C3G by Src increased the GEF activity, albeit to a lesser extent than the stimulation by CrkL, suggesting that phosphorylation releases at least partially the autoinhibition. It is possible that phosphorylation of Tyr<sup>504</sup> or other tyrosine residues might disrupt, directly or indirectly, the blockage of the Cdc25HD activity by the AIR-IT, while having a minor effect on the binding of the AIR-CBR to the Cdc25HD. This mechanism is compatible with the apparently independent effects of CrkL-binding and Src-phosphorylation on the activation. In addition, phosphorylation of C3G could stabilize the interaction with CrkL.

The key role of Tyr<sup>554</sup> in the autoinhibition apparently contrasts with the moderate activation caused by tyrosine phosphorylation of C3G. Tyr<sup>554</sup> is within a sequence (551-MLAYMQL-557) that does not match the favored motifs of Src substrates, which have strong preference for Ile/Leu/Val preceding (position -1) the target Tyr residue, and moderate preference for Glu in positions -3 and -2, Gly/Glu in position +1, and Phe/Ile/Leu in position +3 (51-53). In addition, phosphorylation of Tyr<sup>561</sup>, Tyr<sup>570</sup>, and Tyr<sup>579</sup>, but not of Tyr<sup>554</sup> has been reported (publicly available data at Phosphosite.org portal). Collectively, these data suggest that Tyr<sup>554</sup> is not a substrate for Src, which is in agreement with the negligible impact of Src phosphorylation on the AIR/Cdc25HD interaction.

The isolated Cdc25HD has lower GEF activity than activated C3G full-length (Fig. 6C), revealing that activation requires the release of the autoinhibition and the concurrence of positive signals. The presence of the REM domain is not sufficient to fully activate the Cdc25HD. Instead, the REM domain acts as a mediator that channels the activation boost caused by the NTD/REM interaction. The NTD interacts with helices  $\alpha$ 3 and  $\alpha$ 5 of the REM domain, which are not part of the predicted Cdc25HD-interaction surface, and the NTD/REM interaction does not require the Cdc25HD. Thus, we propose that the NTD enhances the GEF activity by inducing a high-activity conformation of the REM-Cdc25HD tandem. Such allosteric stimulation of C3G is reminiscent of the positive-feedback activation of SOS, where binding of a Ras-GTP molecule to a regulatory site, at the rim of the REM/Cdc25HD interaction surface, induces a high-activity conformation of the Cdc25HD (54). Yet, the NTD-interaction site in the REM of C3G is different from the Ras-GTP allosteric site of SOS.

Despite no other interactors of the REM domain of C3G have been described so far, binding of other molecules to the REM domain could also modulate the GEF activity, suggesting that C3G might be regulated by feedback mechanisms as observed in other GEFs (26). Collectively, C3G works as a “dimmer switch” that can adopt multiple activity levels, which allows fine-tuning of its signaling.

Rap signaling contributes to the progression of several types of lymphomas and leukemias. Rap activation in the murine B-lymphoma cell line A20 increases their dissemination *in vivo*, favoring their adhesion to endothelial cells and their transendothelial migration (55). Constitutive activation of Rap proteins due to high expression of CD38 in human chronic lymphocytic leukemia (CLL) B-cells stimulates cell adhesion and migration (56). Mice deficient in the Rap1 GTPase-activating protein SIPA1 (also known as SPA-1), which have increased levels of Rap1-GTP, develop disorders that resemble human CML (57) and CLL (58). Transplantation in mice of hematopoietic progenitor cells with constitutively activated Rap1, by the expression of a membrane-targeted engineered form of C3G, causes T-cell acute lymphoblastic leukemia (59, 60). CML cell lines and primary cells from CML patients overexpress the variant p87C3G that lacks the NTD, and the CML-causing oncoprotein Bcr-Abl signals through p87C3G (22). Since activation of p87C3G is weakened by the lack of the NTD/REM interaction, CML cells might overexpress p87C3G to compensate for its attenuated signaling. We have also shown that two missense mutations in C3G detected in non-Hodgkin lymphomas disrupt the autoinhibitory mechanism and cause constitutive activation of C3G and Rap1, suggesting that mutant C3G may favor the progression of those lymphomas. Expression of deregulated C3G mutants also caused increased activation of the integrin LFA-1 in Ba/F3. Given the crucial role of LFA-1 in trafficking and homing of leukocytes and in lymphoma dissemination (61, 62), abnormal C3G-induced activation of LFA-1 might contribute to the dissemination of B-cell lymphomas. This is the first description of a catalytic deregulation of C3G linked to a disease. Further work is required to fully understand the impact of C3G alterations in lymphomas and other type of cancers. The detailed characterization of the regulatory mechanisms presented here will pave the way to identify other potentially activating mutations in C3G and to understand their contribution to diseases. This opens to a potential use of C3G mutations as prognostic markers, as well as to the consideration of C3G as a novel therapeutic target.

## **Materials and methods**

### **DNA constructs and mutagenesis**

All constructs of human C3G refer to the isoform a (Uniprot Q13905-1). Constructs for expression in mammalian cells are listed in table S3. Constructs with an N-terminal HA-tag or FLAG-tag, were created in the pCEFLHA and pCEFLAG vectors (63), respectively. cDNA fragments were amplified by the polymerase chain reaction (PCR) using primers that added BglIII and NotI sites at the 5' and 3' ends, respectively, and were cloned into those sites in the vectors. Primers used to create constructs of

C3G are shown in table S4. Correctness of these and all other constructs was confirmed by DNA sequencing.

Constructs of C3G FRET sensors were created in the vector pcDNA3-CFP-Venus that codes for N-terminal CFP (between HindIII and BamHI sites) and C-terminal Venus (between NotI and XhoI sites) tags. The cDNA fragments of C3G were amplified by PCR using primers that add BglII and NotI sites at the 5' and 3' ends, respectively, and do not contain stop codon. The PCR products were cloned in between the BamHI and NotI sites of the vector. C3G-Venus was created in the vector pcDNA3-Venus-N in a similar manner; in this case the C3G cDNA was amplified using a direct primer that added a BglII site and a Kozak sequence. The construct CFP-C3G was created in the vector pcDNA3-CFP-C; in this case the reverse primer used to amplify C3G included a stop codon.

Constructs of C3G with a C-terminal mEGFP tag for transient expression were created in the vector pEF1-mEGFP, which is a modification of pEF1/V5-HisA (Invitrogen), in a similar manner as the pcDNA3 constructs. C3G cDNA fragments were amplified by PCR that added BglII and NotI sites at the 5' and 3' ends, respectively, and were cloned into pEF1-mEGFP digested with BamHI and NotI.

Lentiviral constructs with C3G cDNA were created in a modified version of the pLenti-C-HA-IRES-BSD vector (Origene), which codes for a C-terminal mEGFP tag. To create this vector, hereafter pLenti-C-mEGFP-IRES-BSD, the region coding for the Kozak sequence and mEGFP tag of pEF1-mEGFP was PCR-amplified using primers that added BglII and PspOMI sites at the 5' and 3' ends, respectively. The PCR product was digested with BglII and PspOMI and was ligated into the compatible cohesive ends produced by digestion of pLenti-C-HA-IRES-BSD with BamHI and NotI, which destroyed these restriction sites upon ligation. The resulting pLenti-C-mEGFP-IRES-BSD vector contains new BamHI and NotI sites upstream the mEGFP cDNA. Next, the cDNA of C3G WT and mutants were amplified by PCR adding BamHI and NotI sites at the 5' and 3' ends, respectively, and were ligated into the pLenti-C-mEGFP-IRES-BSD vector using the same sites.

C3G constructs for expression in bacteria are listed in table S5. The cDNAs of C3G full-length, C3G- $\Delta$ NTD, C3G-454-1077, and C3G-530-1077 were PCR-amplified using primers that added 5'-NcoI and 3'-XhoI sites. The PCR products were cloned using NcoI/XhoI in a modified pGEX-4T3 vector (pGEX-4T3-2xTEV-cHis) that codes for an N-terminal GST followed by a sequence recognized by the tobacco etch virus protease (TEV site) and a C-terminal TEV site and a 6xHis tag. The cDNA of the Cdc25HD was amplified by PCR adding NcoI and BamHI sites, and was inserted in a variant of the pETEV15b vector (64) with unique NcoI and BamHI sites for cloning.

The GST-fusion constructs of the SH3b and its fragments were created in the vector pGEX-4T3-TEV that codes for a TEV site downstream the GST and has NdeI, NcoI, and BamHI sites for cloning. The cDNA of the SH3b fragments 274-646, 372-646, 501-646, and 579-646, were PCR-amplified using primers that added NcoI and BamHI sites, which were used for ligation in pGEX-4T3-TEV. cDNA coding for fragments 537-588, 537-646, and 545-646, were PCR-amplified with NdeI and BamHI sites at the 5' and 3' ends, and were cloned in pGEX-4T3-TEV using these sites. The

cDNA of the fragment 537-646 was also cloned in pETEV15b using the same restriction sites. Constructs coding for fragments 274-371, 274-500, 274-578, 501-536, 501-578, 537-560, 537-569, 537-578, 545-560, and 545-569, in pGEX-4T3-TEV were created by introducing a stop codon after the desired positions of constructs coding longer fragments. Stop codons and other point mutations in C3G constructs were introduced by PCR using the QuikChange method. Oligonucleotides for site directed mutagenesis are shown in table S6.

The cDNA coding for residues 1-167 of human Rap1b was amplified from the IMAGE clone 2900837 by PCR. Primers added NdeI and BamHI sites at the 5' and 3' ends, respectively. The amplified DNA was cloned in between the NdeI and BamHI sites of the pETEV15b vector. The cDNA of human CrkL was amplified by PCR, adding NdeI and BglII sites at the 5' and 3' ends, respectively, and was cloned into the NdeI and BamHI sites of pETEV15b. The cDNAs of the kinase domain of human Src (SrcKD, residues 254-536) and the phosphatase YopH were cloned respectively in the first and second multicloning regions of a modified pETDuet-1 vector that codes for an N-terminal His-tag and TEV site in the first expression region. The cDNA of YopH was PCR-amplified using as template the plasmid pEF5HA-YopH (a kind gift from Andrés Alonso) (65); the primers added NdeI and XhoI sites at the 5' and 3' ends, respectively, that were used for cloning to create pETDuet-1-YopH. The cDNA of the SrcKD was amplified by PCR from the plasmid pcDNA3-c-SRC (a gift from Robert Lefkowitz, Addgene #42202) (66); the primers added EcoRI and HindIII sites at the 5' and 3' ends, respectively, and the DNA was ligated into pETDuet-1-YopH using these sites to create pETDuet-1-SrcKD-YopH. Oligonucleotides for Rap1b, CrkL, Src, and YopH are shown in table S7.

### **Expression and purification of recombinant proteins**

All proteins were produced in *E. coli* strain BL21(DE3) grown in Terrific Broth medium and expression was induced with 0.2 mM IPTG at 15 °C overnight.

Proteins C3G full-length, C3G- $\Delta$ NTD, C3G-454-1077, and C3G-530-1077, which carry GST- and His-tags, were purified by two-step affinity chromatography. In a first step they were purified by immobilized metal affinity chromatography (IMAC) using 5 ml His-Trap Ni<sup>2+</sup>-chelating columns (GE Healthcare) as described (67). Subsequently, the sample was loaded in a 6 ml glutathione-agarose column (ABT) and was washed with buffer A (20 mM Tris-HCl, 300 mM NaCl, pH 7.5) first supplemented with 0.1% Triton X-100 and secondly without detergent. The GST-C3G-His proteins were eluted with 20 mM reduced glutathione in buffer A. The GST- and His-tags were removed by digestion with recombinant TEV protease (rTEV) for 3 hours at room temperature and the sample was dialyzed against buffer A. Free GST was removed by a reverse glutathione-affinity chromatography. C3G proteins eluted in the flow-through, were concentrated, and further purified by size exclusion chromatography (SEC) in a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in buffer

A. Proteins were concentrated by ultrafiltration, flash frozen in liquid nitrogen and stored at -80 °C until they were used.

His-tagged Cdc25HD was purified by IMAC in a His-Trap column as above with minor changes. Bacteria were lysed in 20 mM HEPES, 500 mM NaCl, 10 mM imidazole, pH 7.0 (buffer H). In the IMAC, after loading the sample, the column was washed with buffer H and the Cdc25HD was eluted with 500 mM imidazole in the same buffer. Immediately after the IMAC, the buffer was changed to 20 mM citrate, 250 mM NaCl, pH 6.0 (buffer C) using a Sephadex G25 desalting column. The His-tag was removed by digestion with rTEV and the Cdc25HD was further purified by SEC in a Superdex 200 10/300 GL column equilibrated in buffer C.

GST-SH3b fragments and GST-RalGDS Ras-binding domain (RBD, residues 701-851) were purified by affinity chromatography using glutathione-agarose columns (ABT). Proteins were eluted with phosphate-buffered saline (PBS) supplemented with 20 mM reduced glutathione and were extensively dialyzed against buffer B (20 mM Tris-HCl, 150 mM NaCl, pH 7.5).

His-tagged C3G-537-646 and CrkL were purified by IMAC and the His-tag removed by rTEV digestion as described (67). Rap1b (1-167) was produced by co-expression with the GroEL/GroES chaperones (plasmid pBB541, a gift from Bernd Bukau, Addgene #27394) (68). Rap1b was purified by IMAC, the His-tag was cleaved with rTEV, and the protein was subjected to SEC in a Superdex 200 10/300 GL column equilibrated in in buffer B with 10 mM MgCl<sub>2</sub>. SrcKD was produced in *E. coli* by co-expression with the phosphatase YopH as described (69).

### **Western Blotting**

Proteins were separated by Tris-Glycine SDS-PAGE and transferred onto Immobilon-P membranes (Millipore). The following primary antibodies were used at 1/1000 dilution unless otherwise indicated: mouse monoclonal antibody (mAb) against HA (HA11, Covance), rabbit polyclonal antibody (pAb) against HA Y11 (sc-805, Santa Cruz Biotechnology), mouse mAb M2 against FLAG (F1804, Sigma-Aldrich), mouse mAb against polyHis (H1029, Sigma-Aldrich), mouse mAb G9 against C3G (sc-393836, Santa Cruz Biotechnology), mouse mAb B-2 against GFP (sc-9996, Santa Cruz Biotechnology), mouse mAb 4G10 against pTyr (05-321, Millipore), rabbit pAb 121 against Rap1 (sc-65, Santa Cruz Biotechnology), mouse mAb against  $\beta$ -actin (AC-15, dilution 1/2000, Sigma-Aldrich).

The following secondary antibodies were used (1/5000 dilution): horseradish peroxidase (HRP)-conjugated goat antibody against rabbit IgG (sc-2004; Santa Cruz Biotechnology), HRP-linked sheep antibody against mouse IgG (NXA931, GE Healthcare), goat pAb DyLight 680 against mouse IgG, goat pAb DyLight 680 against rabbit IgG, goat pAb DyLight 800 against mouse IgG, and goat pAb DyLight 800 against rabbit IgG (Thermo Fisher Scientific). Signal of DyLight secondary antibodies were detected by infra-red fluorescence using an Odyssey Imaging System (LI-COR Biosciences) and were quantified using the Odyssey Application Software. HRP-conjugated secondary antibodies were detected by enhanced chemiluminescence (ECL) using Clarity Western ECL

Substrate (Bio-Rad) and Super RX-N films (Fujifilm); films were scanned using an Epson VT700-PHOTO scanner and the intensities of the bands were quantified using the program ImageJ (70).

### **Cell lines.**

HEK293T cells and COS-1 cells were grown in Dulbecco's modified Eagle's medium (DMEM; Sigma-Aldrich) with 10% fetal bovine serum (FBS; Life Technologies), 100 U/ml penicillin, and 100 µg/ml streptomycin (Life Technologies), at 37 °C 5% CO<sub>2</sub>. Cells were transfected with plasmids using polyethylenimine (PEI; Polysciences Inc) (PEI:DNA ratio 2.5:1). IL-3-dependent mouse Ba/F3 cells were grown in RPMI 1640 medium (Sigma-Aldrich) with 10% FBS, 100 U/ml penicillin, 100 µg/ml streptomycin, supplemented with 20% conditioned medium of the WEHI-3B cell line as a source of IL-3.

### **Co-immunoprecipitation (coIP) assays**

COS-1 cells were transfected with HA-C3G constructs and FLAG-NTD or the empty pGEMFLAG vector. After 36 to 48 hours, cells were lysed in lysis buffer consisting of 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5 % Triton X-100, 1 mM sodium orthovanadate, 25 mM NaF, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 1x protease inhibitor cocktail (cOmplete, Roche). Debris was removed by centrifugation (16000 xg, 15 min, 4 °C) and the supernatants were incubated with monoclonal anti-FLAG M2 agarose (Sigma) at 4 °C for 1 hour. The beads were washed three times with lysis buffer. Bound proteins were extracted in sample buffer for electrophoresis and boiled. Samples of the cell lysates and the IPs were analyzed by western blot.

### **Affinity pull-down (PD) assays**

Purified GST-fusion proteins produced in *E. coli* were used as baits in PD assays against proteins either expressed in HEK293T cells, or produced in bacteria. In the first case, HEK293T cells were transfected with plasmids expressing HA- or mEGFP-tagged C3G constructs and 24-48 hours later were lysed as above. Samples were clarified by centrifugation (16000 xg, 15 min, 4 °C) and the total protein content was estimated by Bradford assay. Typically, lysates containing 1-2 mg of total protein were incubated with 30 µg of GST-SH3b fragments, 0.01% BSA, and 20 µl of glutathione-agarose resin at 4 °C for 20 min. The resin was washed four times with buffer B with 0.1% Triton X-100 and proteins were released by adding Laemmli sample buffer. HA- and GFP-tagged proteins were detected in cell lysates and PD fractions by western blot with specific antibodies; the GST-SH3b proteins in the PD fraction were detected by Coomassie staining of the gels. PDs of proteins produced in bacteria were done in a similar manner. In this case ~30 µg of recombinant proteins were incubated in the presence of 0.01% BSA with 30 µg of the GST-fusion proteins pre-bound to 20 µl of glutathione-agarose resin. In competition experiments increasing amounts of CrkL were added to the reaction samples. Incubation, washing, and release of bound proteins were done as above. The PD fractions



were analyzed by SDS-PAGE and the proteins were detected either by Coomassie staining (for example CrkL) or by western blot.

### **Fluorescence FRET measurements**

Fluorescence of C3G conformational FRET biosensors was measured in COS-1 cells transiently expressing fluorescently-tagged C3G constructs. Cells were trypsinized 48 hours after transfection, were collected by centrifugation, and were suspended in PBS pre-warmed at 37 °C. Fluorescence was measured in a Fluoromax-3 fluorometer (HORIBA Jobin Yvon) at 37 °C using a 5x5 mm quartz cuvette. Samples were excited at 435 nm and emission was acquired from 450 to 600 nm. The FRET signal was quantified as the ratio between the fluorescence intensities at 527 nm and 475 nm, which are the emission maxima of Venus and CFP, respectively. For representation the spectra were normalized to the intensity at 475 nm.

### **Protein phosphorylation with Src**

For in vitro phosphorylation, C3G proteins were prepared at 20 to 100  $\mu$ M in kinase buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT, 1 mM ATP). Recombinant SrcKD was added at 1  $\mu$ M and the reaction was carried out at 30 °C for 20 min unless otherwise specified. Phosphorylation was analyzed by western blot against pTyr.

### **Nucleotide exchange activity assays**

The nucleotide exchange activity of C3G toward Rap1b was determined in vitro using the GDP fluorescent analogue 2'-deoxy-3'-O-(N'-methylanthraniloyl)guanosine-5'-O-diphosphate (mant-dGDP) (Biolog), as described (71). The fluorescence intensity of mant-dGDP bound to the GTPase is approximately twice the intensity of the free form in aqueous solution. Rap1b (200  $\mu$ M) was loaded with mant-dGDP by incubation in 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, 4 mM ethylenediaminetetraacetic acid (EDTA) with a 10-molar excess of mant-dGDP for 90 min at 4 °C. Afterwards, 10 mM MgCl<sub>2</sub> was added and the protein was separated from the excess of mant-dGDP by SEC using a Superdex 200 10/300 GL column equilibrated in buffer G (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM MgCl<sub>2</sub>). Exchange reactions were done at 25 °C using 200 nM Rap1b:mant-dGDP in buffer G, and were initiated by adding 40  $\mu$ M GDP. Unless otherwise indicated, experiments were done in the presence of 1  $\mu$ M of C3G proteins, which is within a range where the  $k_{obs}$  are directly proportional to the concentration of C3G (fig. S1). Fluorescence was measured in a Fluoromax-3 fluorometer (HORIBA Jobin Yvon) using 3x3 mm quartz cuvettes. The dissociation of mant-dGDP was monitored exciting at 370 nm (1 nm bandwidth) and registering the emission at 430 nm (10 nm bandwidth). The sample was illuminated only during data acquisition (typically 600 data points per time course) to prevent photobleaching. Data were recorded until the signal reached a plateau. The

time-dependent decay of the fluorescence intensity could be described by the following single exponential decay model:

$$I_t = A_0 e^{-k_{obs}t} + B$$

Where  $I_t$  is the fluorescence intensity at each time point,  $B$  is the fluorescence of free mant-dGDP,  $A_0+B$  is the initial intensity before the reaction is started, and  $k_{obs}$  is the apparent dissociation rate constant. Fitting was done using the program SigmaPlot. Data were normalized to  $A_0+B = 1$  and  $B = 0$  for representation.

### Analysis of dose response experiments

Data of the inhibition of the Cdc25HD by multiple concentrations of the AIR fragment were analyzed fitting the following logistic Hill equation:

$$k_{[AIR]} = k_{min} + \frac{(k_{max} - k_{min})}{1 + \left( \frac{IC_{50}}{[AIR]} \right)^{HillSlope}}$$

Where,  $k_{[AIR]}$  is the  $k_{obs}$  determined at a given concentration of the AIR;  $k_{max}$  and  $k_{min}$  are the maximum and minimum values of  $k_{obs}$ ;  $IC_{50}$  is the concentration of AIR that produces the half maximal inhibition; and  $HillSlope$  is the parameter that describes the steepness of the curve. Fitting was done with the program SigmaPlot.

### Lentiviral production and cell transduction

Lentiviral particles were produced by transient transfection of HEK293T cells. WT and mutant C3G constructs in pLenti-C-mEGFP-IRES-BSD lentiviral transfer vector, or the empty vector, were co-transfected with plasmids pMD2.G (kindly provided by Didier Trono; Addgene #12259) and pCMV-deltaR8.91 (Lifescience Market). Lentiviral particles were collected from the medium at 24 and 48 h by centrifugation (2 h, 30.000 xg). Ba/F3 cells were infected with the recombinant viruses in the presence of 8 µg/mL polybrene (hexadimethrine bromide, Sigma). After incubation at 37 °C for 15 hours, infected cells were selected with 20 µg/ml blasticidin (InvivoGen) during at least 3 weeks before analysis; which resulted in more than 90% of the cells expressing GFP.

### Rap1 activation assay in cells

Rap1-GTP was detected in HEK293T and Ba/F3 cells by pull-down using GST-RalGDS-RBD, which binds selectively to the active conformation of Rap1. HEK293T cells of a 9 cm plate were lysed in 0.5 ml of buffer MLB (25 mM HEPES pH 7.5, 150 mM NaCl, 1% Igepal CA-630, 10 mM MgCl<sub>2</sub>, 1 mM EDTA, 2% (v/v) glycerol, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 1 mM PMSF, 1x protease inhibitor cocktail (Roche) and 1 mM DTT) supplemented with 30 µg of purified GST-RalGDS-RBD protein. Samples were clarified by centrifugation (16000 xg, 15 min, 4 °C). 20 µl of glutathione-agarose resin were added to each

sample and were incubated 10 min at 4 °C. The resin was washed three times with 1 mL of MLB. Bound proteins were extracted and denatured with Laemmli buffer supplemented with 10 mM DTT for 12-20 hours at room temperature. Proteins were analyzed by SDS-PAGE and western blot using anti-Rap1 antibody and HRP-conjugated secondary antibody. Total Rap1 was detected in cell lysates, and Rap1-GTP was detected in the PD. GST-RalGDS-RBD was detected in the PD samples by staining with Ponceau S (membranes) or coomassie (gels). At least two independent experiments were performed. The levels of Rap1-GTP in cells expressing C3G-mEGFP-CAAX, which is constitutively attached to the membrane and produced the highest activation of Rap1, were used to normalize data from multiple experiments. The levels of Rap1-GTP in Ba/F3 cells were analyzed in a similar manner using  $5 \times 10^6$  cells. For time course experiments of IL-3 stimulation, Ba/F3 cells were starved for 12 h in media lacking FBS and IL-3, then cells were stimulated by adding 20% WEHI-3B-conditioned media, with IL-3 but without FBS.

### **Flow cytometry analysis of integrin LFA-1 activation**

The percentage of Ba/F3 cells expressing C3G-mEGFP, WT or mutants, or mEGFP alone, with activated integrin LFA-1 was determined by flow cytometry using the antibody mAb24 (ab13219, Abcam), which recognizes the activated conformation of this integrin. Ba/F3 cells were washed in ice-cold PBS containing 3% BSA, and were stained with mAb24 (1/100 dilution, 60 min, 4 °C). Cells were washed, and stained with Cy5-conjugated goat-anti-mouse-IgG secondary antibody (1/200 dilution, 30 min, 4 °C) (Jackson ImmunoResearch). Cells were washed, resuspended in cold PBS, 3% BSA, 1% sodium azide, and were analyzed using an Accuri C6 flow cytometer (BD Biosciences). Cells processed in a similar manner but without the primary antibody were used as a control of unspecific labeling by the secondary antibody. Data were analyzed using the BD Accuri C6 software.

### **Protein sequence analysis**

Search for C3G orthologs was done using the HMMER server (72). Initially, a search using as bait a segment of the NTD (residues 89-250) of human C3G identified sequences from 119 species, which exhibit a C3G-like domain architecture that contains NTD, REM, and Cdc25H domains. These sequences were used to create a hidden Markov model profile that was used in a second search, which identified sequences from 90 additional species. The 209 sequences were aligned with Clustal Omega (73) and conservation scores were calculated using the ConSurf server (74). Sequence-based secondary structure predictions were done with the methods Porter (75), PsiPred (76), and JPred4 (77).

### **Molecular modeling**

The structure of the REM domain of C3G was modelled with Modeller (78) using the ProtMod server (<http://protmod.godziklab.org/>). The structure of the REM of SOS (PDB entry 3KSY) (46), which

shares 18% sequence identity with the REM of C3G, was used as template. Molecular figures were created with PyMol (79).

### Quantification and statistical analysis

Data are presented as scatter plots with lines marking the mean  $\pm$  standard deviation (SD) of independent experiments. Statistical significances of two-group comparisons were analyzed with two-tailed unpaired Welch's t-test. Comparisons of several groups with respect to a control group were done with one-way ANOVA with Dunnett's test when variances were equivalent or Brown-Forsythe ANOVA with Dunnett's T3 test when variances were not equivalent. Differences in mean values were considered statistically significant at  $P < 0.05$ . Significant levels are: \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ ; \*\*\*\*  $P < 0.0001$ . Statistical analyses were done with the program GraphPad Prism 8.

### Supplementary Materials

Fig. S1. Concentration-dependence of the nucleotide exchange activity of C3G proteins.

Fig. S2. Secondary structure prediction for the SH3b domain of C3G.

Fig. S3. Effect of phosphorylation and CrkL on the GEF activity of C3G proteins.

Fig. S4. GEF activity of C3G WT and mutants with the NTD/REM interaction absent or weakened.

Table S1. List of C3G orthologs identified with HMMER.

Table S2. Human cancer-associated somatic missense mutations in the AIR-CBR (residues 548-584) of C3G described in the COSMIC database.

Table S3. Constructs of human C3G for expression in mammalian cells.

Table S4. Oligonucleotides used to create constructs of C3G.

Table S5. Constructs of human C3G for expression in *E. coli*.

Table S6. Oligonucleotides used for site directed mutagenesis of C3G.

Table S7. Oligonucleotides used to create constructs of Rap1b, CrkL, Src, and YopH for expression in *E. coli*.

### REFERENCES

1. T. Gotoh, S. Hattori, S. Nakamura, H. Kitayama, M. Noda, Y. Takai, K. Kaibuchi, H. Matsui, O. Hatase, H. Takahashi, Identification of Rap1 as a target for the Crk SH3 domain-binding guanine nucleotide-releasing factor C3G. *Mol Cell Biol* **15**, 6746-6753 (1995).
2. N. van den Berghe, R. H. Cool, G. Horn, A. Wittinghofer, 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).
3. S. H. Chiang, C. A. Baumann, M. Kanzaki, D. C. Thurmond, R. T. Watson, C. L. Neudauer, I. G. Macara, J. E. Pessin, A. R. Saltiel, Insulin-stimulated GLUT4 translocation requires the CAP-dependent activation of TC10. *Nature* **410**, 944-948 (2001).
4. Y. Ohba, N. Mochizuki, S. Yamashita, A. M. Chan, J. W. Schrader, S. Hattori, K. Nagashima, M. Matsuda, Regulatory proteins of R-Ras, TC21/R-Ras2, and M-Ras/R-Ras3. *J Biol Chem* **275**, 20020-20026 (2000).

5. Y. Ohba, K. Ikuta, A. Ogura, J. Matsuda, N. Mochizuki, K. Nagashima, K. Kurokawa, B. J. Mayer, K. Maki, J. Miyazaki, M. Matsuda, Requirement for C3G-dependent Rap1 activation for cell adhesion and embryogenesis. *EMBO J* **20**, 3333-3341 (2001).
6. A. Arai, Y. Nosaka, H. Kohsaka, N. Miyasaka, O. Miura, CrkL activates integrin-mediated hematopoietic cell adhesion through the guanine nucleotide exchange factor C3G. *Blood* **93**, 3713-3722 (1999).
7. N. Uemura, J. D. Griffin, The adapter protein Crkl links Cbl to C3G after integrin ligation and enhances cell migration. *J Biol Chem* **274**, 37525-37532 (1999).
8. V. Radha, A. Rajanna, A. Mitra, N. Rangaraj, G. Swarup, C3G is required for c-Abl-induced filopodia and its overexpression promotes filopodia formation. *Exp Cell Res* **313**, 2476-2492 (2007).
9. A. Gutierrez-Uzquiza, M. Arechederra, I. Molina, R. Banos, V. Maia, M. Benito, C. Guerrero, A. Porras, C3G down-regulates p38 MAPK activity in response to stress by Rap-1 independent mechanisms: involvement in cell death. *Cellular signalling* **22**, 533-542 (2010).
10. R. Shivakrupa, V. Radha, C. Sudhakar, G. Swarup, Physical and functional interaction between Hck tyrosine kinase and guanine nucleotide exchange factor C3G results in apoptosis, which is independent of C3G catalytic domain. *J Biol Chem* **278**, 52188-52194 (2003).
11. V. Maia, M. Sanz, J. Gutierrez-Berzal, A. de Luis, A. Gutierrez-Uzquiza, A. Porras, C. Guerrero, C3G silencing enhances STI-571-induced apoptosis in CML cells through p38 MAPK activation, but it antagonizes STI-571 inhibitory effect on survival. *Cellular signalling* **21**, 1229-1235 (2009).
12. A. K. Voss, D. L. Krebs, T. Thomas, C3G regulates the size of the cerebral cortex neural precursor population. *EMBO J* **25**, 3652-3663 (2006).
13. K. Sasi Kumar, A. Ramadhas, S. C. Nayak, S. Kaniyappan, K. Dayma, V. Radha, C3G (RapGEF1), a regulator of actin dynamics promotes survival and myogenic differentiation of mouse mesenchymal cells. *Biochimica et biophysica acta* **1853**, 2629-2639 (2015).
14. S. Ortiz-Rivero, C. Baquero, L. Hernandez-Cano, J. J. Roldan-Etcheverry, S. Gutierrez-Herrero, C. Fernandez-Infante, V. Martin-Granado, E. Anguita, J. M. de Pereda, A. Porras, C. Guerrero, C3G, through its GEF activity, induces megakaryocytic differentiation and proplatelet formation. *Cell communication and signaling : CCS* **16**, 101 (2018).
15. V. Martin-Granado, S. Ortiz-Rivero, R. Carmona, S. Gutiérrez-Herrero, M. Barrera, L. San-Segundo, C. Sequera, P. Perdiguero, F. Lozano, F. Martín-Herrero, J. R. González-Porras, R. Muñoz-Chápuli, A. Porras, C. Guerrero, C3G promotes a selective release of angiogenic factors from activated mouse platelets to regulate angiogenesis and tumor metastasis. *Oncotarget* **8**, 110994-111011 (2017).
16. V. Radha, A. Mitra, K. Dayma, K. Sasikumar, Signalling to actin: role of C3G, a multitasking guanine-nucleotide-exchange factor. *Bioscience reports* **31**, 231-244 (2011).
17. C. Hogan, N. Serpente, P. Cogram, C. R. Hosking, C. U. Bialucha, S. M. Feller, V. M. Braga, W. Birchmeier, Y. Fujita, Rap1 regulates the formation of E-cadherin-based cell-cell contacts. *Mol Cell Biol* **24**, 6690-6700 (2004).
18. B. S. Knudsen, S. M. Feller, H. Hanafusa, Four proline-rich sequences of the guanine-nucleotide exchange factor C3G bind with unique specificity to the first Src homology 3 domain of Crk. *J Biol Chem* **269**, 32781-32787 (1994).
19. X. Wu, B. Knudsen, S. M. Feller, J. Zheng, A. Sali, D. Cowburn, H. Hanafusa, J. Kuriyan, Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure* **3**, 215-226 (1995).
20. K. H. Kirsch, M. M. Georgescu, H. Hanafusa, Direct binding of p130(Cas) to the guanine nucleotide exchange factor C3G. *J Biol Chem* **273**, 25673-25679 (1998).
21. V. Maia, S. Ortiz-Rivero, M. Sanz, J. Gutierrez-Berzal, I. Alvarez-Fernandez, S. Gutierrez-Herrero, J. M. de Pereda, A. Porras, C. Guerrero, C3G forms complexes with Bcr-Abl and p38alpha MAPK at the focal adhesions in chronic myeloid leukemia cells: implication in the regulation of leukemic cell adhesion. *Cell communication and signaling : CCS* **11**, 9 (2013).
22. J. Gutierrez-Berzal, E. Castellano, S. Martin-Encabo, N. Gutierrez-Cianca, J. M. Hernandez, E. Santos, C. Guerrero, Characterization of p87C3G, a novel, truncated C3G isoform that is overexpressed in chronic myeloid leukemia and interacts with Bcr-Abl. *Exp Cell Res* **312**, 938-948 (2006).

23. T. Ichiba, Y. Hashimoto, M. Nakaya, Y. Kuraishi, S. Tanaka, T. Kurata, N. Mochizuki, M. Matsuda, Activation of C3G guanine nucleotide exchange factor for Rap1 by phosphorylation of tyrosine 504. *J Biol Chem* **274**, 14376-14381 (1999).
24. M. Popović, Utrecht University, (2013).
25. T. Ichiba, Y. Kuraishi, O. Sakai, S. Nagata, J. Groffen, T. Kurata, S. Hattori, M. Matsuda, 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).
26. J. Cherfils, M. Zeghouf, Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiological reviews* **93**, 269-309 (2013).
27. L. Smit, G. van der Horst, J. Borst, Sos, Vav, and C3G participate in B cell receptor-induced signaling pathways and differentially associate with Shc-Grb2, Crk, and Crk-L adaptors. *J Biol Chem* **271**, 8564-8569 (1996).
28. J. C. Nolz, L. P. Nacusi, C. M. Segovis, R. B. Medeiros, J. S. Mitchell, Y. Shimizu, D. D. Billadeau, 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).
29. R. D. York, H. Yao, T. Dillon, C. L. Ellig, S. P. Eckert, E. W. McCleskey, P. J. Stork, Rap1 mediates sustained MAP kinase activation induced by nerve growth factor. *Nature* **392**, 622-626 (1998).
30. A. Arai, Y. Nosaka, E. Kanda, K. Yamamoto, N. Miyasaka, O. Miura, Rap1 is activated by erythropoietin or interleukin-3 and is involved in regulation of beta1 integrin-mediated hematopoietic cell adhesion. *J Biol Chem* **276**, 10453-10462 (2001).
31. M. Tamada, M. P. Sheetz, Y. Sawada, Activation of a signaling cascade by cytoskeleton stretch. *Developmental cell* **7**, 709-718 (2004).
32. M. Matsuda, Y. Hashimoto, K. Muroya, H. Hasegawa, T. Kurata, S. Tanaka, S. Nakamura, S. Hattori, 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).
33. R. B. Birge, C. Kalodimos, F. Inagaki, S. Tanaka, Crk and CrkL adaptor proteins: networks for physiological and pathological signaling. *Cell communication and signaling : CCS* **7**, 13 (2009).
34. V. Radha, A. Rajanna, G. Swarup, Phosphorylated guanine nucleotide exchange factor C3G, induced by pervanadate and Src family kinases localizes to the Golgi and subcortical actin cytoskeleton. *BMC cell biology* **5**, 31 (2004).
35. A. Mitra, V. Radha, 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).
36. K. Okino, H. Nagai, H. Nakayama, D. Doi, K. Yoneyama, H. Konishi, T. Takeshita, Inactivation of Crk SH3 domain-binding guanine nucleotide-releasing factor (C3G) in cervical squamous cell carcinoma. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **16**, 763-771 (2006).
37. T. Hirata, H. Nagai, K. Koizumi, K. Okino, A. Harada, M. Onda, T. Nagahata, I. Mikami, K. Hirai, S. Haraguchi, E. Jin, O. Kawanami, K. Shimizu, M. Emi, Amplification, up-regulation and over-expression of C3G (CRK SH3 domain-binding guanine nucleotide-releasing factor) in non-small cell lung cancers. *Journal of human genetics* **49**, 290-295 (2004).
38. H. Yanagi, L. Wang, H. Nishihara, T. Kimura, M. Tanino, T. Yanagi, S. Fukuda, S. Tanaka, CRKL plays a pivotal role in tumorigenesis of head and neck squamous cell carcinoma through the regulation of cell adhesion. *Biochem Biophys Res Commun* **418**, 104-109 (2012).
39. Y. L. Che, S. J. Luo, G. Li, M. Cheng, Y. M. Gao, X. M. Li, J. M. Dai, H. He, J. Wang, H. J. Peng, Y. Zhang, W. Y. Li, H. Wang, B. Liu, H. Linghu, The C3G/Rap1 pathway promotes secretion of MMP-2 and MMP-9 and is involved in serous ovarian cancer metastasis. *Cancer letters* **359**, 241-249 (2015).
40. V. De Falco, M. D. Castellone, G. De Vita, A. M. Cirafici, J. M. Hershman, C. Guerrero, A. Fusco, R. M. Melillo, M. Santoro, RET/papillary thyroid carcinoma oncogenic signaling through the Rap1 small GTPase. *Cancer Res* **67**, 381-390 (2007).
41. C. Sequera, S. Manzano, C. Guerrero, A. Porras, How Rap and its GEFs control liver physiology and cancer development. C3G alterations in human hepatocarcinoma. *Hepatic oncology* **5**, HEP05 (2018).

42. A. Jin, T. Kurosu, K. Tsuji, D. Mizuchi, A. Arai, H. Fujita, M. Hattori, N. Minato, O. Miura, BCR/ABL and IL-3 activate Rap1 to stimulate the B-Raf/MEK/Erk and Akt signaling pathways and to regulate proliferation, apoptosis, and adhesion. *Oncogene* **25**, 4332-4340 (2006).
43. P. J. Stork, T. J. Dillon, Multiple roles of Rap1 in hematopoietic cells: complementary versus antagonistic functions. *Blood* **106**, 2952-2961 (2005).
44. N. Engels, M. Merchant, R. Pappu, A. C. Chan, R. Longnecker, J. Wienands, Epstein-Barr virus latent membrane protein 2A (LMP2A) employs the SLP-65 signaling module. *The Journal of experimental medicine* **194**, 255-264 (2001).
45. Y. Guo, B. L. Updegraff, S. Park, D. Durakoglugil, V. H. Cruz, S. Maddux, T. H. Hwang, K. A. O'Donnell, Comprehensive Ex Vivo Transposon Mutagenesis Identifies Genes That Promote Growth Factor Independence and Leukemogenesis. *Cancer Res* **76**, 773-786 (2016).
46. J. Gureasko, O. Kuchment, D. L. Makino, H. Sondermann, D. Bar-Sagi, J. Kuriyan, Role of the histone domain in the autoinhibition and activation of the Ras activator Son of Sevenless. *Proc Natl Acad Sci U S A* **107**, 3430-3435 (2010).
47. M. R. Green, A. J. Gentles, R. V. Nair, J. M. Irish, S. Kihira, C. L. Liu, I. Kela, E. S. Hopmans, J. H. Myklebust, H. Ji, S. K. Plevritis, R. Levy, A. A. Alizadeh, Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood* **121**, 1604-1611 (2013).
48. R. D. Morin, M. Mendez-Lago, A. J. Mungall, R. Goya, K. L. Mungall, R. D. Corbett, N. A. Johnson, T. M. Severson, R. Chiu, M. Field, S. Jackman, M. Krzywinski, D. W. Scott, D. L. Trinh, J. Tamura-Wells, S. Li, M. R. Firme, S. Rogic, M. Griffith, S. Chan, O. Yakovenko, I. M. Meyer, E. Y. Zhao, D. Smailus, M. Mokska, S. Chittaranjan, L. Rimsza, A. Brooks-Wilson, J. J. Spinelli, S. Ben-Neriah, B. Meissner, B. Woolcock, M. Boyle, H. McDonald, A. Tam, Y. Zhao, A. Delaney, T. Zeng, K. Tse, Y. Butterfield, I. Birol, R. Holt, J. Schein, D. E. Horsman, R. Moore, S. J. Jones, J. M. Connors, M. Hirst, R. D. Gascoyne, M. A. Marra, Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298-303 (2011).
49. K. Katagiri, M. Hattori, N. Minato, S. Irie, K. Takatsu, T. Kinashi, Rap1 is a potent activation signal for leukocyte function-associated antigen 1 distinct from protein kinase C and phosphatidylinositol-3-OH kinase. *Mol Cell Biol* **20**, 1956-1969 (2000).
50. J. S. Iwig, Y. Vercoulen, R. Das, T. Barros, A. Limnander, Y. Che, J. G. Pelton, D. E. Wemmer, J. P. Roose, J. Kuriyan, Structural analysis of autoinhibition in the Ras-specific exchange factor RasGRP1. *eLife* **2**, e00813 (2013).
51. Y. Deng, N. L. Alicea-Velazquez, L. Bannwarth, S. I. Lehtonen, T. J. Boggon, H. C. Cheng, V. P. Hytonen, B. E. Turk, Global analysis of human nonreceptor tyrosine kinase specificity using high-density peptide microarrays. *Journal of proteome research* **13**, 4339-4346 (2014).
52. N. H. Shah, M. Lobel, A. Weiss, J. Kuriyan, Fine-tuning of substrate preferences of the Src-family kinase Lck revealed through a high-throughput specificity screen. *eLife* **7**, e35190 (2018).
53. Z. Songyang, K. L. Carraway, 3rd, M. J. Eck, S. C. Harrison, R. A. Feldman, M. Mohammadi, J. Schlessinger, S. R. Hubbard, D. P. Smith, C. Eng, M. J. Lorenzo, B. A. J. Ponder, B. J. Mayer, L. C. Cantley, Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* **373**, 536-539 (1995).
54. S. M. Margarit, H. Sondermann, B. E. Hall, B. Nagar, A. Hoelz, M. Pirruccello, D. Bar-Sagi, J. Kuriyan, Structural evidence for feedback activation by Ras.GTP of the Ras-specific nucleotide exchange factor SOS. *Cell* **112**, 685-695 (2003).
55. K. B. Lin, P. Tan, S. A. Freeman, M. Lam, K. M. McNagny, M. R. Gold, The Rap GTPases regulate the migration, invasiveness and in vivo dissemination of B-cell lymphomas. *Oncogene* **29**, 608-615 (2010).
56. S. Mele, S. Devereux, A. G. Pepper, E. Infante, A. J. Ridley, Calcium-RasGRP2-Rap1 signaling mediates CD38-induced migration of chronic lymphocytic leukemia cells. *Blood advances* **2**, 1551-1561 (2018).
57. D. Ishida, K. Kometani, H. Yang, K. Kakugawa, K. Masuda, K. Iwai, M. Suzuki, S. Itohara, T. Nakahata, H. Hiai, H. Kawamoto, M. Hattori, N. Minato, Myeloproliferative stem cell disorders by deregulated Rap1 activation in SPA-1-deficient mice. *Cancer cell* **4**, 55-65 (2003).
58. D. Ishida, L. Su, A. Tamura, Y. Katayama, Y. Kawai, S. F. Wang, M. Taniwaki, Y. Hamazaki, M. Hattori, N. Minato, Rap1 signal controls B cell receptor repertoire and generation of self-reactive B1a cells. *Immunity* **24**, 417-427 (2006).

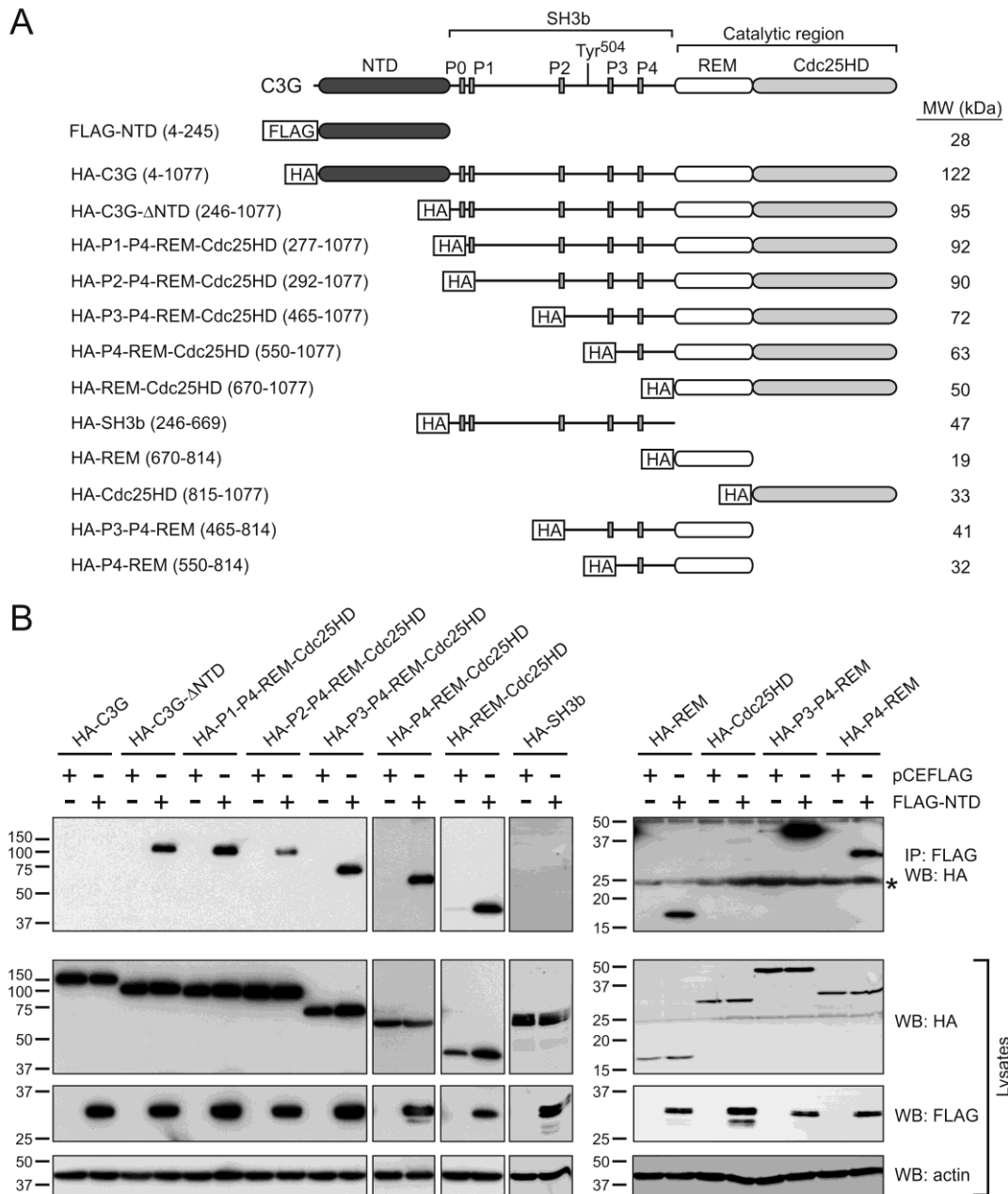
59. K. Doi, T. Imai, C. Kressler, H. Yagita, Y. Agata, M. Vooijs, Y. Hamazaki, J. Inoue, N. Minato, Crucial role of the Rap G protein signal in Notch activation and leukemogenicity of T-cell acute lymphoblastic leukemia. *Scientific reports* **5**, 7978 (2015).
60. S. F. Wang, M. Aoki, Y. Nakashima, Y. Shinozuka, H. Tanaka, M. Taniwaki, M. Hattori, N. Minato, Development of Notch-dependent T-cell leukemia by deregulated Rap1 signaling. *Blood* **111**, 2878-2886 (2008).
61. F. F. Roossien, D. de Rijk, A. Bikker, E. Roos, Involvement of LFA-1 in lymphoma invasion and metastasis demonstrated with LFA-1-deficient mutants. *J Cell Biol* **108**, 1979-1985 (1989).
62. S. T. Pals, D. J. de Gorter, M. Spaargaren, Lymphoma dissemination: the other face of lymphocyte homing. *Blood* **110**, 3102-3111 (2007).
63. M. Chiariello, M. J. Marinissen, J. S. Gutkind, Multiple mitogen-activated protein kinase signaling pathways connect the cot oncoprotein to the c-jun promoter and to cellular transformation. *Mol Cell Biol* **20**, 1747-1758 (2000).
64. N. Alonso-García, A. Ingles-Prieto, A. Sonnenberg, J. M. De Pereda, Structure of the Calx-beta domain of the integrin beta4 subunit: insights into function and cation-independent stability. *Acta Crystallogr D Biol Crystallogr* **65**, 858-871 (2009).
65. A. Alonso, N. Bottini, S. Bruckner, S. Rahmouni, S. Williams, S. P. Schoenberger, T. Mustelin, Lck dephosphorylation at Tyr-394 and inhibition of T cell antigen receptor signaling by Yersinia phosphatase YopH. *J Biol Chem* **279**, 4922-4928 (2004).
66. L. M. Luttrell, S. S. Ferguson, Y. Daaka, W. E. Miller, S. Maudsley, G. J. Della Rocca, F. Lin, H. Kawakatsu, K. Owada, D. K. Luttrell, M. G. Caron, R. J. Lefkowitz, Beta-arrestin-dependent formation of beta2 adrenergic receptor-Src protein kinase complexes. *Science* **283**, 655-661 (1999).
67. J. A. Manso, I. Garcia Rubio, M. Gomez-Hernandez, E. Ortega, R. M. Buey, A. M. Carballido, A. Carabias, N. Alonso-Garcia, J. M. de Pereda, Purification and Structural Analysis of Plectin and BPAG1e. *Methods Enzymol* **569**, 177-196 (2016).
68. A. de Marco, E. Deuerling, A. Mogk, T. Tomoyasu, B. Bukau, Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. *BMC biotechnology* **7**, 32 (2007).
69. M. A. Seeliger, M. Young, M. N. Henderson, P. Pellicena, D. S. King, A. M. Falick, J. Kuriyan, High yield bacterial expression of active c-Abl and c-Src tyrosine kinases. *Protein Sci* **14**, 3135-3139 (2005).
70. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nature methods* **9**, 671-675 (2012).
71. H. Rehmann, Characterization of the activation of the Rap-specific exchange factor Epac by cyclic nucleotides. *Methods Enzymol* **407**, 159-173 (2006).
72. S. C. Potter, A. Luciani, S. R. Eddy, Y. Park, R. Lopez, R. D. Finn, HMMER web server: 2018 update. *Nucleic Acids Res* **46**, W200-W204 (2018).
73. F. Sievers, D. G. Higgins, Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* **27**, 135-145 (2018).
74. H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **44**, W344-350 (2016).
75. C. Mirabello, G. Pollastri, Porter, PaleAle 4.0: high-accuracy prediction of protein secondary structure and relative solvent accessibility. *Bioinformatics* **29**, 2056-2058 (2013).
76. D. W. A. Buchan, D. T. Jones, The PSIPRED Protein Analysis Workbench: 20 years on. *Nucleic Acids Res* **47**, W402-W407 (2019).
77. A. Drozdetskiy, C. Cole, J. Procter, G. J. Barton, JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389-394 (2015).
78. B. Webb, A. Sali, Protein Structure Modeling with MODELLER. *Methods Mol Biol* **1654**, 39-54 (2017).
79. L. Schrödinger. (2015).
80. N. Cancer Genome Atlas, Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330-337 (2012).
81. M. Giannakis, X. J. Mu, S. A. Shukla, Z. R. Qian, O. Cohen, R. Nishihara, S. Bahl, Y. Cao, A. Amin-Mansour, M. Yamauchi, Y. Sukawa, C. Stewart, M. Rosenberg, K. Mima, K. Inamura, K. Nosho, J. A. Nowak, M. S. Lawrence, E. L. Giovannucci, A. T. Chan, K. Ng, J. A. Meyerhardt, E. M. Van Allen, G.



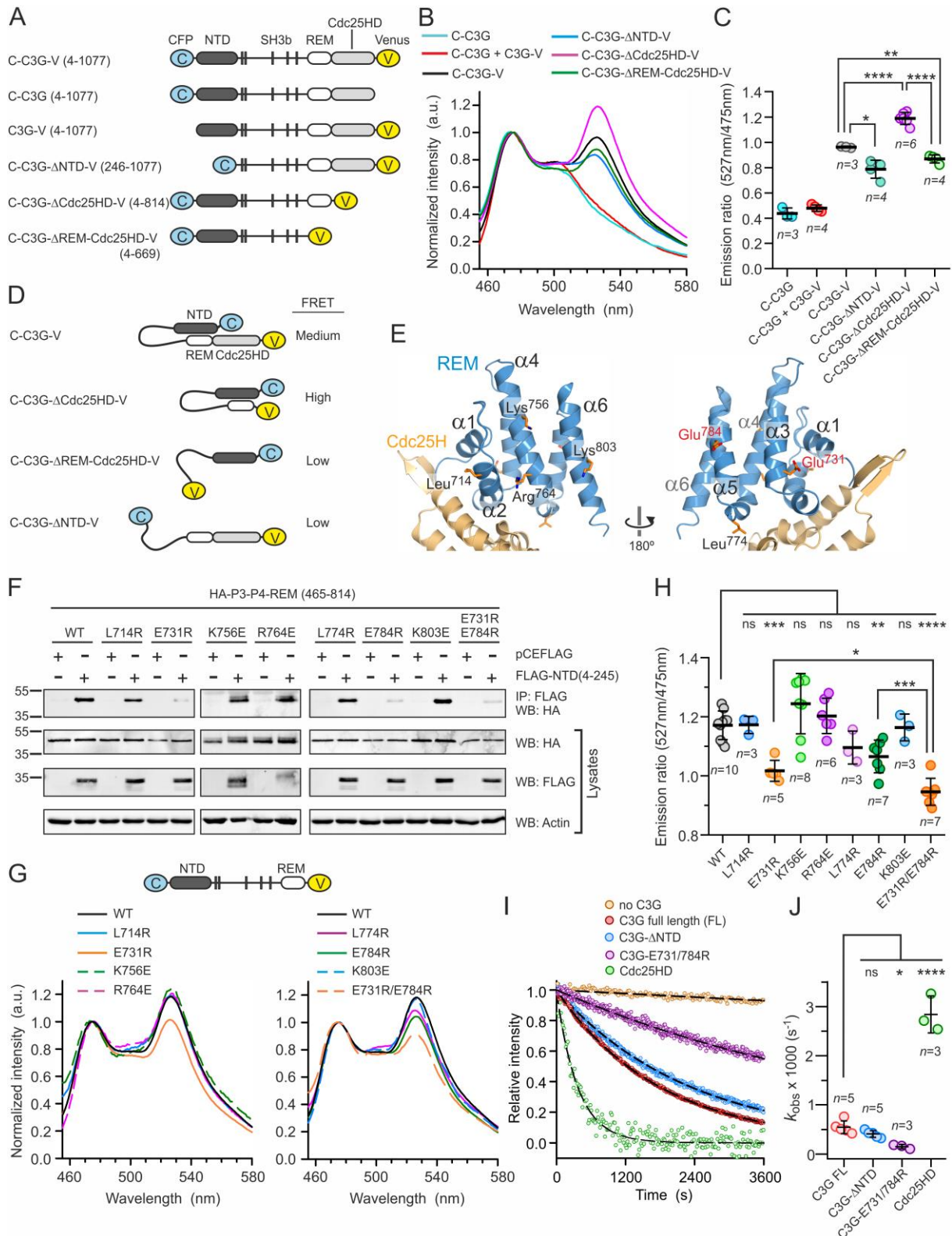
- Getz, S. B. Gabriel, E. S. Lander, C. J. Wu, C. S. Fuchs, S. Ogino, L. A. Garraway, Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell reports* **15**, 857-865 (2016).
82. J. George, J. S. Lim, S. J. Jang, Y. Cun, L. Ozretic, G. Kong, F. Leenders, X. Lu, L. Fernandez-Cuesta, G. Bosco, C. Muller, I. Dahmen, N. S. Jahchan, K. S. Park, D. Yang, A. N. Karnezis, D. Vaka, A. Torres, M. S. Wang, J. O. Korbel, R. Menon, S. M. Chun, D. Kim, M. Wilkerson, N. Hayes, D. Engelmann, B. Putzer, M. Bos, S. Michels, I. Vlastic, D. Seidel, B. Pinther, P. Schaub, C. Becker, J. Altmuller, J. Yokota, T. Kohno, R. Iwakawa, K. Tsuta, M. Noguchi, T. Muley, H. Hoffmann, P. A. Schnabel, I. Petersen, Y. Chen, A. Soltermann, V. Tischler, C. M. Choi, Y. H. Kim, P. P. Massion, Y. Zou, D. Jovanovic, M. Kontic, G. M. Wright, P. A. Russell, B. Solomon, I. Koch, M. Lindner, L. A. Muscarella, A. la Torre, J. K. Field, M. Jakopovic, J. Knezevic, E. Castanos-Velez, L. Roz, U. Pastorino, O. T. Brustugun, M. Lund-Iversen, E. Thunnissen, J. Kohler, M. Schuler, J. Botling, M. Sandelin, M. Sanchez-Cespedes, H. B. Salvesen, V. Achter, U. Lang, M. Bogus, P. M. Schneider, T. Zander, S. Ansen, M. Hallek, J. Wolf, M. Vingron, Y. Yatabe, W. D. Travis, P. Nurnberg, C. Reinhardt, S. Perner, L. Heukamp, R. Buttner, S. A. Haas, E. Brambilla, M. Peifer, J. Sage, R. K. Thomas, Comprehensive genomic profiles of small cell lung cancer. *Nature* **524**, 47-53 (2015).
  83. S. Yachida, L. D. Wood, M. Suzuki, E. Takai, Y. Totoki, M. Kato, C. Luchini, Y. Arai, H. Nakamura, N. Hama, A. Elzawahry, F. Hosoda, T. Shirota, N. Morimoto, K. Hori, J. Funazaki, H. Tanaka, C. Morizane, T. Okusaka, S. Nara, K. Shimada, N. Hiraoka, H. Taniguchi, R. Higuchi, M. Oshima, K. Okano, S. Hirono, M. Mizuma, K. Arihiro, M. Yamamoto, M. Unno, H. Yamaue, M. J. Weiss, C. L. Wolfgang, T. Furukawa, H. Nakagama, B. Vogelstein, T. Kiyono, R. H. Hruban, T. Shibata, Genomic Sequencing Identifies ELF3 as a Driver of Ampullary Carcinoma. *Cancer cell* **29**, 229-240 (2016).
  84. M. Imielinski, A. H. Berger, P. S. Hammerman, B. Hernandez, T. J. Pugh, E. Hodis, J. Cho, J. Suh, M. Capelletti, A. Sivachenko, C. Sougnez, D. Auclair, M. S. Lawrence, P. Stojanov, K. Cibulskis, K. Choi, L. de Waal, T. Sharifnia, A. Brooks, H. Greulich, S. Banerji, T. Zander, D. Seidel, F. Leenders, S. Ansen, C. Ludwig, W. Engel-Riedel, E. Stoelben, J. Wolf, C. Goparju, K. Thompson, W. Winckler, D. Kwiatkowski, B. E. Johnson, P. A. Janne, V. A. Miller, W. Pao, W. D. Travis, H. I. Pass, S. B. Gabriel, E. S. Lander, R. K. Thomas, L. A. Garraway, G. Getz, M. Meyerson, Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107-1120 (2012).
  85. S. Seshagiri, E. W. Stawiski, S. Durinck, Z. Modrusan, E. E. Storm, C. B. Conboy, S. Chaudhuri, Y. Guan, V. Janakiraman, B. S. Jaiswal, J. Guillory, C. Ha, G. J. Dijkgraaf, J. Stinson, F. Gnad, M. A. Huntley, J. D. Degenhardt, P. M. Haverty, R. Bourgon, W. Wang, H. Koeppen, R. Gentleman, T. K. Starr, Z. Zhang, D. A. Largaespada, T. D. Wu, F. J. de Sauvage, Recurrent R-spondin fusions in colon cancer. *Nature* **488**, 660-664 (2012).

**Acknowledgments:** We thank Andrés Alonso (Instituto de Biología y Genética Molecular, Valladolid, Spain), Bernd Bukau (EMBL, Heidelberg, Germany), Robert Lefkowitz (Duke University Medical Center, Durham, NC), and Didier Trono (Ecole Polytechnique Fédérale de Lausanne, Switzerland) for providing plasmids and José M Sánchez Santos (University of Salamanca, Spain) for assessing the statistical analysis. **Funding:** This work was supported by the Spanish Ministry of Economy, Industry and Competitiveness (MINECO), Agencia Estatal de Investigación (AEI), and the European Regional Development Fund (ERDF) (grant BFU2015-69499-P to JMdP, SAF2016-76588-C2-2-R to CG); and from the Consejería de Educación, Junta de Castilla y León (grant SA017U16 to CG and JMdP). AC received funding from MINECO (FPU14/06259). MGH and ARB were funded by Banco Santander and University of Salamanca. The authors' institution is supported by the Programa de Apoyo a Planes Estratégicos de Investigación de Estructuras de Investigación de Excelencia co-

funded by Castilla y León autonomous government and ERDF (CLC-2017-01). **Author contributions:** A.C., M.G.H., and J.M.d.P. conceived the study and designed experiments. C.G. designed and contributed to the interpretation of assays in cells. M.G.H. characterized the NTD/REM interaction. A.C. performed most of the other experiments. S.d.C. purified functional C3G-Cdc25HD. A.R.B. performed in vitro activity assays. A.M.V. performed assays in cells. P.G.S. performed part to the analysis of the NTD/REM interaction. A.C., C.G. and J.M.d.P. wrote the paper with input from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The data generated in this study that support the findings are available from the corresponding author upon request.

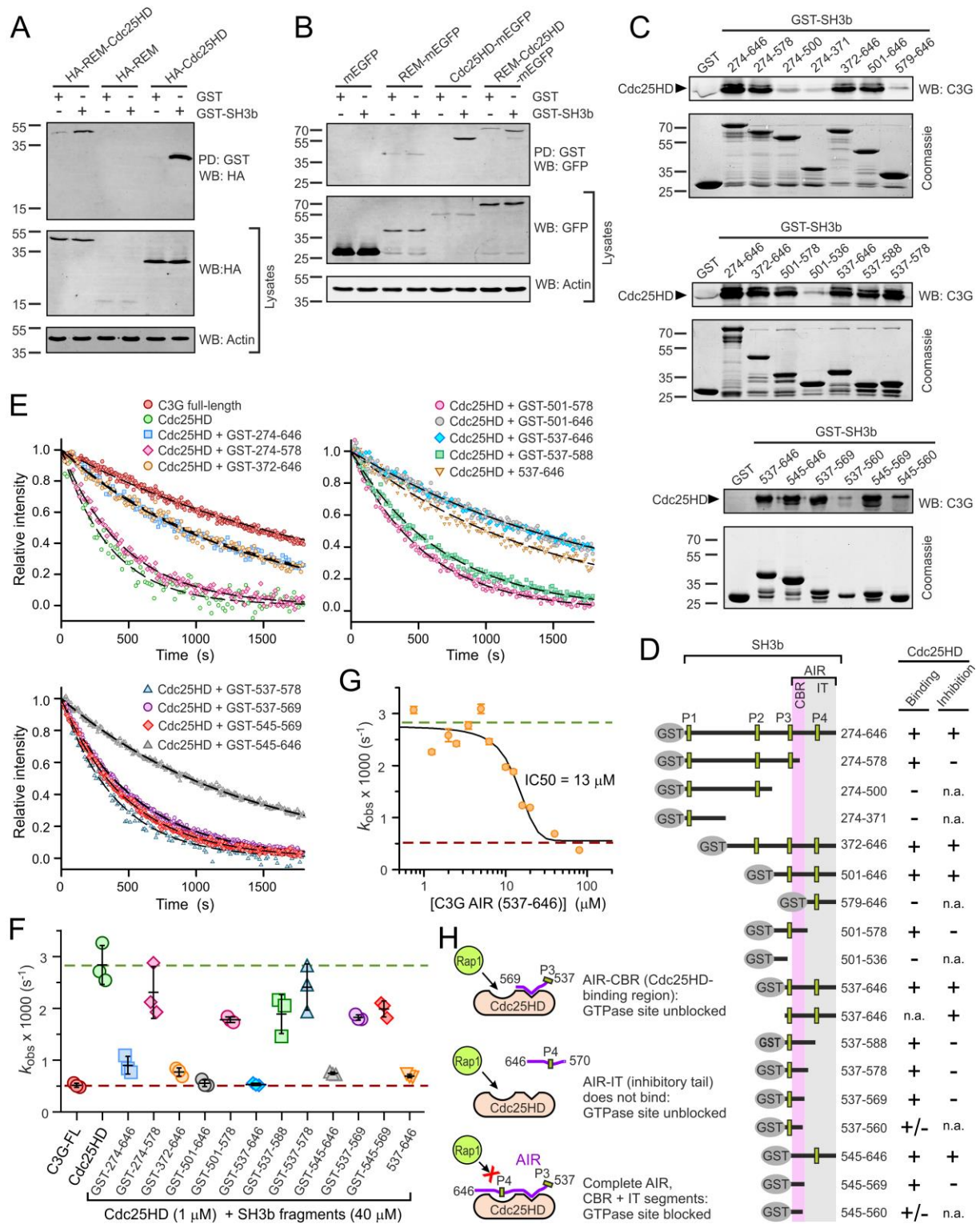


**Fig. 1. The NTD of C3G interacts with the REM domain.** (A) Schematic representation of the domain structure of C3G. The constructs used to study the interaction mediated by the NTD are shown aligned underneath. Molecular weights (MW) calculated from the sequence are indicated. (B) Analysis by co-immunoprecipitation (coIP) of the interaction of FLAG-NTD with other regions of C3G (HA-tagged) co-expressed in COS-1 cells. Cell lysates were immunoblotted for HA, FLAG, and  $\beta$ -actin. Proteins were IP with agarose-immobilized antibody against FLAG and were immunoblotted for HA. The asterisk indicates an unspecific band. Blots are representative of at least two independent experiments.



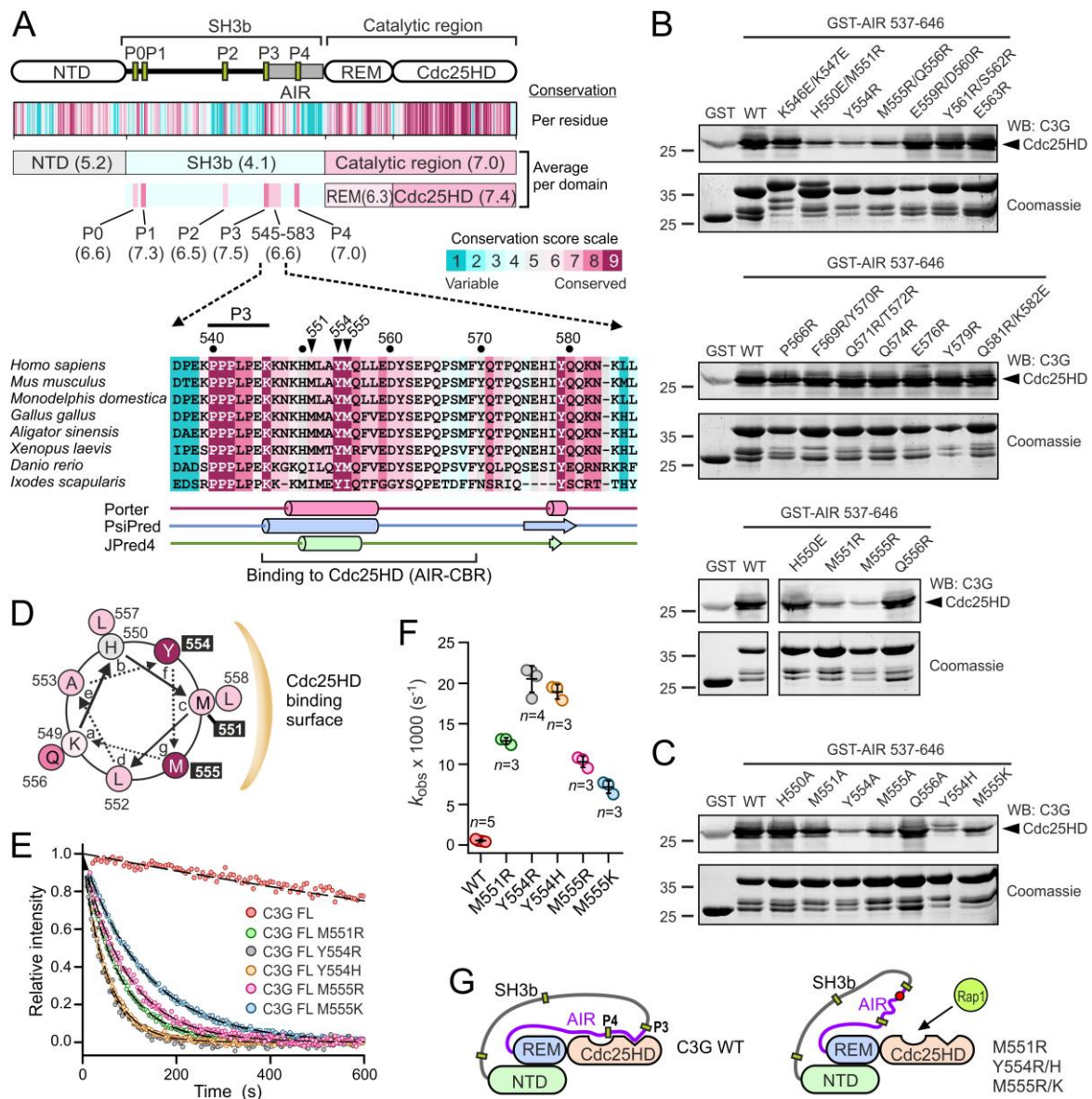
**Fig. 2. The NTD/REM interaction is intramolecular and dispensable for autoinhibition of C3G.** (A) Schematic representation of C3G fusion proteins used as FRET-based conformational sensors. (B) Normalized fluorescence emission spectra (excitation 435 nm) of COS-1 cells expressing C3G constructs with CFP and Venus. (C) Fluorescence intensity ratios at Venus and CFP emission maxima used as a measurement of the relative FRET efficiency.  $n = 3-6$  independent experiments as indicated, one shown in B. Mean  $\pm$  standard deviation (SD). Welch's t-test was used for pairwise comparisons. (D) Schematic representation of the conformations of the C3G sensors based on the FRET data. (E) Structural model of the REM domain of C3G.

The side chains of the mutated residues are shown as sticks. **(F)** Interaction between the FLAG-NTD and HA-P3-P4-REM, WT and point mutants, assayed by co-immunoprecipitation (coIP) using an antibody against FLAG. Co-transfection with empty pCEFLAG vector was used as negative control. Tagged proteins and  $\beta$ -actin in cell lysates and HA-P3-P4-REM in the IP were analyzed by western blot. Blots are representative of two independent experiments. **(G)** Representative emission spectra (excitation 435 nm) of COS-1 cells expressing C-C3G- $\Delta$ Cdc25HD-V WT or point mutants. **(H)** Ratios of the fluorescence intensity of C-C3G- $\Delta$ Cdc25HD-V WT and mutants. Mean  $\pm$  SD.  $n = 3-10$  independent experiments as indicated, one shown in G. Statistical comparison were analyzed using ANOVA followed by Dunnett's test. **(I)** Representative nucleotide exchange reactions of Rap1:mant-dGDP alone and in the presence of C3G proteins (1  $\mu$ M). Single exponential decays were fitted (dashed lines) to determine the apparent nucleotide exchange rates ( $k_{obs}$ ). **(J)** Exchange rates ( $k_{obs}$ ) produced by C3G full-length (FL), C3G- $\Delta$ NTD, C3G-E731R/E784R, and the isolated Cdc25HD. Mean  $\pm$  SD.  $n = 3-5$  independent experiments, as indicated, one shown in I. For statistical analysis ANOVA followed by Dunnett's test was performed. The linear dependence of the  $k_{obs}$  with the concentration of C3G FL and Cdc25HD is shown in fig S1. In all statistical analysis \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ ; \*\*\*\*  $P < 0.0001$ .



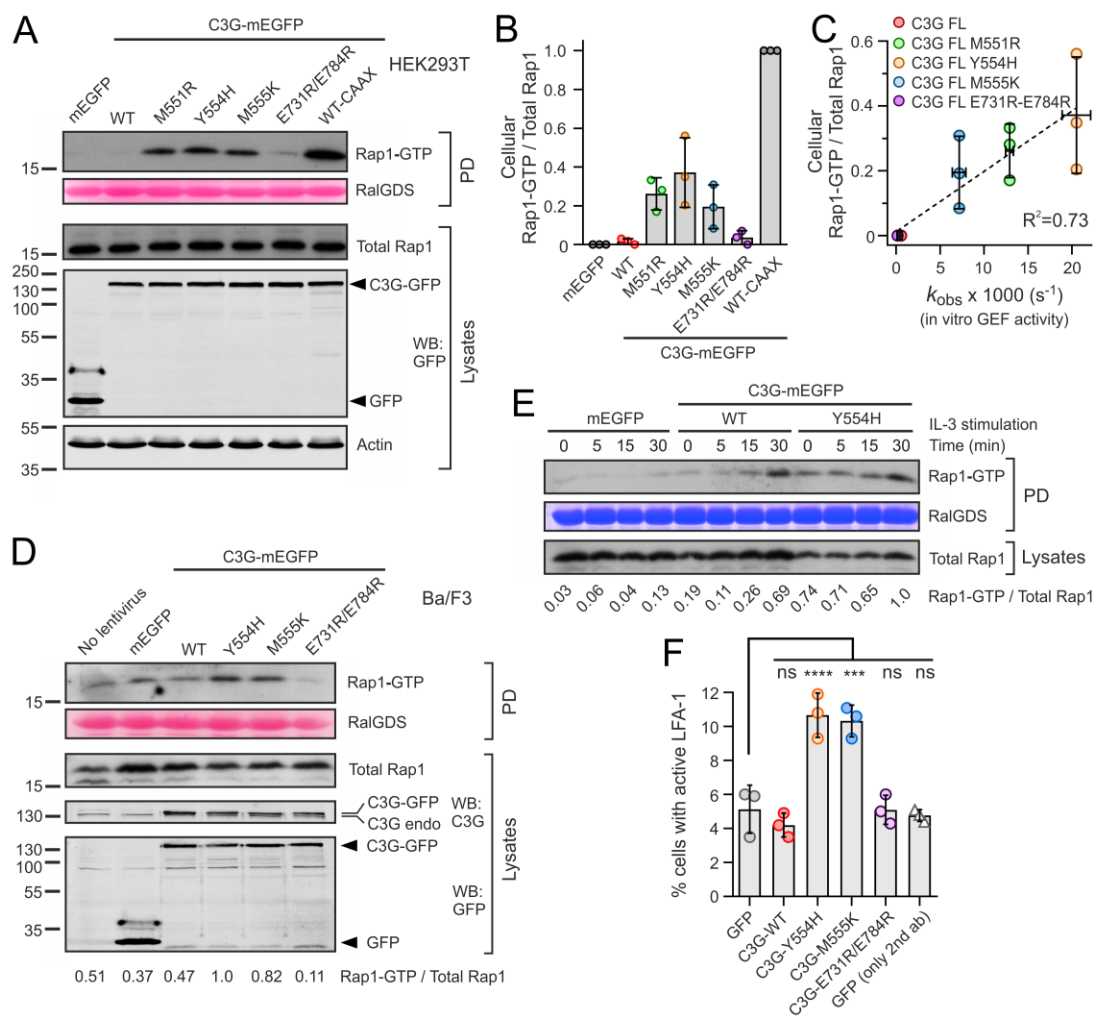
**Fig. 3. The final segment of the SH3b domain binds to the Cdc25HD and blocks its GEF activity.** (A and B) Analysis by pull-down (PD) of the interaction of the SH3b domain (274-646) with the catalytic region. The region REM-Cdc25HD and its individual domains were expressed in HEK293T cells with HA (A) or mEGFP (B) tags, and were subjected to PD with GST-SH3b. Tagged C3G fragments and  $\beta$ -actin in the cell lysates and C3G fragments in the PD were detected by western blot. (C) Binding of the Cdc25HD to GST-SH3b fragments analyzed by PD. The Cdc25HD produced in *E. coli* was detected in the pull-down by western blot with an antibody against C3G. GST-SH3b fragments were visualized by coomassie staining. Pull-downs are

representative of two independent experiments. **(D)** Schematic representation of the fragments of the SH3b domain analyzed and summary of their binding to the Cdc25HD and inhibition of the GEF activity (n.a. not analyzed). **(E)** Exchange reactions of Rap1:mant-dGDP catalyzed by C3G full-length (1  $\mu$ M), the Cdc25HD alone (1  $\mu$ M), and the Cdc25HD (1  $\mu$ M) in the presence of fragments of the SH3b domain (40  $\mu$ M). The  $k_{\text{obs}}$  were determined by fitting a single exponential decay model (dashed lines). Data are representative reactions of three independent experiments. **(F)** GEF activity ( $k_{\text{obs}}$ ) of C3G full-length (FL) and Cdc25HD alone and in the presence of fragments of the SH3b domain. Lines represent mean  $\pm$  SD,  $n = 3$ , independent experiments, one shown in E. Dashed lines mark the exchange rates of autoinhibited full-length C3G and the uninhibited isolated Cdc25HD. **(G)** Dose-dependent effect of untagged AIR on the activity of the Cdc25HD (1  $\mu$ M). A sigmoidal function (solid line) was fitted to estimate the IC50. **(H)** Schematic representation of the two segments of the AIR and their contribution to the repression of the GEF activity of the Cdc25HD. The model depicts the behavior observed using independently purified AIR and Cdc25HD.

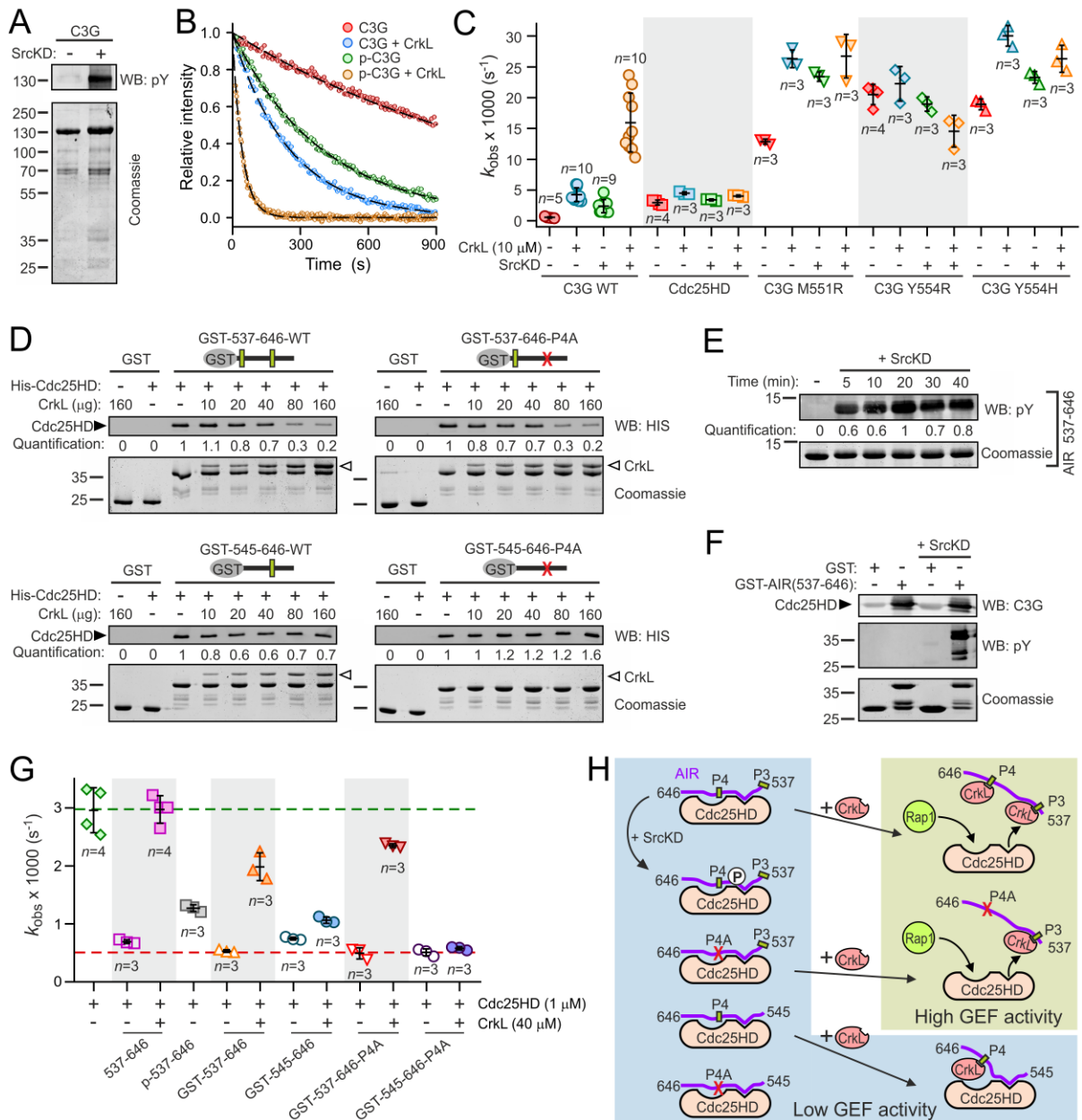


**Fig. 4. Mutations that disrupt the AIR/Cdc25HD interaction activate C3G constitutively.** (A) Conservation scores of C3G per residue (colored bars) and average values per domain (boxes). Alignment of the sequences of the P3 motif and part of the AIR from representative species; positions colored according to their conservation. Secondary structure predictions by three methods are shown under the sequences. Secondary structure prediction for the complete SH3b domain is shown in fig. S2. (B and C) Binding of the Cdc25HD to GST-AIR, WT and mutants, analyzed by pull-down (PD) assays. Cdc25HD in the PD was detected by western blot. Two types of substitutions were assayed: reverse-charge replacements (B), and changes to alanine and mutations described in lymphomas (C). Pull-downs are representative of two independent experiments. (D) Helical wheel representation of the predicted helix in the AIR-CBR. Met<sup>551</sup>, Tyr<sup>554</sup>, and Met<sup>555</sup> define the putative binding site for the Cdc25HD. (E) Representative nucleotide exchange reactions of Rap1:mant-dGDP catalyzed by full-length C3G WT and the indicated point mutants (1  $\mu$ M). (F) Exchange rates ( $k_{obs}$ ) catalyzed by C3G full-length WT and point mutants.  $n = 3$ -5 independent experiments, as indicated, one shown in E. (G) Schematic representation of autoinhibited C3G and the uninhibited conformation induced by mutations that destabilize the AIR/Cdc25HD interaction.



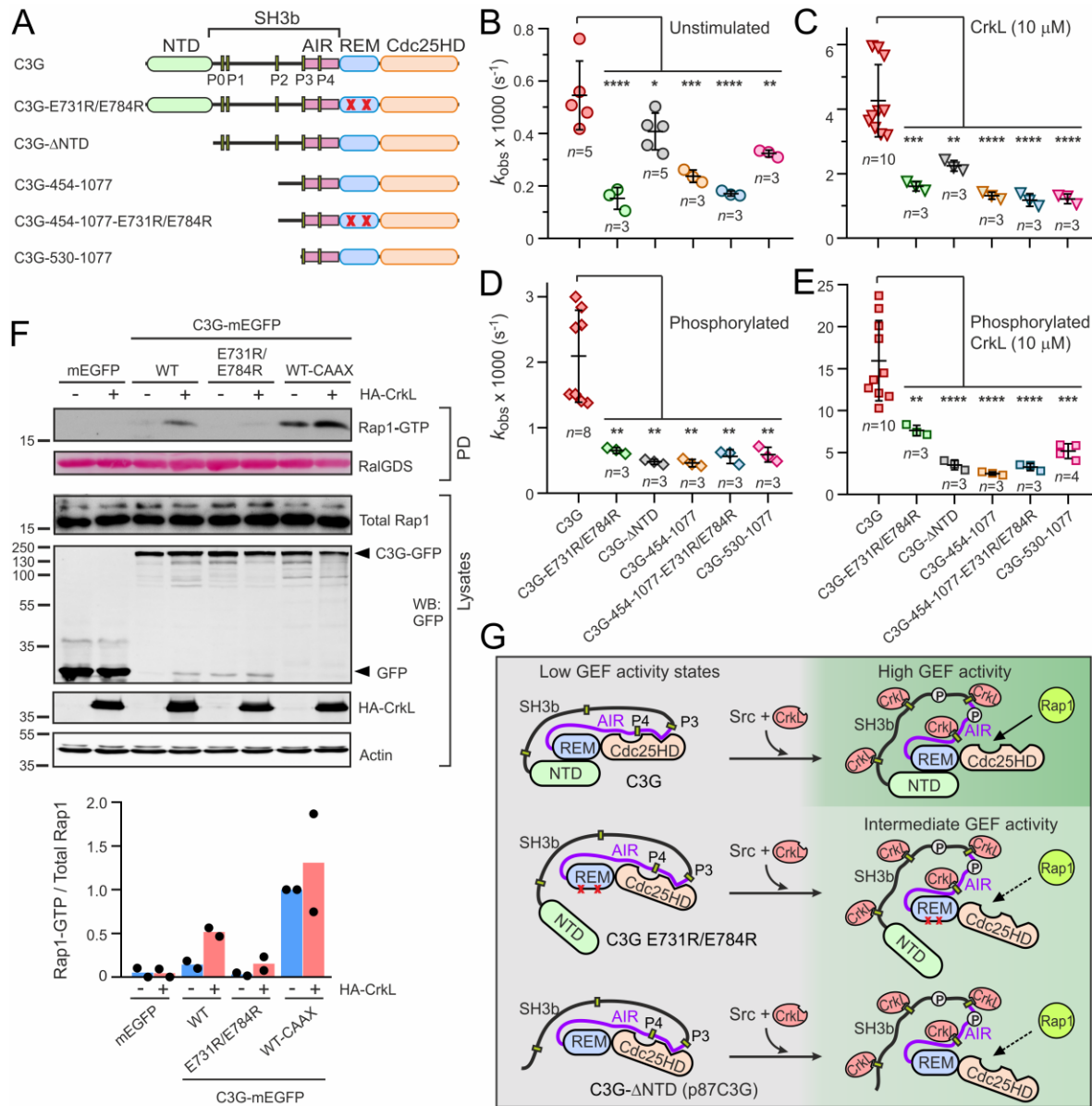


**Fig. 5. Lymphoma-related mutations in C3G activate Rap1 and LFA-1 in cells.** (A) Analysis of Rap1-GTP in HEK293T cells expressing mEGFP, C3G-mEGFP wild type, or the mutants M551R, Y554H, or M555K, or the membrane-targeted WT with a CAAX sequence. Expression of mEGFP and endogenous Rap1 and  $\beta$ -actin in cell lysates were analyzed by western blot. In the pull-down (PD) Rap1-GTP was detected by western blot and GST-RalGDS-RBD was detected by Ponceau S staining. PD is representative of three independent experiments. (B) Scatter plot and bar chart of three independent measurements of Rap1 activation in HEK293T cells, one shown in A. The Rap1-GTP levels in cells expressing C3G-mEGFP-CAAX were used to normalize the data from different experiments. (C) Correlation between the exchange activity in vitro ( $k_{obs}$ ) of C3G WT and mutants (Fig. 4F), and the Rap1-GTP levels that they induced in HEK293T cells (data in B). (D) Rap1-GTP levels in Ba/F3 cells expressing C3G-mEGFP, WT or indicated mutants, or mEGFP alone, and in uninfected cells. Proteins were detected in the cell lysates and in the GST-RalGDS-RBD pull-downs as in A. In addition, C3G (endogenous and mEGFP-tagged) was also detected using an antibody against C3G. PD is representative of two independent experiments. (E) Time course activation of Rap1 after stimulation with IL-3 of Ba/F3 cells expressing mEGFP, C3G-mEGFP WT or Y554H. Rap1-GTP was detected as in A and D. PD is representative of two independent experiments. (F) Activation of the integrin LFA-1 in Ba/F3 cells expressing mEGFP, C3G-mEGFP WT or mutants, determined by flow cytometry ( $n = 3$ , biological replicates, mean  $\pm$  SD). Statistical comparison with respect to cells expressing mEGFP was analyzed using ANOVA followed by Dunnett's test. \*\*\*  $P < 0.001$ ; \*\*\*\*  $P < 0.0001$ .



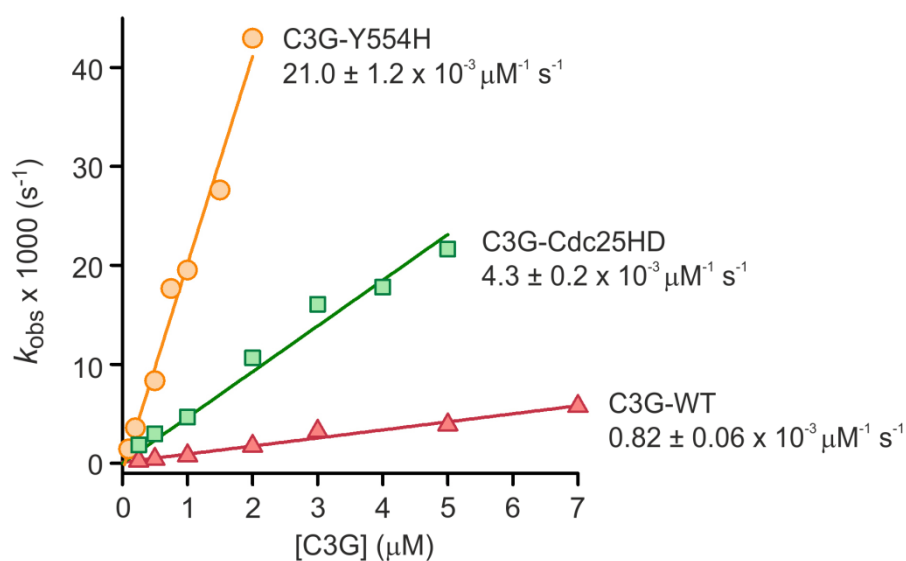
**Fig. 6. The AIR/Cdc25HD interaction is the main autoinhibitory mechanism and is disrupted by CrkL for activation.** (A) Phosphorylation of purified C3G-WT with SrcKD, analyzed by western blot. Similar analysis of Cdc25HD and the C3G mutants M551R, Y554R, and Y554H are shown in fig. S3, A to D. (B) Representative dissociation reactions of Rap1:mant-dGDP catalyzed by C3G (1  $\mu$ M), unmodified or phosphorylated with SrcKD (p-C3G), alone and in the presence of CrkL (10  $\mu$ M). (C) Nucleotide exchange rates ( $k_{obs}$ ) catalyzed by C3G full-length WT and point mutants, and the isolated Cdc25HD. Unphosphorylated and SrcKD-phosphorylated samples were analyzed alone and in the presence of CrkL.,  $n = 3-10$  independent experiments as indicated. Representative dissociation reactions are shown in B (WT) and in fig. S3 A to D (Cdc25HD and mutants). (D) Analysis by pull-down of the competition of CrkL with the Cdc25HD for binding to four constructs of the AIR that contain the P3 and P4 (GST-537-646-WT), only the P3 (GST-537-646-P4A), only the P4 (GST-545-646-WT), or none of these proline-rich motifs (GST-545-646-P4A). Pull-downs are representative of two independent experiments. (E) Time-course of the in vitro phosphorylation of the AIR (537-646) by the SrcKD, analyzed by western blot; representative of two independent experiments. (F) Binding of Cdc25HD to GST-AIR (537-646) phosphorylated with SrcKD in pull-down assays. Pull-down representative of two independent experiments. (G) Exchange activity ( $k_{obs}$ ) of the Cdc25HD alone and in the presence of AIR (untagged or GST-fusion) or the AIR mutants in D.  $n = 3-4$  independent experiments as indicated. Representative dissociation experiments are shown in fig. S3E. (H) Schematic representation of the effect of

phosphorylation and CrkL binding to the AIR on the release of the inhibitory interaction, when the AIR and the Cdc25HD are assayed as individual proteins.

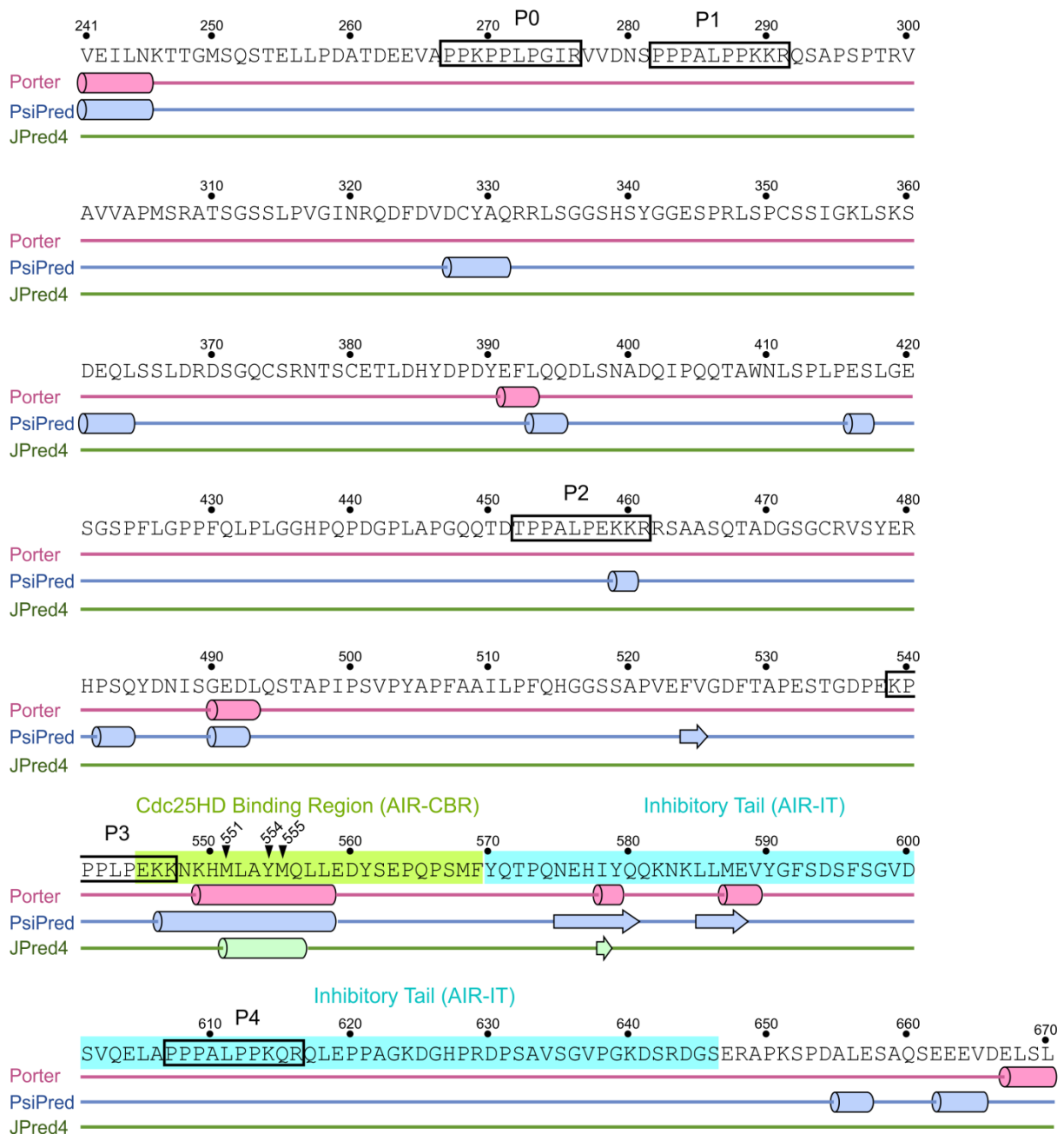


**Fig. 7. The NTD/REM interaction stimulates the GEF activity of the Cdc25HD.** (A) Schematic representation of C3G full-length WT, truncation fragments, and the mutant E731R/E784R in which the NTD/REM interaction is destabilized. (B to E) Nucleotide exchange activity ( $k_{obs}$ ) of the proteins depicted in A, in the absence of stimuli (B), in the presence of CrkL (C), phosphorylated with SrcKD (D), and when CrkL and phosphorylation were combined (E). The number of independent experiments ( $n$ ) is indicated. Lines are mean  $\pm$  SD. Representative nucleotide dissociation reactions are shown in fig. S4, A, B, D and E. One-way ANOVA followed by Dunnett's test was used in B and C; unequal variance Brown-Forsythe ANOVA followed by Dunnett's test was used in D and E. \*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ ; \*\*\*\*  $P < 0.0001$ . (F) Analysis of Rap1 activation in HEK293T cells expressing C3G-mEGFP WT, the mutant E731R/E784R, or C3G with the membrane targeting CAAX-tag. Quantitation of two independent experiments is shown below. (G) Schematic illustration of the contribution of the NTD/REM interaction to the activation by CrkL and phosphorylation of C3G WT and mutants.

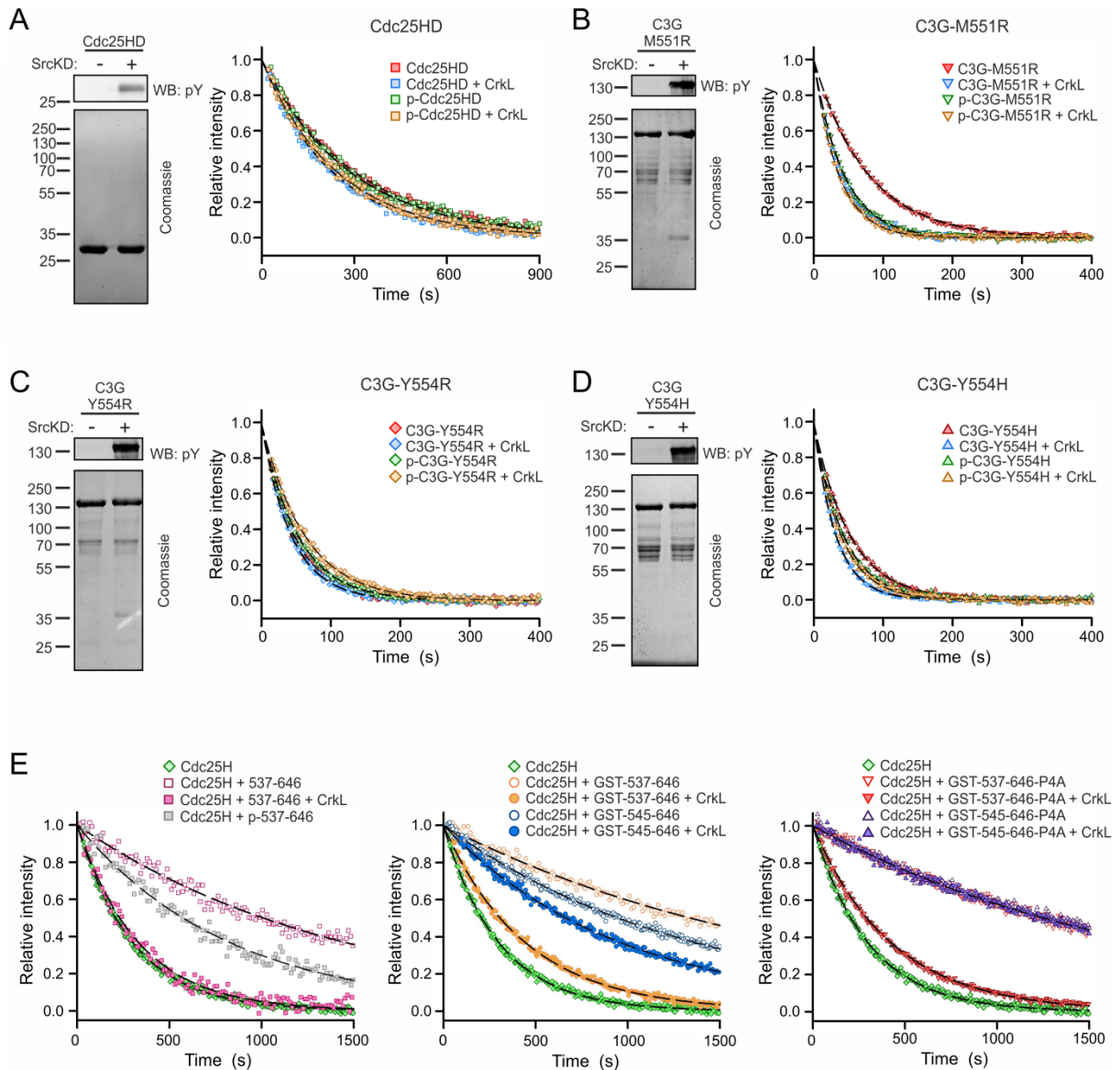
## SUPPLEMENTARY FIGURES AND TABLES



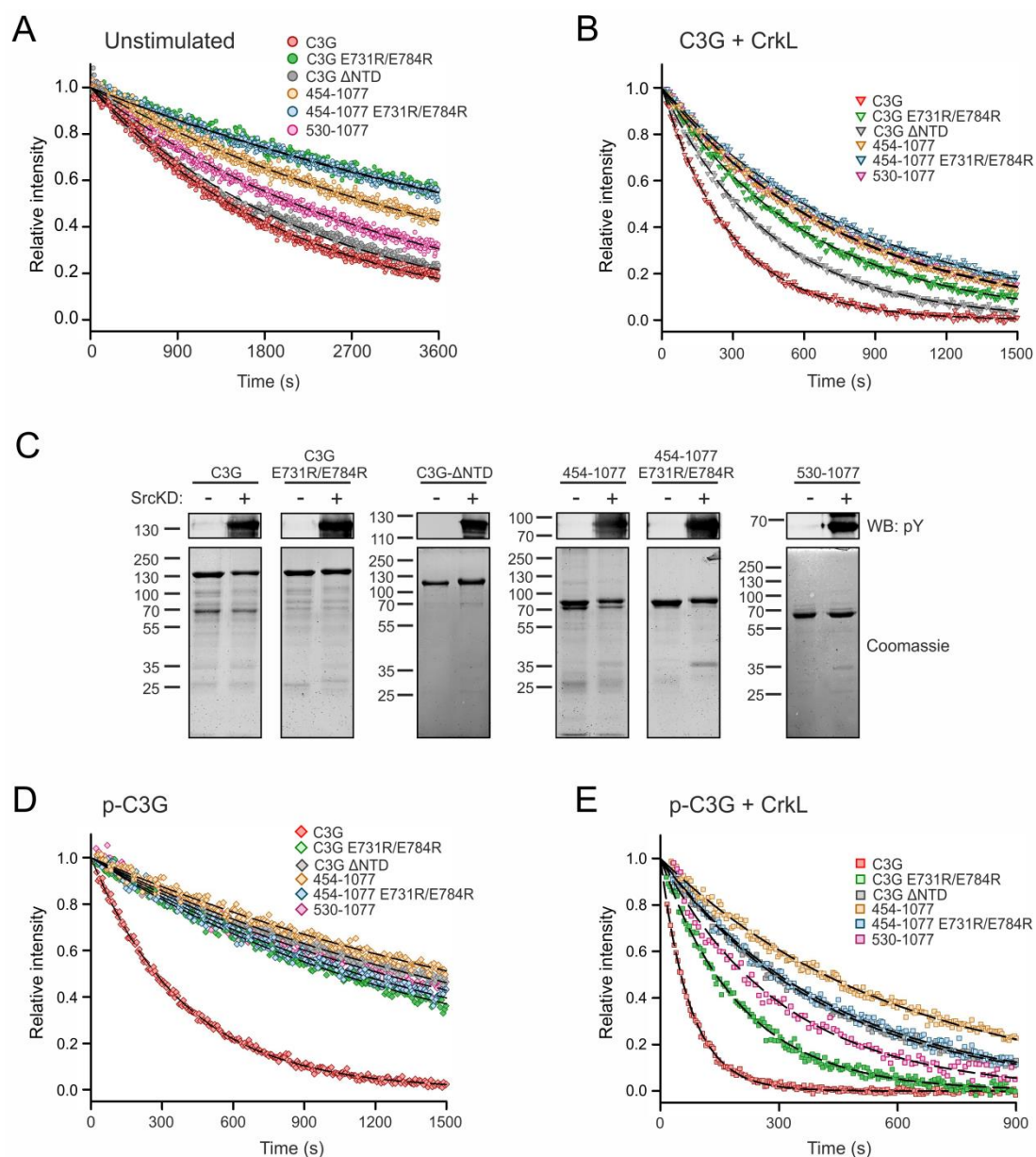
**Fig. S1. Concentration-dependence of the nucleotide exchange activity of C3G proteins.** Apparent dissociation rates ( $k_{obs}$ ) determined for in vitro nucleotide exchange reactions of Rap1b:mant-dGDP (200 nM) catalyzed by C3G proteins at several concentrations. The reactions were triggered by adding an excess of GDP (40 μM). Data points are the apparent dissociation rates ( $k_{obs}$ ) determined for each reaction by fitting a single exponential decay model. Numbers are the specific activity of each protein, which were estimated from the simple linear regression (lines) fit to the data,  $\pm$  the asymptotic standard errors that measure the uncertainties in the estimated coefficients.



**Fig. S2. Secondary structure prediction for the SH3b domain of C3G.** Prediction was done with three independent methods: Porter (75), PsiPred (76), and JPred4 (77). Predicted  $\alpha$ -helices,  $\beta$ -strands, and random coils are indicated by cylinders, arrows, and lines, respectively. Proline-rich motifs are framed in boxes, and the two segments of the AIR are highlighted by colored rectangles.



**Fig. S3. Effect of phosphorylation and CrkL on the GEF activity of C3G proteins.** (A to D) Analysis by western blot of the *in vitro* phosphorylation of C3G proteins by SrcKD (left), and representative nucleotide dissociation kinetics of Rap1b:mant-dGDP catalyzed by the Cdc25HD (A), C3G-M551R (B), C3G-Y554R (C), and C3G-Y554H (D). The GEF activity was analyzed for the unmodified and phosphorylated C3G proteins (1  $\mu$ M), alone and in the presence of CrkL (10  $\mu$ M).  $n = 3-10$  independent experiments as indicated in Fig. 6C. (E) Representative nucleotide dissociation kinetics of Rap1b:mant-dGDP catalyzed by the Cdc25HD (1  $\mu$ M) in the presence of fragments and mutants of the AIR (40  $\mu$ M), alone and in the presence of CrkL (40  $\mu$ M), and in the presence of phosphorylated AIR (p-537-646) (40  $\mu$ M).  $n = 3-4$  independent experiments as indicated in Fig. 6G



**Fig. S4. GEF activity of C3G WT and mutants with the NTD/REM interaction absent or weakened.** (A-B) Representative nucleotide dissociation kinetics of Rap1b:mant-dGDP catalyzed by the indicated C3G proteins (1  $\mu$ M) unstimulated (A) and in the presence of 10  $\mu$ M CrkL (B).  $n = 3-10$  independent experiments as indicated in Fig 7 B and C. (C) Phosphorylation in vitro of C3G proteins by SrcKD.  $\sim 40 \mu$ M of C3G proteins were incubated with 1  $\mu$ M SrcKD in kinase buffer for 20 min at 30  $^{\circ}$ C. Samples before and after phosphorylation were analyzed by western blot using an antibody against pTyr (pY). Blots are representative of two independent experiments. (D-E) Representative nucleotide dissociation kinetics of Rap1b:mant-dGDP catalyzed by C3G proteins (1  $\mu$ M) phosphorylated with SrcKD (D), and phosphorylated proteins in the presence of CrkL (E).  $n = 3-10$  independent experiments as indicated in Fig 7 D and E.

**Table S1. List of C3G orthologs identified with HMMER.**

Clade	Subclade	Species	E-value	UniProt ID
Chordata	Mammalian	<i>Sarcophilus harrisii</i>	9.30E-72	G3WZS0_SARHA
Chordata	Mammalian	<i>Loxodonta africana</i>	3.00E-71	G3SSW6_LOXAF
Chordata	Mammalian	<i>Bos taurus</i>	6.10E-71	F1MJY2_BOVIN
Chordata	Mammalian	<i>Ovis aries</i>	6.30E-71	W5P986_SHEEP
Chordata	Mammalian	<i>Equus caballus</i>	7.80E-71	F6V778_HORSE
Chordata	Mammalian	<i>Equus caballus</i>	7.80E-71	F7D9U2_HORSE
Chordata	Mammalian	<i>Ursus maritimus</i>	9.40E-71	A0A384CRI6_URSMA
Chordata	Mammalian	<i>Physeter catodon</i>	1.20E-70	A0A2Y9S3E4_PHYCD
Chordata	Mammalian	<i>Delphinapterus leucas</i>	1.40E-70	A0A2Y9NJ55_DELLE
Chordata	Mammalian	<i>Leptonychotes weddellii</i>	1.90E-70	A0A2U3XTA8_LEPWE
Chordata	Mammalian	<i>Ailuropoda melanoleuca</i>	2.00E-70	G1MDI0_AILME
Chordata	Mammalian	<i>Erinaceus europaeus</i>	2.00E-70	A0A1S3A6Y2_ERIEU
Chordata	Mammalian	<i>Tursiops truncatus</i>	2.00E-70	A0A2U4A417_TURTR
Chordata	Mammalian	<i>Balaenoptera acutorostrata scammoni</i>	2.80E-70	A0A384B489_BALAS
Chordata	Mammalian	<i>Odobenus rosmarus divergens</i>	2.90E-70	A0A2U3ZI88_ODORO
Chordata	Mammalian	<i>Felis catus</i>	3.30E-70	A0A2I2U2J0_FELCA
Chordata	Mammalian	<i>Mus musculus</i>	4.20E-70	Q91ZZ2_MOUSE
Chordata	Mammalian	<i>Cricetulus griseus</i>	4.90E-70	G3IFG2_CRIGR
Chordata	Mammalian	<i>Tarsius syrichta</i>	5.20E-70	A0A3Q0DUA8_TARSY
Chordata	Mammalian	<i>Trichechus manatus latirostris</i>	9.90E-70	A0A2Y9QVW7_TRIMA
Chordata	Mammalian	<i>Canis lupus familiaris</i>	1.00E-69	F1PH97_CANLF
Chordata	Mammalian	<i>Mustela putorius furo</i>	1.10E-69	M3YMC6_MUSPF
Chordata	Mammalian	<i>Enhydra lutris kenyonii</i>	1.30E-69	A0A2Y9J7C9_ENHLU
Chordata	Mammalian	<i>Rattus norvegicus</i>	1.40E-69	F1M8L9_RAT
Chordata	Mammalian	<i>Dipodomys ordii</i>	1.50E-69	A0A1S3FLW8_DIPOR
Chordata	Mammalian	<i>Ornithorhynchus anatinus</i>	1.50E-69	F7EGC8_ORNAN
Chordata	Mammalian	<i>Cavia porcellus</i>	1.90E-69	H0VFR1_CAVPO
Chordata	Mammalian	<i>Ictidomys tridecemlineatus</i>	2.00E-69	A0A287D9I6_ICTTR
Chordata	Mammalian	<i>Monodelphis domestica</i>	3.20E-69	F6QKH4_MONDO
Chordata	Mammalian	<i>Cebus capucinus imitator</i>	4.40E-69	A0A2K5RV62_CEBCA
Chordata	Mammalian	<i>Myotis lucifugus</i>	4.50E-69	G1PC22_MYOLU
Chordata	Mammalian	<i>Saimiri boliviensis boliviensis</i>	4.90E-69	A0A2K6UXE9_SAIBB
Chordata	Mammalian	<i>Callithrix jacchus</i>	5.00E-69	F7HFD4_CALJA
Chordata	Mammalian	<i>Fukomys damarensis</i>	5.00E-69	A0A091DLG7_FUKDA
Chordata	Mammalian	<i>Aotus nancymaae</i>	5.10E-69	A0A2K5DVJ3_AOTNA
Chordata	Mammalian	<i>Homo sapiens</i>	5.30E-69	RPGF1_HUMAN
Chordata	Mammalian	<i>Gorilla gorilla gorilla</i>	7.10E-69	G3QYV5_GORGO
Chordata	Mammalian	<i>Pan paniscus</i>	7.10E-69	A0A2R9C0A4_PANPA
Chordata	Mammalian	<i>Nomascus leucogenys</i>	7.40E-69	G1RQQ6_NOMLE
Chordata	Mammalian	<i>Propithecus coquereli</i>	7.50E-69	A0A2K6EGH8_PROCO



**Table S1. List of C3G orthologs identified with HMMER (continuation).**

Clade	Subclade	Species	E-value	UniProt ID
Chordata	Mammalian	<i>Pan troglodytes</i>	8.10E-69	A0A2I3SBV7_PANTR
Chordata	Mammalian	<i>Pongo abelii</i>	9.10E-69	H2PTR4_PONAB
Chordata	Mammalian	<i>Pteropus alecto</i>	9.90E-69	L5K9X4_PTEAL
Chordata	Mammalian	<i>Cercocebus atys</i>	1.40E-68	A0A2K5L7S3_CERAT
Chordata	Mammalian	<i>Macaca fascicularis</i>	1.40E-68	G7PRA8_MACFA
Chordata	Mammalian	<i>Macaca mulatta</i>	1.40E-68	F7HBH8_MACMU
Chordata	Mammalian	<i>Macaca nemestrina</i>	1.40E-68	A0A2K6BHW4_MACNE
Chordata	Mammalian	<i>Mandrillus leucophaeus</i>	1.40E-68	A0A2K5Z0K2_MANLE
Chordata	Mammalian	<i>Papio anubis</i>	1.40E-68	A0A2I3LVR3_PAPAN
Chordata	Mammalian	<i>Chlorocebus sabaeus</i>	1.80E-68	A0A0D9RQV8_CHLSB
Chordata	Mammalian	<i>Heterocephalus glaber</i>	7.50E-68	G5AT49_HETGA
Chordata	Mammalian	<i>Oryctolagus cuniculus</i>	7.80E-68	G1SCW1_RABIT
Chordata	Mammalian	<i>Sus scrofa</i>	1.70E-67	F1S0W0_PIG
Chordata	Mammalian	<i>Colobus angolensis palliatus</i>	1.80E-67	A0A2K5HW20_COLAP
Chordata	Mammalian	<i>Rhinopithecus bieti</i>	3.20E-67	A0A2K6LIA8_RHIBE
Chordata	Mammalian	<i>Rhinopithecus roxellana</i>	3.30E-67	A0A2K6RLQ3_RHIRO
Chordata	Mammalian	<i>Mesocricetus auratus</i>	2.50E-54	A0A3Q0DGG2_MESAU
Chordata	Mammalian	<i>Otolemur garnettii</i>	1.70E-46	H0WPA5_OTOGA
Chordata	Aves	<i>Columba livia</i>	1.10E-71	A0A2I0M3R2_COLL1
Chordata	Aves	<i>Egretta garzetta</i>	1.50E-71	A0A091JHC9_EGRGA
Chordata	Aves	<i>Ficedula albicollis</i>	1.50E-71	U3JFG8_FICAL
Chordata	Aves	<i>Manacus vitellinus</i>	1.50E-71	A0A093S2Q0_9PASS
Chordata	Aves	<i>Patagioenas fasciata monilis</i>	1.50E-71	A0A1V4J7G7_PATFA
Chordata	Aves	<i>Corvus brachyrhynchos</i>	1.60E-71	A0A091F9Y5_CORBR
Chordata	Aves	<i>Lonchura striata domestica</i>	1.60E-71	A0A218V6J9_9PASE
Chordata	Aves	<i>Charadrius vociferus</i>	1.90E-71	A0A0A0AYN1_CHAVO
Chordata	Aves	<i>Opisthocomus hoazin</i>	2.10E-71	A0A091V8B4_OPIHO
Chordata	Aves	<i>Anas platyrhynchos</i>	2.30E-71	U3J7C7_ANAPL
Chordata	Aves	<i>Aptenodytes forsteri</i>	2.30E-71	A0A087RGZ7_APTFO
Chordata	Aves	<i>Struthio camelus australis</i>	2.30E-71	A0A093HB10_STRCA
Chordata	Aves	<i>Cuculus canorus</i>	2.60E-71	A0A091GJC1_9AVES
Chordata	Aves	<i>Colinus virginianus</i>	3.10E-71	A0A226P0F5_COLVI
Chordata	Aves	<i>Calypte anna</i>	3.20E-71	A0A091HS62_CALAN
Chordata	Aves	<i>Nipponia nippon</i>	3.20E-71	A0A091UPS4_NIPNI
Chordata	Aves	<i>Dryobates pubescens</i>	7.10E-71	A0A093GB51_DRYPU
Chordata	Aves	<i>Callipepla squamata</i>	7.80E-71	A0A226NP31_CALSU
Chordata	Aves	<i>Tinamus guttatus</i>	9.80E-71	A0A099ZFB2_TINGU
Chordata	Aves	<i>Meleagris gallopavo</i>	1.60E-70	G1N2I3_MELGA
Chordata	Aves	<i>Taeniopygia guttata</i>	1.70E-70	H0Z3S0_TAEGU
Chordata	Aves	<i>Gallus gallus</i>	1.80E-70	E1BS61_CHICK

**Table S1. List of C3G orthologs identified with HMMER (continuation).**

Clade	Subclade	Species	E-value	UniProt ID
Chordata	Aves	<i>Amazona aestiva</i>	2.90E-35	A0A0Q3X9J0_AMAAE
Chordata	Crocodylia	<i>Alligator mississippiensis</i>	2.70E-67	A0A151MFQ4_ALLMI
Chordata	Crocodylia	<i>Alligator sinensis</i>	3.40E-67	A0A3Q0GRB5_ALLSI
Chordata	Testudines	<i>Pelodiscus sinensis</i>	4.60E-67	K7FMX2_PELSI
Chordata	Amphibia	<i>Xenopus tropicalis</i>	9.40E-65	F6YE36_XENTR
Chordata	Amphibia	<i>Xenopus laevis</i>	4.30E-64	Q6GP24_XENLA
Chordata	Actinopteri	<i>Pygocentrus nattereri</i>	3.20E-65	A0A3B4ED60_PYGNA
Chordata	Actinopteri	<i>Paramormyrops kingsleyae</i>	6.80E-65	A0A3B3QYD7_9TELE
Chordata	Actinopteri	<i>Ictalurus punctatus</i>	1.10E-64	A0A2D0T177_ICTPU
Chordata	Actinopteri	<i>Lepisosteus oculatus</i>	1.60E-63	W5M9X1_LEPOC
Chordata	Actinopteri	<i>Stegastes partitus</i>	3.80E-63	A0A3B4ZZA1_9TELE
Chordata	Actinopteri	<i>Salmo salar</i>	4.00E-63	A0A1S3M1U1_SALSA
Chordata	Actinopteri	<i>Danio rerio</i>	4.10E-63	E7F4B2_DANRE
Chordata	Actinopteri	<i>Oreochromis niloticus</i>	8.40E-62	I3JKA7_ORENI
Chordata	Actinopteri	<i>Pundamilia nyererei</i>	1.10E-61	A0A3B4F2N7_9CICH
Chordata	Actinopteri	<i>Scophthalmus maximus</i>	1.40E-61	A0A2U9CIG2_SCOMX
Chordata	Actinopteri	<i>Seriola dumerili</i>	1.50E-61	A0A3B4TXG4_SERDU
Chordata	Actinopteri	<i>Seriola lalandi dorsalis</i>	1.50E-61	A0A3B4XWR2_SERLL
Chordata	Actinopteri	<i>Takifugu rubripes</i>	1.50E-61	H2VCG7_TAKRU
Chordata	Actinopteri	<i>Tetraodon nigroviridis</i>	2.20E-61	H3CPY5_TETNG
Chordata	Actinopteri	<i>Austrofundulus limnaeus</i>	1.70E-60	A0A2I4B2N0_9TELE
Chordata	Actinopteri	<i>Xiphophorus maculatus</i>	2.70E-60	A0A3B5R6T1_XIPMA
Chordata	Actinopteri	<i>Poecilia latipinna</i>	3.20E-60	A0A3B3UA39_9TELE
Chordata	Actinopteri	<i>Poecilia formosa</i>	3.60E-60	A0A087XL47_POEFO
Chordata	Actinopteri	<i>Poecilia mexicana</i>	1.10E-59	A0A3B3XWA7_9TELE
Chordata	Actinopteri	<i>Xiphophorus couchianus</i>	2.20E-59	A0A3B5L1T6_9TELE
Chordata	Actinopteri	<i>Oryzias melastigma</i>	4.00E-59	A0A3B3DDP5_ORYME
Chordata	Actinopteri	<i>Oryzias latipes</i>	7.30E-59	H2LGG7_ORYLA
Chordata	Actinopteri	<i>Oncorhynchus mykiss</i>	3.50E-57	A0A060WGT3_ONCMY
Chordata	Actinopteri	<i>Periophthalmus magnuspinnatus</i>	1.70E-55	A0A3B3ZLX1_9GOBI
Chordata	Actinopteri	<i>Scleropages formosus</i>	1.80E-46	A0A0P7V340_9TELE
Chordata	Actinopteri	<i>Astyanax mexicanus</i>	6.10E-41	A0A3B1KBI6_ASTMX
Chordata	Actinopteri	<i>Gasterosteus aculeatus</i>	1.90E-40	G3P9F6_GASAC
Chordata	Coelacanthiformes	<i>Latimeria chalumnae</i>	2.90E-58	H3BDJ7_LATCH
Arthropoda	Insecta	<i>Ooceraea biroi</i>	1.70E-58	A0A026WC94_OOCBI
Arthropoda	Insecta	<i>Cyphomyrmex costatus</i>	2.20E-58	A0A195CM95_9HYME
Arthropoda	Insecta	<i>Trachymyrmex zeteki</i>	2.80E-58	A0A151XAU3_9HYME
Arthropoda	Insecta	<i>Acromyrmex echinatio</i>	3.00E-58	F4WKD0_ACREC
Arthropoda	Insecta	<i>Trachymyrmex cornetzi</i>	3.00E-58	A0A195DIY2_9HYME
Arthropoda	Insecta	<i>Atta cephalotes</i>	3.10E-58	A0A158NQI2_ATTCE

**Table S1. List of C3G orthologs identified with HMMER (continuation).**

Clade	Subclade	Species	E-value	UniProt ID
Arthropoda	Insecta	<i>Atta colombica</i>	3.10E-58	A0A195BSS5_9HYME
Arthropoda	Insecta	<i>Trachymyrmex septentrionalis</i>	3.30E-58	A0A195EQG7_9HYME
Arthropoda	Insecta	<i>Dufourea novaeangliae</i>	4.30E-58	A0A154PJF2_9HYME
Arthropoda	Insecta	<i>Nasonia vitripennis</i>	4.30E-58	K7IZN6_NASVI
Arthropoda	Insecta	<i>Melipona quadrifasciata</i>	5.40E-58	A0A0N0BBJ6_9HYME
Arthropoda	Insecta	<i>Camponotus floridanus</i>	8.20E-58	E2AFG9_CAMFO
Arthropoda	Insecta	<i>Apis mellifera</i>	1.00E-57	A0A088ACI4_APIME
Arthropoda	Insecta	<i>Habropoda laboriosa</i>	3.20E-57	A0A0L7QYX3_9HYME
Arthropoda	Insecta	<i>Culex quinquefasciatus</i>	2.20E-55	B0XIC5_CULQU
Arthropoda	Insecta	<i>Zootermopsis nevadensis</i>	4.10E-55	A0A067QTZ8_ZOONE
Arthropoda	Insecta	<i>Aedes aegypti</i>	1.10E-54	Q16I23_AEDAE
Arthropoda	Insecta	<i>Pediculus humanus subsp. corporis</i>	7.20E-54	E0VHU1_PEDHC
Arthropoda	Insecta	<i>Anopheles albimanus</i>	2.50E-53	A0A182FT62_ANOAL
Arthropoda	Insecta	<i>Anopheles maculatus</i>	3.40E-53	A0A182SYB4_9DIPT
Arthropoda	Insecta	<i>Anopheles melas</i>	6.00E-53	A0A182U6F6_9DIPT
Arthropoda	Insecta	<i>Anopheles arabiensis</i>	7.80E-53	A0A182IDK1_ANOAR
Arthropoda	Insecta	<i>Anopheles gambiae</i>	1.20E-52	Q7PSV2_ANOGA
Arthropoda	Insecta	<i>Anopheles funestus</i>	1.30E-52	A0A182RVX9_ANOFN
Arthropoda	Insecta	<i>Anopheles minimus</i>	1.30E-52	A0A182W8D2_9DIPT
Arthropoda	Insecta	<i>Anopheles epiroticus</i>	2.40E-52	A0A182PL47_9DIPT
Arthropoda	Insecta	<i>Anopheles merus</i>	2.70E-52	A0A182USQ8_ANOME
Arthropoda	Insecta	<i>Anopheles quadriannulatus</i>	4.30E-52	A0A182WRA8_ANOQN
Arthropoda	Insecta	<i>Anopheles culicifacies</i>	6.20E-52	A0A182MK97_9DIPT
Arthropoda	Insecta	<i>Anopheles dirus</i>	6.90E-52	A0A182NVF5_9DIPT
Arthropoda	Insecta	<i>Anopheles atroparvus</i>	7.20E-52	A0A182ZS8_9DIPT
Arthropoda	Insecta	<i>Anopheles stephensi</i>	1.20E-51	A0A182Y4W6_ANOST
Arthropoda	Insecta	<i>Drosophila sechellia</i>	4.40E-51	B4I052_DROSE
Arthropoda	Insecta	<i>Acyrtosiphon pisum</i>	6.20E-51	J9K728_ACYPI
Arthropoda	Insecta	<i>Drosophila virilis</i>	3.00E-50	B4M6Z9_DROVI
Arthropoda	Insecta	<i>Anopheles farauti</i>	3.10E-50	A0A182QIF9_9DIPT
Arthropoda	Insecta	<i>Drosophila mojavensis</i>	7.70E-50	A0A0Q9XPE3_DROMO
Arthropoda	Insecta	<i>Drosophila melanogaster</i>	2.10E-49	C3G_DROME
Arthropoda	Insecta	<i>Dendroctonus ponderosae</i>	5.20E-49	U4UA26_DENPD
Arthropoda	Insecta	<i>Drosophila ananassae</i>	8.10E-49	A0A0P9BUM9_DROAN
Arthropoda	Insecta	<i>Rhodnius prolixus</i>	8.50E-49	T1HVL0_RHOPR
Arthropoda	Insecta	<i>Drosophila willistoni</i>	3.00E-48	B4NCC9_DROWI
Arthropoda	Insecta	<i>Glossina austeni</i>	6.70E-48	A0A1A9V0K4_GLOAU
Arthropoda	Insecta	<i>Drosophila ficusphila</i>	8.70E-48	A0A1W4VEY8_DROFC
Arthropoda	Insecta	<i>Glossina pallidipes</i>	1.70E-47	A0A1B0A8G7_GLOPL
Arthropoda	Insecta	<i>Drosophila busckii</i>	1.80E-47	A0A0M4ESF0_DROBS

**Table S1. List of C3G orthologs identified with HMMER (continuation).**

Clade	Subclade	Species	E-value	UniProt ID
Arthropoda	Insecta	<i>Glossina morsitans morsitans</i>	1.80E-47	A0A1B0G205_GLOMM
Arthropoda	Insecta	<i>Tribolium castaneum</i>	2.80E-47	A0A139WN71_TRICA
Arthropoda	Insecta	<i>Glossina fuscipes fuscipes</i>	3.60E-47	A0A1A9XVI3_GLOFF
Arthropoda	Insecta	<i>Glossina palpalis gambiensis</i>	3.60E-47	A0A1B0B3W8_9MUSC
Arthropoda	Insecta	<i>Phlebotomus papatasi</i>	1.80E-45	A0A1B0D3V9_PHLPP
Arthropoda	Insecta	<i>Anopheles coluzzii</i>	2.20E-45	A0A182LAB5_9DIPT
Arthropoda	Insecta	<i>Glossina brevipalpis</i>	2.80E-45	A0A1A9W973_9MUSC
Arthropoda	Insecta	<i>Drosophila pseudoobscura pseudoobscura</i>	1.40E-44	A0A0R3NYS3_DROPS
Arthropoda	Insecta	<i>Megaselia scalaris</i>	3.10E-44	T1GPF1_MEGSC
Arthropoda	Insecta	<i>Stomoxys calcitrans</i>	1.20E-43	A0A118PIT5_STOCA
Arthropoda	Insecta	<i>Drosophila persimilis</i>	1.30E-43	B4H2P5_DROPE
Arthropoda	Insecta	<i>Musca domestica</i>	1.40E-42	A0A118M2D7_MUSDO
Arthropoda	Insecta	<i>Diaphorina citri</i>	2.80E-41	A0A1S3D9U7_DIACI
Arthropoda	Insecta	<i>Clunio marinus</i>	5.90E-41	A0A1J1ILN1_9DIPT
Arthropoda	Insecta	<i>Aedes albopictus</i>	3.20E-29	A0A182H652_AEDAL
Arthropoda	Insecta	<i>Heliothis virescens</i>	5.30E-28	A0A2A4K8S8_HELVI
Arthropoda	Insecta	<i>Danaus plexippus plexippus</i>	9.20E-28	A0A212F3Y3_DANPL
Arthropoda	Insecta	<i>Papilio machaon</i>	6.70E-27	A0A0N1IGL6_PAPMA
Arthropoda	Insecta	<i>Papilio xuthus</i>	1.00E-26	A0A194PKK0_PAPXU
Arthropoda	Insecta	<i>Operophtera brumata</i>	9.00E-21	A0A0L7LK86_9NEOP
Arthropoda	Insecta	<i>Harpegnathos saltator</i>	7.30E-18	E2C8C0_HARSA
Arthropoda	Insecta	<i>Drosophila simulans</i>	7.10E-09	B4R5P5_DROSI
Arthropoda	Insecta	<i>Cryptotermes secundus</i>	4.60E-07	A0A2J7PJ17_9NEOP
Arthropoda	Arachnida	<i>Stegodyphus mimosarum</i>	4.70E-49	A0A087TED5_9ARAC
Arthropoda	Arachnida	<i>Ixodes scapularis</i>	1.10E-37	B7QKE9_IXOSC
Arthropoda	Arachnida	<i>Tetranychus urticae</i>	1.30E-31	T1JUM2_TETUR
Arthropoda	Arachnida	<i>Tropilaelaps mercedesae</i>	2.30E-28	A0A1V9XBN6_9ACAR
Arthropoda	Branchiopoda	<i>Daphnia magna</i>	1.30E-20	A0A164RHE6_9CRUS
Arthropoda	Chilopoda	<i>Strigamia maritima</i>	9.90E-53	T1J7M1_STRMM
Nematoda	Nematoda	<i>Soboliphyme baturini</i>	4.70E-12	A0A183J2N6_9BILA
Nematoda	Nematoda	<i>Trichuris muris</i>	2.90E-07	A0A0N5E0P3_TRIMR
Nematoda	Nematoda	<i>Trichinella pseudospiralis</i>	2.60E-06	A0A0V0XNF5_TRIPS
Nematoda	Nematoda	<i>Trichinella britovi</i>	2.90E-06	A0A0V1CRX8_TRIBR
Nematoda	Nematoda	<i>Trichinella murrelli</i>	3.10E-06	A0A0V0TPS4_9BILA
Nematoda	Nematoda	<i>Trichinella spiralis</i>	3.30E-06	E5S397_TRISP
Nematoda	Nematoda	<i>Trichinella patagoniensis</i>	3.40E-06	A0A0V0ZN21_9BILA
Nematoda	Nematoda	<i>Trichinella nativa</i>	3.90E-06	A0A0V1KZS7_9BILA
Nematoda	Nematoda	<i>Trichinella nelsoni</i>	7.50E-06	A0A0V0S5D3_9BILA
Nematoda	Nematoda	<i>Trichinella papuae</i>	2.70E-05	A0A0V1M5Z5_9BILA

**Table S1. List of C3G orthologs identified with HMMER (continuation).**

Clade	Subclade	Species	E-value	UniProt ID
Nematoda	Nematoda	<i>Trichinella zimbabwensis</i>	4.30E-05	A0A0V1HAM0_9BILA
Brachiopoda	Brachiopoda	<i>Lingula unguis</i>	3.30E-35	A0A1S3KDD6_LINUN
Mollusca	Gastropoda	<i>Biomphalaria glabrata</i>	2.80E-27	A0A2C9JWL9_BIOGL
Mollusca	Bivalvia	<i>Mizuhopecten yessoensis</i>	2.30E-44	A0A210QQ99_MIZYE
Mollusca	Bivalvia	<i>Crassostrea gigas</i>	2.80E-41	K1RBH7_CRAGI
Mollusca	Cephalopoda	<i>Octopus bimaculoides</i>	6.80E-43	A0A0L8FY11_OCTBM
Cnidaria	Cnidaria	<i>Stylophora pistillata</i>	9.60E-11	A0A2B4RQI3_STYPI
Echinodermata	Echinodermata	<i>Strongylocentrotus purpuratus</i>	7.00E-41	W4YR90_STRPU
Echinodermata	Echinodermata	<i>Stichopus japonicus</i>	7.30E-28	A0A2G8KM08_STIJA
Tardigrada	Tardigrada	<i>Ramazzottius varieornatus</i>	2.30E-17	A0A1D1W6U7_RAMVA

**Table S2. Human cancer-associated somatic missense mutations in the AIR-CBR (residues 548-584) of C3G described in the COSMIC database <sup>a</sup>.**

Mutation (Amino acid) <sup>b</sup>	Primary Tissue	Cancer Type	sample ID	References
N548S	Large intestine	Colorectal adenocarcinoma	TCGA-AA-A00N-01	(80)
Y554H	Lymphoid	Follicular lymphoma	LPJ021	(47)
M555K	Lymphoid	Diffuse large B-cell lymphoma	DLBCL-PatientM	(48)
P564L	Large intestine	Colorectal adenocarcinoma	T352983	(81)
Y570N	Lymphoid	Diffuse large B-cell lymphoma	DLBCL-PatientM	(48)
Q571H	Lung	Small cell carcinoma	S02139	(82)
T572M	Small intestine	Adenoma	66	(83)
T572M	Lung	Adenocarcinoma	LUAD-RT-S01702	(84)
E576K	Large intestine / Colon	Adenocarcinoma	587332	(85)
Y579C	Thyroid	Neoplasm	PTC-77C	n/a
K584N	Large intestine / Colon	Adenocarcinoma	TCGA-AA-3966-01	(80)
K584N	Stomach	Adenocarcinoma	TCGA-BR-6454-01	n/a

<sup>a</sup> <https://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=RAPGEF1>

<sup>b</sup> Residue numbers according to the Uniprot entry Q13905.

**Table S3. Constructs of human C3G for expression in mammalian cells.**

Construct name	AA limits <sup>a</sup>	Vector	N-term tag	C-term tag
HA-C3G	4-1077	pCEFLHA	HA	-
HA-C3G- $\Delta$ NTD	246-1077	pCEFLHA	HA	-
HA-P1-P4-REM-Cdc25HD	277-1077	pCEFLHA	HA	-
HA-P2-P4-REM-Cdc25HD	292-1077	pCEFLHA	HA	-
HA-P3-P4-REM-Cdc25HD	465-1077	pCEFLHA	HA	-
HA-P4-REM-Cdc25HD	550-1077	pCEFLHA	HA	-
HA-REM-Cdc25HD	670-1077	pCEFLHA	HA	-
HA-SH3b	246-669	pCEFLHA	HA	-
HA-REM	670-814	pCEFLHA	HA	-
HA-Cdc25HD	815-1077	pCEFLHA	HA	-
HA-P3-P4-REM	465-814	pCEFLHA	HA	-
HA-P4-REM	550-814	pCEFLHA	HA	-
FLAG-NTD	4-245	pCEFLAG	FLAG	-
C-C3G	4-1077	pcDNA3	CFP	-
C3G-V	4-1077	pcDNA3	-	Venus
C-C3G-V	4-1077	pcDNA3	CFP	Venus
C-C3G- $\Delta$ NTD-V	246-1077	pcDNA3	CFP	Venus
C-C3G- $\Delta$ Cdc25HD-V	4-814	pcDNA3	CFP	Venus
C-C3G- $\Delta$ REM-Cdc25HD-V	4-669	pcDNA3	CFP	Venus
C3G-mEGFP	4-1077	pEF1-mEGFP-N	-	mEGFP
C3G-mEGFP-CAAX	4-1077	pEF1-mEGFP-CAAX-N	-	mEGFP-CAAX
REM-Cdc25HD-mEGFP	670-1077	pEF1-mEGFP-N	-	mEGFP
REM-mEGFP	670-814	pEF1-mEGFP-N	-	mEGFP
Cdc25HD-mEGFP	815-1077	pEF1-mEGFP-N	-	mEGFP
C3G-mEGFP (Lenti)	4-1077	pLenti-C-mEGFP-IRES-BSD	-	mEGFP

<sup>a</sup> Numbers according to the Uniprot entry Q13905

**Table S4. Oligonucleotides used to create constructs of C3G.**

Name <sup>a</sup>	Sequence (5' to 3')
C3Gh-004-KKB-For-BamHI	TGAGGTACCGCCGCCACCATGGGATCCGACTCTCAGCGTTCTCATC
C3Gh-004-For-BglII	TGAGAATTCAGATCTGACTCTCAGCGTTCTCATC
C3Gh-004-For -NcoI	CGGGGTACCATGGACTCTCAGCGTTCTCAT
C3Gh-246For -BglII	TGAGAATTCAGATCTAAGACGACTGGGATGTCACAG
C3Gh-246For-NcoI	CGGGGTACCATGGGCAAGACGACTGGGATGTCACAGTC
C3Gh-274-For-NcoI	CGGGGTACCATGGGCATTCGGGTGGTTGATAATGG
C3Gh-277-For-BglII	TGAGAATTCAGATCTGTGGTTGATAATGGTCTCTC
C3Gh-292-For-BglII	TGAGAATTCAGATCTCAGTCGGCGCCGTCC
C3Gh-372-For-NcoI	CATATGGGCAGCTCCATGGGGCAGTGTCTCCCGG
C3Gh-454-For-NdeI	TGAGAATTCATATGCCCTGCTCTCCCGAGAAG
C3Gh-465-For-BglII	TGAGAATTCAGATCTGCCCTCCAGACGGCGG
C3Gh-501-For-NcoI	CATATGGGCAGCTCCATGGTCCGTCCCCACGCGC
C3Gh-530-For-NdeI	TGAGAATTCATATGGCTCCTGAGTCAACCGGT
C3Gh-537-For-NdeI	CGGGGTACCCATATGCCAGAAAACCACCTCCTCTACC
C3Gh-545-For-NdeI	TTTTTCAGGGATCACATATGGAGAAGAAAAACAAACACATG
C3Gh-550-For-BglII	TGAGAATTCAGATCTCACATGCTGGCCTACATGC
C3Gh-579-For-NcoI	CATATGGGCAGCTCCATGGTACCAGCAGAAGAACAAGCTC
C3Gh-602-For-NdeI	TGAGAATTCATATGGTGCAGGAGCTGGCCCC
C3Gh-670-For-NcoI	CGGGGTACCATGGCCCTCATTTGACCACAACG
C3Gh-670-For-BglII	TGAGAATTCAGATCTCTCATTTGACCACAACGAAATTATG
C3Gh-815-For-NcoI	CGGGGTACCATGGCACTCAGGTGTGCCACCTCC
C3Gh-815-For-BglII	TGAGAATTCAGATCTCTCAGGTGTGCCACCTCC
C3Gh-245-Stop-Rev-NotI	GAATTCGCGGCCGCTAGTTCAGGATCTCTACCTCGCG
C3Gh-646-Stop-Rev-BamHI	GCCGAATTCGGATCCCTAACTGCCGTCTCTGCTGTCCTTC
C3Gh-669-Stop-Rev-NotI	GAATTCGCGGCCGCTAGGACAGCTCGTCCACTTCTCT
C3Gh-814-Stop-Rev-NotI	GAATTCGCGGCCGCTATAGCTTCTTCTGGTCCACC
C3Gh-814-noStop-Rev-NotI	GAATTCGCGGCCGCTAGCTTCTTCTGGTCCAC
C3Gh-1077-Stop-Rev-BamHI	CGGGGTACCGGATCCTAGGTCTTCTCTTCCCGG
C3Gh-1077-Stop-Rev-NotI	GAATTCGCGGCCGCTAGGTCTTCTCTTCCCGGTC
C3Gh-1077-noStop-Rev-NotI	GAATTCGCGGCCGCGGTCTTCTCTTCCCGGTC
C3Gh-1077-noStop-Rev-XhoI	GCGAATTCCTCGAGGGTCTTCTCTTCCCGGTC

<sup>a</sup> Numbers denote the first or last amino acid position coded by the forward (For) or reverse (Rev) primers, respectively.

**Table S5. Oligonucleotides used to create constructs of C3G (continuation).**

<b>Name<sup>a</sup></b>	<b>Sequence (5' to 3')</b>
C3Gh-371-Stop-For <sup>b</sup>	GGACAGGGACAGTTAGCAGTGCTCCCGG
C3Gh-371-Stop-Rev <sup>b</sup>	CCGGGAGCACTGCTAACTGTCCCTGTCC
C3Gh-500-Stop-For <sup>b</sup>	CAGCCCCGATCCCATAGGTCCCCTACGCGC
C3Gh-500-Stop-Rev <sup>b</sup>	GCGCGTAGGGGACCTATGGGATCGGGGCTG
C3Gh-536-Stop-For <sup>b</sup>	GAGTCAACCGGTGACTAAGAAAAACCACCTCCTC
C3Gh-536-Stop-Rev <sup>b</sup>	GAGGAGGTGGTTTTTCTTAGTCACCGGTTGACTC
C3Gh-560-Stop-For <sup>b</sup>	CAGTTGCTGGAGGACTAGTCGGAGCCGCAGCCCTC
C3Gh-560-Stop-Rev <sup>b</sup>	GAGGGCTGCGGCTCCGACTAGTCCTCCAGCAACTG
C3Gh-569-Stop-For <sup>b</sup>	GCAGCCCTCTATGTTCTAGCAGACGCCACAGAACGAG
C3Gh-569-Stop-Rev <sup>b</sup>	CTCGTTCTGTGGCGTCTGCTAGAACATAGAGGGCTGC
C3Gh-578-Stop-For <sup>b</sup>	GAACGAGCACATCTAGCAGCAGAAGAACAAG
C3Gh-578-Stop-Rev <sup>b</sup>	CTTGTTCTTCTGCTGCTAGATGTGCTCGTTC

<sup>a</sup> Numbers denote the first or last amino acid position coded by the forward (For) or reverse (Rev)primers, respectively.

<sup>b</sup> These primers were used to introduce stop codons after the indicated positions by site directed mutagenesis.



**Table S5. Constructs of human C3G for expression in *E. coli*.**

Construct name	AA limits	Vector	N-term tag	C-term tag
C3G full-length	4-1077	pGEX-4T3-2xTEV-cHis	GST-TEV	TEV-His
C3G- $\Delta$ NTD	246-1077	pGEX-4T3-2xTEV-cHis	GST-TEV	TEV-His
C3G-454-1077	454-1077	pGEX-4T3-2xTEV-cHis	GST-TEV	TEV-His
C3G-530-1077	530-1077	pGEX-4T3-2xTEV-cHis	GST-TEV	TEV-His
Cdc25HD	815-1077	pETEV15b-NcoI	His-TEV	-
GST-SH3b	274-646	pGEX-4T3-TEV	GST-TEV	-
GST-274-578	274-578	pGEX-4T3-TEV	GST-TEV	-
GST-274-500	274-500	pGEX-4T3-TEV	GST-TEV	-
GST-274-371	274-371	pGEX-4T3-TEV	GST-TEV	-
GST-372-646	372-646	pGEX-4T3-TEV	GST-TEV	-
GST-501-646	501-646	pGEX-4T3-TEV	GST-TEV	-
GST-579-646	579-646	pGEX-4T3-TEV	GST-TEV	-
GST-501-578	501-578	pGEX-4T3-TEV	GST-TEV	-
GST-501-536	501-536	pGEX-4T3-TEV	GST-TEV	-
GST-537-646	537-646	pGEX-4T3-TEV	GST-TEV	-
His-537-646	537-646	pETEV15b	His-TEV	-
GST-537-588	537-588	pGEX-4T3-TEV	GST-TEV	-
GST-537-578	537-578	pGEX-4T3-TEV	GST-TEV	-
GST-537-569	537-569	pGEX-4T3-TEV	GST-TEV	-
GST-537-560	537-560	pGEX-4T3-TEV	GST-TEV	-
GST-545-646	545-646	pGEX-4T3-TEV	GST-TEV	-
GST-545-569	545-569	pGEX-4T3-TEV	GST-TEV	-
GST-545-560	545-560	pGEX-4T3-TEV	GST-TEV	-

**Table S6. Oligonucleotides used for site directed mutagenesis of C3G.**

Name	Sequence (5' to 3')
C3Gh-PP4A-For	GGCCCCG <b>GCGGCCG</b> CCCGCAGCCCCCGCGCAGCGGCAG
C3Gh-PP4A-Rev	CTGCCGCTGCGCGGGGGCTGCGGCGGCCCGGGGCC
C3Gh-K546E/K547E-For	CACCTCCTCTACCA <b>GAGG</b> AGGAAAACAAACACATGCTG
C3Gh-K546E/K547E-Rev	CAGCATGTGTTTGTTCCTCCTCTGGTAGAGGAGGTG
C3Gh-H550E/M551R-For	CCAGAGAAGAAAAACAA <b>AGAG</b> GGCTGGCCTACATGCAG
C3Gh-H550E/M551R-Rev	CTGCATGTAGGCCAGCCTCTCTTTGTTTTCTTCTCTGG
C3Gh-H550E-For	CCAGAGAAGAAAAACAA <b>AGAT</b> GCTGGCCTACATGCAG
C3Gh-H550E-Rev	CTGCATGTAGGCCAGCATCTCTTTGTTTTCTTCTCTGG
C3Gh-H550A-For	CCAGAGAAGAAAAACAA <b>AGC</b> ATGCTGGCCTACATGCAG
C3Gh-H550A-Rev	CTGCATGTAGGCCAGCATGGCTTTGTTTTCTTCTCTGG
C3Gh-M551R-For	CCAGAGAAGAAAAACAAACACA <b>GG</b> GCTGGCCTACATGCAG
C3Gh-M551R-Rev	CTGCATGTAGGCCAGCCTGTGTTTGTTCCTTCTCTGG
C3Gh-M551A-For	CCAGAGAAGAAAAACAAACAC <b>GC</b> GCTGGCCTACATGCAG
C3Gh-M551A-Rev	CTGCATGTAGGCCAGCGCGTGTTCCTTCTCTCTGG
C3Gh-Y554R-For	CAAACACATGCTGGCC <b>CG</b> CATGCAGTTGCTGGAG
C3Gh-Y554R-Rev	CTCCAGCAACTGCATGCGGGCCAGCATGTGTTTG
C3Gh-Y554A-For	CAAACACATGCTGGCC <b>GC</b> CATGCAGTTGCTGGAG
C3Gh-Y554A-Rev	CTCCAGCAACTGCATGGCGGCCAGCATGTGTTTG
C3Gh-Y554H-For	CAAACACATGCTGGCC <b>C</b> ACATGCAGTTGCTGGAG
C3Gh-Y554H-Rev	CTCCAGCAACTGCATGTGGGCCAGCATGTGTTTG
C3Gh-M555R/Q556R-For	CATGCTGGCCTACAG <b>GGC</b> GTTGCTGGAGGAC
C3Gh-M555R/Q556R-Rev	GTCCTCCAGCAACCGCCTGTAGGCCAGCATG
C3Gh-M555R-For	CATGCTGGCCTACAG <b>GC</b> AGTTGCTGGAGGAC
C3Gh-M555R-Rev	GTCCTCCAGCAACTGCCTGTAGGCCAGCATG
C3Gh-M555A-For	CATGCTGGCCTAC <b>GC</b> AGTTGCTGGAGGAC
C3Gh-M555A-Rev	GTCCTCCAGCAACTGCGCGTAGGCCAGCATG
C3Gh-M555K-For	CATGCTGGCCTAC <b>AA</b> GCAGTTGCTGGAGGAC
C3Gh-M555K-Rev	GTCCTCCAGCAACTGCTTGTAGGCCAGCATG
C3Gh-Q556R-For	CATGCTGGCCTACATGC <b>G</b> GTTGCTGGAGGAC
C3Gh-Q556R-Rev	GTCCTCCAGCAACCGCATGTAGGCCAGCATG
C3Gh-Q556A-For	CATGCTGGCCTACATGC <b>GC</b> GTTGCTGGAGGAC
C3Gh-Q556A-Rev	GTCCTCCAGCAACGCCATGTAGGCCAGCATG

Nucleotides changed are highlighted in red in the sequence of the forward primer

**Table S6. Oligonucleotides used for site directed mutagenesis of C3G (continuation).**

Name	Sequence (5' to 3')
C3Gh-E559R/D560R-For	CTACATGCAGTTGCTG <b>CGGCG</b> CTACTCGGAGCCGCAG
C3Gh-E559R/D560R-Rev	CTGCGGCTCCGAGTAGCGCCGCAGCAACTGCATGTAG
C3Gh-Y561R/S562R-For	GTTGCTGGAGGAC <b>CGCCG</b> GGAGCCGCAGCCC
C3Gh-Y561R/S562R-Rev	GGGCTGCGGCTCCCGGCGGTCTCCAGCAAC
C3Gh-E563R/P564R-For	GGAGGACTACTCG <b>CGGCC</b> GCAGCCCTCTATG
C3Gh-E563R/P564R-Rev	CATAGAGGGCTGCGGCGCGAGTAGTCTCTCC
C3Gh-P566R-For	CTCGGAGCCGCAGC <b>G</b> CTCTATGTTCTACC
C3Gh-P566R-Rev	GGTAGAACATAGAGCGCTGCGGCTCCGAG
C3Gh-F569R/Y570R-For	CCGCAGCCCTCTATG <b>CGCCG</b> CCAGACGCCACAGAAC
C3Gh-F569R/Y570R-Rev	GTTCTGTGGCGTCTGGCGGCGCATAGAGGGCTGCGG
C3Gh-Q571R/T572R-For	CTCTATGTTCTACC <b>GGA</b> GCCACAGAACGAGC
C3Gh-Q571R/T572R-Rev	GCTCGTTCTGTGGCCTCCGGTAGAACATAGAG
C3Gh-Q574R-For	CTACCAGACGCCAC <b>GGA</b> ACGAGCACATC
C3Gh-Q574R-Rev	GATGTGCTCGTTCCGTGGCGTCTGGTAG
C3Gh-E576R-For	CAGACGCCACAGAAC <b>CGGC</b> ACATCTACCAGC
C3Gh-E576R-Rev	GCTGGTAGATGTGCCGTTCTGTGGCGTCTG
C3Gh-Y579R-For	CAGAACGAGCACATC <b>CGCC</b> CAGCAGAAGAACAAG
C3Gh-Y579R-Rev	CTTGTCTTCTGCTGGCGGATGTGCTCGTTCTG
C3Gh-Q581R/K582E-For	GAGCACATCTACCAGC <b>GGG</b> AGAACAAGCTCCTCATG
C3Gh-Q581R/K582E-Rev	CATGAGGAGCTTGTCTCCCGCTGGTAGATGTGCTC
C3Gh-L714R-For	CAGGAAAGATTTGGTG <b>CGG</b> TACTGCGAGGCATTCCTG
C3Gh-L714R-Rev	CAGGAATGCCTCGCAGTACCGCACCAAATCTTTCCTG
C3Gh-E731R-For	CATCTCCCAGAG <b>CGG</b> CTCATCAAGAAGCTGCAGTAC
C3Gh-E731R-Rev	GTACTGCAGCTTCTTGATGAGCCGCTCTGGGGAGATG
C3Gh-F760R-For	GTCAGCAAGAACACGTT <b>CCG</b> CGTGCTGGTACGGGTGG
C3Gh-F760R-Rev	CCACCCGTACCAGCACGCGGAACGTGTTCTTGCTGAC
C3Gh-L774R-For	CTGCCTGGTGGAG <b>CGG</b> ACAGAAGAGATCCTGAAGC
C3Gh-L774R-Rev	GCTTCAGGATCTCTTCTGTCCGCTCCACCAGGCAG
C3Gh-E784R-For	GAAGCTGCTGATG <b>AG</b> ACTGGTCTTCCGCCTGGTGTGC
C3Gh-E784R-Rev	GCACACCAGGCGGAAGACCAGTCTCATCAGCAGCTTC
C3Gh-K803E-For	GTGTGCTCCGG <b>G</b> AGAACATCCTGGACAAGGTGG
C3Gh-K803E-Rev	CCACCTTGTCAGGATGTTCTCCCGGAGCACAC

Nucleotides changed are highlighted in red in the sequence of the forward primer

**Table S7. Oligonucleotides used to create constructs of Rap1b, CrkL, Src, and YopH for expression in *E. coli*.**

<b>Name <sup>a</sup></b>	<b>Sequence (5' to 3')</b>
Rap1b-U001-NdeI	TGAGAATTCCATATGCGTGAGTATAAGCTAGTCGTT
Rap1b-L167-Stop-BamHI	GCCGGATCCCTATCTGTAAATTTGCCGCACTAGGTC
CrkL-U001-NdeI	TGACCATGGCATATGTCTCCGCCAGGTTC
CrkL-L303-Stop-BglII	GCCGTGACAGATCTACTCGTTTTTCATCTGGGTTTTGAG
YopH-U001-NdeI	TGAGAATTCCATATGAACTTATCATTAAAGCGATCTTCATC
YopH-L468-Stop-XhoI	GCCGAATTCCTCGAGCTAGCTATTTAATAATGGTCGCC
Src-U254-EcoRI	GCCTGAGAATTCAGACTCAGGGCCTGGCCAAG
Src-L536-Stop-HindIII	TGAGGCGGCCGCAAGCTTCTAGAGGTTCTCCCCGGGCTGG

<sup>a</sup> Numbers denote the first or last amino acid position coded by the forward (U) or reverse (L) primers, respectively.