Structure of the DOCK2–ELMO1 complex provides insights into regulation of the auto-inhibited state

Leifu Chang1,7, Jing Yang1, Chang Hwa Jo2, Andreas Boland1,8, Ziguo Zhang1, Stephen H. McLaughlin1, Afnan Abu-Thuraia3, Ryan C. Killoran2, Matthew J. Smith2,4,9, Jean-Francois Côté3,5,6,9 & David Barford1

DOCK (dedicator of cytokinesis) proteins are multidomain guanine nucleotide exchange factors (GEFs) for RHO GTPases that regulate intracellular actin dynamics. DOCK proteins share catalytic (DOCKDHR2) and membrane-associated (DOCKDHR1) domains. The structurally-related DOCK1 and DOCK2 GEFs are specific for RAC, and require ELMO (engulfment and cell motility) proteins for function. The N-terminal RAS-binding domain (RBD) of ELMO (ELMORBD) interacts with RHOG to modulate DOCK1/2 activity. Here, we determine the cryo-EM structures of DOCK2–ELMO1 alone, and as a ternary complex with RAC1, together with the crystal structure of a RHOG–ELMO2RBD complex. The binary DOCK2–ELMO1 complex adopts a closed, auto-inhibited conformation. Relief of auto-inhibition to an active, open state, due to a conformational change of the ELMO1 subunit, exposes binding sites for RAC1 on DOCK2DHR2, and RHOG and BAI GPCRs on ELMO1. Our structure explains how up-stream effectors, including DOCK2 and ELMO1 phosphorylation, destabilise the auto-inhibited state to promote an active GEF.

1 MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK. 2 Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, Québec H3T 1J4, Canada. 3 Montreal Institute of Clinical Research (IRCM), Montréal, QC H2W 1R7, Canada. 4 Department of Pathology and Cell Biology, Faculty of Medicine, Université de Montréal, Montréal, QC H3T 1J4, Canada. 5 Department of Biochemistry and Molecular Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC H3C 3J7, Canada. 6 Department of Anatomy and Cell Biology, McGill University, Montréal, QC H3A OC7, Canada. 7 Present address: Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA. 8 Present address: Department of Molecular Biology, Science III, University of Geneva, Geneva, Switzerland. 9 These authors contributed equally: Matthew J. Smith, Jean-Francois Côté. *Email: dbarford@mrc-lmb.cam.ac.uk

https://doi.org/10.1038/s41467-020-17271-9
RO family small GTPases are critical regulators of cell motility, polarity, adhesion, cytoskeletal organization, proliferation, gene expression and apoptosis. These diverse functions are stimulated by the active GTP-bound state of RHO proteins that engage a diverse array of effector proteins, thereby triggering downstream signal transduction pathways. Conversion of these biomolecular switches to the GTP-bound state is controlled by two families of guanine nucleotide exchange factors (GEFs); the Dbl family and the DOCK family. GEFs catalyse the release of bound GDP in exchange for GTP.

Dbl family proteins are a large group of RHO GEFs comprising a catalytic Dbl homology (DH) domain with an adjacent PH domain, within the context of functionally diverse signaling modules. The evolutionary distinct and smaller family of DOCK proteins activates CDC42 and RAC to control cell migration, morphogenesis and phagocytosis, and have been implicated as important components of tumour cell movement and invasion. DOCK proteins exhibit high specificity, in contrast to the Dbl GEFs that often stimulate nucleotide exchange on multiple GTPases in vitro. In humans, the eleven DOCK proteins are organized into four subfamilies encoding multidomain proteins of ~2000 amino acids. DOCK A (DOCK1/DOCK180, DOCK2 and DOCK5) and DOCK B (DOCK3 and DOCK4) subfamilies activate RAC, whereas the DOCK D subfamily (DOCK9/Zizimin1, DOCK10 and DOCK11) activates CDC42, with only DOCK10 also activating RAC. DOCK6 and DOCK7 of the DOCK C (DOCK6, DOCK7 and DOCK8) subfamily are dual specificity GEFs with activity towards both RAC and CDC42 in vivo, a finding recently confirmed for DOCK7 in vitro, whereas DOCK8 was shown to be a CDC42 GEF.

All DOCK proteins contain a catalytic DHR2 domain of ~450 residues situated within their C-terminal region (DOCK DHR2). The DHR2 domain is divergent across the family, with the DHR2 domains of DOCK1 (RAC specific) and DOCK9 (CDC42 specific) sharing only 22% sequence identity. A second region of common similarity is the ~200 residue DHR1 domain located towards the N-terminus (DOCK DHR1). The DHR1 domain of DOCK proteins adopts a C2-like conformation.

**Fig. 1 Overall structure of the DOCK2 – ELMO1 – RAC1 complex.**

(a) Schematic of the domain structures of DOCK2, ELMO1 and RAC1. (b, c) Two views of the cryo-EM map of the DOCK2 – ELMO1 – RAC1 complex with each domain colour-coded. Ribbon representation of the structural models were placed in the cryo-EM density map. The two DOCK2 – ELMO1 – RAC1 protomers of the dimeric complex are indicated by light blue and light yellow backgrounds. The open-conformation of the DOCK2 – ELMO1 – RAC1 complex is shown.
architecture and interacts with PI(3,4,5)P₃ to mediate signalling and membrane localization. DOCK proteins locate primarily to the cytosol, but recruitment to the cell membrane is critical for their roles in cytoskeleton reorganization. DOCK A and B subfamilies incorporate an N-terminal SH3 domain and an extreme C-terminal poly-proline sequence (Fig. 1a). In contrast, the DOCK D subfamily incorporates an N-terminal PH domain, whereas the DOCK C subfamily lacks recognizable SH3 or PH domains.

Stimulation of RAC-induced cytoskeletal reorganization by the DOCK A and B subfamilies is dependent on interactions with ELMO proteins. The DOCK A and B subfamilies have a C-terminal poly-proline sequence (Fig. 1a). ELMONTD itself is a C-terminal poly-proline sequence (Fig. 1a). ELMO proteins are involved in the regulation of GTPase activity. ELMO-mediated GEF activity requires the presence of nucleotide-free RAC1, and nucleotide-free RAC1 binds to the DOCK1–ELMO complex more efficiently and shows higher GEF activity.

Among DOCK proteins, DOCK2 plays multiple roles in regulating immune responses. It is a haematopoietic cell protein which functions downstream of chemokine receptors to control actin reorganization, thus, regulating lymphocyte activation, migration, and morphology. In T cells, DOCK2 acts downstream of the T cell receptor, and recently DOCK2 was implicated in mediating signalling from the FLT3 protein tyrosine kinase and identified as a leukaemia drug target. Through binding to the C-terminal polybasic region of DOCK2, phosphatidic acid stabilizes the recruitment of DOCK2 to the cell membrane to mediate neutrophil chemotaxis.

In vivo, RHOG–GTP interacts with ELMO1 RBD, an interaction that is required for DOCK1-mediated RAC1 activation. Furthermore, the Arl family GTPase Arl4A promotes actin cytoskeleton remodelling, dependent on ELMO1 acting as a membrane localization signal for ELMO1. The Arl family GTPase Arl4A promotes actin cytoskeleton remodelling, dependent on ELMO1 acting as a membrane localization signal for ELMO1. The Arl family GTPase Arl4A promotes actin cytoskeleton remodelling, dependent on ELMO1 acting as a membrane localization signal for ELMO1. The adhesion-type BAI (brain-specific angiogenesis inhibitor) subfamily of GPCRs are up-stream regulators of the DOCK–ELMO–RAC signalling module, controlling engulfment and degradation of apoptotic cells by phagocytosis and promoting myoblast fusion. BAI receptors bind ELMO through their conserved C-terminal region, proposed to become exposed for ELMO-binding upon receptor activation. A recent crystal structure of ELMO2 NTD (lacking ELMO2 PH) revealed that this region of BAII forms an α-helical segment that interacts with ELMO2 EID.

The molecular basis for the GEF activity of DOCK proteins was elucidated from crystal structures of DOCK2 DHR2 and DOCK9 DHR2 in complex with their cognate GTPases RAC and CDC42, respectively, showing that GDP release and discharge of the activated GTP-bound CDC42 and RAC are catalysed by a universal invariant valine residue that functions as a nucleotide sensor. However, the regulation of DOCK proteins and the signal transduction events responsible for their activation are poorly understood. It is not yet known how ELMO proteins regulate DOCK-mediated GEF activity or how effector proteins such as RHOG and BAI receptors regulate DOCK–ELMO GEF activity. To understand the molecular architecture and mechanism of DOCK proteins, we have used cryo-electron microscopy to determine structures of DOCK2–ELMO1 alone and as a ternary complex with nucleotide-free RAC1, and combined this with a crystal structure of ELMO2 RBD in complex with RHOG. The binary DOCK2–ELMO1 complex adopts a closed, auto-inhibited conformation, whereas in the DOCK2–ELMO1–RAC1 ternary complex, DOCK2–ELMO1 adopts an open, active conformation through a conformational change of the ELMO1 subunit. This exposes binding sites for RAC1 on DOCK2 DHR2, with ELMO1 also directly contacting RAC1 to promote binding to DOCK2–ELMO1. Binding sites for RHOG and BAI GPCRs on ELMO1 also become accessible in the active, open conformation. Our study suggests a model for how upstream regulators control DOCK2 activation through a conformational change of ELMO1, thereby relieving autoinhibition.

Results

Cryo-EM structure determination of DOCK2–ELMO1–RAC1. We prepared the human DOCK2–ELMO1 complex using the baculovirus-insect cell expression system (Supplementary Fig. 1a). The molecular weight of the complex of ~600 kDa estimated by size exclusion chromatography (Supplementary Fig. 1b) indicated the formation of a tetramer formed of two DOCK2–ELMO1 protomers. This is consistent with the observation that DOCK2 homo-dimerization is required for DOCK2-mediated RAC activation in vivo and for lymphocyte migration. Examination of raw particle images and reference-free 2D class averages indicated a considerable degree of structural flexibility (Supplementary Fig. 1c–e), limiting the resolution of the overall 3D reconstruction. To improve the EM density map quality, we generated two particle sets, each with one protomer subtracted from the original images. Refinement of the combined particles resulted in an EM density map with an overall resolution of 4.1 Å (Supplementary Fig. 2a, b). The DHR2 domain was the best-resolved feature, with other domains resolved at lower resolution (Supplementary Fig. 2c). To improve the resolution of individual structural segments, we performed focussed 3D classification on individual rigid modules, followed by focussed refinement, and obtained reconstructed maps of the (i) DOCK2 DHR2–RAC1 catalytic module at 3.8 Å resolution, (ii) DOCK2 ARM domain at 4.2 Å, (iii) DOCK2 SH3–ELMO1PH at 4.1 Å, (iv) DOCK2 DHR1–DOCK2 C at 4.6 Å and (v) ELMO1 NTD at 6.2 Å resolution (Supplementary Fig. 2b, c and Supplementary Tables 1–3). From the whole dataset, we selected ~2% of the particles to reconstruct a map representing the entire complex at a resolution of 7.8 Å with twofold symmetry (Supplementary Fig. 3). Reconstruction without imposed symmetry resulted in a very similar map, although at lower resolution of 9.1 Å (Supplementary Fig. 3e).

Figure 1b, c shows a composite map consisting of the individual maps of each domain together with a fitted model.

Structure of the DOCK2–ELMO1–RAC1 ternary complex. The ternary DOCK2–ELMO1–RAC1 dimeric complex adopts an elongated “S”-like shape with pseudo twofold symmetry, measuring 320 Å and 65 Å in the longest and shortest dimensions, respectively (Figs. 1b, c and 2a). Both the DOCK2 and ELMO1 subunits are assembled from a series of modular domains (Fig. 1a). In the complex, these two subunits are arranged in a roughly parallel manner, but because DOCK2 adopts a hook-like structure (Fig. 2b), its N-terminal DOCK2 SH3 domain is positioned to interact with the C-terminal PxxP motif of ELMO1 (Figs. 1b and 2a). Located at the centre of DOCK2–ELMO1–RAC1 is DOCK2 DHR2, dimerized through its A lobe, and associated with nucleotide-free RAC1 through its B and C lobes (Figs. 1b and 2a, c). Side-chains are visible...
Fig. 2 Structure of the DOCK2–ELMO1–RAC1 protomer. a Ribbon representation of the structural assembly of the DOCK2–ELMO1 protomer. DOCK2, ELMO1 and RAC1 are highlighted in cyan, yellow-orange and red backgrounds, respectively. Insert shows the major stable interface between DOCK2 and ELMO1 involving the six α-helical bundle. b Ribbon representation of DOCK2. Domains are labelled and colour-coded according to the domain scheme in Fig. 1a. c Close-up view of the dimer interface formed by lobe A of the DHR2 domains of each DOCK2 monomer (shown in purple and grey-blue). Lobe A, B and C of the DOCK2 DHR2 domains are labelled. d DOCK2ARM is rainbow-coloured from blue (N-terminus) to red (C-terminus). e Close-up view of the interaction interface between DOCK2DHR2, ELMO1PH domain and RAC1. The switch 1 and switch 2 loops of RAC1 are coloured in yellow and green, respectively. f Ribbon representation of ELMO1 colour-coded according to Fig. 1a.
in this region (Supplementary Fig. 4a), allowing atomic model building. The polybasic region C-terminal of DOCK2DHR2 (Fig. 1a) is disordered. The region immediately N-terminal of DOCK2DHR2, and connected to the lobe A TPR repeats of DOCK2DHR2, is a right-handed α-solenoid domain. This is formed of 28 α-helices composed of a mixture of ARM and HEAT repeats, termed DOCK2ARM (Figs. 1 and 2b, d). A DALI search43 revealed the vacuolar protein 8, an ARM repeat protein46, as the closest match (Z score of 11.8, RMSD of 4.4 Å based on 290/568 aligned residues). A short α-helix inserted within the ARM domain, projects out from the main α-solenoid (Fig. 2b). This α-helical insert provides a major site of interaction between DOCK2ARM and DOCK2SH3 (Fig. 2b). Connected to DOCK2ARM is a globular density feature composed of two closely packed domains. A homology model of DOCK2DHR1 based on the crystal structure of DOCK1DHR123 was confidently fitted to the domain closest to DOCK2ARM, consistent with the connectivity of DOCK2DHR1 and DOCK2ARM in the protein sequence (Figs. 1a, c and 2b). The other globular domain (residues 219–389), located between DOCK2DHR1 and DOCK2SH3 is mainly composed of β-strands, consistent with secondary structure predictions. A β-sandwich structure composed of two β-sheets each with four anti-parallel β-strands, predicted using Rosetta47, was fitted to this β-sheet-like cryo-EM density (Figs. 1b, c and 2b). A DALI search45 revealed this β-sandwich domain to be a C2-domain, with the highest structural similarity to the C2A domain of Otoferlin48 (Z score of 4.8). Thus, DOCK2DHR123 and DOCK2C2 share related C2-fold architectures. DOCK2C2 is connected to the N-terminal DOCK2SH3 and its adjacent α-helical domain by a flexible region found to be phosphorylated in multiple mass spectrometry studies (PhosphoSitePlus). We termed this the phosphorylation linker (Fig. 1a, c).

The structure of the DOCK2DHR2–RAC1 module in the context of the entire DOCK2–ELMO1–RAC1 complex (Fig. 2e) is virtually identical to the crystal structure of the isolated DOCK2DHR2–RAC1 complex43, with a RMSD between them of less than 1 Å. DOCK2DHR2 forms the dimer interface of the two DOCK2–ELMO1–RAC1 protomers (Fig. 2c), similar to other DOCK proteins revealed by structures of their DHR2 domains that are dimers, for example DOCK9DHR2 (Ref. 44), and as we previously published for DOCK2DHR2 (Ref. 43). Lobe A contains three TPR motifs with TPR3 forming the main DOCK2 dimerization interface. We produced a monomeric complex of DOCK2–ELMO1 by mutating the dimerization interface of lobe A of DOCK2DHR2 (Supplementary Fig. 5a–c). Cryo-EM examination showed that monomers are less stable. About 90% of particles were disassembled, with the remaining particles reconstructed to generate a low-resolution map (Supplementary Fig. 5d, e). This result indicates that DOCK2 dimerization stabilizes the complex. Although there are no clear functional explanations for why DOCK proteins dimerize, one possibility would be to enable cooperativity between the two catalytic sites of the DOCK dimer.

ELMO1 comprises ELMO1NTD, ELMO1PH and C-terminal PxxP motif (Fig. 1a). ELMO1NTD and ELMO1PH are connected by a single α-helix termed the hinge helix (Fig. 2f). A flexible elbow hinge at the C-terminus of the hinge helix allows mobility of ELMO1NTD to adopt open- and closed-conformations (Fig. 3a and Supplementary Fig. 3a, b). 3D classification showed that in the DOCK2–ELMO1–RAC1 ternary complex, ELMO1NTD is predominantly in the open-conformation (Supplementary Fig. 3a–d). In this conformation, ELMO1NTD is situated above DOCK2SH3 (Figs. 1b and 3a – left panel), whereas in the closed-conformation, ELMO1NTD is poorly ordered, with its tip projected towards DOCK2DHR2 (Fig. 3a – right panel). In the ternary DOCK2–ELMO1–RAC1 complex (open-conformation), ELMO1 adopts an elongated, gently curved shape (Fig. 2f). We fitted the crystal structure of residues 1–520 of ELMO2 (ELMO2NTD, PDB:6IDX)42 (74% sequence identity to ELMO1NTD) to cryo-EM density assigned to ELMO1NTD, ELMO2NTD superimposes closely with ELMO1NTD within the context of the DOCK2–ELMO1–RAC1 complex. ELMO1ED of ELMO1NTD is composed of five pairs of anti-parallel α-helices (Fig. 2f). Helix 5B is characterized by its length and protrusion into the ELMO domain of ELMO1NTD (Fig. 2f). The EID and ELMO domains together are common to six human proteins; ELMO1, ELMO2, ELMO3, ELMOD1, ELMOD2, and ELMOD3. ELMO1ED at the N-terminus of ELMO1 adopts a ubiquitin-like fold that is exposed to solvent and accessible for interactions with RHOG35 and Arl437 (Figs. 1b, c and 2a).

The major and stable DOCK2–ELMO1 interface is generated by the PH domain and C-terminal PxxP motif of ELMO1 associated with the N-terminal SH3 domain and adjacent α-helical segment of DOCK2 (Figs. 1a, b and 2a). This is essentially identical to the crystal structure of the isolated DOCK2SH3–ELMO1PH assembly31. This interface comprises two segments. First, a six-α-helical bundle produced by the accretion of an α-helical segment adjacent to DOCK2SH3 with the two α-helices that flank ELMO1PH31 (Fig. 2a, insert). Second, the PxxP motif immediately C-terminal to ELMO1PH engages DOCK2SH3 as also previously defined31 (Figs. 1b and 2a – insert). In the cryo-EM structure of the DOCK2–ELMO1–RAC1 complex (as for the DOCK2–ELMO1 binary complex described below), there are no direct contacts between DOCK2SH3 and DOCK2DHR2, although such interactions may exist in DOCK2 alone as suggested by biochemical data for DOCK1 and DOCK31,49.

In an interaction that is important for DOCK2 GEF activity12,30–53, ELMO1PH is positioned in a groove created between lobe B of DOCK2DHR2 and the nucleotide-free RAC1 (Figs. 1b, c and 2a, e). This is consistent with an earlier study showing that DOCK1 and ELMO1PH interact directly50. In contrast to the PH domains of Dbl family GEFs such as Sos1, the non-canonical ELMO1PH lacks critical phosphoinositide-binding residues, and is not involved in membrane attachment30. However, it is critical for optimal DOCK GEF activity as shown in the nucleotide exchange assay (Supplementary Fig. 6a), consistent with previous observations both in vitro and in vivo12,30–53. A possible mechanism is that ELMO1 stabilizes DOCK2 by interacting with DOCK2SH3 through ELMO1PH (Fig. 2a). This is consistent with our observation that the purified DOCK2 protein alone (without ELMO1) elutes from a size-exclusion column as a broad extended peak (Supplementary Fig. 6b), and that in negative stain EM micrographs it appears highly heterogeneous with a tendency to aggregate (Supplementary Fig. 6c). In addition, our structure suggests that ELMO1PH is directly involved in DOCK2 GEF activity by interacting simultaneously with DOCK2DHR2 and nucleotide-free RAC1, thereby stabilizing DOCK2DHR2–RAC1 interactions (Fig. 2e).

**DOCK2–ELMO1 structure shows conformational change of ELMO1.** To understand potential regulatory mechanisms that underlie DOCK2–ELMO1 functions, we also analysed the structure of the DOCK2–ELMO1 binary complex by cryo-EM (Supplementary Fig. 7 and Supplementary Tables 1 and 2). 3D classification showed that in the DOCK2–ELMO1 binary complex, ELMO1NTD adopts two conformations: the open and closed-conformations (Fig. 3b and Supplementary Fig. 7). In ~12% of DOCK2–ELMO1 particles, ELMO1NTD adopts the open-conformation, resembling that of the DOCK2–ELMO1
A ternary complex with RAC1 (Fig. 3a, b – left panels and Supplementary Fig. 7d). However, in the majority (~80%) of DOCK2−ELMO1 binary complex particles, ELMO1NTD adopts a closed-conformation (Fig. 3b – right panel). This conformation differs slightly from the closed-conformation of the ternary DOCK2−ELMO1−RAC1 complex. In the DOCK2−ELMO1 binary complex, ELMO1NTD is more ordered through interactions with DOCK2DHR2, whereas ELMO1NTD adopts multiple conformations in the DOCK2−ELMO1−RAC1 complex. Interconversion between the open DOCK2−ELMO1−RAC1 ternary complex and the closed DOCK2−ELMO1 binary complex involves ELMO1NTD undergoing a rigid-body rotation of ~120°, centred on the elbow hinge that connects ELMO1NTD and ELMO1PH (Fig. 3c–e and Supplementary Movie 1). The closed-conformation is stabilized by ELMO1 intra-domain contacts. Specifically, the long flexible loop, inserted between α6 and α7 of ELMO1ELMO, together with the neighbouring ELMO1EID, form intimate contacts with ELMO1PH (Fig. 3e – right panel and Fig. 3f). The tip of the 16-residue α6/α7 loop of ELMO1ELMO contacts the positively charged surface of ELMO1PH, with the invariant Glu436 of the ELMO1ELMO α6/α7 loop inserting into a groove created by the basic β3/β4 loop of ELMO1PH. This β3/β4 loop in
In the closed-conformation of the DOCK2–ELMO1 binary complex, ELMO1NTD also forms extensive interactions with DOCK2 (Fig. 3b, e). Importantly, ELMO1RBD is positioned to directly contact the RAC1-binding site of DOCK2DHR2. In addition, ELMO1 ELMO contacts DOCK2C2 (Fig. 3b, e), in contrast with the open-conformation of ELMO1NTD which forms no contacts with DOCK2 (Fig. 1b, c). Superimposition of binary and ternary structures (on DOCK2DHR2) indicates that ELMO1RBD in the closed-conformation of the DOCK2–ELMO1 binary complex, and RAC1 in the DOCK2–ELMO1–RAC1 ternary complex, overlap (Fig. 3e). Thus, in the binary state with ELMO1NTD in the closed-conformation, ELMO1RBD sterically occludes RAC1 engagement, suggesting an auto-inhibited conformation (Fig. 3e).

Another conformational difference between the binary and ternary complexes is located in lobe C of DOCK2DHR2. In the binary complex, lobe C is rotated away from lobes A and B of DOCK2DHR2, with a maximum shift of ~10 Å. This allows ELMO1RBD to engage DOCK2DHR2 (Supplementary Fig. 4b and Supplementary Movie 1). In this conformation, as suggested by the weaker, less well-defined EM density, DOCK2DHR2 is more flexible than in the ternary complex. This change is mainly due to the absence of RAC1 because lobe C adopts similar conformations in both the open- and closed-conformations of the DOCK2–ELMO1 binary complex (Fig. 3b).

Crystal structure of RHOG–GMP–PNP complexed with ELMO2RBD. Binding of RHOG to ELMO may regulate RAC1 activation by spatially restricting DOCK-ELMO complexes in cells, or through direct molecular regulation of GEF activity. Whether ELMO2RBD uses a similar binding mode as RBDs found in RAS GTPase effectors is unknown. To answer these questions, we determined a crystal structure of activated RHOG complexed with the RBD domain of ELMO2 (ELMO2RBD). The crystals diffracted to 2.4 Å resolution with a single molecule in the asymmetric unit (Fig. 4a and Supplementary Table 4). The RHOG structure consists of six β-strands and six α-helices and is similar in structure to other RHO family small GTPases. Characteristically, RHOG nucleotide binding is coordinated by two key regions: switch 1 and switch 2 loops which bind the terminal γ-phosphate of GTP and a divalent magnesium ion (Fig. 4a, c). ELMO2RBD demonstrates a classical ubiquitin-like fold consisting of four β-strands and two α-helices, but interacts with RHOG at a non-canonical binding interface (Fig. 4b, c).

RAS is archetypally complexed with effector RBDs through an intermolecular, anti-parallel β-sheet consisting of β2 and β3 (switch 1) of the GTPase and β1, β2 and α1 of the RBD34. In contrast, the interface between RHOG and ELMO2RBD comprises both switch 1 and switch 2 regions of the GTPase (Fig. 4b, c). Major interactions include a hydrogen bond between ELMO2 Lys9 and the main chain at RHOG Phe37, as well as a salt bridge between Arg66 of RHOG and Glu13 of the RBD (Fig. 4c). In addition, there are extensive hydrophobic interactions between side chains of five RHOG residues (Val36, Phe37, Tyr64, Leu67 and Leu70) and four ELMO2 residues (Ala11, Ala19, Leu21 and Ile74) (Fig. 4c). Each of these amino acids are highly evolutionarily conserved in vertebrates (Supplementary Fig. 8). There is no direct ortholog of RHOG in simple organisms, the most homologous being RAC orthologs, but all key binding residues are conserved in these GTPases. Interestingly, the side chain of Arg66 is surface exposed in every available structure of a RAC GTPase. This residue is nonetheless invariant through evolution, suggesting a potential role in mediating protein interactions, as observed in our structure (Fig. 4c). For ELMO, many orthologs in less complex organisms have substitutions at Glu13, but key hydrophobic residues and particularly Lys9 are well conserved. Thus, the crystal structure of RHOG complexed with ELMO2 elucidates a novel GTPase-RBD binding interface dependent on several key residues that are highly conserved evolutionarily.

Mutations disrupt RHOG–ELMO2RBD interactions. To validate the RHOG–ELMO2RBD crystal structure and identify mutations that can disrupt formation of the RHOG–ELMO2RBD complex, we performed binding assays using isothermal titration calorimetry (ITC). We first determined the nucleotide dependency of this interaction by performing ITC with GDP- or GMPNNP-loaded RHOG and the ELMO2RBD. As shown in Fig. 5a, GDP-bound RHOG does not interact with ELMO2RBD, while GTP analogue-bound RHOG interacts with an affinity of 7.8 μM. This result is consistent with the paradigm of small GTPase signalling, and with our structural data revealing the interaction involves the nucleotide-sensitive switch regions. We were unable to crystallize RHOG in the GDP bound state, but a model based on available RHO family GDP-bound structures demonstrates that RHOG switch 1 would move significantly outward upon hydrolysis of GTP (Fig. 4b and Supplementary Fig. 9a). This would likely disrupt hydrogen bonding between ELMO2 Lys9 and the backbone of RHOG (Fig. 4c shows the position of this residue at the binding interface). Indeed, ITC analysis of an ELMO2 Lys9 to Ala mutant shows the importance of this Lys side chain, as binding to RHOG is completely abrogated (Fig. 5b). Our crystal structure revealed this is a backbone interaction with Phe37 of RHOG switch 1 rather than a salt bridge with RHOG Asp38 (Fig. 4c). Supporting
this, a RHOG Asp38 to Ala mutant binds ELMO2RBD with an affinity of 9.0 µM, comparable to wild-type (Fig. 5c, left). A key RHOG side chain at the binding interface is Arg66 in the switch 2 region (Fig. 4c), and ITC analysis revealed a RHOG R66A mutant is unable to bind ELMO2RBD (Fig. 5c, right). Size exclusion chromatography and NMR verified that the two mutants that disrupt the RHOG-ELMO2RBD interaction, RHOGD38A and ELMO2K9A, were correctly folded and comparable to wild type.
RASSF5/NORE1A55 and AF6/AFDN56 (Supplementary Fig. 9f, the binding site on ELMO1EID for the C-terminal helix of the complex, with ELMO1 NTD in the closed-conformation, is DOCK2 binding site of DOCK2DHR2 suggested the possibility binding to DOCK2−ELMO1RBD, engagement of BAI1 to ELMO1EID would be facilitated DOCK2-dependent RAC1 binding or activation. Using Boyden migration and Matrigel-invasion assays, we found that co-expression of DOCK2YYS/EEE with ELMO1 and CRKII in HeLa cells led to higher cell migration and invasion compared with DOCK2WT (Fig. 7b), suggesting that DOCK2 phosphorylation on these sites can lead to increasing DOCK2 RAC1 GEF activity in HEK293T cells. We then reasoned that if DOCK2YYS/EEE can promote RAC1 GFP-loading, then it should enhance cell migration and invasion. Using Boyden migration and Matrigel-invasion assays, we found that co-expression of DOCK2YYS/EEE with ELMO1 and CRKII in HeLa cells led to higher cell migration and invasion compared with DOCK2WT (Fig. 7c, d). This was further confirmed in wound healing as well as time-lapse live imaging assays of cells co-expressing ELMO1 and CRKII with either DOCK2YYS/EEE or DOCK2WT (Fig. 7c, f, Supplementary Movie 2). CRKII was included because in functional assays, the RAC1-dependent activities of DOCK1 and DOCK2 are maximal when co-expressed with both ELMO1 and CRKII12,30. Collectively, these data reveal a mechanism whereby phosphorylation of DOCK2 in the phosphorylation linker likely alleviates DOCK2−ELMO1 auto-inhibition to catalyse RAC1 activation to enhance cell migration and invasion.

ELMO1NTD is essential for RAC1 signalling by DOCK2 in cells. In the auto-inhibited DOCK2−ELMO1 structure, the region of ELMO1NTD, in particular ELMO1BD, occludes the DOCK2DHR2 RAC1-binding site. Removal of ELMO1NTD increased the activity of the DOCK2−ELMO1 complex in vitro (Fig. 4a, b). We aimed to determine if removing ELMO1NTD would facilitate DOCK2-dependent RAC1 binding or activation in cells. We truncated the first 529 residues from ELMO1 to generate a mutant that lacks the N-terminal domain (ELMO1ANTD) (Fig. 8a). We tested whether the expression of ELMO1ANTD in comparison to ELMO1WT has an effect on DOCK2 GEF activity by performing a GST-Pak pulldown assay to assess the levels of active RAC1. Co-expression of DOCK2 with ELMO1WT led to a significant increase in active RAC1 levels whereas co-expressing DOCK2 with ELMO1ANTD led to a small but significant decrease in the levels of active RAC1 in comparison to the conditions with ELMO1WT (Fig. 8b). These results are consistent with a view suggesting that ELMO1NTD is required to target the DOCK2−ELMO1 complex to the membrane for efficient DOCK2-mediated RAC1 activation. To directly test if DOCK2DHR2 is more accessible to bind RAC1 in the conditions where ELMO1ANTD is co-expressed, we conducted nucleotide-
free RAC1<sup>G15A</sup> pulldowns. GEFs form a stable complex with their target GTPases when in a nucleotide-free state<sup>60</sup>. Such an interaction is supported by structural evidence for DOCK2 DHR2–RAC1<sup>44</sup>. Hence, we generated the nucleotide-free RAC1<sup>G15A</sup> mutant, the equivalent of RHOA<sup>G15A</sup>. We found that RAC1<sup>G15A</sup> binding to DOCK2 was minimally but significantly decreased upon expression of ELMO1<sup>ΔNTD</sup> compared with conditions with ELMO1<sup>WT</sup> (Fig. 8c). Hence, this suggests that the decreased RAC1 activation upon expression of ELMO1<sup>ΔNTD</sup> might be due to decreased levels of the DOCK2–RAC1 complex. Functionally,
co-expression of DOCK2 with ELMO1ΔNTD and CRKII led to a decrease in the migration and invasion of HeLa cells, using Boyden migration and Matrigel-invasion assays, respectively (Fig. 8d, e). These results were further confirmed in wound healing and time-lapse live imaging assays, where cells expressing ELMO1ΔNTD were less motile when compared with cells expressing ELMO1WT (Fig. 8f, g and Supplementary Movie 2). Collectively, these data further confirm the importance of...
**Discussion**

Our structure of DOCK2–ELMO1 reveals how the modular organization of two multiple domain subunits associate to form the binary complex. Conformational changes within ELMO1 relieves auto-inhibition suggesting a model for how interdependent ligand binding to multiple sites on both DOCK2 and ELMO1, and DOCK2–ELMO1 phosphorylation, regulates DOCK2 GEF activity (Fig. 9a). Contacts between domains within subunits create fairly rigid structures, with the only major conformational changes in ELMO1NTD for optimal RAC1 activation and for the induction of RAC1-mediated cell migration and invasion in cells.
variability resulting from rotation about the hinge elbow connecting ELMO1\textsuperscript{NTD} and ELMO1\textsuperscript{PH}. Interactions between the two subunits are centred on stable and invariant contacts between the N-terminal SH3 and helical domains of DOCK2 and the PH domain and poly-Pro segment at the C-terminus of ELMO1. The binary DOCK2–ELMO1 complex adopts two conformational states. In the closed, auto-inhibited state of the DOCK2–ELMO1 binary complex (the dominate state for DOCK2–ELMO1), rotation of ELMO1\textsuperscript{NTD} by 120° about an elbow hinge connecting ELMO1\textsuperscript{NTD} with ELMO1\textsuperscript{PH} causes ELMO1\textsuperscript{NTD} to form contacts with DOCK2\textsuperscript{DHR2} and ELMO1\textsuperscript{PH}. These new interfaces create an auto-inhibited state that occludes binding sites for RAC1, RHOG and BAI. Phosphorylation of ELMO1 on either the phosphorylation linker (Fig. 1c) or Tyr 18 of ELMO1\textsuperscript{RBD} (Fig. 4d) would disrupt the closed state. The view is similar to Figs. 3a, b and 6d.

Membrane attachment sites for DOCK2 and RAC1 are situated on the same face of the DOCK2–ELMO1–RAC1 complex. PIP\textsubscript{3} binds to a site defined by the L1 loop of DOCK2\textsuperscript{DHR1} incorporating Lys437, Lys440 and Lys444, whereas RAC1 attaches to the membrane through a prenyl group attached to Cys189. Leu177 denotes the last ordered residue of the RAC1 crystal structure.

**Fig. 9 Schematic of conformational changes of DOCK2–ELMO1 and membrane attachment model.** a In the DOCK2–ELMO1–RAC1 ternary complex, ELMO1\textsuperscript{NTD} adopts the open conformation with binding sites for RAC1, RHOG and BAI exposed on DOCK2\textsuperscript{DHR2}, ELMO1\textsuperscript{RBD} and ELMO1\textsuperscript{EID} (on ELMO1\textsuperscript{NTD}), respectively. In the closed, auto-inhibited state of the DOCK2–ELMO1 binary complex (the dominate state for DOCK2–ELMO1), rotation of ELMO1\textsuperscript{NTD} by 120° about an elbow hinge connecting ELMO1\textsuperscript{NTD} with ELMO1\textsuperscript{PH} causes ELMO1\textsuperscript{NTD} to form contacts with DOCK2\textsuperscript{DHR2} and ELMO1\textsuperscript{PH}. These new interfaces create an auto-inhibited state that occludes binding sites for RAC1, RHOG and BAI. Phosphorylation of ELMO1 on either the phosphorylation linker (Fig. 1c) or Tyr 18 of ELMO1\textsuperscript{RBD} (Fig. 4d) would disrupt the closed state. The view is similar to Figs. 3a, b and 6d. b Membrane attachment sites for DOCK2 and RAC1 are situated on the same face of the DOCK2–ELMO1–RAC1 complex. PIP3 binds to a site defined by the L1 loop of DOCK2\textsuperscript{DHR1} incorporating Lys437, Lys440 and Lys444, whereas RAC1 attaches to the membrane through a prenyl group attached to Cys189. Leu177 denotes the last ordered residue of the RAC1 crystal structure.
ELMO1NTD. We propose that this could be regulated by two mechanisms. First, the binding of upstream regulators to distinct domains of ELMO, whose binding sites are blocked in the auto-inhibited state, for example RHOG or Arl4A to ELMOBD and BAI receptors to ELMOBD, would promote the active conformation. Second, phosphorylation of DOCK2 within its phosphorylation linker region, and ELMO1BD, whose sites are buried in the auto-inhibited state, would also promote the active conformation. Conceivably, prior phosphorylation of DOCK2—ELMO1 relieves auto-inhibition to expose binding sites for both BAI receptors and RHOG, recruiting DOCK2—ELMO1 to the cell membrane where RAC is localized. Because ELMO interacts with numerous additional proteins, such as the membrane protein ClipR-59, required for myoblast fusion, the model we propose here for release of ELMO-mediated auto-inhibition may be general for numerous regulators of the DOCK—ELMO—RAC signalling module.

Methods

Cloning and mutagenesis. The coding sequence for human DOCK2 and ELMO1 were amplified by PCR using primers 35–46 (see Supplementary Table 5 for a list of primers used in this study) and cloned into pF1 and pU1 plasmids, respectively. cDNAs were ligated together with a His-tag integrative vector pET17B (Novagen) containing the T7 promoter and the T7 terminator. The coding sequence for human ClipR-59 was amplified by PCR using primers 2–7 and 2–529 with primers 27–30. Human RHOG (GeneID 391); amino acids 1–179 was cloned into the pDEST17 bacterial expression vector using Gateway technology, with a thrombin cleavage site inserted between the poly-His tag and the RHOG coding sequence using primers 15 and 16 (Supplementary Table 5). A sequence encoding the N-terminal RBD domain of murine ELMO2 (GeneID 140579; amino acids 1–80) was cloned into a bacterial expression vector (pBR322) with an N-terminal glutathione S-transferase (GST) tag. To confirm the interaction interface of the RHOG and ELMO2 complex, RHOGΔRD38, RHOGβ68A and ELMO2K79R were prepared by the site directed mutagenesis method using primers 17–22. The plasmid pU1-ELMO1 was used to generate pCS-6Myc-hELMO1 and pCS-6Myc-hELMO1NTD by the Gateway cloning system using primers 1–4 (Supplementary Table 5). pCN2X-FLAG-hDOCK2 was a kind gift from Dr. Michiyuki Matsuda (Kyoto University) and was used to generate the pCN2X-FLAG-hDOCK2YYS/EEE mutant using the HiFi Gibson Assembly technology using primers 5–14 (Supplementary Table 5).

Expression and purification. To express and purify the DOCK2—ELMO1 complexes, DOCK2—ELMO1, monomeric DOCK2—ELMO1, DOCK2—ELMO1ΔNTD and DOCK2—ELMO1ΔNTD complexes were expressed and purified in negative-stain electron microscopy (EM) grids using the Talon-His pET17B-TN-5B1–4) (ThermoFisher) for 2 days. The DOCK2—ELMO1 complex was purified in a buffer of 50 mM Tris HCl (pH 8.0), 200 mM NaCl, 2 mM DTT, 2 mM benzamidine, 1 mM EDTA, 0.2 mM PMSF and then loaded onto a Streptactin Column (Qiagen) and the Strept-tagged complex was eluted with 5 mM desthiobiotin. The GST-pUL22 fragment was used to generate pCS-6Myc-hELMO1 and pCS-6Myc-hELMO1NTD, pCS-6Myc-hELMO1 and pCS-6Myc-hELMO1NTD by the Gateway cloning system using primers 1–4 (Supplementary Table 5). The RAC1-mant GTP complex was prepared by incubating 2 μM RAC1 with 10 μM GDP and 5 μM GTPγS in a solution of 20 mM Tris HCl (pH 8.0), 200 mM NaCl, 10 mM MgCl2, 1 mM DTT. For the GST-pUL22 fragment, pUL22 was expressed and purified in a buffer containing 0.03% glutaraldehyde for 10 min on ice. After removing the glutaraldehyde, the GST-pUL22 fragment was incubated with nickel-nitrilotriacetic acid or GSH resin for 1–2 h at 4 °C. After washing in high-salt buffer (20 mM Tris HCl (pH 7.5), 500 mM NaCl, 5 mM MgCl2, and 1 mM DTT or 5 mM β-mercaptoethanol), His-tagged proteins were eluted with 250 mM imidazole followed by thrombin cleavage. GST fusions were cleaved with thrombin directly on GSH resin overnight at 4 °C. For nucleotide exchange, RHOG proteins were incubated at 37 °C in the presence of 10 mM EDTA and a tetanol fold excess of nucleotide-protein of GMPPNP (Sigma-Aldrich) for 10 min. 20 mM MgCl2 was added and the sample was placed on ice. Cleaved and nucleotide-loaded proteins were purified on a HiLoad 26/600 Superdex 75 prep-grade column with a buffer consisting of 20 mM Tris HCl (pH 7.5), 100 mM NaCl, 5 mM MgCl2, and 1 mM dithiothreitol. Fractions containing purified protein were pooled, concentrated to 10 mg/mL and stored at −80 °C prior to crystallization.

For RHOG purification, recombinant plasmids were transfected into Escherichia coli BL21 (DE3) CodonPlus cells (Agilent Technologies) and grown at 30 °C in Luria-Bertani medium with 100 μg ampicillin up to optical density (OD) at 600 nm of 0.6. Protein expression was induced by the addition of 0.5 mM isopropyl-D-thiogalactopyranoside (IPTG) at 20 °C and cultures were grown for 18 h. Cells were lysed by sonication in 20 ml Tris HCl (pH 7.5), 130 mM NaCl, 5 mM MgCl2, 10% (v/v) glycerol, 0.4% Nonidet P-40, protease inhibitors, and 1 μM pepstatin A. Lysates were centrifuged and incubated with nickel-nitrilotriacetic acid or GSH resin for 1–2 h at 4 °C. After washing in high-salt buffer (20 mM Tris HCl (pH 7.5), 500 mM NaCl, 5 mM MgCl2, and 1 mM DTT) or 5 mM β-mercaptoethanol, His-tagged proteins were eluted with 250 mM imidazole followed by thrombin cleavage. GST fusions were cleaved with thrombin directly on GSH resin overnight at 4 °C. For nucleotide exchange, RHOG proteins were incubated at 37 °C in the presence of 10 mM EDTA and a tetanol fold excess of nucleotide-protein of GMPPNP (Sigma-Aldrich) for 10 min. 20 mM MgCl2 was added and the sample was placed on ice. Cleaved and nucleotide-loaded proteins were purified on a HiLoad 26/600 Superdex 75 prep-grade column with a buffer consisting of 20 mM Tris HCl (pH 7.5), 100 mM NaCl, 5 mM MgCl2, and 1 mM dithiothreitol. Fractions containing purified protein were pooled, concentrated to 10 mg/mL and stored at −80 °C prior to crystallization.

Fluorescence kinetics analysis. The RAC1-mant GTP complex was prepared by incubating 2 μM RAC1 with 10 μM 2′,3′-O-(N-methylanthraniloyl)guanosine 5′-O-triphosphate GTP (mANT GTP) for 15 min at 20 °C in reaction buffer (50 mM Tris HCl (pH 8.0), 200 mM NaCl, 10 mM MgCl2, 1 mM DTT). Fluorescence experiments were performed using a PHERAstar FS plate reader (BMG LabTech). In the absence and presence of 0.2 μM GEF, varying concentrations of RAC1-mant GTP (from 0.2 to 1.6 μM) were incubated in reaction buffer (50 mM Tris HCl (pH 8.0), 200 mM NaCl, 10 mM MgCl2, 2 mM DTT) for 10 min at 22 °C. The release reaction was initiated by injection of 100 mM GDP. The fluorescence signals (excitation wavelength 350 nm and emission wavelength 450 nm) were monitored every 0.6 s. The initial rates of exchange were fitted and the average from triplicate assays were plotted against concentration of substrate (RAC1-mant-GDP) and fitted to a Michaelis–Menten equation. Data were analysed using PRISM 8.2.1 (GraphPad Software). Graphs were plotted after subtraction of the uncatalysed nucleotide release rate.

Electron microscopy. Freshly purified DOCK2 samples (DOCK2—ELMO1—RAC1, DOCK2—ELMO1—or monomeric DOCK2—ELMO1—RAC1 mutant) were first immobilized on HiperGrid Filter grid preparation. DOCK2 samples were treated to be 0.052% glutaraldehyde for 10 min on ice before a size exclusion chromatography purification using a Superose 6 Increase column (GE Healthcare) to remove aggregates. Without cross-linking treatment, most particles were found disassembled on cryo-EM grids. Aliquots of 3 μl samples at −120 °C were applied onto glow-discharged Quantifoil R1/2/3.1/3 holey carbon grids. The grids were incubated for 30 s at 4 °C and 100% humidity and then blotted for 8 s and plunged into liquid ethane using a Vitrobot III (Thermo Fisher).

DOCK2—ELMO1—RAC1 was imaged by a Thermo Fisher Scientific Titan Krios electron microscope at the MRC Laboratory of Molecular Biology (MRC-LMB) that was operated at an acceleration voltage of 300 kV and at a nominal magnification of 81,000 (resulting a calibrated physical pixel size of 1.43 Å/pixel), and a Gatan K2 summit detector that was installed after a GIF Quantum energy filter and operated with a slit width of 20 eV. In total, 1,338 micrographs were collected manually using K2 super-resolution mode. Each micrograph was exposed for 16 s at a dose rate of 5 electrons/pixel/sec and saved as 20 movie frames. Calculated defocus values are in a range of −1.5 to −3.2 μm. A second batch of 576 micrographs was collected on the same microscope at a nominal magnification of 64,000 (resulting calibrated pixel size of 1.76 Å/pixel).
EM image processing. All movie frames were aligned using MotionCor2 before subsequent processing. The contrast transfer function parameters were calculated using e2initialmodel.py program in EMAN2 package. For DOCK2−ELMO1−RAC1, particles in 282 pixels x 282 pixels were selected by the automatic particle picking module in RELION 1.4, using reference-free 2D class averages from manually selected particles as templates. The following steps were performed to exclude bad particles from the dataset: (1) Automatically picked particles were flagged for alignment and those with low Z-scores were deleted; (2) 2D-classification was performed and particles in bad class were marked by manual inspection. Finally, by screening manually with the automatic particle picking module in RELION 1.4, using reference-free 2D class averages, 245,763 good particles were selected from two distinct datasets (Fig. 2a). For the second dataset collected at a lower magnification (1.76 Å/px), particles were rescaled to 1.43 Å/px when performing particle extraction in RELION 1.4. Specifically, a box size of 230 pixels and a re-scaled box size of 282 pixels were used to scale the particles in Fourier space (230 pixels x 1.76 Å/px; 282 pixels = 1.435 Å/px).

To generate an initial model for 3D refinement, e2initialmodel.py program in EMAN2 package was used with 10 selected 2D average images from RELION 2D classification. Initial refinement with a subset of 80,411 particles resulted in a map with 12−15 Å resolution. When a mask with ordered regions comprising the DOCK2SH3−RAC1 dimer, DOCK2ARM, DOCK2DHR1, DOCK2DHR2 and DOCKS313−ELMO1 PH was used (Supplementary Fig. 1e), focussed refinement resulted in a significantly improved map at 6.6 Å resolution. By making a mask based on the improved map that only included strong densities, the refinement resulted in a 6.1 Å resolution map. Refinement using all particles resulted in a map at an overall resolution of 4.6 Å. The results indicated that flexibility of individual parts/domains. For instance, DOCK2 ARM and the DOCK2SH3−ELMO1 PH assembly can be distinguished into two conformations by making the dimerization interface of DOCK2DHR2 (Y1315A/L1322A) due to the DOCK2 SH3 interaction. Furthermore, 2D-classification with particle features excluded (245,763 good particles) was performed using the automatic particle picking module in RELION 1.4, using reference-free 2D class averages, 245,763 good particles were selected from two distinct datasets (Fig. 2a).

The structure model was built ab initio using ROSETTA and RAC1 complex by mutating the dimerization interface of DOCK2DHR2 (Y1315A/L1322A). Although we could purify monomeric complex, the complex was very unstable on cryo-EM grids and we were not able to improve the resolution of this approach (Supplementary Fig. 6).

To improve the resolution, we separated two monomers computationally using signal subtraction. With a mask that includes one monomer of DOCK2−ELMO1−RAC1 and the DOCK2DHR2−RAC1 region from the other monomer, we performed a refinement to align all particles. We then subtracted the aligned monomer (all of the monomer but not DOCK2DHR2−RAC1 region) from original particles to create a subtracted dataset A. With dataset A, we then performed a refinement with the mask of the other monomer map. We then used the alignment parameters to subtract the other monomer from the original dataset to create subtracted dataset B. By combining datasets A and B, a map was reconstructed to 4.1 Å resolution (Supplementary Fig. 2a, b).

From local resolution map and further 3D classification, we observed flexibility of individual parts/domains. For instance, DOCK2 ARM and the DOCK2SH3−ELMO1 PH assembly can be distinguished into two conformations by making the dimerization interface of DOCK2DHR2 (Y1315A/L1322A) due to the DOCK2 SH3 interaction. Furthermore, 2D-classification with particle features excluded (245,763 good particles) was performed using the automatic particle picking module in RELION 1.4, using reference-free 2D class averages, 245,763 good particles were selected from two distinct datasets (Fig. 2a).

For DOCK2−ELMO1, data processing followed a similar procedure. The resolutions were generally lower compared with DOCK2−ELMO1−RAC1, likely due to the DOCK2−ELMO1 sample being less stable. However, the ELMO1NTD region was improved and showed more detailed structure.

Map visualization. Figures were generated using Pymol and Chimera.

DOCK2−ELMO1 model building. Initial model building for each domain was based on maps from DOCK2−ELMO1−RAC1, except for ELMO1NTD that was based on a map from DOCK2−ELMO1. For DOCK2DHR2−RAC1, the crystal structure (PDB: 2YN1) was fitted to the density map using “fit to map” program in Chimera and then rebuilt in COOT guided by side-chain densities (Supplementary Fig. 4a). For the ARM domain, 3D structure predictions from PHYRE2 and I-TASSER were used as a guide for determining conformations of helices, but the final map was built with poly-alanine. For the DHR1 domain, a structure model was firstly built using PHYRE2 based on the crystal structure of DOCK1 SH3 and then fitted to the cryo-EM density map. For DOCK2C2, a structure model was built ab initio using ROSETTA and fitted to the density map. For the DOCK2SH3−ELMO1 PH domain, the crystal structure of this region (PDB: 3A98) was fitted as a rigid-body. The N-terminal extension of DOCK2 (a loop and α-helix) and ELMO1 (α-helix) beyond the crystal structure were built in COOT. ELMO1NTD was then fitted to the cryo-EM density based on a crystal structure of ELMO2 (PDB: 6DIX). All models were refined with PHENIX.

RHOG−ELMO2 complex crystal structure determination. Initial crystal screens for the RHOG and ELMO2 complex (1:1 molar ratio) were performed by sitting-drop vapour diffusion at 22 °C. After 2 days, microcrystals were obtained in 0.1 M CHES pH 9.5 and 1.0 M sodium citrate. These crystalization conditions were optimized by the hanging-drop vapour diffusion method at 22 °C. Suitable crystals of native and selenium-methionine-derivatized versions were grown in 0.1 M CHES pH 8.8 and 0.95 M sodium citrate, respectively. The crystals were with 3.5 M sodium citrate as a cryoprotectant. X-ray data were collected at McGill Chemistry Characterization Facility of McGill University (Montreal, Canada) using a Bruker D8 Venture single crystal X-ray diffractometer, wavelength was 1.3417 Å, temp 100 K. Raw data were indexed, integrated, and scaled using the Phaser software. Crystallographic statistics of data collection are provided in Supplementary Table 4.

The structure was solved by molecular replacement using Phaser in the PHENIX software package. Model building and refinement were performed using COOT and PHENIX. The structure was validated with MolProbity. The statistics of structure refinement are provided in Supplementary Table 4. Coordinates and structure factors of the RHOG−ELMO2 RBD complex are deposited in the Protein Data Bank (PDB) with the accession code 6UKA.

Isothermal titration calorimetry. RHOG interactions with ELMO2 RBD domains were measured using a MicroCal ITC200 (Malvern). Stock solutions were diluted into filtered and degassed 20 mM Tris-HCl (pH 7.5), 100 mM NaCl and 1 mM DTT. Experiments were carried out at 25 °C. Wild-type or mutant RHOG was injected into a reaction cell containing RHOG wild-type or mutants. Fifty individual injections at 150 s intervals were performed. Data were fit using the Origin (version 7.0) software (OriginLab Corporation).

Cell culture and transfections. HEK293T (293T) and HeLa cells (both cell lines obtained from ATCC) were cultured in DME supplemented with 10% foetal bovine serum (Gibco) and 1% Penicillin/streptomycin antibiotics (Wisent). HeLa cells were transfected using Lipofectamine 2000 (Invitrogen) according to manufacturer instructions. 293T cells were transfected using the calcium phosphate method.

GTP-pulldown assays and immunoblotting. For the RAC activation assays, 293 T cells were lysed as described. For GST-RAC1GSTA pulldowns, 293T cells were lysed using 1% NP-40 (15 mM NaCl, 50 mM Tris pH 7.5, 1% Nonidet P-40, 10 mM NaF, 1 mM Na3PO4, 1 mM Na3VO4, 1X complete protease inhibitor) buffer and cell lysates were cleaved by centrifugation and incubated with the corresponding GST-tagged proteins on beads for 2 h at 4 °C (GST-RAC1GSTA Pulldown) or 30 min at 4 °C (GST-PAK-PBD Pulldown). Lysates and GST-fusion bound complexes, were run on SDS-electrophoresis acrylamide gels at 180 V and transferred on nitrocellulose for 3 h at 4 °C at 50 V or overnight at 4 °C at 20 V. Immunoblots were then blocked with 1% BSA and incubated with the indicated primary antibodies overnight at 4 °C or room temperature. Immunoblots were then washed with 0.01% TBST three times and incubated with the corresponding secondary antibody for 30 min at room temperature. Protein signals were revealed via Clarity western ECL substrate (BioRad). Antibodies used: anti-MyC (Santa Cruz – SC40) dilution factor: 2,000; anti-RAC1 (EMD - Millipore 05389), dilution factor: 1,000; anti-FLAG (Sigma – AB8592), dilution factor: 10,000. RAC activation and DOCK2−RAC1 binding levels were quantified by densitometry analysis using the ImageJ software program.

Boyden migration and invasion assay. Boyden assays were performed using 8 μm pores Boyden Chambers (24-well, Costar). For the invasion assays, the upper chamber was coated with 6 μL of Matrigel (BD Biosciences) dissolved in 100 μL of DMEM. HeLa cells were detached and washed with DMEM 0.1% BSA. In total, 100,000 cells were seeded in the top chamber and allowed to migrate for 6 h (migration) or 24 h (invasion) toward the bottom chamber containing 10% FBS. Upper and lower chambers were then washed with 1x PBS and cells on the bottom side of the chamber were fixed with 4% PFA. Cells in the upper chambers were removed using cotton swabs and the membrane was mounted on a glass slide using SlowFade Gold reagent (Invitrogen). The average number of migrating cells in 10 independent 20X microscope fields were evaluated, and each experiment was performed in triplicate.

Time-lapse cell imaging. HeLa cells plated on fibronectin-coated plates (1000 cells/well in a 12-well plate) were transfected with 1 μg Myc-ELMO1/Myc-Hel-MO2NTD, 3 μg Flag-DOCK2/Flag-DOCK2NTD and 0.5 μg GFP-CR2. Then 48 h later, cells were imaged using Time-lapse microscopy at 10 min intervals for 6 h (Speed tracking) or 24 h (for wound healing), using phase contrast brightfield. Videos and images were obtained using Velocity and analysed via Image J software for their speed and distance measurements.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.
Data availability
EM maps are deposited in the Electron Microscopy Data Bank under accession codes: 10498 (DOCK2–ELMO1–RAC1 ternary complex, open conformation), 10497 (DOCK2–ELMO1 binary complex, closed conformation). Protein coordinates are deposited in the Protein Data Bank under accession codes: 6TGC (DOCK2–ELMO1–RAC1 ternary complex, open conformation), 6TGB (DOCK2–ELMO1–RAC1 ternary complex, open conformation) and 6UKA (ELMO2RBD–RHOG). Plasmids and cell lines that were generated for and used in this study are available upon request from the authors. Source data are provided with this paper.

Received: 26 November 2019; Accepted: 17 June 2020; Published online: 10 July 2020

References
1. Heasman, S. J. & Ridley, A. J. Mammalian Rho GTPases: new insights into their functions from in vivo studies. Nat. Rev. Mol. Cell Biol. 9, 690–701 (2008).
2. Bacia, H. et al. Mapping the proximity interaction network of the Rho-family GTPases reveals signalling pathways and regulatory mechanisms. Nat. Cell Biol. 22, 120–134 (2020).
3. Bos, J. L., Rehmman, H. & Wittinghofer, A. GEFs and GAPs: critical elements in the control of small G proteins. Cell 129, 865–877 (2007).
4. Rosomak, K. L., Der, C. J. & Sondel, J. GEFs mean go: turning on RHO GTPases with guanine nucleotide-exchange factors. Nat. Rev. Mol. Cell Biol. 6, 167–180 (2005).
5. Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. Science 294, 1299–1304 (2001).
6. Laurin, M. & Cote, J. F. Insights into the biological functions of Dock family guanine nucleotide exchange factors. Genes Dev. 28, 533–547 (2014).
7. Cote, J. F. & Vuori, K. GEF what? Dock180 and related proteins help Rac to polarize cells in new trends. Trends Cell Biol. 17, 383–393 (2007).
8. Gadea, G. & Blangy, A. Dock-family exchange factors in cell migration and disease. Eur. J. Cell Biol. 93, 466–477 (2014).
9. Gadea, G., Sanz-Moreno, V., Self, A., Godi, A. & Marshall, C. J. Dock10-mediated Cdc42 activation is necessary for aneboid invasion of melanoma cells. Curr. Biol. 18, 1456–1465 (2008).
10. Meller, N., Merlot, S. & Guda, C. C2H2 proteins: a new family of Rho-GEFs. J. Cell Sci. 118, 4937–4946 (2005).
11. Sanz-Moreno, V. et al. Rac activation and inactivation control plasticity of tumor cell movement. Cell 133, 510–523 (2008).
12. Brugnera, E. et al. Unconventional Rac-GEF activity is mediated through the Dock180–ELMO complex. Nat. Cell Biol. 4, 574–582 (2002).
13. Cote, J. F. & Vuori, K. Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. J. Cell Biol. 179, 466–477 (2007).
14. Cote, J. F. & Vuori, K. In vitro guanine nucleotide exchange activity of DHR–CED-12–ELMO1, a novel member of the CRII/DOCK180 family required for phagocytosis and cell migration. Cell 107, 27–41 (2001).
15. Wu, Y. C., Tsai, M. C., Cheng, L. C., Chou, C. J. & Weng, N. Y. C. Eclipses CED-12 acts in the conserved crkII/DOCK180/Rac pathway to control cell migration and cell morphogenesis. Dev. Cell 1, 491–502 (2001).
16. Zhou, Z., Cano, E., Hartwig, E., Hall, A. & Horvitz, H. R. The C. elegans PH domain protein CED-12 regulates cytoskeletal reorganization via a Rho/Rac GTPase signaling pathway. Dev. Cell 1, 477–489 (2001).
17. Sanui, T. et al. Dock2 regulates Rac activation and cytoskeletal reorganization through interaction with ELMO1. Blood 102, 2948–2950 (2003).
18.5.0.0.0Dock2 and Dock180: a bipartite Rac1 guanine nucleotide exchange factor, promote human glioma cell invasion. J. Biol. Chem. 286, 38966–38979 (2011).
19. Park, D. et al. BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module. Nature 450, 430–434 (2007).
20. Harnoud, N., Tran, V., Croteau, L. P., Kania, A. & Cote, J. F. G-protein coupled receptor BAII3 promotes myoblast fusion in vertebrates. Proc. Natl Acad. Sci. USA 111, 3745–3750 (2014).
21. Hochreuter-Hulfder, A. E. et al. Phosphatidylinositol receptor BAII1 and apoptotic cells as new promoters of myoblast fusion. Nature 497, 263–267 (2013).
22. Hamoud, N. et al. Spatiotemporal regulation of the GPCR activity of BAII3 by Cnx4 and Stabilin-2 controls myoblast fusion. Nat. Commun. 9, 4479 (2018).
23. Weng, Z. et al. Structure of BAII1/ELMO2 complex reveals an action mechanism of adhesion GPCRs via ELMO family scaffolds. Nat. Commun. 10, 51 (2019).
24. Kulkarni, K., Yang, J., Zhang, Z. & Barford, D. Multiple factors confer specific Cdc42 and Rac protein activation by detergent of cytostokines (DOCK) superfamily exchange factors. Curr. Biol. 16, 1358–1362 (2006).
25. Yang, J., Zhang, Z., Roe, S. M., Marshall, C. J. & Barford, D. Activation of Rac GTPases by DOCK exchange factors is mediated by a nucleotide sensor. Science 325, 1398–1402 (2009).
26. Helm, I., Kaaraiyen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DALiLite v.3. Bioinform. (Oxf., Engl.) 24, 2780–2781 (2008).
27. Jeong, H. et al. Mechanistic insight into the nucleus-vacuole junction based on the Vac8p-Nyi1p crystal structure. Proc. Natl Acad. Sci. USA 114, E4539–E4548 (2017).
28. Fleishman, S. J. et al. RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. PloS ONE 6, e21661 (2011).
29. Helfmann, S. M. et al. The crystal structure of the C(2)A domain of otferlin reveals an unconventional top loop region. J. Mol. Biol. 406, 479–490 (2011).
30. Lu, M. et al. A Steric-inhibition model for regulation of nucleotide exchange via the Dock180 family of GEFs. Curr. Biol. 15, 371–377 (2005).
31. Grimsley, C. M. et al. Dock180 and ELMO1 proteins cooperate to promote evolutionarily conserved Rac-dependent cell migration. J. Biol. Chem. 279, 6087–6097 (2004).
32. Lu, M. & Ravichandran, K. S. Dock180–ELMO1 cooperation in Rac activation. Methods Enzymol. 406, 388–402 (2006).
33. Jarzynka, M. J. et al. ELMO1 and Dock180, a bipartite Rac1 guanine nucleotide exchange factor, promote human glioma cell invasion. Cancer Res. 67, 7203–7211 (2007).
34. Mott, H. R. & Owen, D. Structures of Ras superfamily effector complexes: What have we learnt in two decades? Crit. Rev. Biochem Mol. Biol. 50, 85–133 (2015).
35. Stiegler, B. et al. Novel type of Rac effector interaction established between tumour suppressor Nore1A and Ras switch II. EMBO J. 27, 1995–2005 (2008).

ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-020-17271-9 | www.nature.com/naturecommunications
56. Smith, M. J. et al. Evolution of AF6-RAS association and its implications in mixed-lineage leukemia. Nat. Commun. 8, 1099 (2017).
57. Abu-Tharaiya, A. et al. Axl phosphorylates Elmo scaffold proteins to promote Rac activation and cell invasion. Mol. Cell Biol. 35, 76–87 (2015).
58. Feng, H. et al. Activation of Rac1 by Src-dependent phosphorylation of Dock180(Y1811) mediates PDGFRalpha-stimulated glioma tumorigenesis in mice and humans. J. Clin. Investig. 121, 4670–4684 (2011).
59. Nakane, T., Kimanius, D., Lindahl, E. & Scheres, S. H. Characterisation of Rac-specific guanine nucleotide exchange factor DOCK1 is a critical regulator of HER2-mediated breast cancer metastasis. Proc. Natl Acad. Sci. USA 110, 7434–7439 (2013).
60. Pacold, M. E. et al. Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma. Cell 103, 931–943 (2000).
61. Sun, Y. et al. ClipR-59 interacts with Elmo2 and modulates myoblast fusion. J. Biol. Chem. 290, 6130–6140 (2015).
62. Zhang, Z., Yang, J. & Barford, D. Recombinant expression and reconstitution of multiprotein complexes by the USER cloning method in the insect cell-baculovirus expression system. Methods 95, 13–25 (2016).
63. Bitinaite, J. et al. USER friendly DNA engineering and cloning method by uracil excision. Nucleic Acids Res. 35, 1992–2002 (2007).
64. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332 (2017).
65. Zhang, K. Gctf: real-time CTF determination and correction. J. Struct. Biol. 193, 1–12 (2016).
66. Fernandez-Leiro, R. & Scheres, S. H. W. A pipeline approach to single-particle processing in RELION. Acta crystallographica. Section D. Struct. Biol. 73, 496–502 (2017).
67. Nakane, T., Kimanius, D., Lindahl, E. & Scheres, S. H. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. eLife 7, 10.7554/eLife.36861 (2018).
68. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180, 519–530 (2012).
69. Scheres, S. H. Semi-automated selection of cryo-EM particles in RELION-1.3. J. Struct. Biol. 189, 114–122 (2015).
70. Bai, X. C., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. Sampling the conformational space of the catalytic subunit of human gamma-secretase. eLife 4, 10.7554/eLife.11182 (2015).
71. Yang, Z. et al. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. J. Struct. Biol. 179, 269–278 (2012).
72. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr D. Biol. Crystallogr 66, 486–501 (2010).
73. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10, 845–858 (2015).
74. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5, 725–738 (2010).
75. Adams, P. D. et al. PHENIX: a comprehensive python-based system for macromolecular structure solution. Acta Crystallogr D. Biol. Crystallogr 66, 213–221 (2010).

Acknowledgements
This work was funded by the MRC Laboratory of Molecular Biology and MRC (MC_UP_1201/6) and Cancer Research UK (C576/A14109) grants to D.B. We are grateful to members of the Barford group for discussion; S. Chen, C. Savva and G. McMullan for help with the EM data collection; J. Grimmett and T. Darling for computing. This work was supported by grants from the Canadian Institutes for Health Research (CIHR) (to M.J.S. and I.F.C.) and the National Science and Engineering Council of Canada (NSERC) (M.J.S.). A.A.T and R.K. were supported by research studentships and fellowships, respectively, from the Fonds de recherche du Québec – Santé (FRQS). M.J.S holds a Canada Research Chair in Cancer Signaling and Structural Biology. I.F.C holds the TRANSAT Chair in Breast Cancer Research. We acknowledge Diamond Light Source for access and support of the cryo-EM facilities at the UK national electron bio-imaging centre (eBIC), proposal EM13708, funded by the Wellcome Trust, MRC and BBSRC.

Author contributions
The project was conceived by D.B., A.A.T, I.F.C, M.J.S. Experiments were performed by D.B., L.C., J.Y., Z.Z., A.B., S.H.M., C.H.J., R.C.K., M.J.S., A.A.T. The paper was written by D.B., C.H.J., L.C., A.B., A.A.T, I.F.C., M.J.S.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-17271-9.

Correspondence and requests for materials should be addressed to D.B.

Peer review information Nature Communications thanks Jacqueline Cherfilgs with Agata Nawrot and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.