Article

Structure of the C9orf72 ARF GAP complex that is haploinsufficient in ALS and FTD

Ming-Yuan Su[1,2], Simon A. Fromm[1,2], Roberto Zoncu[1,2] & James H. Hurley[1,2,3 ✉]

Mutation of *C9orf72* is the most prevalent defect associated with amyotrophic lateral sclerosis and frontotemporal degeneration[1]. Together with hexanucleotide-repeat expansion[2,3], haploinsufficiency of *C9orf72* contributes to neuronal dysfunction[4–6]. Here we determine the structure of the C9orf72–SMCR8–WDR41 complex by cryo-electron microscopy. C9orf72 and SMCR8 both contain longin and DENN (differentially expressed in normal and neoplastic cells) domains[7], and WDR41 is a β-propeller protein that binds to SMCR8 such that the whole structure resembles an eye slip hook. Contacts between WDR41 and the DENN domain of SMCR8 drive the lysosomal localization of the complex in conditions of amino acid starvation. The structure suggested that C9orf72–SMCR8 is a GTPase-activating protein (GAP), and we found that C9orf72–SMCR8–WDR41 acts as a GAP for the ARF family of small GTPases. These data shed light on the function of C9orf72 in normal physiology, and in amyotrophic lateral sclerosis and frontotemporal degeneration.

Expansion of hexanucleotide GGGGCC repeats in the first intron of *C9orf72* is the most prevalent genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal degeneration (FTD), and accounts for approximately 40% of familial ALS, 5% of sporadic ALS and 10–50% of FTD[1]. Two hypotheses—which are not mutually exclusive—could explain how the mutation leads to a progressive loss of neurons. The toxic gain-of-function hypothesis suggests that toxic molecules, including RNA and dipeptide-repeat aggregates, disrupt neural function and lead to their destruction. The loss-of-function hypothesis is based on the observation of a reduction in *C9orf72* mRNA and protein levels in patients. The endogenous function of C9orf72 is essential for microglia[4] and for the normal dynamics of axonal actin in motor neurons[5], and restoring normal expression of C9orf72 rescues function in *C9orf72*-mutant model neurons[6].

C9orf72 contains longin and DENN domains[7] (Fig. 1a), and exists as a stable complex with another protein that contains these domains, Smith–Magenis syndrome chromosome region, candidate 8 (SMCR8), as well as the WD repeat-containing protein 41 (WDR41)[8–13] (Fig. 1a). WDR41 targets C9orf72–SMCR8 to lysosomes[14] via an interaction with the transporter PQ loop repeat-containing 2 (PQLC2)[15]. Previously proposed functions of C9orf72–SMCR8 include the regulation of RAB-positive endosomes[16], regulation of RAB8A and RAB39B in membrane transport[8,12], regulation of the ULK1 complex in autophagy[9,12,13,17] and regulation of mTORC1 at lysosomes[10,11,18]. Thus far it has been difficult to deconvolute which of these roles are direct and which are indirect. To gain more insight, we reconstituted and purified the complex, determined its structure and assessed its function as a purified complex.

We expressed and purified full-length human C9orf72–SMCR8 and C9orf72–SMCR8–WDR41 (Extended Data Fig. 1a–c). We determined the structure of C9orf72–SMCR8–WDR41 at a resolution of 3.8 Å by cryo-electron microscopy (cryo-EM) (Fig. 1b, c, Extended Data

Figs. 2, 3, Extended Data Table 1). We were able to visualize the ordered, approximately 120-kDa portion of the complex, which corresponds to about 60% of the total mass of the complex. Portions of the density—notably, in the DENN domains of both C9orf72 and SMCR8—were very well-resolved, such that side-chain density was clear. Other regions (particularly the longin domains of C9orf72 and SMCR8, and the portion of WDR41 most distal to SMCR8) were less well-resolved, and were not clear enough for side-chain placement. The structure has the shape of an eye slip hook with a long dimension of about 140 Å (Fig. 1c). The ring of the hook was straightforward to assign to WDR41 by its appearance as an eight-bladed β-propeller. The remainder of the density showed evidence of two longin domains at the tip of the hook, with the bulk of the hook made up of two DENN domains. The DENN domain of SMCR8 is in direct contact with WDR41, whereas C9orf72 has no direct contact with WDR41. We assigned the hook-tip portion of the longin domain of SMCR8 to residues I165–A219, which were predicted to comprise a long helical extension unique to this domain. The longin and DENN domains of SMCR8 are near each other but not in direct contact, and are connected by a helical linker that consists of residues K320–V383. Both the longin and the DENN domain of C9orf72 are positioned between the longin and DENN domains of SMCR8. This linear arrangement of domains gives the overall complex an elongated shape.

To map the interactions of WDR41 and to facilitate the interpretation of the less well-resolved portions of the cryo-EM structure, we subjected C9orf72–SMCR8 and C9orf72–SMCR8–WDR41 to hydrogen deuterium exchange mass spectrometry (HDX-MS) for 0.5, 5, 50, 500 and 50,000 seconds, and compared them to each other (Fig. 2, Extended Data Figs. 1d–f, 4, 5, Supplementary Data 1). We achieved excellent peptide coverage (89, 87 and 80% for SMCR8, C9orf72 and WDR41, respectively), and consistent patterns were observed across experimental time points. Several regions in SMCR8—including the N-terminal 54 residues, and residues V104–V118, E212–I230, P257–F315, V378–I714 and

[1]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA, USA. [2]California Institute for Quantitative Biosciences, University of California, Berkeley, CA, USA. [3]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ✉e-mail: jimhurley@berkeley.edu
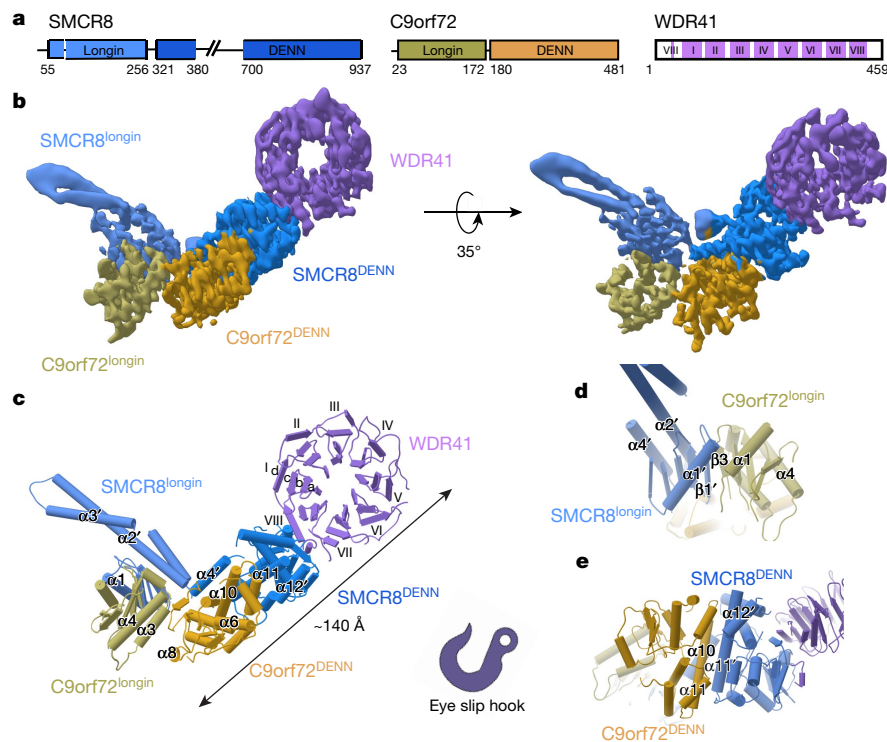
Fig. 1 | Cryo-EM structure of the C9orf72–SMCR8–WDR41 complex.
a, Schematic of the domain structure of the C9orf72–SMCR8–WDR41 complex.
b, **c**, Cryo-EM density map (local filter map, *b* factor, −50 Å²) (**b**) and the refined coordinates (**c**) of the complex shown as pipes and planks for α-helices and β-sheets, respectively. The domains are colour-coded as follows: longin domain of SMCR8 (SMCR8^{longin}), cornflower blue; DENN domain of SMCR8 (SMCR8^{DENN}), Dodger blue; longin domain of C9orf72 (C9orf72^{longin}), olive; DENN domain of C9orf72 (C9orf72^{DENN}), goldenrod; and WDR41, medium purple. **d**, **e**, Organizations of SMCR8^{longin}–C9orf72^{longin} (**d**) and SMCR8^{DENN}–C9orf72^{DENN} (**e**) arrangement.

V788–Y806—showed more than 50% deuterium uptake at 0.5 seconds, which indicates these regions are intrinsically disordered regions—consistent with sequence-based predictions. Nearly all of C9orf72 was protected from exchange, except for the N-terminal 21 residues and the C terminus. For WDR41, the N-terminal 24 residues, and the loops that connect blades II and III (R128–C131), blades V and VI (R260–D270 and L277–I284), the internal loop of blade VII and the loop connecting to blade VIII (R352–L357 and M369–E396) were flexible.

Difference heat maps for C9orf72 and SMCR8 (Fig. 2a, b) showed that, in the presence of WDR41, regions of the DENN domain of SMCR8—including K363–F372 (M1 region), P763–Q778 (M2 region), S729–V735 (M3 region), T807–D811 (M4 region) and the C-terminal K910–Y935 (M5 region)—were protected from exchange (Figs. 2, 3, Extended Data Figs. 4–6), consistent with the structure. There was no substantial exchange in C9orf72, with the exception of K388–R394 (the M1 region of C9orf72) (Figs. 2, 3). We mutagenized the regions that showed protection against exchange, and tested these mutants in co-expression and pulldown experiments (Fig. 2c, d; see 'Protein expression and purification' in Methods for details of the mutants). Except for the helical linker mutant in the M1 region of SMCR8, the SMCR8 mutants abolished the interaction with WDR41. When WDR41 did not pull down SMCR8 mutants, wild-type C9orf72 was not detected either. This confirms the structural finding that SMCR8 bridges the other two components. Because alterations of the M1 region of C9orf72 did not prevent interaction with SMCR8–WDR41, we concluded that this region was protected by a conformational change induced upon WDR41 binding, consistent with the lack of direct interaction in the cryo-EM structure. The interface between SMCR8 and C9orf72 is extensive, mediated by longin–longin and DENN–DENN dimerization (Fig. 1d, e). Substitutions in the C9orf72(F397E/T411W) double mutant disrupt the interaction with SMCR8, as shown by co-expression and pulldown experiments (Extended Data Fig. 7a, b). The cryo-EM structure showed that SMCR8 bound to blade VIII and the C-terminal helix of WDR41 (Fig. 3a, Extended Data Fig. 6). The pulldown experiment showed that the N-terminal residues E35–K40 of blade VIII and the C-terminal helix S442–V459 are required for SMCR8 binding (Extended Data Fig. 7c). Collectively, the HDX-MS and mutational results corroborate the structural interpretation.

WDR41 is responsible for the reversible targeting of C9orf72–SMCR8 to lysosomes under conditions of nutrient depletion[14]. WDR41, in turn, binds to lysosomes via PQLC2[15]. We cotransfected DNA encoding green fluorescent protein-tagged SMCR8 (GFP–SMCR8), C9orf72, WDR41 and PQLC2 tagged with monomeric red fluorescent protein (PQLC2–mRFP) in HEK293A cells. SMCR8 clustered on PQLC2-positive lysosomes in conditions of amino acid depletion and was diffusely localized in the cytosol upon refeeding (Fig. 3b), consistent with previous reports[14,15]. SMCR8 mutants deficient in WDR41 binding in vitro did not colocalize with PQLC2-positive lysosomes, but rather were diffusely localized in the cytosol even under amino acid-starved conditions (Fig. 3b, c). These findings confirm that the WDR41-binding site on SMCR8 as mapped by cryo-EM and HDX-MS is responsible for the lysosomal localization of the complex under conditions of amino acid starvation.

The structure showed that the longin domain of SMCR8 forms a heterodimer with the longin domain of C9orf72 in the same manner as NPRL2–NPRL3 of the GATOR1 complex[19] and FLCN–FNIP2 in the lysosomal folliculin complex[20,21]. The NPRL2 and FLCN subunits of these complexes are GAPs for the lysosomal small GTPases RAGA[22] and RAGC[23], respectively. Structure-based alignment of SMCR8 with FLCN and NPRL2 showed that they shared a conserved arginine finger residue[20,21,24] (Fig. 4a), which corresponds to SMCR8 R147. This arginine residue is exposed on the protein surface near the centre of a large concave surface, which appears suitable for binding a small GTPase (Extended Data Fig. 8). Using a tryptophan fluorescence and high-performance liquid chromatography (HPLC)-based assay, we assayed C9orf72–SMCR8
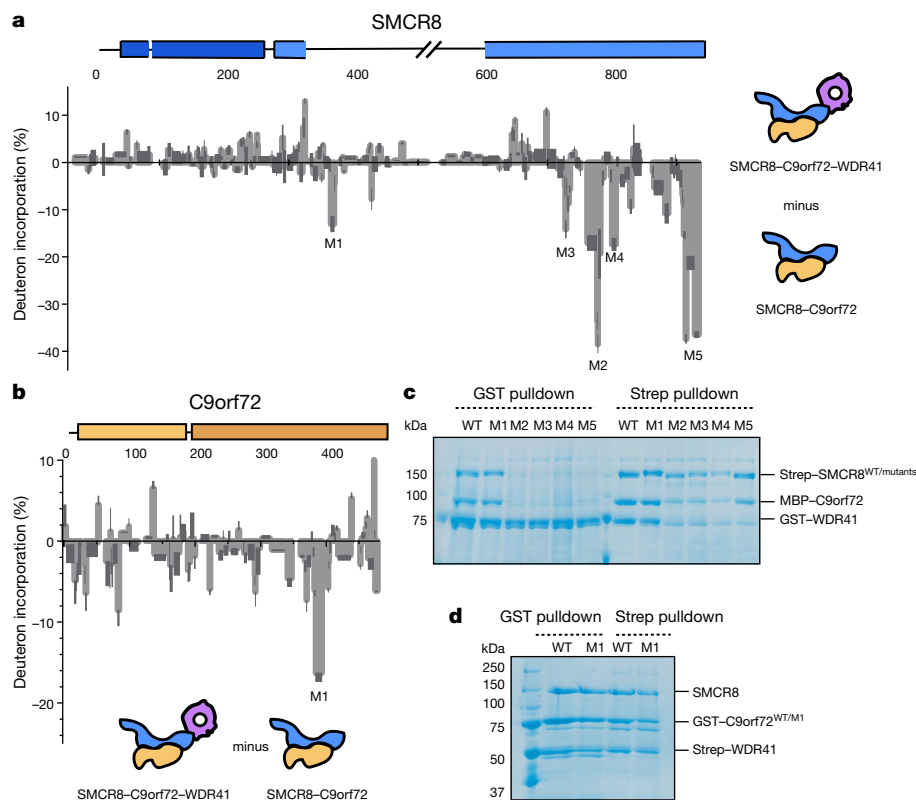
Fig. 2 | HDX-MS of C9orf72–SMCR8 in the absence of WDR41. a, Difference plot of percentage of deuteron incorporation of SMCR8 in the heterotrimer versus the dimer, at the 5-s time point. **b**, Difference plot of percentage of deuteron incorporation of C9orf72 in the heterotrimer versus the dimer, at the 0.5-s time point. **c**, Co-expression and pulldown experiment of Strep-tagged SMCR8 mutants with wild-type MBP–C9orf72 and GST–WDR41. M1–5, mutants of the corresponding M1–5 regions of SMCR8; Strep–SMCR8[WT/mutants]: WT, wild type; mutants, mutants of the M1–5 regions of SMCR8. **d**, Co-expression and pulldown experiments of GST–C9orf72 mutant with wild-type untagged SMCR8 and Strep–WDR41. M1, mutant of the M1 region of C9orf72; GST–C9orf72[WT/M1]: WT, wild type; M1, mutant of the M1 region of C9orf72. The pulldown experiments were repeated at least twice with similar results (**c**, **d**).

for GAP activity with respect to RAGA or RAGC and found none detectable (Extended Data Fig. 9a, b, d). We also assayed for GAP activity with respect to RAB1A[17] and the late endosomal RAB7A[16], and—again—activity was undetectable (Extended Data Fig. 9a, b, d).

It has previously been reported that C9orf72 interacts with the small GTPases ARF1 and ARF6[25] in neurons[5]. We found that C9orf72–SMCR8–WDR41 was an efficient GAP for ARF1 on the basis of both tryptophan fluorescence and HPLC-based assays (Fig. 4). The ARF1(Q71L) GTP-locked mutant had no activity (Fig. 4b, Extended Data Fig. 10); nor did the version of the complex that contained the SMCR8(R147A) finger mutant. FLCN–FNIP2 and GATOR1 had no GAP activity towards ARF1. C9orf72–SMCR8 was as active as C9orf72–SMCR8–WDR41, consistent with the location of WDR41 on the opposite side of the complex from R147. C9orf72–SMCR8–WDR41 has activity against the other members of the ARF family, ARF5 and ARF6 (Extended Data Fig. 9a, c, d)— but not against the lysosomal ARF-like proteins ARL8A and ARL8B (Extended Data Fig. 9a, b, d). These observations clarify the nature of the reported C9orf72–ARF interaction by showing that the role of C9orf72 is to stabilize a complex with SMCR8, which is—in turn—an efficient and selective GAP for ARF GTPases.

RAB5A[26], RAB7A[26], RAB8A[8] and RAB39B[8,12] have all previously been reported to be guanine nucleotide exchange factor (GEF) substrates of C9orf72. We tested the activity of the purified complex with respect to these RAB proteins and another putative C9orf72 interactor, RAB1A[17]. Compared to a RABEX5 and RAB5A positive control, no exchange was observed on any of these upon addition of C9orf72–SMCR8–WDR41 (Extended Data Fig. 11a, b). The structure of RAB35 in complex with the GEF DENND1B[27] was previously used as a basis for modelling[26]. In comparing C9orf72 in our structure with the structure of DENND1B

in complex with RAB35[27], we found that the alignment of the longin domains showed that RAB35 collides with the longin domain of SMCR8, and superimposition of DENN domains indicated that RAB35 collides with the longin domain of C9orf72—consistent with our result that C9orf72–SMCR8 does not have DENND1B-like GEF activity (Extended Data Fig. 11c).

These data shed light on the normal function of C9orf72, which is thought to contribute to neuronal loss-of-function in ALS and FTD[6]. The structure shows that C9orf72 is the central component of its complex with SMCR8. The longin and DENN domains of SMCR8 flank, and are stabilized by, C9orf72. SMCR8 contains the binding site for WDR41, which is responsible for lysosomal localization during amino acid starvation. C9orf72–SMCR8 belongs to the same class of double-longin-domain GAP complexes as GATOR1[19] and FLCN–FNIP2[20,21]. Unlike GATOR1 and FLCN–FNIP2, C9orf72–SMCR8 is inactive against RAG GTPases but is active against ARF GTPases. The GAP active site is located at the opposite end of the complex from the lysosomal targeting site on WDR41.

Our in vitro observation that C9orf72–SMCR8 and C9orf72–SMCR8–WDR41 have comparable GAP activities suggests that—in cells—C9orf72–SMCR8 may regulate ARF GTPases both in full nutrient conditions, when the complex is primarily localized in the cytosol, and under conditions of amino acid starvation, when it relocalizes to the lysosomal membrane via interactions between WDR41 and PQLC2. However, additional factors could limit or augment the ARF GAP activity of C9orf72–SMCR8 in either condition, and restrict or enhance access to the GTP-bound ARF substrate. ARF proteins are not observed on lysosomes, and their closest lysosomal cousins (ARL8A and ARL8B) are not substrates for C9orf72–SMCR8. Thus, sequestration of C9orf72–SMCR8–WDR41 on lysosomes could prevent it from
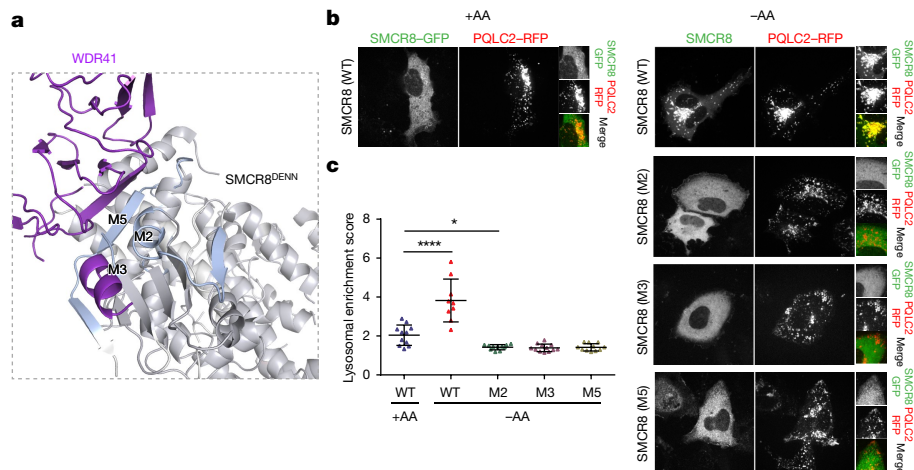
Fig. 3 | SMCR8 mutants do not localize on lysosomes. a, The HDX uptake difference at 0.5 s was mapped on C9orf72–SMCR8. Close-up view of the SMCR8–WDR41 interface, highlighting the SMCR8 mutants. **b**, SMCR8–PQLC2 lysosome colocalization experiment in cells that express the indicated SMCR8 constructs under the indicated nutrient conditions. –AA, cells starved for amino acids for 1 h; +AA, cells starved for amino acids for 1 h, and then restimulated with amino acids for 10 min. The experiment was repeated at least three times independently with similar results. **c**, Quantification of SMCR8 lysosomal enrichment score for fluorescence images shown in **b**. Data are mean ± s.d. From left to right, $n = 11, 9, 11, 12$ and 11 cells were quantified for each condition. *$P < 0.05$, ****$P < 0.0001$, evaluated by one-way analysis of variance.

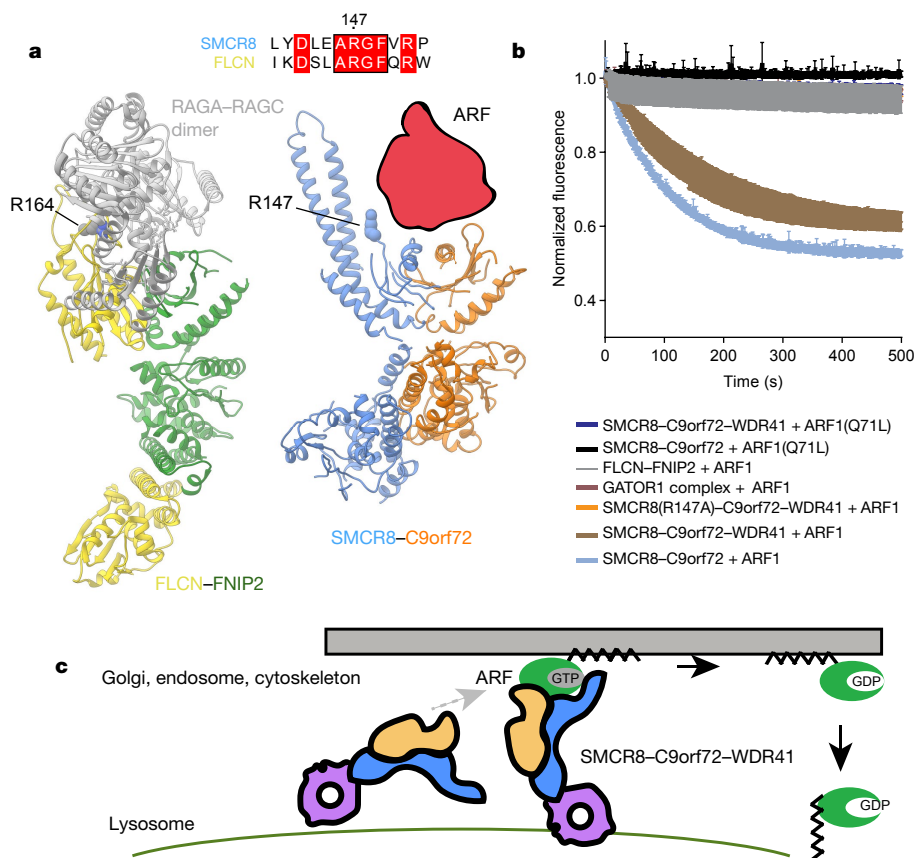**Fig. 4 | C9orf72–SMCR8 is a GAP for ARF proteins. a**, Structure comparison of FNIP2–FLCN and C9orf72–SMCR8, which suggests a potential binding site for substrates. The conserved arginine residue is shown in spherical representation. **b**, Tryptophan fluorescence GTPase signal was measured for wild-type ARF1 or ARF1(Q71L) before and after the addition of the C9orf72–SMCR8–WDR41, C9orf72–SMCR8(R147A)–WDR41, C9orf72–SMCR8, FLCN–FNIP2 or GATOR1 complex. The fluorescence signal upon addition of the GAP was normalized to 1 for each experiment. Plots are mean ± s.d. of triplicate technical experiments. **c**, Model for ARF protein family activation by C9orf72–SMCR8–WDR41.

regulating the ARF proteins in *cis* under unfavourable metabolic conditions. Alternatively, C9orf72–SMCR8–WDR41 could act in *trans* on ARF proteins bound to the membrane of a compartment other than the lysosome. ARF GTPases are found on the Golgi, endosomes, plasma membrane and cytoskeleton, and in the cytosol[25], and function on membranes in their active GTP-bound form. C9orf72 can associate

with endosomes[6,16,28] and the cytoskeleton[5], which could be loci of the ARF substrate of C9orf72–SMCR8. The potential *trans* GAP activity of C9orf72–SMCR8–WDR41 versus endosomal or cytoskeletal ARF would be facilitated by its elongated structure and the distal positioning of the GAP and lysosomal localization sites (Fig. 4c).

Haploinsufficient GAP activity for ARF GTPases could contribute to ALS and FTD in several ways. Defects in actin dynamics in neurons could contribute to problems with endosomal transport[5]. Indeed, several studies connect C9orf72 to endosomal sorting[6,16,28], a process in which the role of ARF proteins is well-established[25]. It has previously been reported that ARF1 promotes mTORC1 activation[29], so the GAP function of C9orf72–SMCR8 with respect to ARF GTPases could explain how this complex antagonizes mTORC1[18]. mTORC1 negatively regulates autophagy, and thus the ARF1–mTORC1 connection could explain how haploinsufficient C9orf72 leads to a decrease in autophagy—which has, in turn, previously been linked to multiple neurodegenerative diseases[30]. While this Article was under review, the cryo-EM structure of a dimeric form of the C9orf72–SMCR8–WDR41 complex was reported and proposed to serve as a GAP for RAB8A and RAB11A[31]. The relative roles of GAP activity with respect to different small GTPases in normal function and disease remain to be determined. The structural and in vitro biochemical data reported here, and previously[31], provide a framework and a foothold for understanding how the normal functions of C9orf72 relate to lysosomal signalling, autophagy and neuronal survival.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2633-x.

1. Majounie, E. et al. Frequency of the *C9orf72* hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* **11**, 323–330 (2012).
2. DeJesus-Hernandez, M. et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9ORF72* causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
3. Renton, A. E. et al. A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 (2011).
4. O'Rourke, J. G. et al. *C9orf72* is required for proper macrophage and microglial function in mice. *Science* **351**, 1324–1329 (2016).
5. Sivadasan, R. et al. C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons. *Nat. Neurosci.* **19**, 1610–1618 (2016).
6. Shi, Y. et al. Haploinsufficiency leads to neurodegeneration in *C9ORF72* ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313–325 (2018).
7. Zhang, D., Iyer, L. M., He, F. & Aravind, L. Discovery of novel DENN proteins: implications for the evolution of eukaryotic intracellular membrane structures and human disease. *Front. Genet.* **3**, 283 (2012).
8. Sellier, C. et al. Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J.* **35**, 1276–1297 (2016).
9. Sullivan, P. M. et al. The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol. Commun.* **4**, 51 (2016).
10. Amick, J., Roczniak-Ferguson, A. & Ferguson, S. M. C9orf72 binds SMCR8, localizes to lysosomes, and regulates mTORC1 signaling. *Mol. Biol. Cell* **27**, 3040–3051 (2016).
11. Ugolino, J. et al. Loss of C9orf72 enhances autophagic activity via deregulated mTOR and TFEB signaling. *PLoS Genet.* **12**, e1006443 (2016).
12. Yang, M. et al. A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Sci. Adv.* **2**, e1601167 (2016).
13. Jung, J. et al. Multiplex image-based autophagy RNAi screening identifies SMCR8 as ULK1 kinase activity and gene expression regulator. *eLife* **6**, e23063 (2017).
14. Amick, J., Tharkeshwar, A. K., Amaya, C. & Ferguson, S. M. WDR41 supports lysosomal response to changes in amino acid availability. *Mol. Biol. Cell* **29**, 2213–2227 (2018).
15. Amick, J., Tharkeshwar, A. K., Talaia, G. & Ferguson, S. M. PQLC2 recruits the C9orf72 complex to lysosomes in response to cationic amino acid starvation. *J. Cell Biol.* **219**, e20190676 (2020).
16. Farg, M. A. et al. C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum. Mol. Genet.* **23**, 3579–3595 (2014).
17. Webster, C. P. et al. The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J.* **35**, 1656–1676 (2016).
18. Lan, Y., Sullivan, P. M. & Hu, F. SMCR8 negatively regulates AKT and MTORC1 signaling to modulate lysosome biogenesis and tissue homeostasis. *Autophagy* **15**, 871–885 (2019).
19. Shen, K. et al. Architecture of the human GATOR1 and GATOR1–Rag GTPases complexes. *Nature* **556**, 64–69 (2018).
20. Lawrence, R. E. et al. Structural mechanism of a Rag GTPase activation checkpoint by the lysosomal folliculin complex. *Science* **366**, 971–977 (2019).
21. Shen, K. et al. Cryo-EM structure of the human FLCN–FNIP2–Rag–Ragulator complex. *Cell* **179**, 1319–1329 (2019).
22. Bar-Peled, L. et al. A tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. *Science* **340**, 1100–1106 (2013).
23. Tsun, Z. Y. et al. The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. *Mol. Cell* **52**, 495–505 (2013).
24. Shen, K., Valenstein, M. L., Gu, X. & Sabatini, D. M. Arg78 of Nprl2 catalyzes GATOR1-stimulated GTP hydrolysis by the Rag GTPases. *J. Biol. Chem.* **294**, 2970–2975 (2019).
25. Sztul, E. et al. ARF GTPases and their GEFs and GAPs: concepts and challenges. *Mol. Biol. Cell* **30**, 1249–1271 (2019).
26. Iyer, S., Subramanian, V. & Acharya, K. R. C9orf72, a protein associated with amyotrophic lateral sclerosis (ALS) is a guanine nucleotide exchange factor. *PeerJ* **6**, e5815 (2018).
27. Wu, X. et al. Insights regarding guanine nucleotide exchange from the structure of a DENN-domain protein complexed with its Rab GTPase substrate. *Proc. Natl Acad. Sci. USA* **108**, 18672–18677 (2011).
28. Corrionero, A. & Horvitz, H. R. A C9orf72 ALS/FTD ortholog acts in endolysosomal degradation and lysosomal homeostasis. *Curr. Biol.* **28**, 1522–1535 (2018).
29. Jewell, J. L. et al. Differential regulation of mTORC1 by leucine and glutamine. *Science* **347**, 194–198 (2015).
30. Nixon, R. A. The role of autophagy in neurodegenerative disease. *Nat. Med.* **19**, 983–997 (2013).
31. Tang, D. et al. Cryo-EM structure of C9ORF72–SMCR8–WDR41 reveals the role as a GAP for Rab8a and Rab11a. *Proc. Natl Acad. Sci. USA* **117**, 9876–9883 (2020).

Article

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Protein expression and purification

Synthetic genes encoding SMCR8 were amplified by PCR and cloned into the pCAG vector coding for an N-terminal twin Strep–Flag tag using KpnI and XhoI restriction sites. The pCAG vector encoding an N-terminal GST followed by a TEV restriction site or uncleaved MBP tag was used for expression of C9orf72. WDR41 was cloned into pCAG vector without a tag or with a GST tag for pulldown experiments. For the mutations of SMCR8 identified from HDX experiments, the M1 region of SMCR8 (K363–F372) was mutated to MSDYDIPTTE, which is a 10-residue linker derived from the pETM11 vector. For lysosome localization experiments, the M2 region of SMCR8 (P771–Q778 or K762–L782) was mutated to GGKGSGGS. Mutants of the M3 (S729–V735) and M4 (T807–D811) regions of SMCR8 were made by mutating these regions to GGKGSGG and GGKGS, respectively. The mutant of the M5 region of SMCR8 was made by truncation after residue K910. The M1 region of C9orf72 (K388–L393) was mutated to polyalanine. The SMCR8 arginine finger mutation (R147A), C9orf72(F397E) and C9orf72(T411W) mutants were made using two-step PCR and cloned into the expression vector.

HEK293-GnTi cells adapted for suspension were grown in Freestyle medium supplemented with 1% FBS and 1% antibiotic–antimycotic at 37 °C, 80% humidity, 5% CO_2 and shaking at 140 rpm. Once the cultures reached 1.5–2 million cells ml^{-1} in the desired volume, they were transfected as followed. For a 1-l transfection, 3 ml PEI (1 mg ml^{-1}, pH 7.4, Polysciences) was added to 50 ml hybridoma medium (Invitrogen) and 1 mg of total DNA (isolated from transformed *Escherichia coli* XL10-gold) in another 50 ml hybridoma medium. One mg of transfection DNA contained an equal mass ratio of C9orf72-complex expression plasmids. PEI was added to the DNA, mixed and incubated for 15 min at room temperature. One hundred ml of the transfection mix was then added to each 1-l culture. Cells were collected after 3 days.

Cells were lysed by gentle rocking in lysis buffer containing 50 mM HEPES, pH 7.4, 200 mM NaCl, 2 mM MgCl$_2$, 1% (v/v) Triton X-100, 0.5 mM TCEP, protease inhibitors (AEBSF, leupeptin and benzamidine) and supplemented with phosphatase inhibitors (50 mM NaF and 10 mM β-glycerophosphate) at 4 °C. Lysates were clarified by centrifugation (15,000g for 40 min at 4 °C) and incubated with 5 ml glutathione sepharose 4B (GE Healthcare) for 1.5 h at 4 °C with gentle shaking. The glutathione sepharose 4B matrix was applied to a gravity column, washed with 100 ml wash buffer (20 mM HEPES, pH 7.4, 200 mM NaCl, 2 mM MgCl$_2$ and 0.5 mM TCEP), and purified complexes were eluted with 40 ml wash buffer containing 50 mM reduced glutathione. Eluted complexes were treated with TEV protease at 4 °C overnight. TEV-treated complexes were purified to homogeneity by injection on Superose 6 10/300 (GE Healthcare) column that was pre-equilibrated in gel filtration buffer (20 mM HEPES, pH 7.4, 200 mM NaCl, 2 mM MgCl$_2$ and 0.5 mM TCEP). For long-term storage, fractions from the gel filtration chromatography were frozen using liquid nitrogen and kept at −80 °C. C9orf72–SMCR8 and C9orf72–SMCR8–WDR41 were expressed and purified using the same protocol (Supplementary Fig. 1).

For expression of human His$_6$-tagged ARF1 (residues E17–K181), ARF1(Q71L), ARF5(Q17–Q180), ARF6(R15–S175), ARF6(Q67L), His$_6$–RAB1A, His$_6$–ARL8A(E20–S186), His$_6$–ARL8B(E20–S186), His$_6$–RAB39B and bovine His$_6$–RABEX5 helix bundle–Vps9 domain (S133–E398), plasmids were transformed into *E. coli* BL21 DE3 star cells and induced with 0.5 mM IPTG at 18 °C overnight. The cells were lysed in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 2 mM MgCl$_2$, 5 mM imidazole, 0.5 mM TCEP and 1 mM PMSF by ultrasonication. The lysate was centrifuged at 15,000g for 30 min. The supernatant was loaded into Ni-NTA resin and washed with 20 mM imidazole and eluted with 300 mM imidazole. The eluate was further purified on a Superdex 75 10/300 (GE Healthcare) column equilibrated in 20 mM HEPES, pH 7.4, 200 mM NaCl, 2 mM MgCl$_2$ and 0.5 mM TCEP. RAG, FLCN–FNIP2 and GATOR1 complex were purified as previously described[20]. GST-tagged human RAB7A, or RAB5A (*Canis familiaris*), was expressed in the same conditions as above and purified with GST resin, eluted in 50 mM reduced glutathione buffer and applied on Superdex 200 column. Twin Strep–Flag tag RAB8A was expressed in HEK293-GnTi cells and purified by Strep resin and eluted in 10 mM desthiobiotin buffer. The eluted protein was applied on Superdex 75 10/300 column.

Hydrogen–deuterium exchange experiment

Sample quality was assessed by SDS–PAGE before each experiment. Amide hydrogen exchange mass spectrometry was initiated by a 20-fold dilution of 10 μM C9orf72–SMCR8–WDR41 or C9orf72–SMCR8 into 95 μl D$_2$O buffer containing 20 mM HEPES pH (pD 8.0), 200 mM NaCl, 1 mM TCEP at 30 °C. Incubations in deuterated buffer were performed at intervals from 0.5, 5, 50, 500 and 50,000 s (0.5 s was carried out by incubating proteins with ice-cold D$_2$O for 5 s). All exchange reactions were carried out in triplicate or quadruplicate. Backbone amide exchange was quenched at 0 °C by the addition of ice-cold quench buffer (400 mM KH$_2$PO$_4$/H$_3$PO$_4$, pH 2.2). The 50,000-s sample served as the maximally labelled control. Quenched samples were injected onto a chilled HPLC setup with in-line peptic digestion and then eluted onto a BioBasic 5 μM KAPPA Capillary HPLC column (Thermo Fisher Scientific), equilibrated in buffer A (0.05% TFA), using 10–90% gradient of buffer B (0.05% TFA, 90% acetonitrile) over 30 min. Desalted peptides were eluted and directly analysed by an Orbitrap Discovery mass spectrometer (Thermo Fisher Scientific). The spray voltage was 3.4 kV and the capillary voltage was 37 V. The HPLC system was extensively cleaned between samples. Initial peptide identification was performed via tandem mass spectrometry experiments. A Proteome Discoverer 2.1 (Thermo Fisher Scientific) search was used for peptide identification and coverage analysis against entire complex components, with precursor mass tolerance ± 10 ppm and fragment mass tolerance of ± 0.6 Da. Mass analysis of the peptide centroids was performed using HDExaminer (Sierra Analytics), followed by manual verification of each peptide. The difference plots were prepared using Origin 6.0.

Cryo-EM grid preparation and data acquisition

The purified C9orf72–SMCR8–WDR41 complex was diluted to 0.8 μM in 20 mM HEPES pH 7.4, 2 mM MgCl$_2$ and 0.5 mM TCEP, and applied to glow-discharged C-flat (1.2/1.3, Au 300 mesh) grids. The sample was vitrified after blotting for 2 s using a Vitrobot Mark IV (FEI) with 42-s incubation, blot force 8 and 100% humidity. The complex was visualized with a Titan Krios electron microscope (FEI) operating at 300 kV with a Gatan Quantum energy filter (operated at 20-eV slit width) using a K2 summit direct electron detector (Gatan) in super-resolution counting mode, corresponding to a super-resolution pixel size of 0.5745 Å on the specimen level. In total, 3,508 movies were collected in nanoprobe mode using Volta phase plate (VPP) with defocus collected at around −60 nm. Movies consisted of 49 frames, with a total dose of 59.8 e$^-$ per Å^2, a total exposure time of 9.8 s and a dose rate of 8.1 e$^-$ per pixel per s. Data were acquired with SerialEM using custom macros for automated single-particle data acquisition. Imaging parameters for the dataset are summarized in Extended Data Table 1.

Cryo-EM data processing

Preprocessing was performed during data collection within Focus[32]. Drift, beam-induced motion and dose weighting were corrected with MotionCor2[33] using 5 × 5 patches and Fourier cropping with a factor of 0.5 after motion correction. CTF fitting and phase-shift estimation were performed using Gctf v.1.06[34], which yielded the characterized pattern of phase-shift accumulation over time for each position. The data were manually inspected and micrographs with excess ice contamination or shooting on the carbon were removed. A total of 4,810,184 particles

from 3,220 micrographs were picked using gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/) and extracted with binning 4. All subsequent classification and reconstruction steps were performed using Relion3-beta[35] or cryoSPARC v.2[36]. The particles were subjected to 3D classification ($K = 5$) using a 60 Å low-pass-filtered ab initio reference generated in cryoSPARC. Around 2.2 million particles from the two best classes were selected for 3D auto-refinement and another round of 3D classification ($K = 8$ classes, $T = 8$, E-step = 8 Å) without alignment. About 1.8 million particles from the best 6 classes were re-extracted with binning 2 and refined to 4.9 Å, and further subjected to 2D classification without alignment for removing contamination and junk particles. After another round of 3D classification ($K = 4$) with alignment, the best class was extracted and imported into cryoSPARC v.2 for another round of 2D classification. The cleaned-up 571,002 particles were subjected to CTF refinement, Bayesian polishing, and further particles at the edges were removed in Relion 3. A final set of 381,450 particles resulted in final resolution of 3.8 Å, with a measured map B-factor of -102 Å^2. More extensive 3D classification and focus classification in Relion3 did not improve the quality of the reconstruction. Local filtering and B-factor sharpening were done in cryoSPARC v.2. All reported resolutions are based on the gold-standard Fourier shell correlation (FSC) 0.143 criterion.

Atomic model building and refinement
The model of WDR41 was generated with I-Tasser[37] and used Protein Data Bank codes (PDB) 5NNZ, 2YMU, 5WLC, 4NSX and 6G6M as starting models. The model of the longin domain of C9orf72 was generated on the basis of the longin domain of NPRL2 (PDB 6CES) in Modeller[38]. The model of the DENN domain of SMCR8 was generated from Modeller and RaptorX[39] using the DENN domain of FLCN (PDB 3V42) or DENND1B (PDB 3TW8) as templates. The longin domain of SMCR8 and DENN domain of C9orf72 were generated with Phyre2[40] using longin domain of FLCN and DENN domain of FNIP2 (PDB 6NZD), respectively, as templates. Secondary structure predictions of each protein were carried out with Phyre2[40] or Psipred[41]. The models were docked into the 3D map as rigid bodies in UCSF Chimera[42]. The coordinates of the structures were manually adjusted and rebuilt in Coot[43]. The resulting models were refined using Phenix.real_space.refine in the Phenix suite with secondary structure restraints and a weight of 0.1[44,45]. Model quality was assessed using MolProbity[46] and the map-versus-model FSC (Extended Data Fig. 3c, Extended Data Table 1). Data used in the refinement excluded spatial frequencies beyond 4.2 Å to avoid overfitting. A half-map cross-validation test showed no indication of overfitting (Extended Data Fig. 3d). Figures were prepared using UCSF Chimera[42] and PyMOL v.1.7.2.1.

Live cell imaging
Eight hundred thousand HEK 293A cells were plated onto fibronectin-coated glass-bottom Mattek dishes and transfected with the indicated wild-type GFP–SMCR8 or mutants, C9orf72, WDR41 and PQLC2–mRFP constructs with transfection reagent Xtremegene. Twenty-four h later, cells were starved for amino acids for 1 h (−AA), or starved and restimulated with amino acids for 10 min (+AA). Cells in the −AA condition were transferred to imaging buffer (10 mM HEPES, pH 7.4, 136 mM NaCl, 2.5 mM KCl, 2 mM CaCl$_2$, 1.2 mM MgCl$_2$) and cells in the +AA condition were transferred to imaging buffer supplemented with amino acids, 5 mM glucose and 1% dialysed FBS (+AA), and imaged by spinning-disc confocal microscopy. Lysosomal enrichment was scored as previously described[20] using a home-built MATLAB script to determine the lysosomal enrichment of GFP–SMCR8. The score was analysed for at least ten cells for each condition. The one-way analyses of variance were calculated using Prism 6 (Graphpad).

HPLC analysis of nucleotides
The nucleotides bound to small GTPases were assessed by heating the protein to 95 °C for 5 min followed by 5 min centrifugation at 16,000g. The supernatant was loaded onto a HPLC column (Eclipse XDB-C18, Agilent). Nucleotides were eluted with HPLC buffer (10 mM tetra-n-butylammonium bromide, 100 mM potassium phosphate pH 6.5, 7.5% acetonitrile). The identity of the nucleotides was compared to GDP and GTP standards.

HPLC-based GAP assay
HPLC-based GTPase assays were carried out by incubating 30 μl of GTPases (30 μM) with or without GAP complex at a 1:50 molar ratio for 30 min at 37 °C. Samples were boiled for 5 min at 95 °C and centrifuged for 5 min at 16,000g. The supernatant was injected onto an HPLC column as described in 'HPLC analysis of nucleotides'. The experiments were carried out in triplicate and one representative plot is shown.

Tryptophan-fluorescence-based GAP assay
Fluorimetry experiments were performed using a FluoroMax-4 (Horiba) instrument and a quartz cuvette compatible with magnetic stirring (Starna Cells), a path length of 10 mm, and were carried out in triplicate. The tryptophan fluorescence signal was collected using 297-nm excitation (1.5-nm slit) and 340-nm emission (20-nm slit). Experiments were performed in gel filtration buffer at room temperature with stirring. Data collection commenced with an acquisition interval of 1 s. Two μM GTPase was added to the cuvette initially. Once the signal was equilibrated, C9orf72–SMCR8–WDR41, C9orf72–SMCR8(R147A)–WDR41, C9orf72–SMCR8, FLCN–FNIP2 or GATOR1 complex was pipetted into the cuvette at a 1:10 molar ratio. Time (t) = 0 corresponds to GAP addition. The fluorescence signal upon GAP addition was normalized to 1 for each experiment. Mean ± s.d. of three replicates per conditions or one representative plot were plotted.

MantGDP loading for GEF assay
To load GTPases for the N-methylanthraniloyl (mant) fluorescence-based GEF assay, purified GTPases were diluted at least 1:10 into PBS buffer without MgCl$_2$ (10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, 137 mM NaCl, 2.7 mM KCl). EDTA was added to a final concentration of 5 mM and incubated at room temperature for 10 min. A tenfold molar excess of mantGDP nucleotide (Millipore Sigma) was added to the GTPases and incubated for 30 min at room temperature. After addition of MgCl$_2$ to a final concentration of 20 mM and incubation at room temperature for 10 min, unbound nucleotides were removed by buffer exchange into gel filtration buffer using a PD-10 column (GE Healthcare).

GEF assay
GEF assays were carried out with the same instrument and cuvette as the tryptophan fluorescence assays (see 'Tryptophan-fluorescence-based GAP assay'). Mant fluorescence was collected using a 360-nm excitation (10-nm slit) and 440-nm emission (10-nm slit). Experiments were performed in gel filtration buffer at room temperature. Five hundred μl of gel filtration buffer was added to the cuvette, and after baseline equilibration, 20 μl of the respective GTPase with or without RABEX5 or C9orf72–SMCR8–WDR41 were added to a final concentration of 350 nM. After signal equilibration, the assay commenced by addition of 20 μl of GTP to a final concentration of 5 μM (about 15-fold molar excess over the respective GTPase) and fluorescence was measured in 1-s intervals for 1,400 s. All experiments were performed in triplicates. Data were baseline-subtracted and normalized to the signal immediately after GTP addition, which also is the 0-s time point in the plots. Plots are mean ± s.d. of each triplicate experiment.

Cell line authentication
Both HEK293 GnTi and HEK 293A cell lines were purchased from the UC Berkeley Cell Culture Facility, and were authenticated by short-tandem repeat analysis and confirmed to be mycoplasma-negative by nuclear staining and fluorescence microscopy screening.

Article

Data availability

The electron microscopy density map has been deposited in the Electron Microscopy Data Bank with accession number EMD-21048. Atomic coordinates for C9orf72–SMCR8–WDR41 have been deposited in the PDB with accession number 6V4U. Source data are provided with this paper.

32. Biyani, N. et al. Focus: the interface between data collection and data processing in cryo-EM. *J. Struct. Biol.* **198**, 124–133 (2017).
33. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
34. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
35. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
36. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
37. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).
38. Webb, B. & Sali, A. Protein structure modeling with MODELLER. *Methods Mol. Biol.* **1137**, 1–15 (2014).
39. Peng, J. & Xu, J. RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins* **79** (Suppl 10), 161–171 (2011).
40. 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. Protocols* **10**, 845–858 (2015).
41. McGuffin, L. J., Bryson, K. & Jones, D. T. The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404–405 (2000).
42. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
43. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
44. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).
45. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
46. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

Author contributions M.-Y.S. designed and carried out all experiments, except for GEF assays and lysosome colocalization experiments, and performed all data analysis. S.A.F. performed GEF assays. R.Z. carried out lysosome colocalization experiments. M.-Y.S. and J.H.H. conceptualized the project and wrote the first draft. All authors contributed to the editing of the manuscript.

a, SMCR8-C9orf72-WDR41

b, SMCR8-C9orf72

c

d, C9orf72 0.5 sec
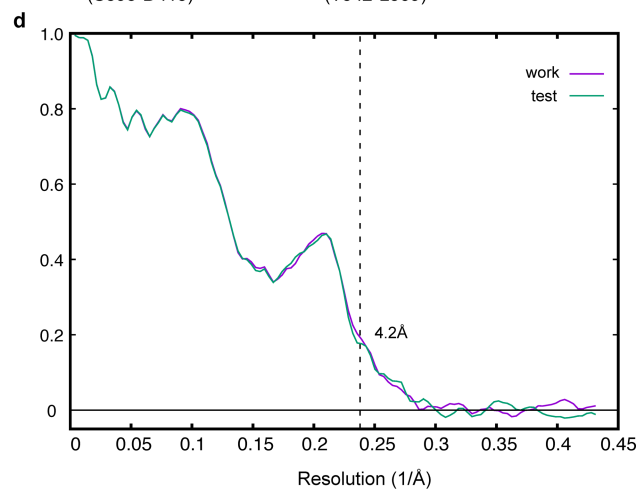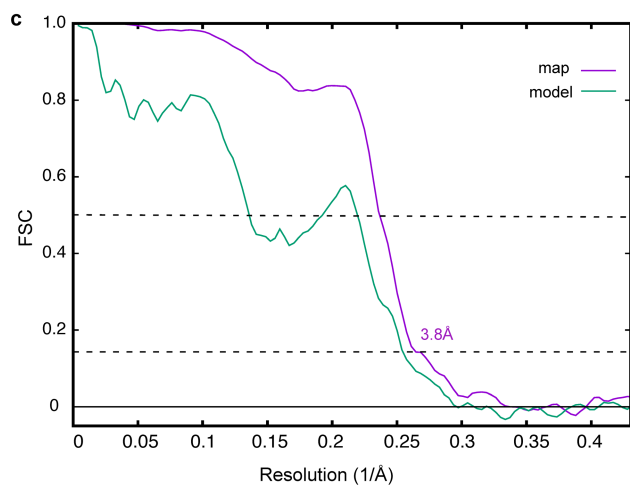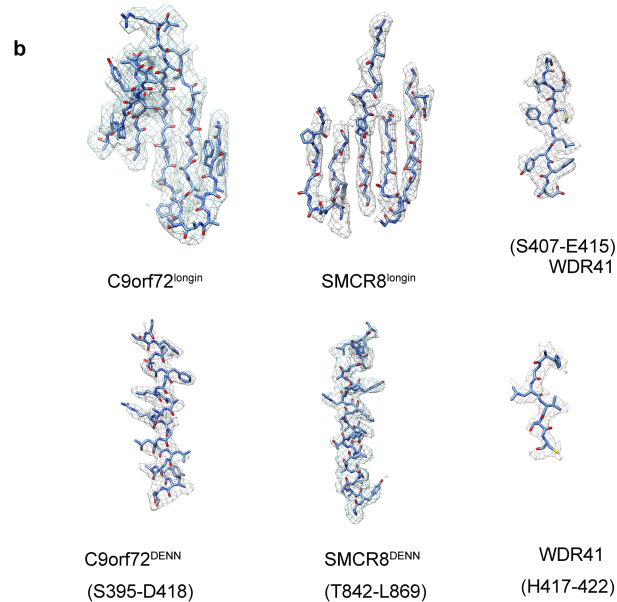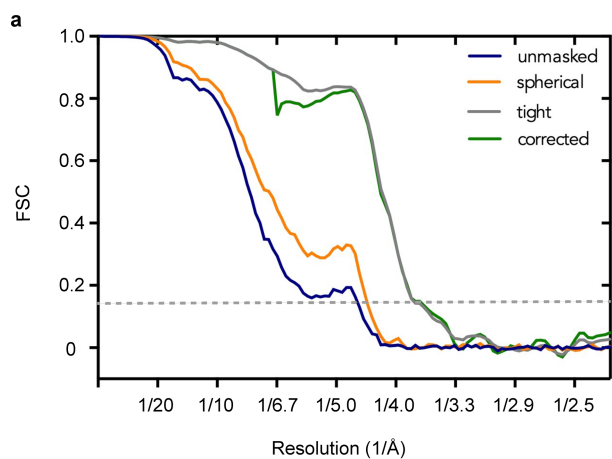
e, SMCR8 0.5 sec

f, WDR41 0.5 sec
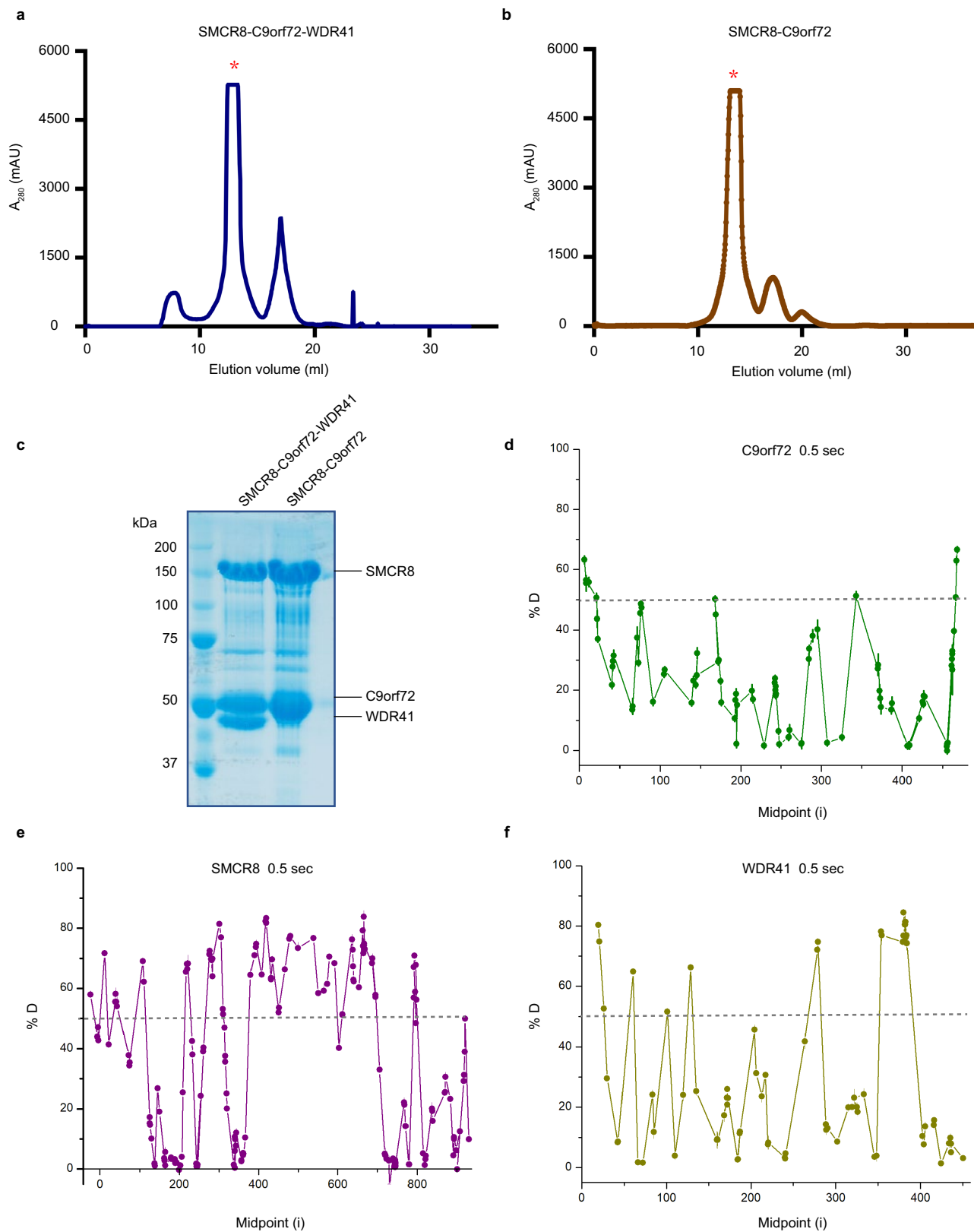
Extended Data Fig. 1 | Purification of the C9orf72–SMCR8–WDR41 and C9orf72–SMCR8 complex, as well as the HDX data for the trimer. a, The Superose-6 gel filtration elution profile for the C9orf72–SMCR8–WDR41 complex. **b**, The Superose-6 gel filtration elution profile for the C9orf72–SMCR8 complex. mAU, milli-absorbance units. **c**, The purified full-length C9orf72–SMCR8–WDR41 and C9orf72–SMCR8 complexes were analysed by SDS–PAGE. The proteins were purified at least five times with similar results (**a**–**c**). **d**–**f**, Deuterium uptake data for the C9orf72–SMCR8–WDR41 complex at the 0.5-s time point, with error bars from triplicate technical measurements. Peptides with more than 50% deuterium uptake are the flexible regions. Y axis represents the average per cent deuteration. X axis demonstrates the midpoint of a single peptic peptide.
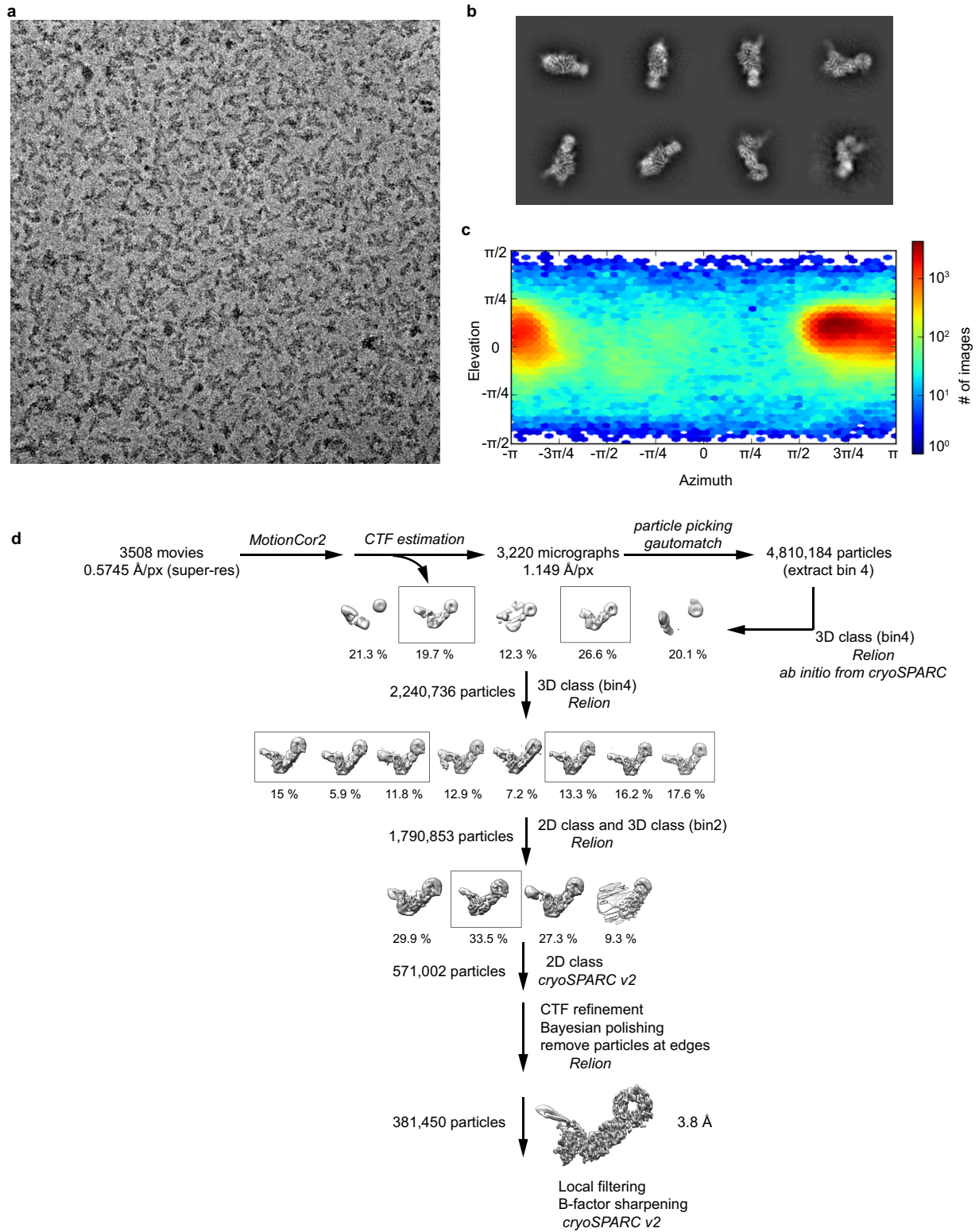
Article

Extended Data Fig. 2 | Cryo-EM data processing. a, A representative cryo-EM micrograph of the C9orf72–SMCR8–WDR41 complex. **b**, Representative 2D classes. **c**, Orientation distribution of the aligned C9orf72–SMCR8–WDR41 particles. **d**, Image processing procedure.

a

b

C9orf72[longin]

SMCR8[longin]

(S407-E415)
WDR41

C9orf72[DENN]
(S395-D418)

SMCR8[DENN]
(T842-L869)

WDR41
(H417-422)

c

3.8Å

d

4.2Å

e

90°   90°   90°
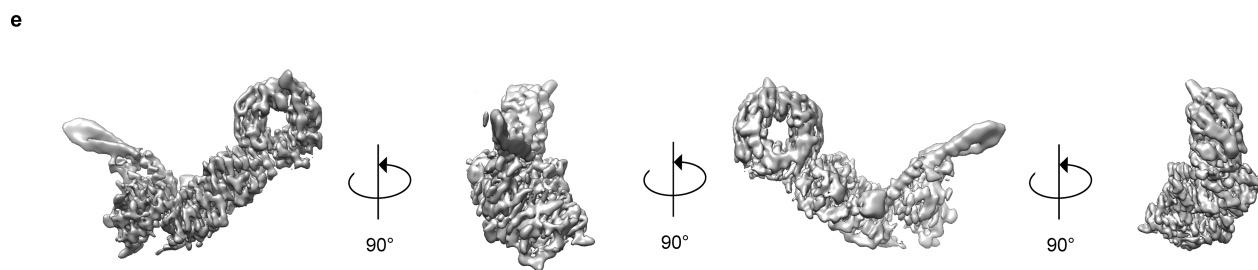
Extended Data Fig. 3 | Resolution estimation of the cryo-EM map, as well as model building and validation. a, Comparison between FSC curves. **b**, Refined coordinate model fit of the indicated region in the cryo-EM density. **c**, Refinement and map-versus-model FSC. **d**, Cross-validation of test FSC curves to assess overfitting. The refinement target resolution (4.2 Å) is indicated. **e**. Different views of the final reconstruction.
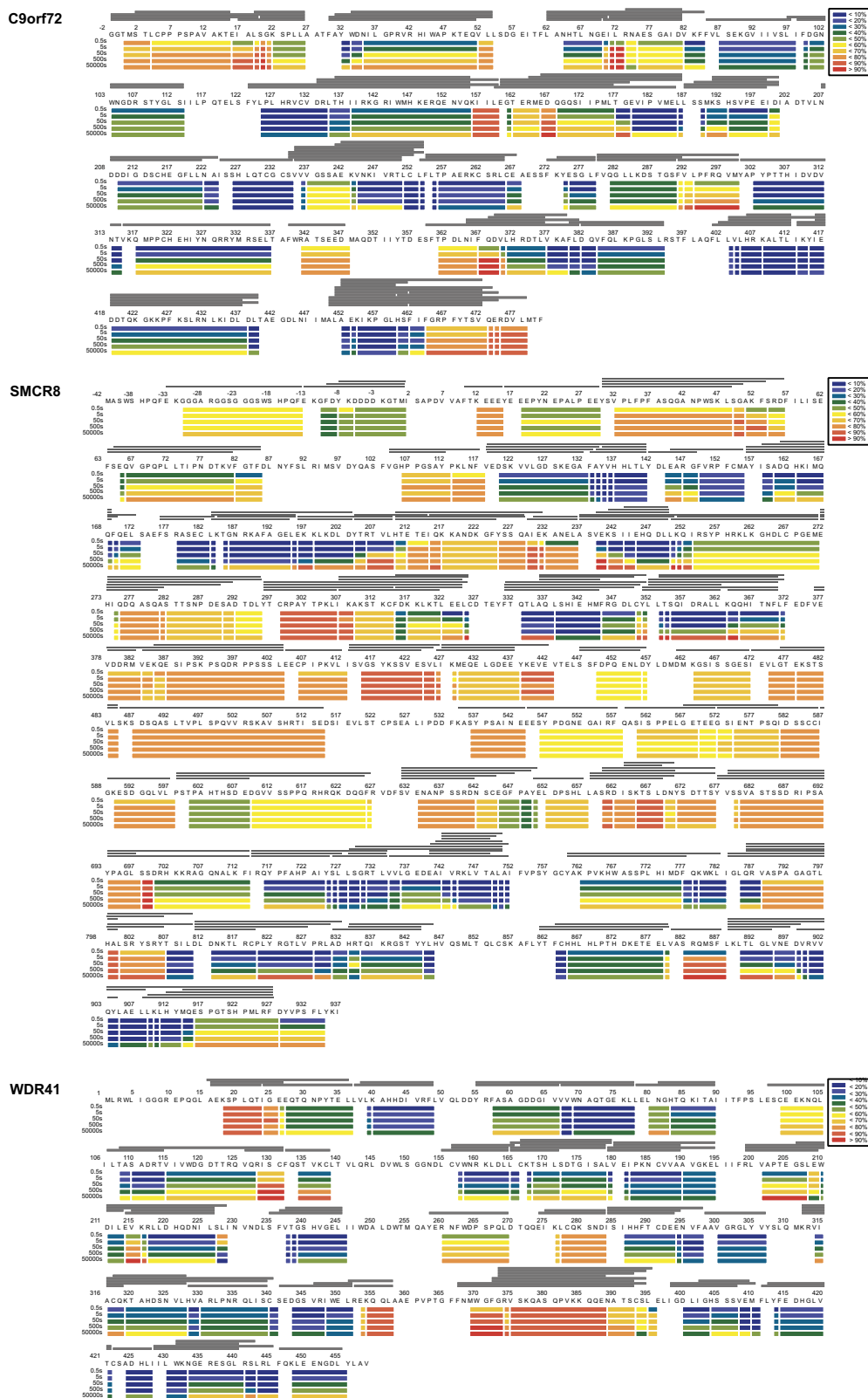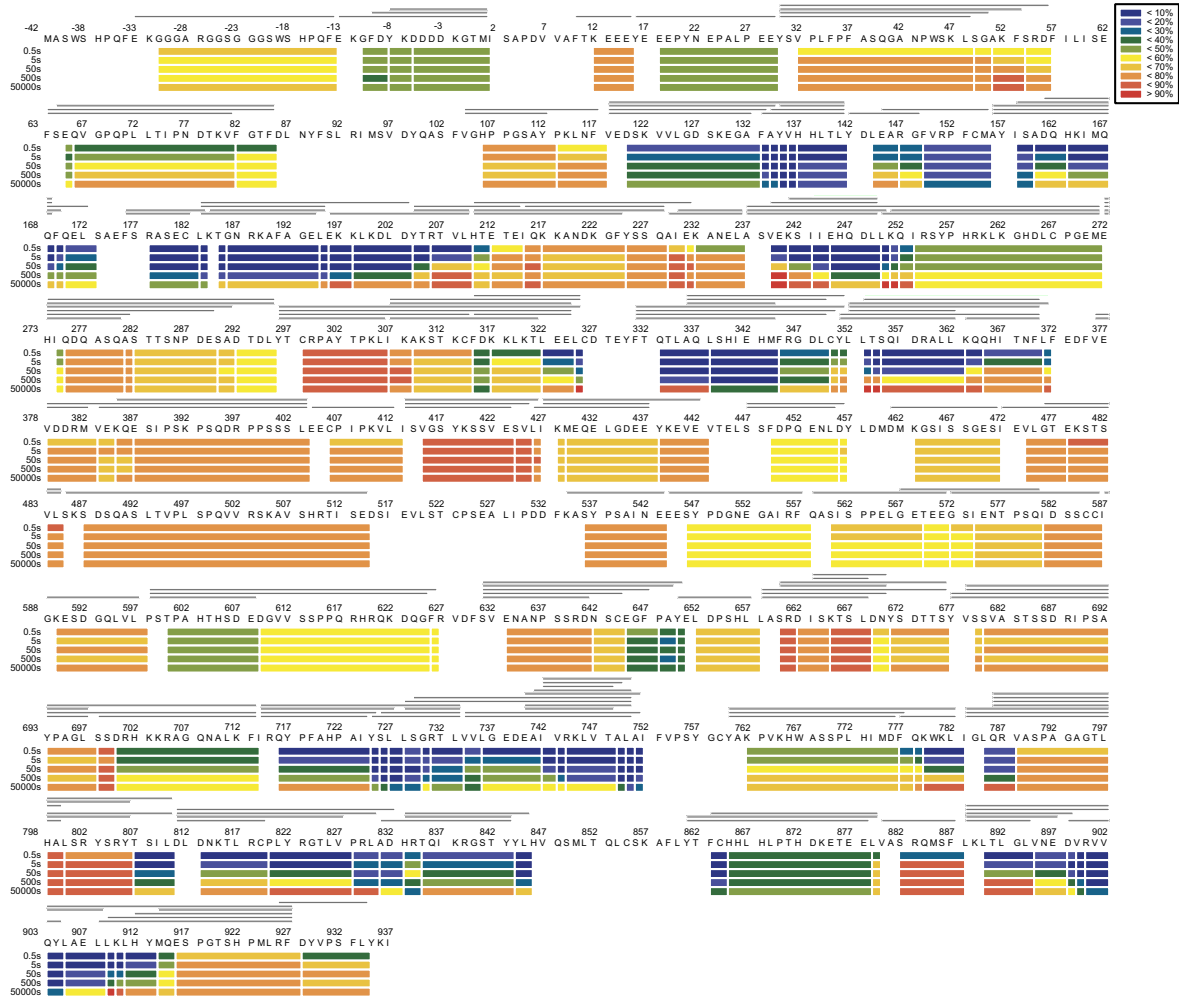
Extended Data Fig. 4 | Deuterium uptake of C9orf72–SMCR8–WDR41.
HDX-MS data are shown in heat map format, in which peptides are represented
by rectangular strips above the protein sequence. Absolute deuterium uptake
after 0.5, 5, 50, 500 and 50,000 s are indicated by a colour gradient below the
protein sequence.
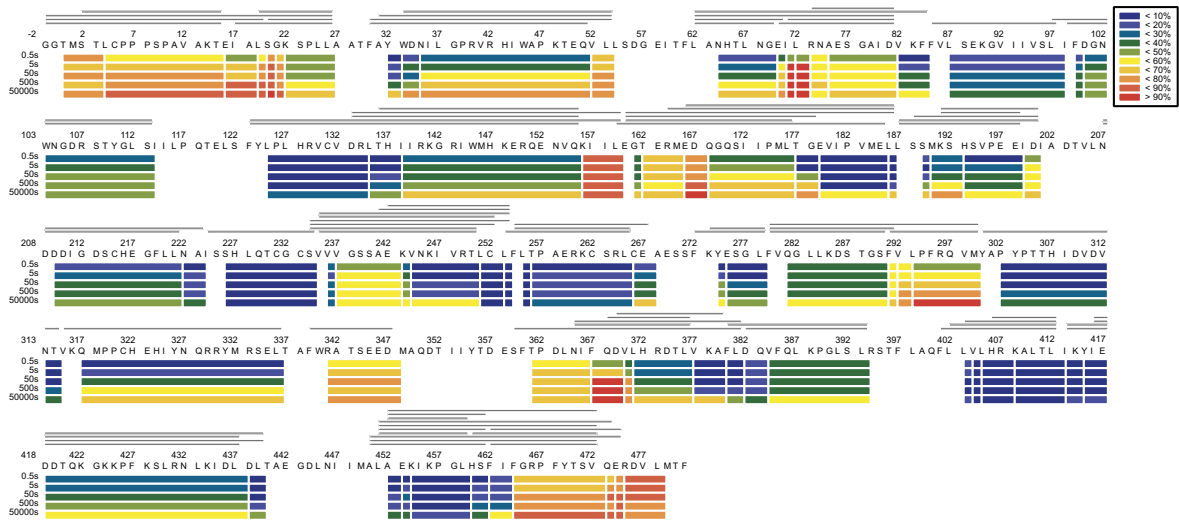
Extended Data Fig. 5 | Deuterium uptake of C9orf72–SMCR8 complex.
HDX-MS data are shown in heat map format, in which peptides are represented
by rectangular strips above the protein sequence. Absolute deuterium uptake
after 0.5, 5, 50, 500 and 50,000 s are indicated by a colour gradient below the
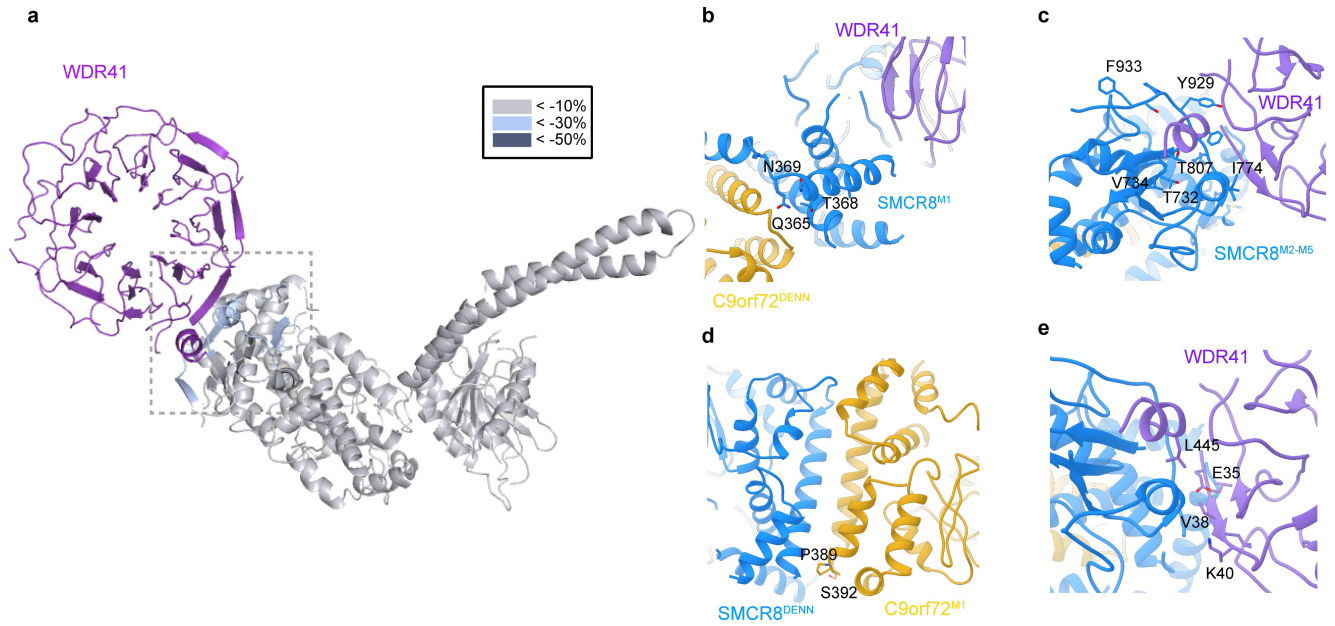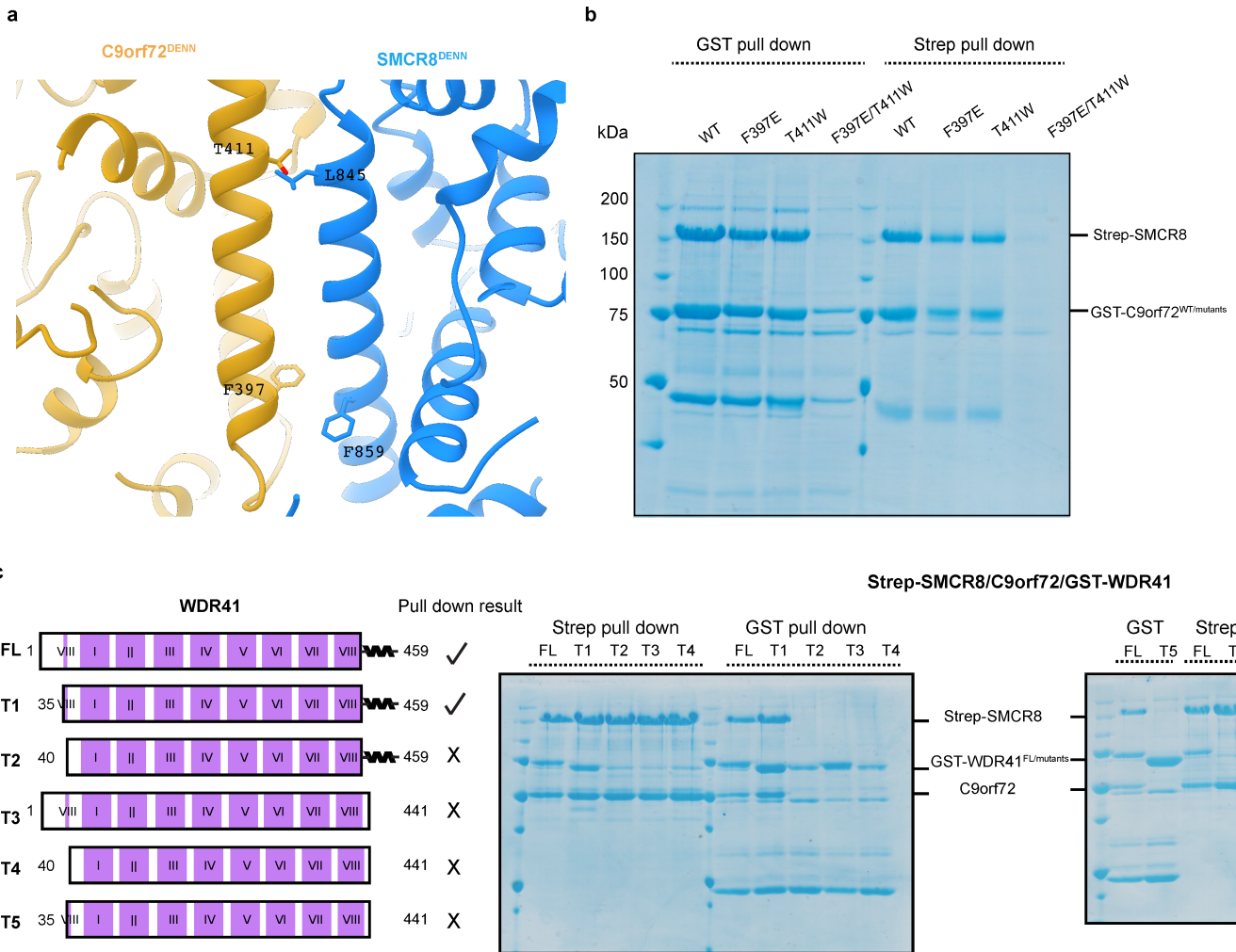protein sequence.

Extended Data Fig. 6 | Mapping the protected region from HDX results onto the SMCR8–C9orf72–WDR41 structure. a, The HDX uptake difference at 0.5 s was mapped on C9orf72–SMCR8. **b**–**d**, Close-up view of M1 (**b**) and M2–M5 (**c**) regions of SMCR8, and the M1 region of C9orf72 (**d**). **e**, Enlarged view of the WDR41 residues in the SMCR8–WDR41 interface.

Extended Data Fig. 7 | Co-expression and pulldown validation of C9orf72–SMCR8 and SMCR8–WDR41 interface. a, Close-up view of the residues that mediate the DENN– DENN dimerization between C9orf72 and SMCR8. **b,** Co-expression and pulldown experiment of Strep-tagged SMCR8 with GST– C9orf72 mutants and WDR41. **c,** Pulldown experiment of GST–WDR41 mutants with C9orf72–SMCR8. The pulldown experiments were carried out at least twice with similar results (**b**, **c**).
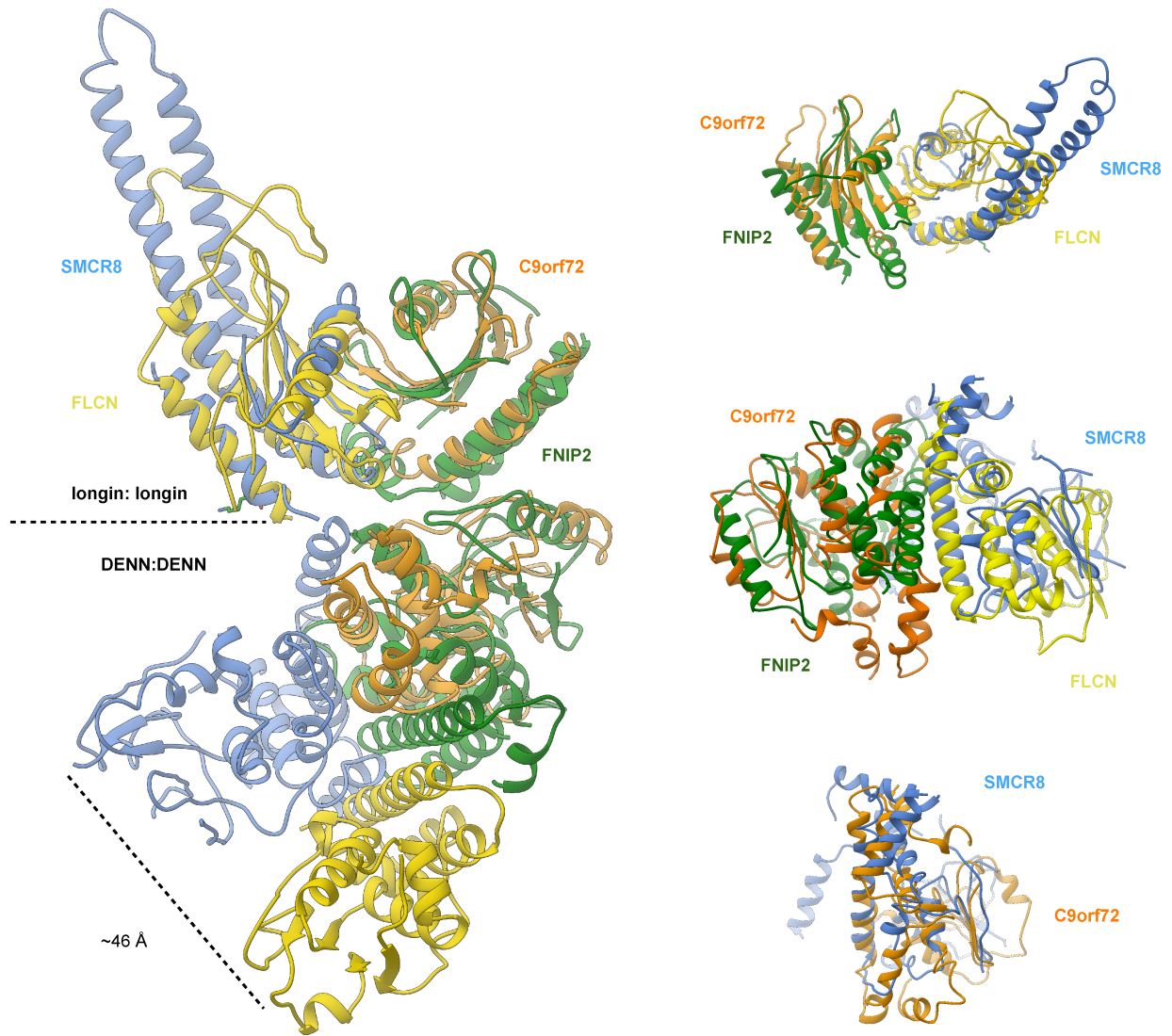
**Extended Data Fig. 8 | Structural comparison between C9orf72–SMCR8 and FNIP2–FLCN.** Left, structural alignment of full-length FNIP2–FLCN and C9orf72–SMCR8. Right, comparison between the longin dimers (top), DENN dimers (middle) and SMCR8 DENN with C9orf72 DENN domain (bottom).
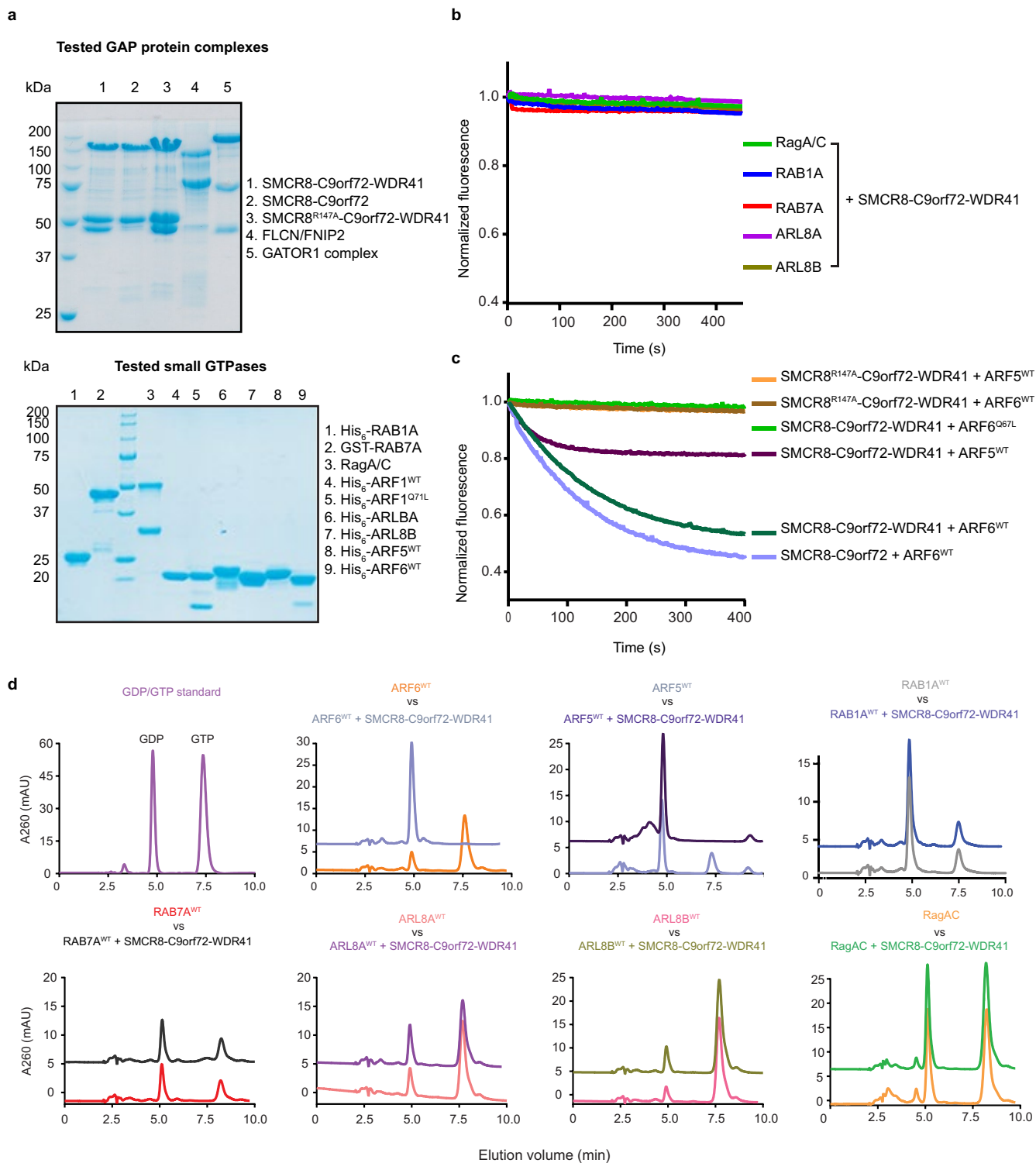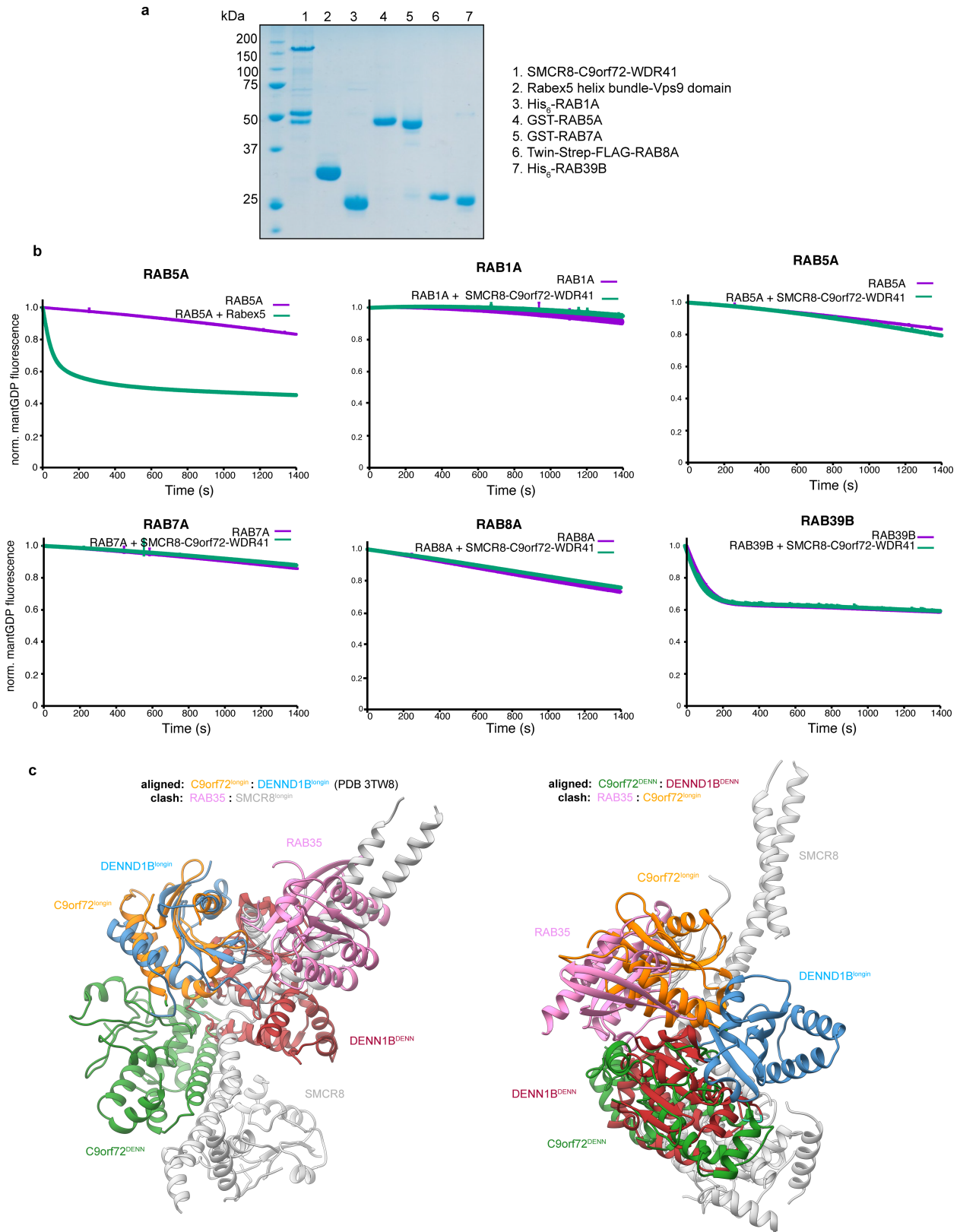
Extended Data Fig. 9 | GTPase assay for different small GTPases with C9orf72–SMCR8–WDR41. a, SDS–PAGE of GAP protein complex (top) and GTPase proteins (bottom) used in the experiments. **b**, Tryptophan fluorescence GTPase signal was measured for purified RAGA–RAGC, ARL8A, ARL8B, RAB1A and RAB7A before and after addition of C9orf72–SMCR8–WDR41. The fluorescence signal upon GAP addition was normalized to 1 for each experiment. The experiments were carried out in triplicate and one representative plot is plotted. **c**, Tryptophan fluorescence GTPase signal was measured for purified ARF6, ARF(Q67L) or ARF5, before and after addition of C9orf72–SMCR8–WDR41, C9orf72–SMCR8(R147A)–WDR41 or C9orf72–SMCR8. **d**, HPLC-based GTPase assay with ARF6, ARF5, RAB1A, RAB7A, ARL8A, ARL8B and RAGA–RAGC proteins in the absence and addition of C9orf72–SMCR8–WDR41 complex, as indicated. The experiments were carried out in triplicate and one representative plot is shown. All experiments were carried out at least three times independently with similar results (**a**–**d**).
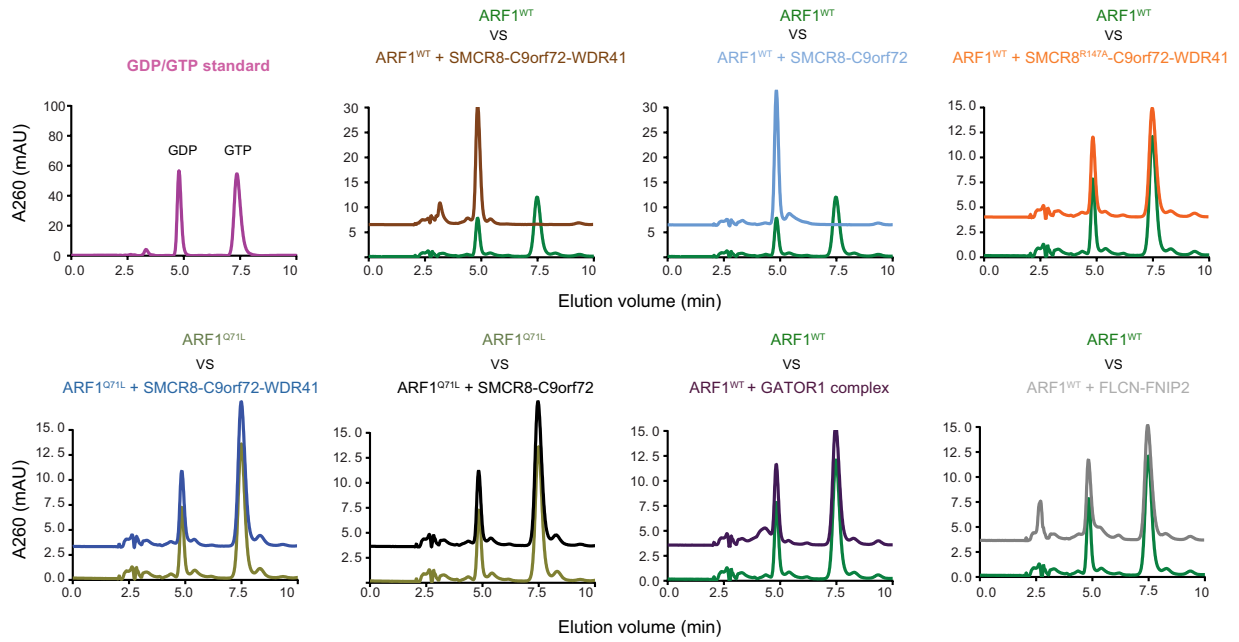
Extended Data Fig. 10 | HPLC-based GTPase assay with ARF1 or ARF1(Q71L) proteins. Assays were performed in the absence and addition of GAP complex, as indicated. The experiments were carried out in triplicate and one representative plot is shown. All experiments were carried out at least three times independently with similar results.

Extended Data Fig. 11 | GEF assay for different small GTPases with C9orf72–SMCR8–WDR41. a, SDS–PAGE of C9orf72–SMCR8–WDR41 complex and GTPase proteins used in the experiments. **b**, GEF assay with mantGDP-reloaded RAB1A, RAB5A, RAB7A, RAB8A and RAB39B proteins in the absence and addition of C9orf72–SMCR8–WDR41 complex, as indicated. RAB5A treated with RABEX5 was used as a positive-control reaction. Data were baseline-subtracted and normalized to the signal immediately after GTP addition, which also is the 0-s time point in the plots. Plots are mean ± s.d. of each technical triplicate experiment. All experiments were carried out at least twice independently with similar results (**a**, **b**). **c**, Structural alignment of C9orf72–SMCR8–WDR41 with DENND1B–RAB35 (PDB 3TW8).

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics

	C9orf72-SMCR8-WDR41 (EMDB-21048) (PDB 6V4U)
Data collection and processing	
Magnification (calibrated)	43,516
Voltage (kV)	300
Electron exposure (e–/Å^2)	59.6
Defocus range (μm)	0.06
Pixel size (Å)	1.149
Symmetry imposed	C1
Initial particle images (no.)	4,810,184
Final particle images (no.)	381,450
Map resolution (Å)	3.80
FSC threshold	0.143
Map resolution range (Å)	3.3-11
Refinement	
Initial model used (PDB code)	-
Model resolution (Å)	4.5
FSC threshold	0.5
Model resolution range (Å)	n.a.
Map sharpening B factor (Å^2)	-50
Model composition	
Non-hydrogen atoms	7,073
Protein residues	1,106
Ligands	0
B factors (Å^2)	
Protein	108.36
Ligand	
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.472
Validation	
MolProbity score	1.60
Clashscore	4.14
Poor rotamers (%)	0
Ramachandran plot	
Favored (%)	93.89
Allowed (%)	6.11
Disallowed (%)	0.00

natureresearch

Corresponding author(s):   James H. Hurley

Last updated by author(s):   2020/05/06

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	SerialEM 3.6.14
Data analysis	For Cryo-EM data processing, the following software were used: Focus, MotionCor2, gctf-v1.06, gautomatch, Relion 3.0-beta, cryoSPARC, cryoSPARCV2, For model building, the following software were used: I-TASSER, Modeller, RaptorX, Phyre2, Coot 0.9 and 0.8.9.1, Phenix , PyMOL v1.7.2.1, UCSF Chimera 1.13 and 1.14, UCSF pyem, Molprobity (as part of Phenix) For HDX analysis, the following software were used: Orbitrap Discovery mass spectrometer, HDExaminer, Proteome Discoverer 2.1 For data display, the following software were used: Prism 6,7,8 and Origin 6.0 Lysosomal enrichment was scored using a home-built Matlab script.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

EM density map has been deposited in the EMDB with accession number EMD-21048. Atomic coordinates for the C9orf72-SMCR8-WDR41 have been deposited in the PDB with accession number 6V4U.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. The cryoEM data was collected in one session with more than 3000 movies. The high particle number included in the final reconstruction (381,450) indicates that based on similar methods and analysis that are widely published the sample size was not limiting.
Data exclusions	Electron microscopy: micrographs were screened manually and bad ones discarded. After CTF estimation, the bad movies which were bad focus or included ice contamination were excluded. HDX-MS: some peptides are with low signal therefore not included in analysis
Replication	HDX experiment was performed technically triplicate or quadruplicate for 0.5, 5, 50, 500 and 50,000 sec timepoints, . GAP and GEF assay were carried out either technical or biological triplicates. All replicates were successful and yielded similar results.
Randomization	Samples were not allocated into groups. Randomization is not relevant to this study.
Blinding	Blinding was not relevant to this study because there was no group allocation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK 293 Gn-Ti and HEK293A cell lines are purchased from UC Berkeley Cell Culture Facility.
Authentication	The cell lines were authenticated using short tandem repeat analysis.
Mycoplasma contamination	Both cell lines were tested negative for mycoplasma contamination using a nuclear stain by fluorescence microscopy.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.