



OPEN

## Function of SYDE C2-RhoGAP family as signaling hubs for neuronal development deduced by computational analysis

Zen Kouchi<sup>1✉</sup> & Masaki Kojima<sup>2</sup>

Recent investigations of neurological developmental disorders have revealed the Rho-family modulators such as Syde and its interactors as the candidate genes. Although the mammalian Syde proteins are reported to possess GTPase-accelerating activity for RhoA-family proteins, diverse species-specific substrate selectivities and binding partners have been described, presumably based on their evolutionary variance in the molecular organization. A comprehensive *in silico* analysis of Syde family proteins was performed to elucidate their molecular functions and neurodevelopmental networks. Predicted structural modeling of the RhoGAP domain may account for the molecular constraints to substrate specificity among Rho-family proteins. Deducing conserved binding motifs can extend the Syde interaction network and highlight diverse but Syde isoform-specific signaling pathways in neuronal homeostasis, differentiation, and synaptic plasticity from novel aspects of post-translational modification and proteolysis.

Numerous adhesion molecules regulate synapse development by recruiting presynaptic and postsynaptic components to the synaptic clefts. Several RhoGTPase-activating proteins (RhoGAPs) have been implicated as critical regulators of these processes. Synapse-defective-1 protein (Syde1) is an essential regulator of the presynaptic developmental organization. Its role was demonstrated by the loss of synaptic recruitment and neuronal path-finding in genetic studies on Syd1, Syde1 ortholog mutants in *C. elegans* and *Drosophila*<sup>1,2</sup>. Scheiffele's group also demonstrated a similar function for Syde1 in vertebrates by identifying abnormal synaptic clustering in knockout mice brains and dissociated neuronal cultures<sup>3</sup>. The N-terminal disordered region of mouse Syde1 is important for the synaptogenic activity and the RhoGAP domain regulates dendritogenesis. The RhoGAP activity of *Drosophila* Syd1 regulates the clustering of Bruchpilot (Brp) as the active zone component<sup>3,4</sup>, suggesting that the domain function and regulation differ among the species. Although mouse and *Drosophila* Syde1 possess RhoGAP activity<sup>3,4</sup>, no GAP activity has been detected in worm Syde1, which binds to GTP-bound MIG-2/Rac<sup>4,5</sup>. It has been suggested that SYDE1 is conserved phylogenetically among worms, *Drosophila*, and mice. Yet, given the molecular organization, the activation of Syde1 by upstream signaling, and the overlap between the functions of Syde2 and Syde1 in mammalian development are unclear<sup>2,3</sup>. Another knockout study of Syde1 in mice showed that it regulates trophoblast migration and invasion during the development of placentas by remodeling the cytoskeleton; it is also involved in the fetal growth regulation during gestation<sup>6</sup>. Regulation of the Syde1 expression by the transcriptional factor glial cell missing 1 (GCM1) during placentation suggests a mammalian-specific developmental signaling cascade that contrasts with the impairment of synaptic vesicle docking commonly observed among their mutant organisms<sup>3,4,7</sup>.

Mouse Syde1 forms a complex with Munc18-1 and liprin- $\alpha$ 2 through the N-terminal disordered domain, the presynaptic organizers necessary for synaptogenesis. The Syde1 binding domain in liprin- $\alpha$ 2 mediates the interaction with CASK<sup>3,8</sup>. Liprin- $\alpha$ 2 organizes type IIa LAR-RPTP complex through selective liprin- $\alpha$  isoforms, implicating that Syde1 is subjected to complex regulation in the presynaptic compartment by the neuronal adhesion machinery<sup>9</sup>. Transcriptomic meta-analysis revealed Syde1 as one of the differentially expressed genes in the cortex of autistic patients<sup>10</sup>. A clinical study indicated that Syde2 is also a causative gene for intellectual disability. Nonsense mutations in Syde2 that theoretically generate a C-terminal-truncated protein devoid of the RhoGAP

<sup>1</sup>Department of Genetics, Institute for Developmental Research, Aichi Developmental Disability Center, 713-8 Kamiya-cho, Kasugai, Aichi 480-0392, Japan. <sup>2</sup>Laboratory of Bioinformatics, School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, 1432-1 Horinouchi, Hachioji 192-0392, Japan. ✉email: zkouchi@inst-hsc.jp

domain might cause its degradation, and dysregulation of axonal guidance<sup>11</sup>; the neuronal function of mammalian Syde2 is unknown. Pleckstrin homology domain-containing family G member 3 (PLEKHG3) possesses a guanidine nucleotide exchange activity and was identified as a Syde2 complex by high-throughput immunofluorescence screening<sup>12</sup>. Furthermore, PLEKHG3, a brain-enriched RhoGEF, is a potential candidate genes for autism<sup>13</sup>.

Syde-family proteins are characterized by the presence of RhoGAP and C2 motif in domain organization. The regions outside the domains have diversity in their sequences, defined mainly by disordered domains that are poorly characterized<sup>14</sup>. The N-terminal region of Syde1 mediates autoinhibition of the RhoGAP activity. However, the mechanism of integration of the Syde network in the cell-surface molecular architecture remains obscure<sup>3,12</sup>. Promiscuity of substrate specificity and the complex regulation of RhoGAP family proteins have been implicated in several studies using cell-based expression systems<sup>6,12,15</sup>.

The structure–function relationship of the SYDE-family proteins was elucidated by in silico modeling of structural domain by combining sequence and phylogeny. Additionally, the regulatory regions of the Syde proteins were identified based on the conserved candidate short linear motifs (SLiMs) located in the disordered domains of soluble proteins. The SLiMs provide regulatory flexible interface prerequisites for the dynamic assembly of protein complexes. Interestingly, predicted interactors highlight the mammalian Syde network involved in neuronal homeostasis, with post-translational modification and RhoGAP selectivity as synaptic scaffold molecules. The hypothesis of Syde function as a neuronal signaling hub will be worth investigating experimentally and clinically in the future.

## Methods

**Sequence feature analysis.** Syde1 and Syde2 (UniProt accession No. Q6ZW31 and Q5VT97, respectively) were downloaded from UniProt and NCBI Gene database. Each protein was aligned using the MAFFT multiple sequence alignment software<sup>16</sup> and visualized with JalView<sup>17</sup>. Domain and sequence features were predicted by using InterproScan<sup>18</sup> and secondary structure was predicted with PSIPRED<sup>19</sup>. Regions outside the predicted domains were assessed for disordered domain including short linear motifs (SLiMs) by MobiDB<sup>20</sup>, ELM<sup>21</sup>, and D<sub>2</sub>P<sub>2</sub><sup>22</sup> (Supplementary Table S1). SLiMs play critical roles in many biological processes and we predicted SLiMs localized within the disordered domains that are commonly mapped among orthologs. Since SLiMs have propensity of random occurrences<sup>23</sup>, binding motifs are also predicted from ANCHOR program<sup>24</sup> analyzing the sequence of disordered region and energy for molecular interaction. Phosphorylation sites and domain organization were also confirmed by Phosphosite in D2P2 and ScanProsite<sup>25</sup>. S-palmitoylation site was searched from SwissPalm and manually curated from PUBMED database<sup>26,27</sup>. Presence of nuclear export signal was examined from sequence and structural aspects by NESdb database<sup>28</sup>.

**Phylogenetic analysis.** Sequences of Syde orthologs were downloaded from UniProt database to reconstruct the phylogeny of the protein family. Seventeen Syde1 and vertebrate Syde2 sequences were retrieved and multiple alignment was conducted with MAFFT using Jalview. Phylogenetic analysis was performed with MEGA11 by Maximum Likelihood based on the JTT model + G (Gamma distributed sites) with 500 bootstrap replicates<sup>29</sup>.

**Homology modeling.** Sequences of RhoGAP and C2 domain of each SYDE proteins were submitted to the homology detection method HHpred and multiple alignment-based detection was conducted with Swiss-model search, taking into account target-template secondary structure similarity<sup>30,31</sup>. Models for each domain was built and their model quality was assessed and estimated with transform-restrained Rosetta<sup>32</sup>, I-Tasser<sup>33</sup>, and QMEAN<sup>34</sup>, respectively and all the graphic images were generated by UCSF Chimera software. Evaluation of predicted model was assessed by Procheck and ERRAT<sup>35</sup> (Supplementary Table S2).

**Molecular docking.** Modelled Syde1 and Syde2-RhoA docking analysis were performed by SwarmDock docking program, one of the best flexible docking performing programs<sup>36</sup>. The ability of the programs for reproducing the RhoGTPase-RhoGAP interaction was checked by 5c2k and 5c2j for face-to-face. Optimization of inter-atom energies and side-chains of modeled Syde structures were performed by EGAD program<sup>37</sup> and adjusted PDB files were subjected to the SwarmDock web server in full blind mode.

**Known Syde interaction analysis and Syde interactors prediction.** BioGrid<sup>38</sup> and IntAct<sup>39</sup> databases were used for compiling a list of experimentally analyzed Syde interactors and manually curated from the literatures (Table 1, Supplementary Table S1). The binding proteins were annotated with the domain architecture and biological processes involved in physiological aspects by retrieving from the UniProt, InterPro<sup>18</sup>, KEGG<sup>40</sup> database. Details in the interaction were inferred from PubMed with selected keywords for neuronal function.

For prediction of Syde interactors we searched for putative domains and linear motifs putatively mediating the interactions detected in commonly cell-based model or in neurodevelopmental processes. We assume that a predicted Syde linear motif may be potential interaction site for a novel interactor if the known binding protein is categorized to a class of protein or contains the motif which is already known to bind the corresponding sequences in the Syde proteins (Table 1).

## Results

Syde proteins have emerged as critical Rho-family regulators in neuronal and embryonic development. However, their physiological roles and molecular mechanisms in vertebrates are not well known. Therefore, a computational analysis of the SYDE-family proteins was performed by comparing the primary structures of *Drosophila* and vertebrate orthologs to explore their regulatory and catalytic functions. Homology search using InterproScan identified the C2 and RhoGAP domains that comprised 118 and 207 residues in SYDE1, and 120 and 216 residues in SYDE2, respectively. These predicted regions are highly conserved between vertebrate Syde1 and Syde2, although each N- or C-terminal sequences is quite diverse (Figs. 1, 2, 3). In addition to this unique evolutionary variance observed in the disordered Syde domains, crystal structure is not available for any Syde-family protein (Fig. 3a). Disordered regions are a particular stretch of amino acid patterns or conserved regions with predicted

Protein	Predicted interactor	Domain architecture	Signaling pathway	Method	Score (ELM)	Predicted SYDE1 or SYDE2 interacting region
SYDE1	ATG8 (GABARAP)	ATG8, LIR domain	Autophagic pathway	ELM: LIG_LIR_Gen1 <sup>55</sup>	3.60e-03	Y_LIR sequence 9TF SRL <sup>13</sup>
SYDE1	Ruk/CIN85	SH3 domain	Synaptic endocytosis	ELM:LIG_SH3_3 <sup>68</sup>	1.32e-02	<sup>207</sup> DSSVGGP <sup>213</sup> <sup>233</sup> GDSPERP <sup>239</sup> <sup>641</sup> PEVVTRP <sup>647</sup>
SYDE1	YWHAZ	14-3-3	Neuronal differentiation	ELM: LIG_14-3-3 14-3-3 binding motif <sup>63</sup>	4.48e-03	Arg containing phospho- motif <sup>177</sup> RRRLSLR <sup>183</sup>
SYDE1	Cyclin B/CDK1	cyclin	Neuronal differentiation	ELM: DOC_Cyclin_RxL S/T-X-(X)-R/K motif <sup>54</sup> (peptide library)	4.21e-03	RxL docking motif <sup>166</sup> RLSIKMKKLP <sup>175</sup>
SYDE1	P38α, p38β	kinase	Neuronal homeostasis	ELM: DOC_MAPK Classical D motif specificity <sup>52</sup>	4.32e-03	<sup>617</sup> RPKRQPPLHL <sup>626</sup>
SYDE1	Casein kinase Iγ	kinase	Endocytosis synaptic vesicles	ELM: MOD_CK1_1, D <sub>2</sub> P <sub>2</sub> Palmitoylation-dependent phosphorylation <sup>49</sup>	1.70e-02	<sup>645</sup> SPPSNRY <sup>651</sup>
SYDE1	Casein kinase II	kinase	Dendritic spine formation	ELM: MOD_CK2_1, D <sub>2</sub> P <sub>2</sub> , PROSITE <sup>25</sup>	1.46e-02	pS/pT-D/E-X-D/E motif <sup>678</sup> VTGSDSEDE <sup>686</sup> (SYDE1)
SYDE1	RhoB	RhoGTPase domain	Synaptic plasticity	Docking by Swarmdock <sup>36</sup>	–	RhoGAP
SYDE2	Pin1	WW domain	Regulation of NMDA receptor	ELM: LIG_WW3 Pin1 consensus motif <sup>61,62</sup>	5.67e-04	S/T-P motif <sup>60</sup> SPPRS <sup>64</sup>
SYDE2	Smek1	Ser/Thr phosphatase4 regulatory subunit (EVH domain)	Neuronal differentiation	ELM: DOC_PP4 PP4 binding motif <sup>60</sup>	1.37e-03	FxxP motif, SLiM region <sup>158</sup> FRDP <sup>161</sup>
SYDE2	Grb2	kinase	Cerebral cortical develop- ment	ELM: LIG_SH2_Grb2 Consensus pYxN motif by pull-down experiment <sup>50</sup>	3.18e-04	SH2 domain binding motif <sup>551</sup> YINS <sup>554</sup>
SYDE2	Fbw7	SCF ubiquitin ligase com- plex, WD-40 motif	Neuronal differentiation	ELM: DEG_SCF_FBW7 Alignment of degron motifs <sup>63</sup> ϕXϕϕϕTPPxS (ϕ: hydro- phobic)	7.14e-04	TPxxS phospho-dependent degron <sup>552</sup> INSPDNTPSLS <sup>562</sup>
SYDE2	Calcineurin	Ser/Thr phosphatase	Ca <sup>2+</sup> -dependent synaptic plasticity	ELM: DOC_PP2B_LxvP1 <sup>41</sup> πϕLxVP, SLiM region	2.30e-03	<sup>179</sup> SFLRPP <sup>184</sup>
SYDE2	Crk, Src	Kinase		ELM:LIG_SH3_3 Class I: RxxPxP <sup>56,57</sup>	1.32e -02	<sup>56</sup> RQVSP <sup>62</sup>
SYDE1,2	Crk	Kinase		ELM:LIG_SH2_Crk/Nck, putative phosphorylation and conserved binding motif <sup>59</sup>	1.52e-03	SH2 domain binding motif <sup>112</sup> YNPIP <sup>116</sup> (SYDE1) <sup>372</sup> YNPIP <sup>376</sup> (SYDE2)
SYDE1,2	GSK3	Kinase	Regulation of synaptic clustering	ELM: MOD_GSK3_1, D <sub>2</sub> P <sub>2</sub> , conserved putative phosphorylation motif with DmRhoGAP100F <sup>58</sup>	2.68e-02	<sup>228</sup> GYLSDGDS <sup>235</sup> (SYDE1) <sup>619</sup> GYLSDGDS <sup>626</sup> (SYDE2)
SYDE1,2	RhoA	RhoGTPase domain	Regulation of dendritogen- esis	Docking by Swarmdock <sup>36</sup>	–	RhoGAP
Protein	Interactor	Domain architecture	Signaling pathway	Experimental condition	Ref	Interacting region
SYDE1	Munc18	Sec1-like domain	Presynaptic differentiation	coIP	<sup>3</sup>	Disordered domain
SYDE1	Liprina2	Coiled-coil, SAM domain	Presynaptic differentiation	coIP	<sup>3</sup>	Disordered domain
SYDE1	Ruk/CIN85	SH3 domain	Control of Rho-family GTPases	Pull-down, LC-MS/MS	<sup>69</sup>	–
SYDE1	XPO1	ImportinβN domain, Nuclear export (exportin1, XPO1 domain)	Nuclear export to the cytoplasm	High through-put AP-MS, putative nuclear export signals*	<sup>80</sup>	L <sup>536</sup> RLVSS <sup>541</sup> * L <sup>713</sup> KDFDALILDLERELS <sup>728</sup> *
SYDE1	ARHGAP28	RhoGAP domain	RhoGAP activity, Actin filament	coIP, High through-put AP-MS	<sup>12</sup>	–
SYDE2	PLEKHG3	DH domain, PH domain	RhoGEF activity	High through-put AP-MS	<sup>12</sup>	–
SYDE2	calcineurin	Serine/Threonine phos- phatase	Activity-dependent dendri- togenesis	Pull-down, ITC	<sup>41</sup>	SYDE2 232–289 <sup>239</sup> RVLSVP <sup>240</sup>
DmSyd1	Mtl	Small GTPase Rho	Rho-GTPase activity	Pull-down	<sup>4</sup>	RhoGAP
DmSyd1	Rac1, Rac2	Small GTPase Rho	Rho-GTPase activity	Pull-down	<sup>4</sup>	RhoGAP
DmSyd1	RhoA, RhoL	Small GTPase Rho	Rho-GTPase activity	Pull-down	<sup>4</sup>	RhoGAP
DmSyd1	Cdc42	Small GTPase Rho	Rho-GTPase activity	Pull-down	<sup>4</sup>	RhoGAP
DmSyd1	Neurexin-1	lamininG domain, Synde- can/Neurexin domain	Neurexin-1 recruitment to active zones Postsynaptic GluR2 localiza- tion	Y2H, Co-IP, FLAP	<sup>7,81</sup>	PDZ

Continued

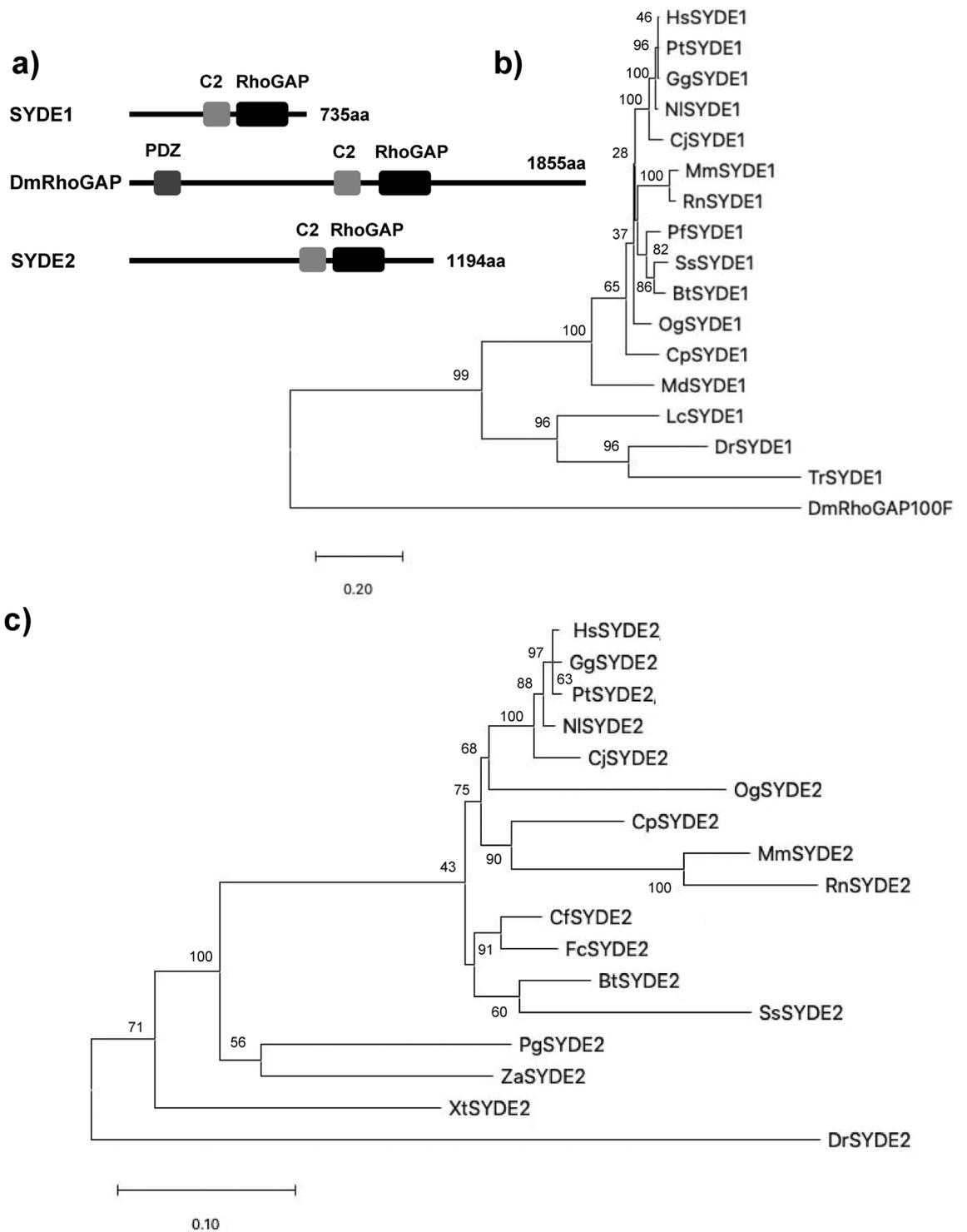
Protein	Interactor	Domain architecture	Signaling pathway	Experimental condition	Ref	Interacting region
DmSyd1	Bruchpilot (ELKS-related protein)	ELKS/CAST domain, coiled-coil	Presynaptic active zone organization Postsynaptic GluR2 localization	Y2H, coIP-MS/MS	82	PDZ, C-terminal (1-320aa, 1301-1844aa)

**Table 1.** SYDE interactors. (Upper panel) The putative domains in linear motifs predicted to mediate SYDE interaction were selected. It is assumed that if the known binding protein is the same class of protein-family or presents the domain or motif known to interact with a predicted SYDE linear motif the molecule is considered to be the putative interactor binding site (SLiM: short linear motif<sup>23</sup>). The consensus motifs were examined by ELM database and the selected sequence patterns localized in SYDE1 and SYDE2 are indicated according to several literatures listed in the references. (Lower panel) The SYDE1, SYDE2, and *Drosophila* Syd1 (DmSyd1) interactors are indicated by categorizing from the literatures as follows: domain architecture obtained from InterPro database, the experimental condition, and interacting region which is experimentally validated (–: not determined). AP-MS affinity purification-mass spectrometry, FLAP fluorescence recovery after photobleaching, Y2H yeast two-hybrid system: coIP: immunoprecipitation. \*Predicted nuclear export signal is indicated by NES database<sup>28</sup>.

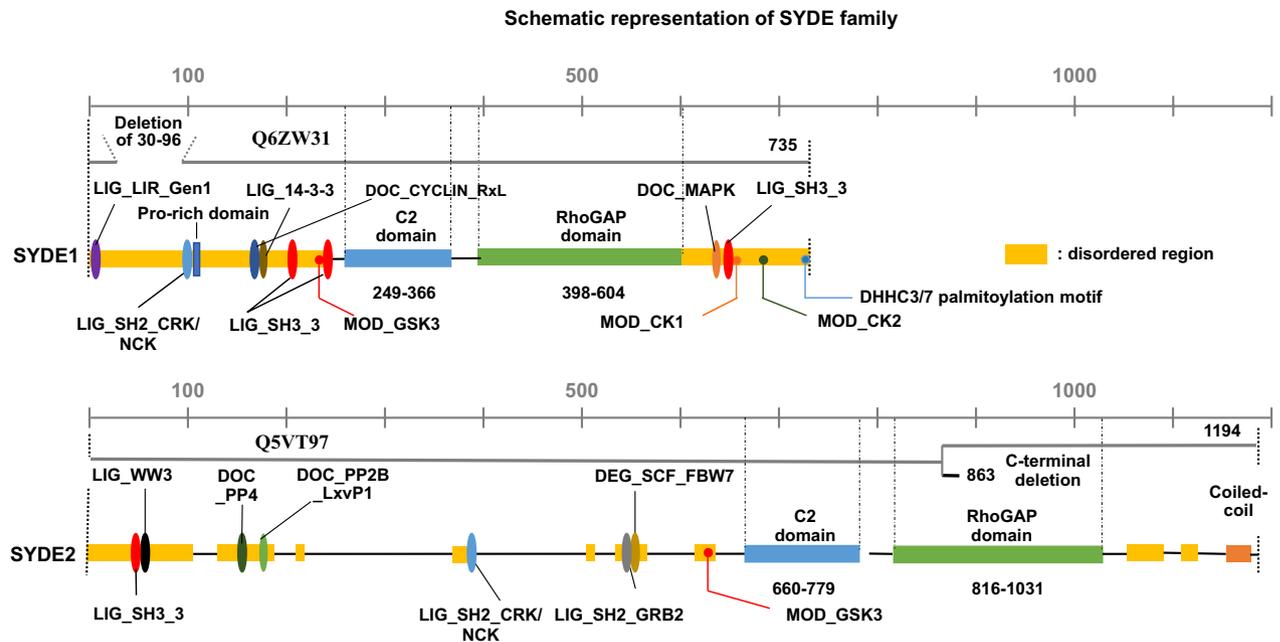
linear motifs that mediate modification by phosphorylation or provide an interface to specific binding partners for signaling (Fig. 2). Syde1- and 2-binding proteins were examined by BioGrid and IntAct and listed by curating additional relevant literature (Table 1, lower table). These findings were used to expand for assembling the Syde interaction network (Table 1). The SYDE interactors were selected as the following procedures. Putative interaction motif scoring e-value lower than  $1.0 \times 10^{-2}$  was selected based on its location within the disordered domain mapped by ELM, D<sub>2</sub>P<sub>2</sub> and Anchor, and phosphorylation site was selected based on the disordered propensity and predicted site assigned by Phosphosite. Then, the motif conservation through mammalian SYDE1/2 orthologs was also taken account to select the functional interaction sites. As for evolutionary aspects, all the listed motifs were conserved and could be functional in MdSYDE1 (*M. domestica*), whilst casein kinase II phosphorylation site (MOD\_CK2) and Grb2 binding motif (LIG\_SH2\_Grb2) lacked in LcSYDE1 and FcSYDE2, respectively as indicated in ELM search (Fig. 3b,c). In contrast, ZaSYDE2 was devoid of most of the interaction and phosphorylation motifs (Figs. 1c and 3c). The putative functional units are described by predicted domains or elements separately as follows.

**N-terminal regulatory regions.** The intrinsically disordered region resides in the N-terminus, especially with a more extended form in Syde1 than in Syde2. This portion of mammalian Syde2 proteins shows a relatively low identity between its paralogs, suggesting functionally divergent roles (Fig. 1). The lower organisms contain a PDZ domain in the middle region with the N-terminal side of C2 domain. However, the corresponding motif does not reside in the vertebrate Syde proteins (Fig. 1). An alignment of the N-terminus of the vertebrate Syde1 highlights the highly conserved regulatory motifs including the SH2 binding region. But vertebrate Syde2 contains evolutionarily conserved motifs with variable regional sequences outside of the motif. The ATG8 family interacting LIR motif and Crk docking motifs are located at the N-terminus of Syde1. The 14-3-3zeta and cyclin B/Cdk1-binding site are located within the C-terminal side of the disordered domain (Fig. 3b), and the GSK3 phosphorylation site resides within the linear motif. The Syde1 isoform2 (Q6ZW31-2) lacks 67 residues in the disordered domain but preserves all of the predicted motifs. In contrast, the N-terminal portion of Syde2 has calcineurin docking LxVP motifs and the motif in the marginal region of the disordered domain was experimentally proven to interact with calcineurin<sup>41</sup> (Figs. 2, 3c). Pin1 specifically binds the WW recognition motif represented by phosphorylated serine or threonine residues preceding a proline, and SYDE2 proteins contain a group III/IV motif and SH3 binding region at the N-terminus. The WW domain for Pin1 recognition and Smek1 binding site for serine/threonine protein phosphatase 4 (PP4)-mediated regulation are conserved among mammalian Syde2 proteins but not in birds, reptiles, amphibians and fish (Fig. 1, *Z. albicollis*, *P. guttatus*, *X. tropicalis*, and *D. rerio*). Clusters of short-length disordered domains are located in the middle of Syde2 (Figs. 2, 3). Crk and Grb2 binding motifs reside separately in the disordered regions where highly conserved sequences among vertebrates and C-terminal variable regions are sequentially organized.

**The RhoGAP domains of SYDE1/2 and *Drosophila* Syd1.** The RhoGAP domain of human SYDE1 and SYDE2 proteins show 57.1% sequence homology with each other. These two RhoGAP domains have similar levels of lower sequence homology to the corresponding domain of *Drosophila* Syd1 (Q9V7SV) (SYDE1: 46.3%; SYDE2: 45.0%) (Fig. 4a). The RhoGAP domains of both Syde1 and Syde2 were modeled separately based on tertiary template selection. The HHPred search and potential template candidates by trRosetta and I-Tasser selected human MgcRacGAP protein (PDB code: 5c2k) for both Syde RhoGAP domains, and ArhGAP2 (N-chimerin PDB code: 3cxl) was selected as the best model for *Drosophila* Syd1 (Fig. 4b, Supplementary Fig. S1, Supplementary Table S2). Evaluation of predicted template-based model was performed by Procheck and ERRAT<sup>35</sup> (Table 2). Despite the insertion between Syde and template sequences, the secondary structural elements of the RhoGAP domains in both Syde proteins were superimposed well to MgcRacGAP, which is organized chiefly with helix and turn structures. MgcRacGAP functions as a GAP activity toward Cdc42 and RhoA and the structures of the docking complexes with MgcRacGAP were experimentally resolved<sup>42</sup>. The structures of the catalytic core were well structurally superimposed. Both RhoGAP-Rho interface showed well-matched structures by forming a conserved arginine finger in the GAP domain of Syde proteins bridging the  $\beta$  and  $\gamma$  phosphate groups



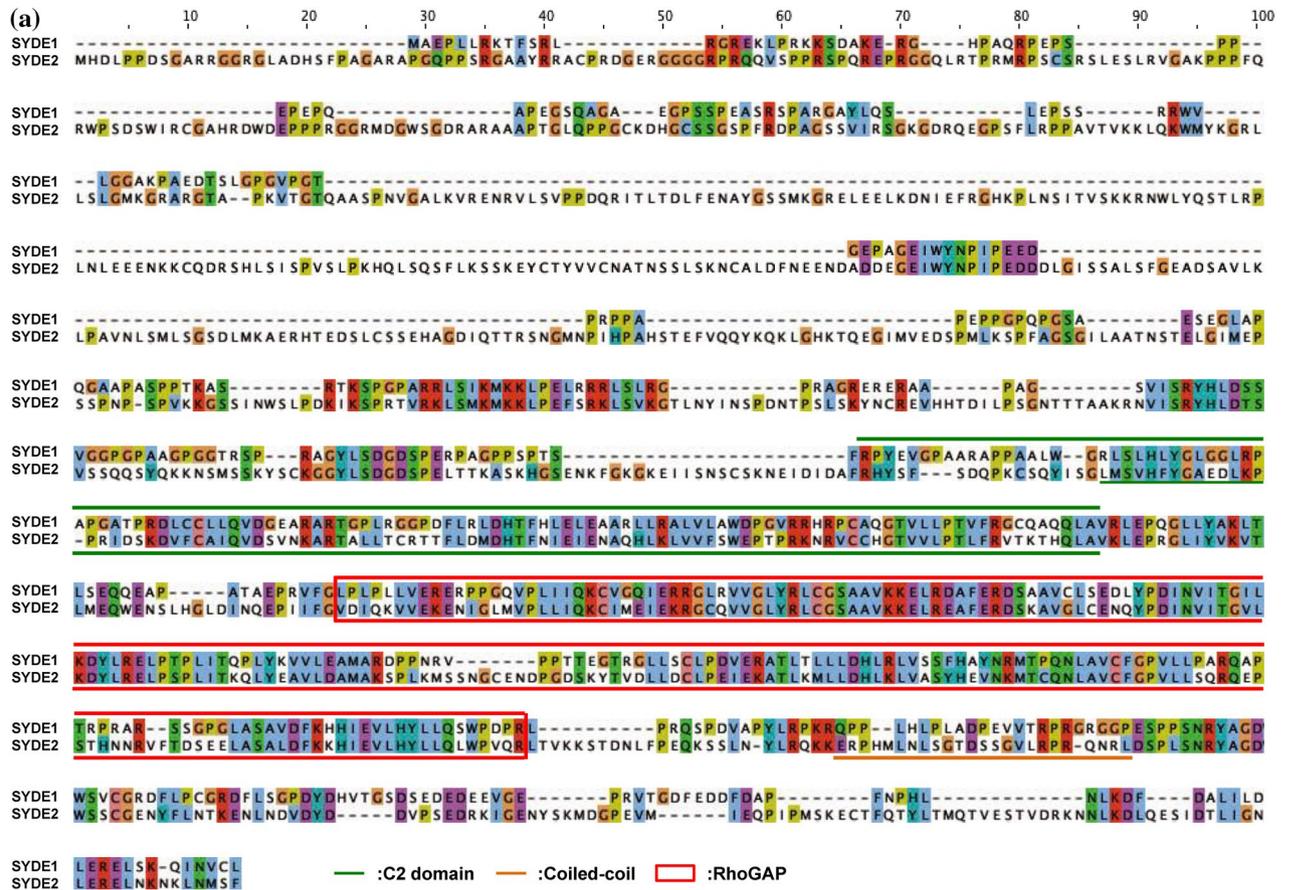
**Figure 1.** Phylogenetic tree construction of SYDE1 and SYDE2. **(a)** Domain organization of Syde proteins. **(b)** The tree is constructed among vertebrate SYDE1 orthologs. Each cluster is organized as the following Syde1 genes: Bt: *B. taurus* (F1MXU4); Cj: *C. jacchus* (U3BZL8); Cp: *C. porcellus* (H0VZ67); Dr: *D. rerio* (E7FA87); Gg: *G. gorilla* (G3R4S4); Hs: *H. sapiens* (Q6ZW31); Lc: *L. chalumnae* (XP006002980); Md: *M. domestica* (F7F6S9); Mm: *M. musculus* (Q9DBZ9); Nl: *N. leucogenys* (XP030676439); Og: *O. garnettii* (H0XAB4); Pf: *M. putorius furo* (M3YXJ5); Pt: *P. troglodytes* (H2QFL6); Rn: *R. norvegicus* (D3ZZN9); Ss: *S. scrofa* (F1SAN7); Tr: *T. rubripes* (XP02968941); DmRhoGAP100F: *D. melanogaster* RhoGAP isoformC. **(c)** For SYDE2 orthologs, each cluster is organized as the following Syde2 genes: Bt: *B. taurus* (NP001179527); Cf: *C. familiaris* (XP025273572); Cj: *C. jacchus* (F6R312); Cp: *C. porcellus* (H0UWH0); Dr: *D. rerio* (A0A0G2KK80); Fc: *F. catus* (M3XAW0); Gg: *G. gorilla* (G3QK10); Hs: *H. sapiens* (Q5VT97); Mm: *M. musculus* (E9PUP1); Nl: *N. leucogenys* (G1RG60); Og: *O. garnettii* (H0WYC5); Pg: *P. guttatus* (A0A6P9DB82); Pt: *P. troglodytes* (H2PZB3); Rn: *R. norvegicus* (NP001305229); Ss: *S. scrofa* (A0A4X1UF43); Xt: *X. tropicalis* (A0A618RMF6); Za: *Z. albicollis* (XP005482373).



**Figure 2.** Domain architecture of SYDE proteins based on their sequence analysis. Both human SYDE proteins are characterized by disorder domain (yellow), C2 (light blue) and RhoGAP domain (green) that are indicated as sequence boundaries of human SYDEs (SYDE1: Q6ZW31; SYDE2: Q5VT97). Conserved linear motifs among vertebrates and mammals for SYDE1 and SYDE2, respectively, are represented as the following according to Dinkel et al. (2016): Canonical LIR motif required for ATG8-mediated autophagy (LIG\_LIR\_Gen1) (purple oval); Crk and Nck SH2 domain binding motif (LIG\_SH2\_CRK/NCK) (light blue oval); 14-3-3 interaction motif (LIG\_14-3-3) (brown oval); Cyclin/CDK binding motif (DOC\_Cyclin\_RxL) (dark blue oval); PPR-specific WW domain (LIG\_WW3) (black oval); a docking motif mediating interaction with Erk1/2 and p38 subfamilies of MAP kinases (DOC\_MAPK) (orange oval); calcineurin docking motif (DOC\_PP2B\_LxvP1) (light green oval); FxxP-like docking motif recognized by PP4 holoenzyme (DOC\_PP4) (dark green oval); Phospho-dependent degenon recognized by FBW7 Fbox proteins (DEG\_SCF\_FBW7) (dark yellow oval); Casein kinase (CK1) phosphorylation site (MOD\_CK1) (orange dot); CK2 phosphorylation site (MOD\_CK2) (dark green dot); DHHHC3/7 palmitoylation site (green dot); GSK3 phosphorylation site (MOD\_GSK3) (red dot); SH3 binding site (LIG\_SH3\_3) (red oval); Grb2-like SH2 domains binding motif (LIG\_SH2\_GRB2) (gray oval). Alternative SYDE isoforms are indicated and the SYDE1 isoform Q6ZW31-2 (668 residues) is missing the N-terminal region (30–96). SYDE2 isoform Q5VT97-2 (863 residues) is truncated at residue 849 with substitution in the region from position 850 to 863 (LSYYGSLLLPLIID).

of GTP required for its hydrolysis and the cysteine residues located at the substrate interfaces of both proteins are substituted by Ser387 in MgcRacGAP<sup>43</sup> (Fig. 4b). Phosphorylation of serine 387 in MgcRacGAP is known to determine the substrate specificity towards RhoA rather than Rac1 and the replacement of this residue by aspartic acid reduces the RacGAP activity<sup>44</sup>. The 3D model quality evaluation of Syde1 and Syde2 revealed closely homologous structures, with QMEAN scores of 0.66 and 0.64, respectively, and *Drosophila* Syd1 is remotely homologous to the ArhGAP2 RhoGAP domain with QMEAN score of 0.74. Lower quality regions are located in both Syde1 and Syde2 as insertion between the sixth and seventh  $\alpha$ -helices due to partially disordered loops; otherwise highly positional conservation is shown in the Syde proteins with MgcRacGAP (Supplementary Fig. S1). Both proteins showed similar  $\alpha$ -helical distribution patterns, which could mediate the similar recognition patterns of Rho-family substrates. The SYDE2 isoform (Q5VT97) lacks residues 864–1194, missing RhoGAP and the coiled-coil domains (Fig. 2).

Highly conserved residues among RhoGAP family proteins (Arg436, Lys476, Arg480, Met548, Asn552, and Pro559 in Syde1 numbering) are conserved with MgcRacGAP<sup>15,42</sup>. Among the vertebrate Syde1 and Syde2 orthologs, these residues in the Rho-family proteins are highly conserved, although some exceptional cases exist (Syde1 in *D. rerio* and Syde2 in *O. garnettii*, E7FA87 and H0WYC5, respectively). Structural analysis with sequence comparison revealed that both Syde1 and Syde2 with MgcRacGAP displayed completely matched orientation of the  $\alpha$ -helices that contained the putative conserved residues interacting with Rho family protein (Fig. 4b, Supplementary Fig. S1). Cdc42-bound MgcRacGAP has different regional structures around the C-terminal  $\alpha$ -helices 2 and 3 sequentially flanking the catalytic domain compared to its RhoA-bound model (PDB code: 5c2k and 5c2j). The conserved residues around the arginine finger and  $\alpha$ -helix 9 in ArhGAP2 for Rho-family protein binding are Gly301/Arg304 and Met413/Asn417/Val421/Pro424, respectively, and these residues are conserved with the corresponding residues of DmSyd1 RhoGAP domain<sup>15</sup>. In the case of DmSyd1, lower quality of matched region exists in the loop region between  $\alpha$ -helix 1–2 and 6–7, which is remotely located at the putative interface of the Rho-family protein (Supplementary Fig. S1). *Drosophila* Syd1 variable residues in the  $\alpha$ -helix 9–10 organizing interface for Rho-family proteins do not match with the corresponding residues in



**Figure 3.** Multiple alignments of SYDE proteins. Color code is based on Clustal $\sigma$  scheme and linear motifs are identified by ELM and D<sub>2</sub>P<sub>2</sub> database Anchor definition<sup>22</sup>. (a) Alignment of SYDE1 and SYDE2. (b) Syde1 N-terminus and C-terminus. (c) Syde2 N-terminus and C-terminus. The signatures indicate: LIG\_SH2\_CRK/NCK: Crk and Nck SH2 domain binding motif; LIG\_14-3-3: 14-3-3 binding motif; DOC\_Cyclin\_RxL: Cyclin/CDK binding motif; LIG\_WW3: PPR-specific WW domain; DOC\_MAPK: a docking motif for interaction with Erk1/2 and p38 subfamilies of MAP kinases; DOC\_PP2B\_LxvP1: calcineurin docking motif; DOC\_PP4: a docking motif recognized by PP4 holoenzyme; MOD\_CK1: CK1 phosphorylation site; MOD\_CK2: CK2 phosphorylation site; MOD\_GSK3: GSK3 phosphorylation recognition site; LIG\_SH2\_GRB2: Grb2-like SH2 binding motif; LIG\_SH3\_3: SH3 binding motif; DEG\_SCF\_FBW7: Phospho-dependent degron recognized by FBW7 F-box proteins; LIG\_LIR\_Gen1: LIR motif for ATG8-mediated autophagy.

(b)  
SYDE1  
N-terminus

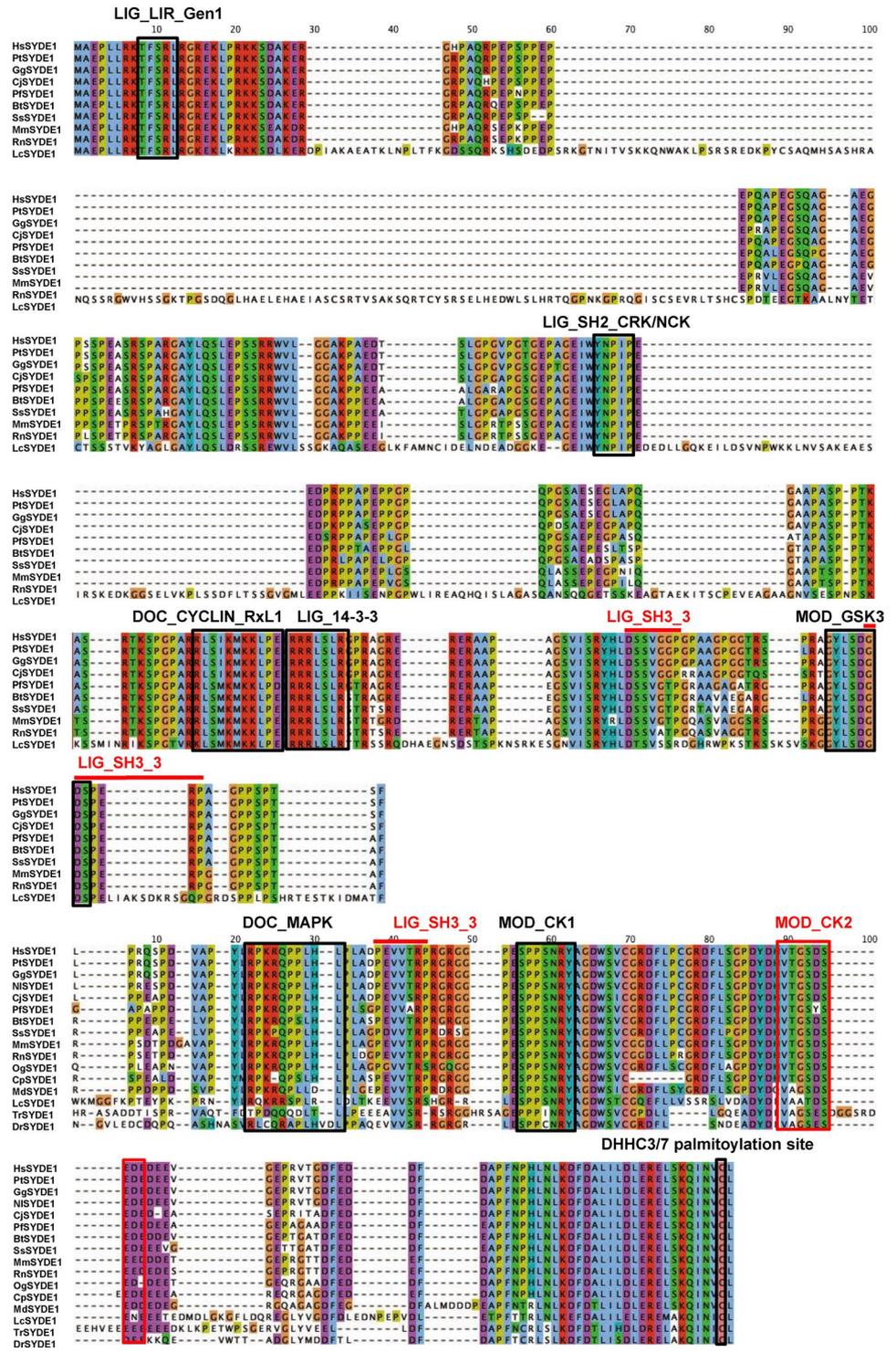
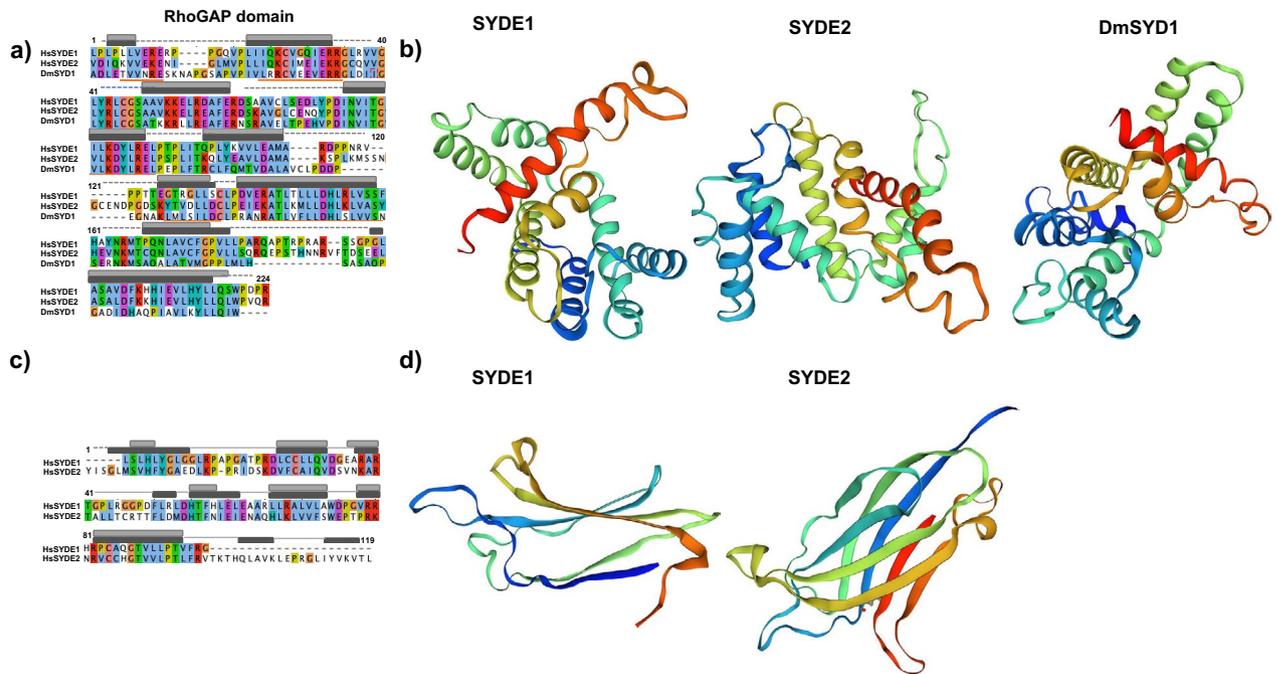


Figure 3. (continued)





**Figure 4.** Overview of SYDE RhoGAP domains and structural RhoGAP models. **(a)** Consensus sequences of SYDE1 RhoGAP domains are shown in the panels. Secondary structures of human SYDE1 are indicated above the sequences ( $\alpha$ -helix: gray column for SYDE1 and dark gray column for SYDE2; coiled-coil: gray dotted line) and  $\alpha$ -helix of DmSyd1 is represented by horizontal orange line below the sequences. **(b)** Cartoons of SYDE RhoGAP domains are colored from the N-terminus (blue) to C-terminus (red). **(c)** Consensus sequences of SYDE C2 domains and secondary structures of human SYDE1/2 used for structure prediction are indicated above the sequences ( $\beta$ -sheet: gray column for SYDE1 and dark gray column for SYDE2). **(d)** Structure of SYDE C2 domains are colored from the N-terminus (blue) to C-terminus (red) by UCSF Chimera. The three-dimensional structures were visualized using UCSF Chimera software version 1.15 (<http://www.cgl.ucsf.edu/chimera>).

	HsSYDE1 RhoGAP (5c2k)	HsSYDE1 C2 (1dix)	HsSYDE2 RhoGAP (5c2k)	HsSYDE2 C2 (5iz5)	DmSyd1 RhoGAP (3cxl)
Ramachandran Plots (favored)	98.2%	100%	97.8%	98.2%	98.3%
Ramachandran Plots (disallowed)	0.6%	0.0%	0.5%	0.9%	0.6%
ERRAT (overall quality factor)	94.09	77.10	98.96	82.69	92.67

**Table 2.** Evaluation of predicted structural models. The model based on template-based prediction was calculated. Ramachandran plots showed several disallowed residues in RhoGAP and HsSYDE2 C2 domain except HsSYDE1 C2 domain with completely favored score in all the predicted region. Each predicted model showed high score in overall quality factor for non-bonded atomic interactions by ERRAT.

Asp708 as opposed to two  $\text{Ca}^{2+}$ . However, the other residues in Syde1 were not conserved with those located in the loop region mediating  $\text{Ca}^{2+}$  binding and PS binding in PLC $\delta$ 1<sup>47</sup>. Both GIVB and GIVD cPLA<sub>2</sub> C2 domains have a high affinity for 1-palmitoyl-2-arachidonyl-sn-glycerol-3-phosphocholine (PAPC) in the presence of  $\text{Ca}^{2+}$ . The basic residues (Lys24/Arg49/Lys52, Lys78 and His44/His82) in the GIVB cPLA<sub>2</sub> C2 domain form the interface for binding to PAPC. Although the residues in GIVB are not conserved with GIVD cPLA<sub>2</sub>, GIVD possessed a higher affinity for PAPC than GIVB cPLA<sub>2</sub> when tested in the in vitro vesicle binding assay<sup>48</sup>. Moreover, the structural interface organized with these residues of Syde2 C2 region was superimposed well to GIVD cPLA<sub>2</sub>, except for the loop region containing Lys24 of GIVD cPLA<sub>2</sub><sup>48</sup>.

**C-terminal regulatory region in Syde1.** Syde1 has a C-terminal disordered domain conserved among mammalian orthologs. The Cys734 palmitoylation site resides at the C-terminus in vertebrate SYDE1 but not in *Drosophila* and worm Syd1<sup>26</sup> (Figs. 2, 3b). Reversible palmitoylation by Golgi-localized DHHC3 allows Syde1 to shuttle between intracellular compartments and plasma membranes. The C-terminal Syde domain contains phosphorylation sites by casein kinases and the p38 subfamily MAPK and calcineurin interaction sites which

are highly conserved among mammalian orthologs. Putative SH3 binding motif is located in the middle of the highly conserved disordered domain.

**Syde1 and Syde2 networks.** Eleven Syde1 and eight Syde2 interactors were manually curated, and four SYDE1 and two SYDE2 binding proteins were directly retrieved from literatures (Fig. 6, Table 1). Most of interactors belonged to a category of proteins that recognized a conserved linear motif located in the disordered regions. XPO1 is indicated as an Syde1 complex by high-throughput evidence, and two regions are predicted as the potential nuclear export signatures by XPO1 in Syde1 (Table 1). S-palmitoylation of Casein kinase 1 $\gamma$  mediates phosphorylation and intracellular transport of Lyn at Golgi apparatus<sup>49</sup>. Golgi-localized DHHC3 specifically palmitoylates Syde1 and casein kinase 1 $\gamma$  might phosphorylate palmitoylated SYDE1<sup>26</sup>. ELM search selected the SH2 binding motif YINS in SYDE2 that was matched with typical Grb2 binding motif validated by pull-down experiment<sup>50,51</sup>. MAPK binding D-motif resides at the C-terminal region in SYDE1, and the sequence has similarity with MKK-type phosphorylation motifs catalyzed by p38<sup>52</sup>. ELM search hit two 14-3-3 binding motifs in SYDE1, and several interactors such as Pctaire1 kinase possess similar RLSLP sequence mediating the interaction<sup>53</sup>. 14-3-3 $\zeta$ s are dimeric with diagonal symmetry wherein two phosphate-binding sites lie in diagonally opposite position when they form a complex with kinases such as the CaMK and AGC family, including PKA. Cin85 is enriched in the brain and localized in the postsynapse, and associated with endocytosis regulators such as endophilins and potential SH3 interaction sites were indicated in SYDE1 (Fig. 3b). SYDE1 possesses RxL docking motif for cyclin binding (DOC\_Cyclin) at the N-terminal disordered domain. The sequence was matched to consensus Cdk1/CyclinB complex binding motif identified by Arg/Lys-scanning oriented peptide libraries<sup>54</sup>. The LIR (LC3-interacting region) motif consists of core W/F/I-X-X- $\phi$  ( $\phi$ : L, I, F) sequence required for selective autophagy. Structural studies indicate the following key features: the aromatic and hydrophobic residues interact with distinct hydrophobic pockets of ubiquitin-like modifiers and an acidic or a phosphorylated Ser/Thr immediately upstream of the core sequence promotes the interaction<sup>55</sup>. The LIR motif in SYDE1 satisfied the definition that was categorized as F-LIR based on the aromatic residue. Non-canonical class I SH3 binding motif RQQVSPP resides at the N-terminus of SYDE2 and brain-specific ArhGAP32 contains the similar motifs that mediate interaction with Crk<sup>56,57</sup>. GSK3 antagonizes Syd1 and liprin ortholog Syd2 pathway in synaptic vesicle clustering in *Drosophila* and sequence alignment suggested that the potential conserved phosphorylation sites lie within the linear motif in both the SYDE1 and SYDE2 as described<sup>58</sup> (Table 1). Conserved Crk binding YNPPIP sites reside in both SYDE1 and SYDE2, and this Tyr phosphorylated motif shows similarity in the patterns with Crk binding sites in Dab1 and other substrates<sup>59</sup>. Calcineurin binding  $\pi\phi$ LxVP motif containing polar and polar/hydrophobic residues at position - 2 ( $\pi$ ) and - 1 ( $\phi$ ), respectively, was selected based on ELM search and its location in SLiM domain<sup>41</sup>. SYDE2 contained SFLRPP and RVLSVP sequence matched the  $\pi\phi$ LxVP motif by ELM search although the motif selected by ELM in SYDE1 did not fulfill the definition. ELM search hit Smek4 as the SYDE2 interactor and centrobilin fragment containing FRVP showed the similar motif patterns and high affinity to PP4 by ITC and colP experiments<sup>60</sup> (<http://slim.icr.ac.uk/motifs/pp4/>). As for Pin1 binding motif, multiple S/T-P motifs mediate the interaction in the case of gephyrin and PSD95, and one S/T-P motif sufficiently functions as Pin1-mediated neuroligin2 binding to gephyrin<sup>61,62</sup>. SYDE2 contains the typical S/T-P motif similar to the Pin1 binding region in the synaptic components at the N-terminus and other putative Pin interaction sites (DOC\_WW\_Pin1\_4) were omitted due to low probability in ELM score. Fbw7 binding consensus degron is represented by  $\phi$ X $\phi\phi\phi$ TTPXS ( $\phi$ : hydrophobic residue, X: any amino acid). Phosphorylated Thr and Ser residues interact with WD domain of Fbw7 and these residues were conserved in Fbw7 binding motif in SYDE2<sup>63</sup>. ArhGAP28 has actin-associated RhoGAP activity with no Rac1 and Cdc42 selectivity; PLEKHG3 functions as RhoGEF in the brain. Both Syde1 and Syde2 have lower GAP activity toward Cdc42 than RhoA and Rac1, which was consistent with our prediction model<sup>12</sup> (Fig. 5). Most Syde1 interactors such as liprina2 and Munc18 are categorized as regulators of synaptogenesis including active zone homeostasis. In comparison, Syde2 binding proteins are primarily involved in neuronal differentiation signaling during embryonic brain development.

## Discussion

Rho GTPases are important for cell adhesion, motility, cytokinesis and contractile responses due to growth factor-induced reorganization of the actin cytoskeleton. Syde1 regulates synaptic exocytosis and its RhoGAP activity is a prerequisite for dendritogenesis. Although the *in vivo* function of mammalian Syde2 remains unknown, Syde2 isoform2 lacks most of the RhoGAP domain at the C-terminus, suggesting complex regulation of the Syde proteins. The homozygous nonsense C.1544C>G mutation (p.(Ser515\*) homo) in Syde2 found in an intellectual disability patient generated a C-terminal truncated transcript devoid of the entire C2 and RhoGAP domain, suggesting a specific role for Syde2 in brain development<sup>11</sup>. It was reported that the C2 domain of Syde1 in addition to the N-terminal disordered domain has an autoinhibitory role for RhoGAP activity<sup>3</sup>; Therefore, the current study may provide the vertebrate-specific Rho-family substrate recognition by the Syde1 RhoGAP domain. Prediction and the structural quality of RhoGAP domains listed in Supplementary Table S2 satisfied the criteria as the potential template listed by trRosetta and I-Tasser. As for the C2 domain of SYDE1 and SYDE2, the potential templates were also selected in trRosetta predicting the structure by multiple sequential alignments as well as calculation of inter-residue orientation, but not in I-Tasser. The predicted model of RhoGAP in Syde1 and Syde2 applied to MgcRacGAP enabled the comparison of the structural interface for its RhoA and Cdc42 (PDB:5c2j), which have different regional conformations that may constrain substrate recognition. A flexible molecular docking search by SwarmDock selected the interaction of both predicted Syde1 and Syde2 with RhoA but not cdc42. The result was consistent with the substantial and extremely low GAP activity toward cdc42 in the RhoGAP assay system using FRET sensors<sup>3,12</sup>. Interestingly, RhoB was also selected as the SYDE1 substrate although the C-terminal regions of RhoA and RhoB show difference in primary sequences to each other.

**Figure 5.** Model of SYDE1 and SYDE2 RhoGAP domain complexed with RhoA. RhoGAP domains of SYDE1(a) and SYDE2 (b) bound to RhoA, and SYDE1 RhoGAP (c) bound to RhoB were modeled by SwarmDock program. Structure of MgcRacGAP/RhoA complex is shown (d). Conserved regions consist of arginine finger (red dotted circle), positively charged interface (purple dotted oval) and hydrophilic interface (orange dotted oval) in RhoGAP domains face to switch II (orange oval) and P-loop (red circle) of RhoA are indicated. Predicted structure of the substrate recognition sites specifically interacting with P-loop and Switch I and II are indicated (e–g). The side-chains of the conserved residues among the RhoGAPs are shown on the cartoon (e–g). (e) Arginine finger locates between second and third  $\alpha$ -helices and cysteine and arginine tether to the interface formed by glycine and glutamic acid of RhoA or RhoB (red-edged circles). (f) Each modeled arginine finger in the RhoGAP complex with RhoA or RhoB is superimposed to the corresponding region of MgcRacGAP structure (PDB: 5c2k). RhoGAP domains of SYDE1/2 and MgcRacGAP are shown as the gold and light blue, respectively. The bound-RhoGTPases to SYDE1/2 and MgcRacGAP are colored as green and pink, respectively. (g) RhoGAP domain and P-loop and SwitchI/II region of RhoGTPases are colored as gold and green, respectively. Side chains of the conserved positively charged residues and conserved residues organizing hydrophilic regions show similar orientation on the ternary helical structure to the corresponding residues of the template by forming the Rho binding interfaces. Cartoon of each model was manually retrieved and indicated as colored molecules separately from the predicted docking structure by Chimera software.

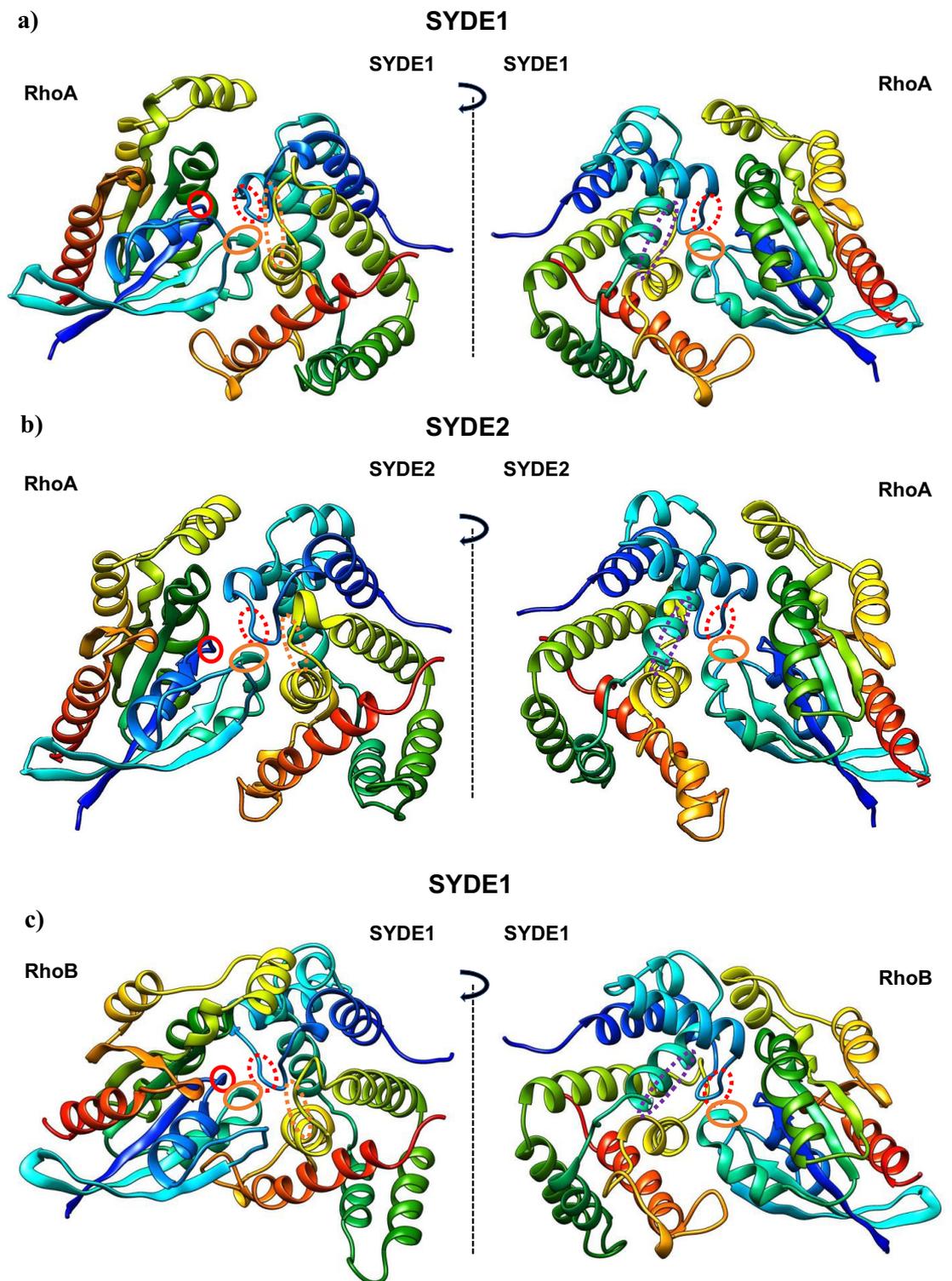
Neuronal regulation of RhoB activity by RhoGAP remains unknown and RhoB is known to regulate dendritic morphology and synaptic plasticity<sup>64</sup>.

The functional Syde protein network was expanded and categorized by integrating structure- and interactor-based prediction with high-throughput evidences: this included inferring the binding partners conserved in lower organisms. Most short linear motifs have a high chance of random occurrence and low specificity. However, stringent criteria were used for selection of interaction sites<sup>65</sup>. A predicted binding site and post-translationally modified residues must be conserved among orthologs and located in a disordered region. Therefore, motif patterns and phosphorylation sites were inferred through several methods<sup>20–22,25</sup>. As for the mammalian Syde proteins, the RhoGAP activity may be highly regulated by phosphorylation/de-phosphorylation and membrane targeting. Both Syde proteins may regulate synaptic plasticity and dendritic morphology with cooperative  $\text{Ca}^{2+}$ -dependent synaptic regulation by Munc18-1 and liprin- $\alpha$ 2, Palmitoylation regulates Syde1-specific membrane targeting and casein kinase I $\gamma$  might cooperatively regulate RhoGTPase's activity at plasma membrane or presynaptic compartment<sup>26,27,49</sup>. Liprin- $\alpha$ 2 and Munc18-1 are known to interact with the Syde1 N-terminal disorder domain and organize active zone complexes in presynaptic regions<sup>3</sup> presumably in a competitive manner with RPTP-liprin $\alpha$  interaction<sup>9</sup>. Optimal 14-3-3 binding sites satisfy basic residues in positions – 3 to – 5 relative to the phosphorylated site and kinases in the AGC family, such as PKA/PKG/PKC and CaMK subfamilies, are most commonly implicated in the phosphorylation of the binding sites<sup>53,66</sup>. Therefore,  $\text{Ca}^{2+}$ -dependent scaffold complex formation might be involved in the Syde1 interaction network. Crk and CIN85 containing SH2-SH3 and three SH3 domains, respectively, are known to regulate pre- and post-synaptic architecture colocalized with PSD95, synaptophysin, and Dock-180<sup>67</sup>. Syde1 and MgcRacGAP were identified by CIN85 interactors through the SH3-C region recognizing the PX(P/A)XXR motif, and Syde1 possesses potential SH3 binding sites as the consensus sequence matched with a variety of sequence patterns<sup>68,69</sup>. The predicted Syde1 interactor cyclin B/Cdk1 is involved in embryonic neurogenesis with the fate competency in opposition to cyclin D function<sup>70</sup>. Cyclins have cross-selectivity and Cyclin D recognizes RxL motif, therefore, SYDE1 might also interact with cyclinD. Casein kinase 2 localizes to the cytosol or membrane in neurons and phosphorylates PACSIN1, thus shutting off Rac1 hydrolysis in the process of dendritic spine formation<sup>71</sup>.

Neuronal Pin1's prolyl cis/trans isomerase activity mediates NMDA dissociation from the PSD95 complex and structural arrangement of voltage-gated  $\text{K}^+$  channels upon neuronal excitability<sup>72</sup>. Syde2's RhoGAP activity might be regulated by Syde2-specific Pin1 binding through the N-terminal disorder region in contrast to the autoinhibitory conformational changes in SYDE1. The Smek1 and PP4 catalytic subunit PP4c are required for neurogenesis<sup>73</sup> and Grb2 is recruited to  $\text{Ca}^{2+}$ -dependent kinase Pyk2 and ErbB2/3 receptors upon Nrg1 signaling during cerebral cortical development<sup>74</sup>. PLEKHG3 is highly expressed in the brain and microdeletion, including the gene, which is known to cause mild mental retardation with spherocytosis<sup>13,75</sup>.

Crk is known to be recruited to Dab1 upon Reelin signaling<sup>76</sup> and may recognize Syde1 and Syde2 proteins through the SH2 domain. Liprin2 converging with PP2A phosphatase are regulated by GSK3 $\beta$  phosphorylation as downstream signaling cascades of Syd1 during neurogenesis in *Drosophila*<sup>58</sup>. Predicted phosphorylation by GSK3 $\beta$  may be required for both mammalian Sydes signaling involved in the exocytosis of synaptic vesicles (Table 1)<sup>58</sup>. The LxVP binding pocket interacts with calcineurin for comprehensive analysis of its substrates<sup>41</sup>. The current search additionally predicted the N-terminal calcineurin binding motif in the Syde2 disordered domain. Linkage of calcineurin with Syde2 implies synaptic activity-dependent AMPA and NMDA receptor function<sup>77</sup>.

Lastly, Syde-family proteins may be regulated by autophagy and proteasomal degradation during neuronal development. LC3/GABARAP protein noncovalently binds to the canonical ATG8 family-interacting motif (AIM), a core motif consisting of W/F/I-X-X-L/I/V, where any two amino acids flank the aromatic and hydrophobic amino acid during the process of phagophore formation in the autophagy pathway<sup>63</sup>. Syde1 might be regulated by ATG8-family-mediated signaling through the AIM motif by coupling with a reversible DHH3/7-mediated palmitoylation. Interestingly Syde2 also contains degradation signals with reflecting the relatively lower Syde2 expression than Syde1 in the developing brain ([www.brainspan.org](http://www.brainspan.org)). Fbw7 is one of the listed SCF-E3 ligases involved in neuronal differentiation<sup>78,79</sup> and the regulation of astrocyte generation, and proteasome-mediated degradation of Syde2 may be involved in its regulation of RhoGTPase during neurodevelopment.



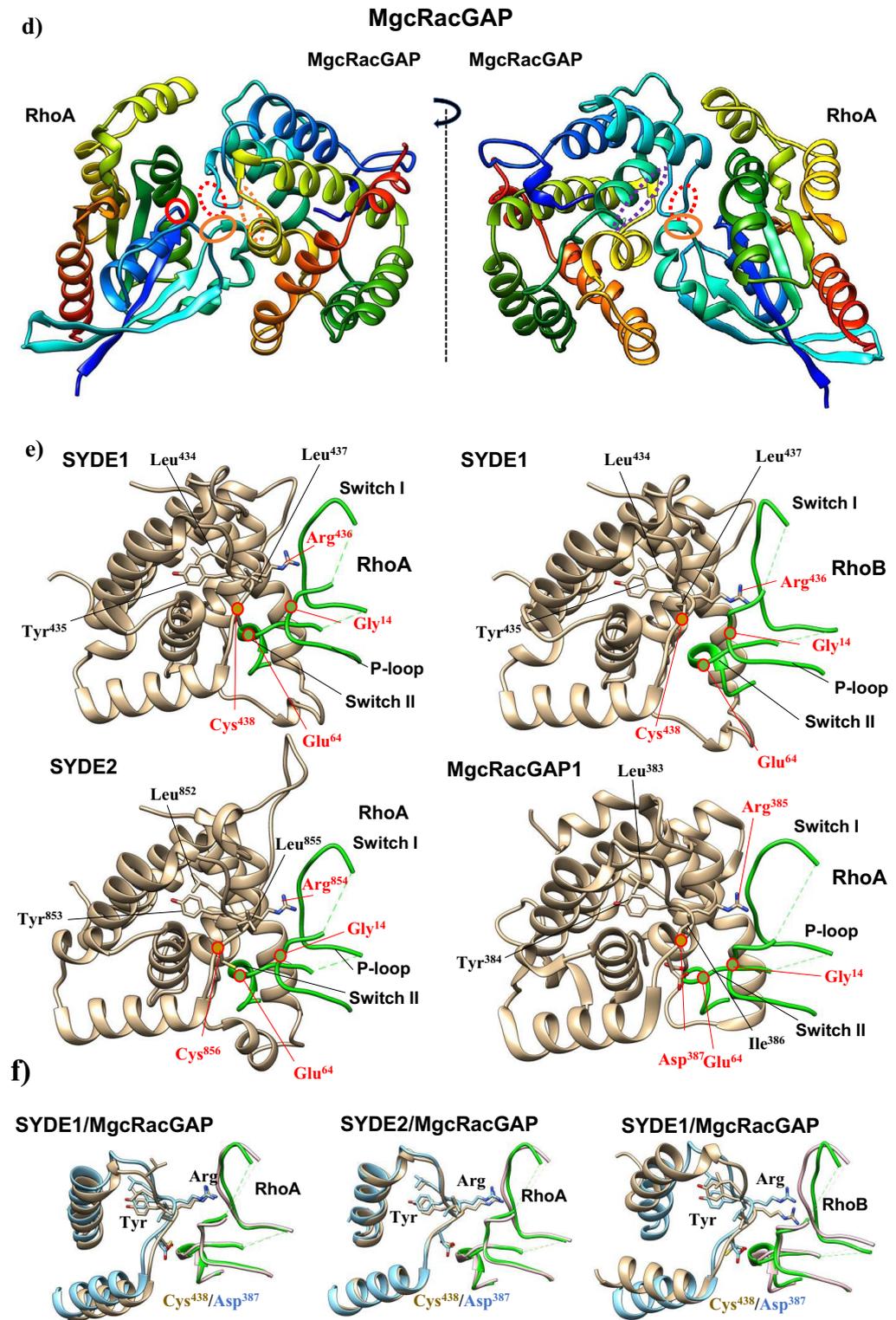


Figure 5. (continued)

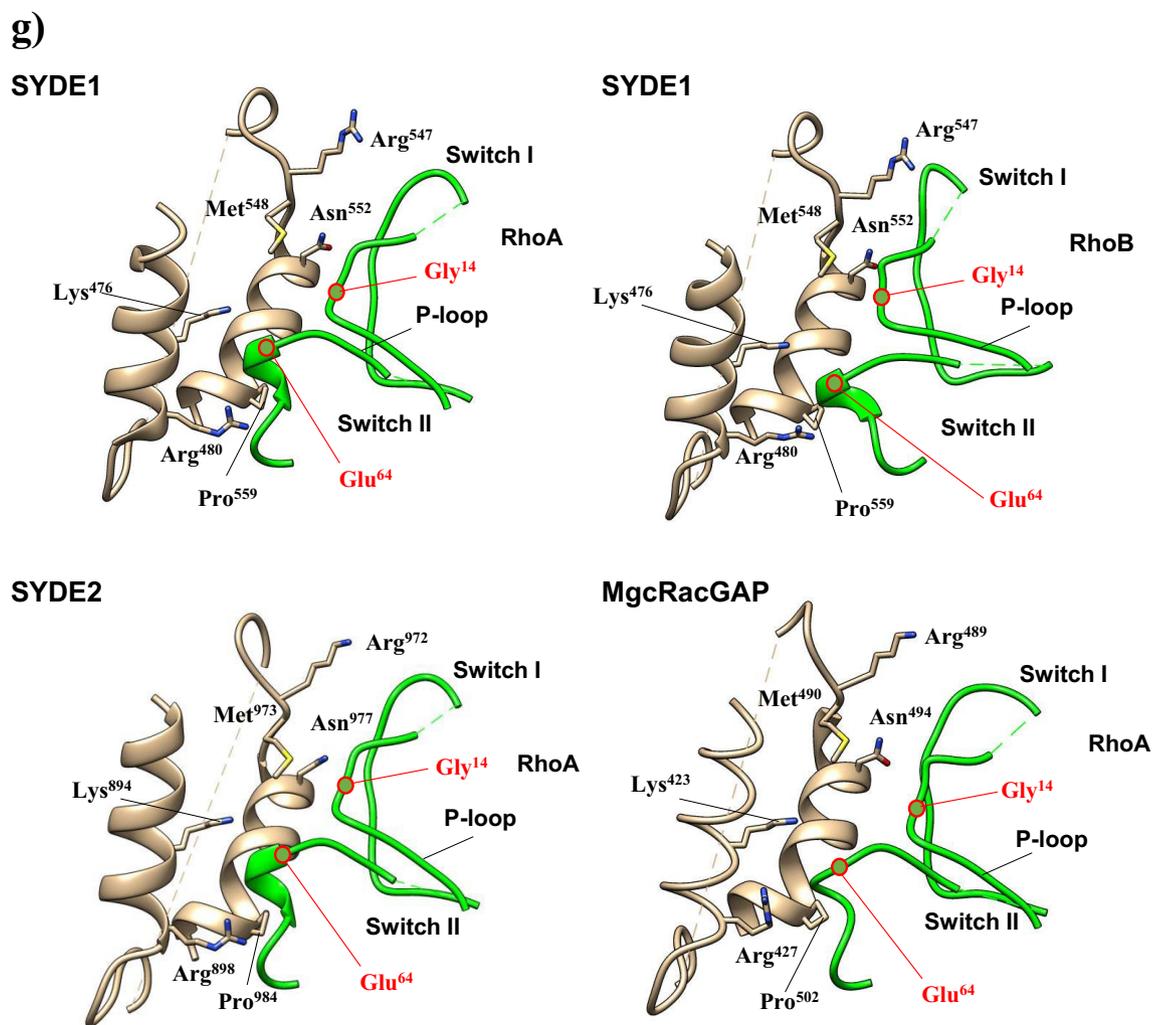
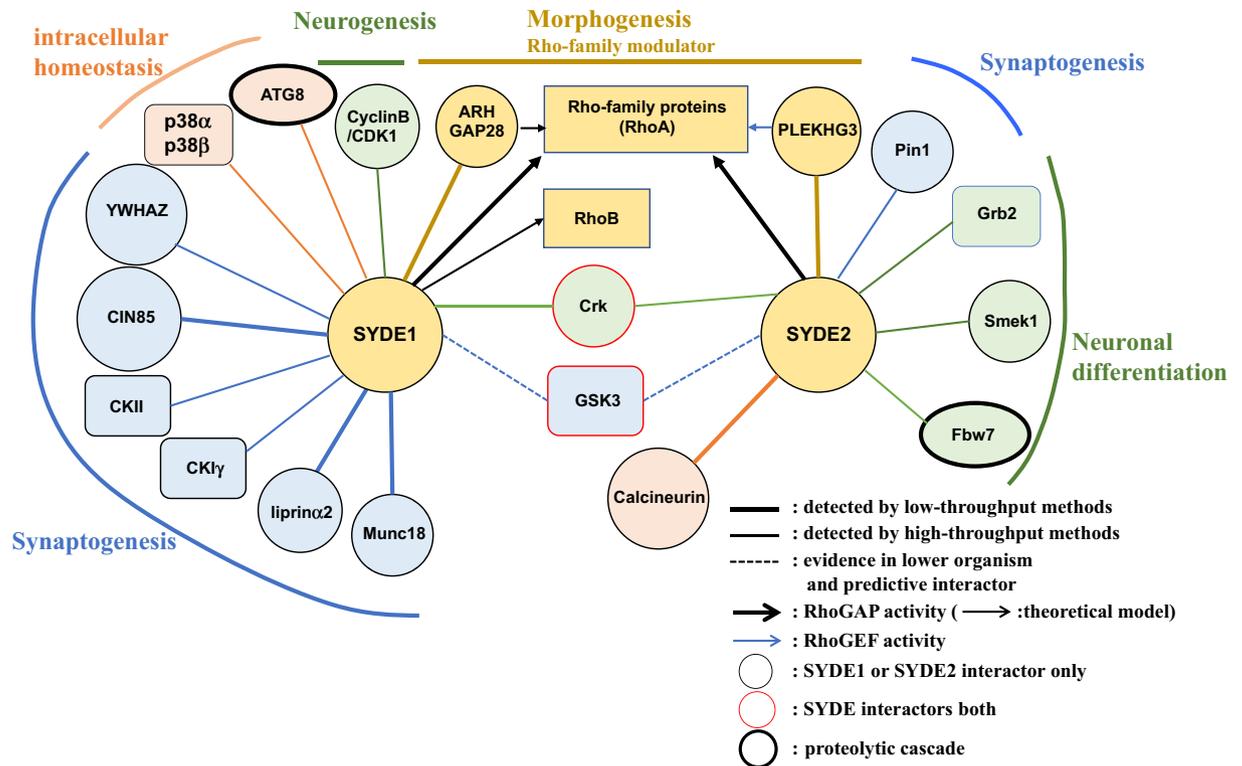


Figure 5. (continued)



**Figure 6.** SYDE protein interaction network. SYDE binding partners identified by low-throughput data (bold line), high-throughput protein–protein interaction data or linear motifs prediction (thin line), and the interactions mediated by conserved sequences as linear motif that are predicted to be associated with binding partner proven in lower organism (dotted line). Interactors are represented by different shapes based on their molecular function: protein kinases (rectangles); degradation machineries (bold-lined ovals); adaptor proteins or phosphatases (circles). Interactor is colored and categorized by neuronal regulatory function as follows: Rho-family modulators (yellow); synaptogenesis (light blue); intracellular homeostasis (light pink); neuronal differentiation (green). SYDE interactor both is marked by red edges and either SYDE1 or SYDE2 only binding partner is represented by black edges. *CKIγ* casein kinase Iγ, *CKII* casein kinase 2.

Received: 19 September 2021; Accepted: 2 March 2022

Published online: 12 March 2022

## References

- Hallam, S. J., Goncharov, A., McEwen, J., Baran, R. & Jin, Y. Syd-1, a presynaptic protein with PDZ, C2 and rhoGAP-like domains specifies axon identity in *C. elegans*. *Nat. Neurosci.* **5**, 1137–1146 (2002).
- Xu, Y. & Quinn, C. C. SYD-1 promotes multiple developmental steps leading to neuronal connectivity. *Mol. Neurobiol.* **53**, 6768–6773 (2016).
- Wentzel, C. *et al.* mSYD1A, a mammalian synapse-defective-1 protein, regulates synaptogenic signaling and vesicle docking. *Neuron* **78**, 1012–1023 (2013).
- Spinner, M. A., Walla, D. A. & Herman, T. G. *Drosophila* Syd-1 has RhoGAP activity that is required for presynaptic clustering of Bruchpilot/ELKS but not Neurexin-1. *Genetics* **208**, 705–716 (2018).
- Xu, Y., Taru, H., Jin, Y. & Quinn, C. C. Syd-1c, Unc-40 (DCC) and Sax-3 (Robo) function interdependently to promote axon guidance by regulating the MIG-2 GTPase. *PLoS Genet.* **11**, e1005185 (2015).
- Lo, H. F. *et al.* Association of dysfunctional synapse defective 1 (SYDE1) with restricted fetal growth—SYDE1 regulates placental cell migration and invasion. *J. Pathol.* **241**, 324–336 (2017).
- Ramesh, N. *et al.* Antagonistic interaction between two neuroligins and coordinate pre- and postsynaptic assembly. *Curr. Biol.* **31**, 1–15 (2021).
- Wei, Z. *et al.* Liprin-mediated large signaling complex organization revealed by the liprina/Cask and liprina/liprinβ complex structures. *Mol. Cell* **43**, 586–598 (2011).
- Wakita, M. *et al.* Structural insights into selective interaction between type IIa receptor protein tyrosine phosphatases and liprin-α. *Nat. Commun.* **11**, 649–658 (2020).
- Rahman, M. R. *et al.* Comprehensive analysis of RNA-seq gene expression profiling of brain transcriptomes reveals novel genes, regulators, and pathways in autism spectrum disorder. *Brain Sci.* **10**, 747 (2020).
- Anazi, S. *et al.* Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield. *Mol. Psychiatry* **22**, 615–624 (2016).
- Müller, P. M. *et al.* Systems analysis of RhoGEF and RhoGAP regulatory proteins reveals spatially organized Rac1 signaling from integrin adhesions. *Nat. Cell Biol.* **22**, 498–511 (2020).
- Griswold, A. J. *et al.* A de novo 1.5Mb microdeletion on chromosome 14q23.2–23.3 in a patient with autism and spherocytosis. *Autism Res.* **4**, 221–227 (2011).
- Mosaddeghzadeh, N. & Ahmadian, M. R. The Rho family GTPases: Mechanisms of regulation and signaling. *Cells* **10**, 1831 (2021).
- Amin, E. *et al.* Deciphering the molecular and functional basis of RhoGAP family proteins. *J. Biol. Chem.* **291**, 20353–20371 (2016).

16. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2019).
17. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
18. Blum, M. *et al.* The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* **49**, D344–354 (2020).
19. Buchan, D. W. A., Minnici, F., Nugent, T. C. O., Bryson, K. & Jones, D. T. Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res.* **41**, W349–W357 (2013).
20. Potenza, E., Di Domenico, T., Walsh, I. & Tosatto, S. C. E. MobiDB 2.0: An improved database of intrinsically disordered and mobile proteins. *Nucleic Acid. Res.* **43**, D315–D320 (2015).
21. Dinkel, H. *et al.* ELM 2016-data update and new functionality of eukaryotic linear motif resource. *Nucleic Acids Res.* **44**, D294–D300 (2016).
22. Oates, M. E. *et al.* D2P2: Database of disordered protein predictions. *Nucleic Acids Res.* **41**, D508–D516 (2012).
23. Tompa, P., Davey, N. E., Gibson, T. J. & Babu, M. M. A million peptide motifs for the molecular biologist. *Mol. Cell* **55**, 161–169 (2014).
24. Mészáros, B., Simon, I. & Dosztányi, Z. Prediction of protein binding regions in disordered proteins. *PLoS Comput. Biol.* **5**, e1000376 (2009).
25. de Castro, E. *et al.* ScanProsite: Detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids. Res.* **34**, W362–365 (2006).
26. Oku, S., Takahashi, N., Fukata, Y. & Fukata, M. In silico screening for palmitoyl substrates reveals a role for DHHC1/3/10(zDHHC1/3/11)-mediated neurochondrin palmitoylation in its targeting to rab5-positive endosomes. *J. Biol. Chem.* **288**, 19816–19829 (2013).
27. Blanc, M. *et al.* SwissPalm: Protein palmitoylation database. *F1000 Res.* **4**, 26 (2015).
28. Xu, D., Farmer, A., Collett, G., Grishin, N. V. & Chook, Y. M. Sequence and structural analyses of nuclear export signals in the NESb database. *Mol. Biol. Cell* **23**, 3677–3693 (2012).
29. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
30. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005).
31. Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res.* **31**, 3381–3385 (2003).
32. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).
33. Yang, J. *et al.* Improved protein structure prediction using predicted interresidue orientations. *Proc. Natl. Acad. Sci.* **117**, 1496–1503 (2020).
34. Benkert, P., Künzli, M. & Schwede, T. QMEAN server for protein model quality estimation. *Nucleic Acids Res.* **37**, W510–W514 (2009).
35. Messaoudi, A., Belguith, H. & Hamida, B. Homology modeling and virtual screening approaches to identify potent inhibitors of VEB-1  $\beta$ -lactamase. *Theor. Biol. Med. Model.* **10**, 22 (2013).
36. Torchala, M., Moal, I. H., Chaleil, R. A. G., Fernandez-Recio, J. & Bates, P. A. SwarmDock: A server for flexible protein-protein docking. *Bioinformatics* **29**, 807–809 (2013).
37. Pokala, N. & Handel, T. M. Protein-protein complex affinities, models for the unfolded state, and negative design of solubility and specificity. *J. Mol. Biol.* **347**, 203–227 (2005).
38. Chattrayamontri, A. *et al.* The BioGRID interaction database: 2015 update. *Nucleic Acids Res.* **43**, D470–D478 (2015).
39. Orchard, S. *et al.* The MintAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42**, D358–363 (2014).
40. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D351 (2017).
41. Sheftic, S., Page, R. & Peti, W. Investigating the human calcineurin interaction network using the  $\pi\phi$ LxVP SLiM. *Sci. Rep.* **6**, 38920 (2016).
42. Matsuura, A. & Lee, H. H. Crystal structure of GTPase-activating domain from human MgcRacGAP. *Biochem. Biophys. Res. Commun.* **435**, 367–372 (2013).
43. Mishima, M. & Glotzer, M. Cytokinesis: A logical GAP. *Curr. Biol.* **13**, R589–R591 (2003).
44. Minoshima, Y. *et al.* Phosphorylation by AuroraB converts MgcRacGAP to a RhoGAP during cytokinesis. *Dev. Cell* **4**, 549 (2003).
45. Wang, H. *et al.* Phospholipase C $\gamma$ /diacylglycerol-dependent activation of  $\beta$ 2-chimaerin restricts EGF-induced Rac signaling. *EMBO J.* **25**, 2062–2074 (2006).
46. Corbalan-García, S. & Gomez-Fernández, J. C. Signaling through C2 domain: More than one lipid target. *Biochim. Biophys. Acta* **1838**, 1536–1547 (2014).
47. Lomansey, J. W., Cheng, H. F., Roffler, S. R. & King, K. Activation of phospholipase C $\delta$ 1 through C2 domain by a Ca<sup>2+</sup>-enzyme-phosphatidylserine ternary complex. *J. Biol. Chem.* **274**, 21995–22001 (1999).
48. Ghomashchi, F. *et al.* Interfacial kinetic and binding properties of mammalian group IVB phospholipase A<sub>2</sub> (cPLA<sub>2</sub> $\beta$ ) and comparison with the other cPLA<sub>2</sub> isoforms. *J. Biol. Chem.* **285**, 36100–36111 (2010).
49. Kinoshita-Kikuta, E. *et al.* Protein N-myristoylation-dependent phosphorylation of serine 13 of tyrosine kinase Lyn by casein kinase 1 $\gamma$  at the Golgi during intracellular protein traffic. *Sci. Rep.* **10**, 16273 (2020).
50. Tinti, M. *et al.* The SH2 domain interaction landscape. *Cell Rep.* **3**, 1293–1305 (2013).
51. Liu, B. A. *et al.* SH2 domains recognize contextual peptide sequence information to determine selectivity. *Mol. Cell. Proteom.* **9**(11), 2391–2404 (2010).
52. Garai, A. *et al.* Specificity of linear motifs that bind to a common mitogen-activated protein kinase docking groove. *Sci. Signal.* **5**, ra74. <https://doi.org/10.1126/scisignal.2003004> (2013).
53. Johnson, C. *et al.* Bioinformatics and experimental survey of 14-3-3-binding sites. *Biochem. J.* **427**, 69–78 (2010).
54. Suzuki, K. *et al.* Identification of non-Ser/Thr-Pro consensus motifs for Cdk1 and their roles in mitotic regulation of C2H2 zinc finger proteins and Ect2. *Sci. Rep.* **5**, 7929 (2015).
55. Rogov, V., Dötsch, V., Johansen, T. & Kirkin, V. Interaction between autophagy receptors and ubiquitin-like proteins. *Mol. Cell* **53**, 167–178 (2014).
56. Teyra, J. *et al.* Comprehensive analysis of the human SH3 domain family reveals a wide variety of non-canonical specificities. *Structure* **25**, 1598–1610 (2017).
57. Moon, S. Y., Zang, H. & Zheng, Y. Characterization of a brain-specific Rho GTPase-activating protein, p200RhoGAP. *J. Biol. Chem.* **278**, 4151–4159 (2003).
58. Li, L., Tian, X., Zhu, M., Bulgari, D. & Böhme, M. A. Drosophila Syd-1, Liprin- $\alpha$ , and protein phosphatase 2A B' subunit wrd function in a linear pathway to prevent ectopic accumulation of synaptic materials in distal axons. *J. Neurosci.* **34**, 8474–8487 (2014).
59. Ballif, B. A. *et al.* Activation of a Dab1/CrkL/C3G/Rap1 pathway in Reelin-stimulated neurons. *Curr. Biol.* **14**, 606–610 (2004).
60. Ueki, Y. *et al.* A consensus binding motif for the PP2 protein phosphatase. *Mol. Cell* **76**, 953–964 (2019).

61. Antonelli, R. *et al.* Pin1-dependent signaling negatively affects GABAergic transmission by modulating neuroligin2/gephyrin interaction. *Nat. Commun.* **5**, 5066 (2014).
62. Zita, M. M. *et al.* Post-phosphorylation proyl isomerization of gephyrin represents a mechanism to modulated glycine receptors function. *EMBO J.* **26**, 1761–1771 (2007).
63. Hao, B., Oehlmann, S., Sowa, M. E., Harper, J. W. & Pavletich, N. P. Structure of a Fbw7-Skp1-Cyclin E complex: Multisite-phosphorylated substrate recognition by SCF ubiquitin ligases. *Mol. Cell* **26**, 131–143 (2007).
64. McNair, K. *et al.* A role for RhoB in synaptic plasticity and the regulation of neuronal morphology. *J. Neurosci.* **30**, 3508–3517 (2010).
65. Gibson, T. J., Dinkel, H., Roey, K. V. & Diella, F. Experimental detection of short regulatory motifs in eukaryotic proteins: tips for good practice as well as for bad. *Cell Commun. Signal.* **13**, 42 (2015).
66. Toyo-oka, K. *et al.* 14-3-3 $\epsilon$  and  $\zeta$  regulate neurogenesis and differentiation of neuronal progenitor cells in the developing brain. *J. Neurosci.* **34**, 12168–12181 (2014).
67. Shimokawa, N. *et al.* CIN85 regulates dopamine receptor endocytosis and governs behavior in mice. *EMBO J.* **29**, 2421–2432 (2010).
68. Rouka, E. *et al.* Differential recognition preferences of the three src homology 3 (SH3) domains from the adaptor CD2-associated protein (CD2AP) and direct association with Ras and Rab interactors 3 (RIN3). *J. Biol. Chem.* **290**, 25275–25292 (2015).
69. Havrylov, S., Rzhpetsky, Y., Malinowska, A., Drobot, L. & Redowicz, M. J. Proteins recruited by SH3 domains of Ruk/CIN85 adaptor identified by LC-MS/MS. *Proteome Sci.* **7**, 21 (2009).
70. Hagey, D. W. *et al.* Cyclin-B1/2 and -D1 act in opposition to coordinate cortical progenitor self-renewal and lineage commitment. *Nat. Commun.* **11**, 2898 (2020).
71. Schael, S. *et al.* Casein kinase 2 phosphorylation of protein kinase C and casein kinase 2 substrate in neurons (PASCIN)1 protein regulates neuronal spine formation. *J. Biol. Chem.* **288**, 9393–9312 (2013).
72. Antonelli, R. *et al.* Pin1 modulates the synaptic content of NMDA receptors via proylisomerization of PSD95. *J. Neurosci.* **36**, 5437–5447 (2016).
73. Lyu, J. *et al.* Protein phosphatase 4 and smek complex negatively regulate Par3 and promote Neuronal Differentiation of neural stem/progenitor cells. *Cell Rep.* **5**, 593–600 (2013).
74. Mei, L. & Xiong, W. C. Neuregulin 1 in neural development, synaptic plasticity and schizophrenia. *Nat. Rev. Neurosci.* **9**, 437–452 (2008).
75. Lybaek, H., Oyen, N., Fauske, L. & Houge, G. A 2.1 Mb deletion adjacent but distal to 14q21q23 paracentric inversion in a family with spherocytosis and severe learning difficulties. *Clin. Genet.* **74**, 553–339 (2008).
76. Sekine, K. *et al.* Reelin controls neuronal positioning by promoting cell-matrix adhesion via inside-out activation of integrin  $\alpha 5 \beta 1$ . *Neuron* **76**, 353–369 (2012).
77. Sanderson, J. L., Gorski, J. A. & Dell'Acqua, M. L. NMDA receptor-dependent LTD requires transient synaptic incorporation of Ca<sup>2+</sup>-permeable AMPARs mediated by AKAP150-anchored PKA and calcineurin. *Neuron* **89**, 1000–1015 (2016).
78. Matsumoto, A. *et al.* Fbxw7-dependent degradation of notch is required for control of “stemness” and neuronal-glial differentiation in neural stem cells. *J. Biol. Chem.* **286**, 13753–13764 (2011).
79. Hoeck, J. D. *et al.* Fbw7 controls neural stem cell differentiation and progenitor apoptosis via Notch and c-Jun. *Nat. Neurosci.* **13**, 1365–1372 (2010).
80. Kirli, K. *et al.* A deep proteomics perspective on CRM1-mediated nuclear export and nucleocytoplasmic partitioning. *Elife* **4**, e11466 (2015).
81. Oswald, D. *et al.* Cooperation of Syd-1 with neuexin synchronizes pre- with postsynaptic assembly. *Nat. Neurosci.* **15**, 1219–1226 (2012).
82. Oswald, D. *et al.* A Syd-1 homologue regulates pre- and postsynaptic maturation in *Drosophila*. *J. Cell Biol.* **188**, 565–579 (2010).

## Author contributions

Z.K. designed the study, constructed the sequence and interaction analysis, wrote the manuscripts. Z.K. and M.K. performed molecular modeling and analyzed the structural data.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-08147-7>.

**Correspondence** and requests for materials should be addressed to Z.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022