

**Supplementary Information**

**for**

**Structural basis for regulation of CELSR1 by a**

**compact module in its extracellular region**

Sumit J. Bandekar<sup>1, 2, 3, 4</sup>, Krassimira Garbett<sup>5</sup>, Szymon P. Kordon<sup>1, 2, 3, 4</sup>, Ethan E. Dintzner<sup>1, 2, 3, 4</sup>, Jingxian Li<sup>1, 2, 3, 4</sup>, Tanner Shearer<sup>5</sup>, Richard C. Sando<sup>5, \*</sup>, and Demet Araç<sup>1, 2, 3, 4, \*</sup>

<sup>1</sup>Department of Biochemistry and Molecular Biology, The University of Chicago,  
Chicago, IL, 60637, USA.

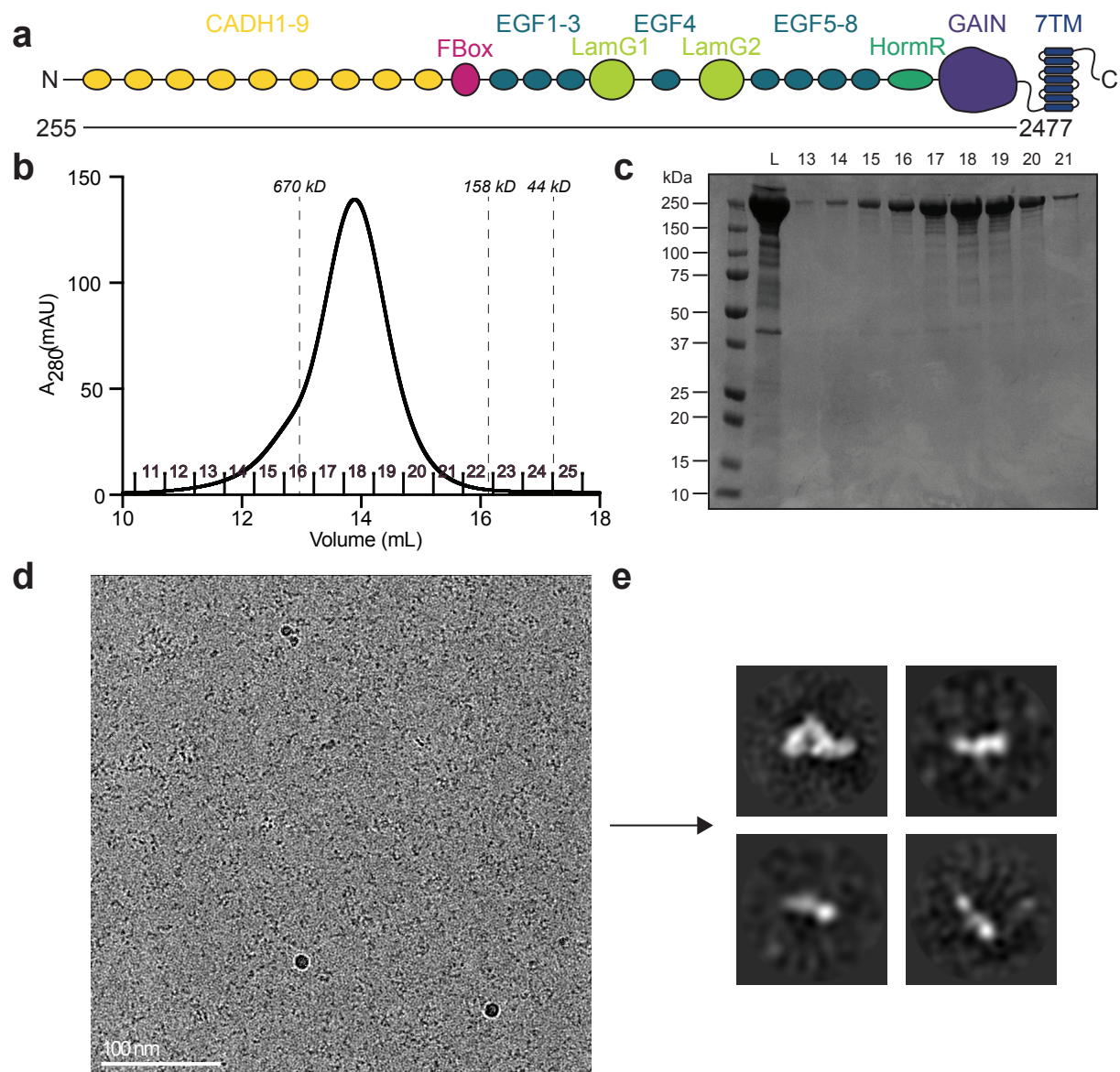
<sup>2</sup>The University of Chicago Neuroscience Institute, The University of Chicago, Chicago,  
IL, 60637, USA.

<sup>3</sup>Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL, 60637,  
USA.

<sup>4</sup>Center for Mechanical Excitability, The University of Chicago, Chicago, IL, 60637, USA.

<sup>5</sup>Department of Pharmacology, Vanderbilt Brain Institute, Vanderbilt University,  
Nashville, TN 37240, USA

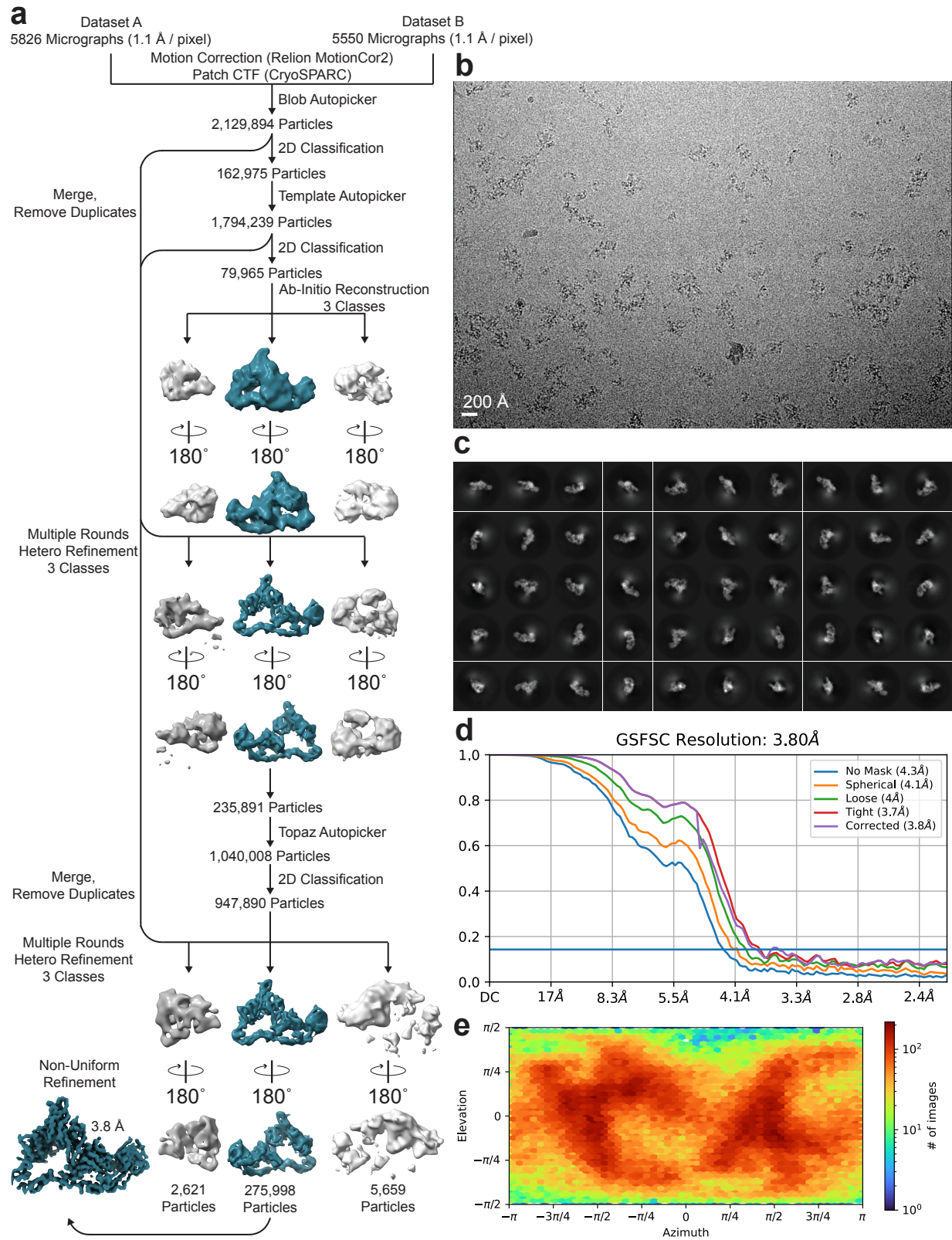
\*Corresponding authors: Demet Araç, [arac@uchicago.edu](mailto:arac@uchicago.edu) and Richard C. Sando, [Richard.sando@vanderbilt.edu](mailto:Richard.sando@vanderbilt.edu)



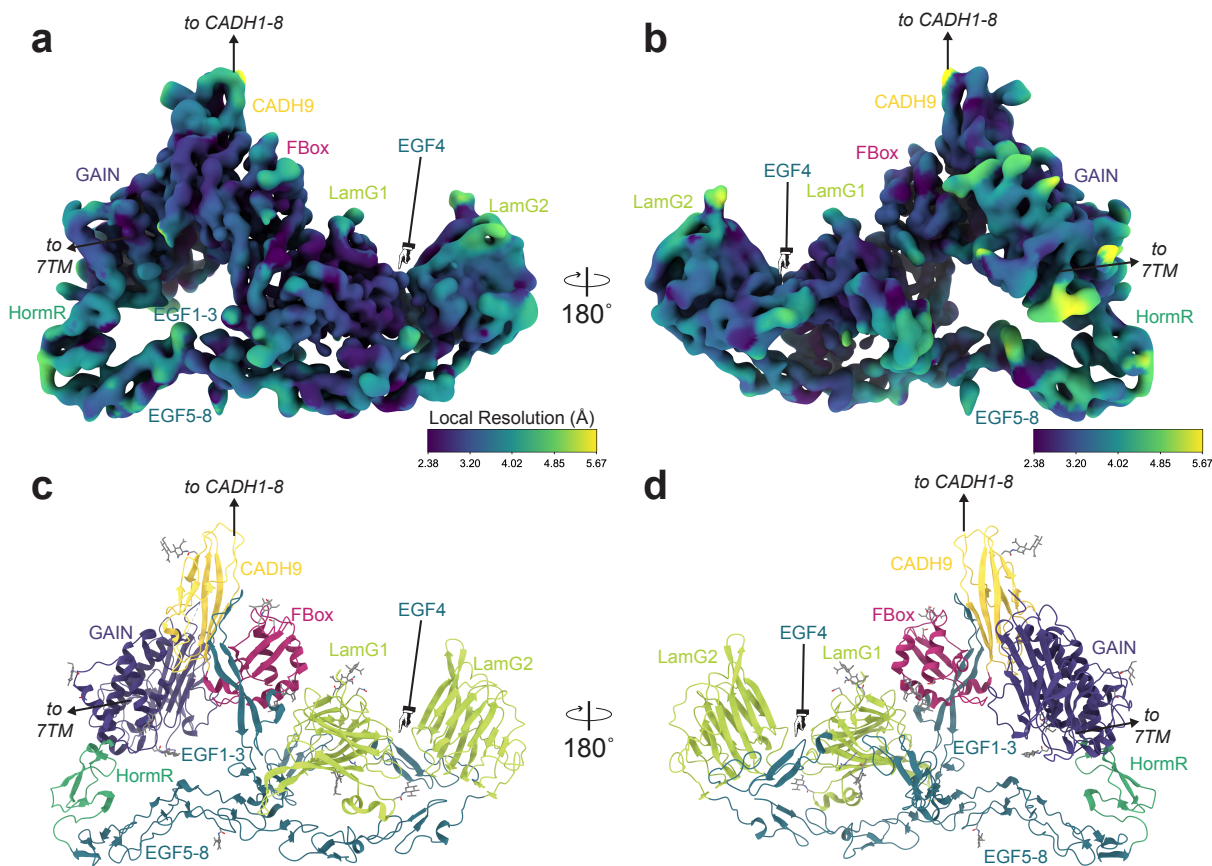


**Supplementary Fig. 1. Purification and cryo-EM screening of the CELSR1 ECR. a**

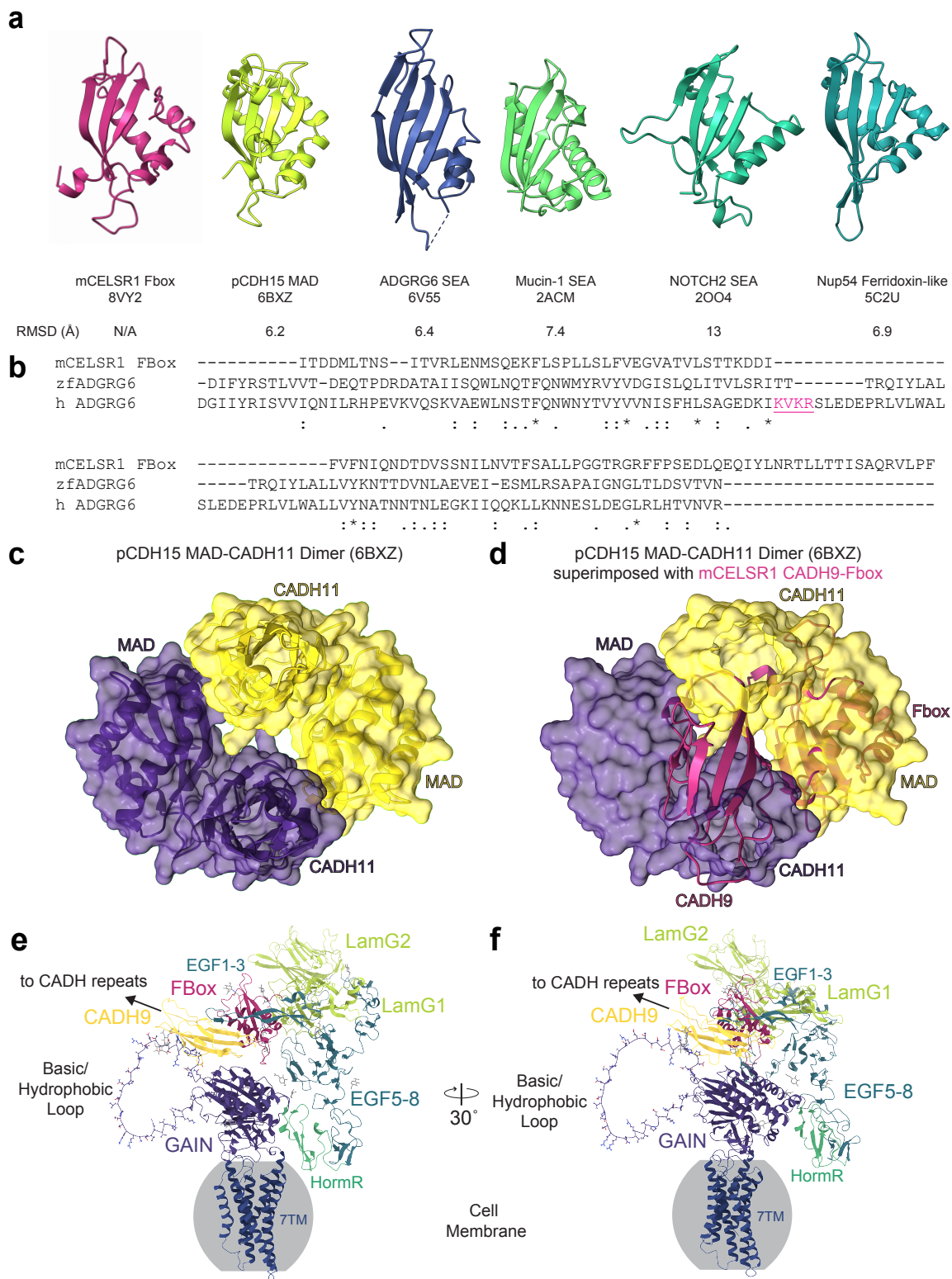
CELSR1 domain diagram with horizontal line showing the extent of the construct used for the structural studies. CADH: cadherin repeat (yellow oval); Fbox: Flamingo Box domain (pink oval); EGF: epidermal growth factor repeat (teal oval); LamG: laminin G repeat (lime green oval); HormR: hormone receptor domain (green oval); GAIN: GPCR autoproteolysis-inducing domain (purple shape); 7TM: seven transmembrane region (blue cylinders); ICR: intracellular region. **b** Size exclusion chromatogram of CELSR1 purification with fraction numbers labeled in purple and size exclusion standards shown as vertical dashed lines. **c** Reducing SDS-PAGE of fractions from SEC trace in **b**. L: fraction loaded onto SEC. **d** Representative screening micrograph of CELSR1 ECR. without added calcium. **e** 2D averages generated from screening micrographs without added calcium. Source data are provided as a Source Data file.



**Supplementary Fig. 2. Cryo-EM data processing workflow for the CELSR1 ECR reconstruction.** The steps used to process the cryo-EM dataset are shown. **a** Flowchart explaining the data processing steps to generate the CELSR1 reconstruction. The blue volumes represent the volumes and particle sets that were carried on to following steps. **b** Example micrograph collected on Krios TEM. **c** Representative 2D averages of the final particle set. **d** Fourier shell correlation curve showing gold-standard Fourier shell correlation-based resolution estimation for the CELSR1 reconstruction. Blue curve, no masking; orange curve, spherical masking; green curve, loose masking; red curve, tight masking; purple curve, corrected masking. **e** Viewing direction distribution plot showing the angular distributions of how particles are assigned to the final 3D volume. Blue points on this plot represent angular values with few particle assignments and red points represent angular values with a high amount of particles assigned.

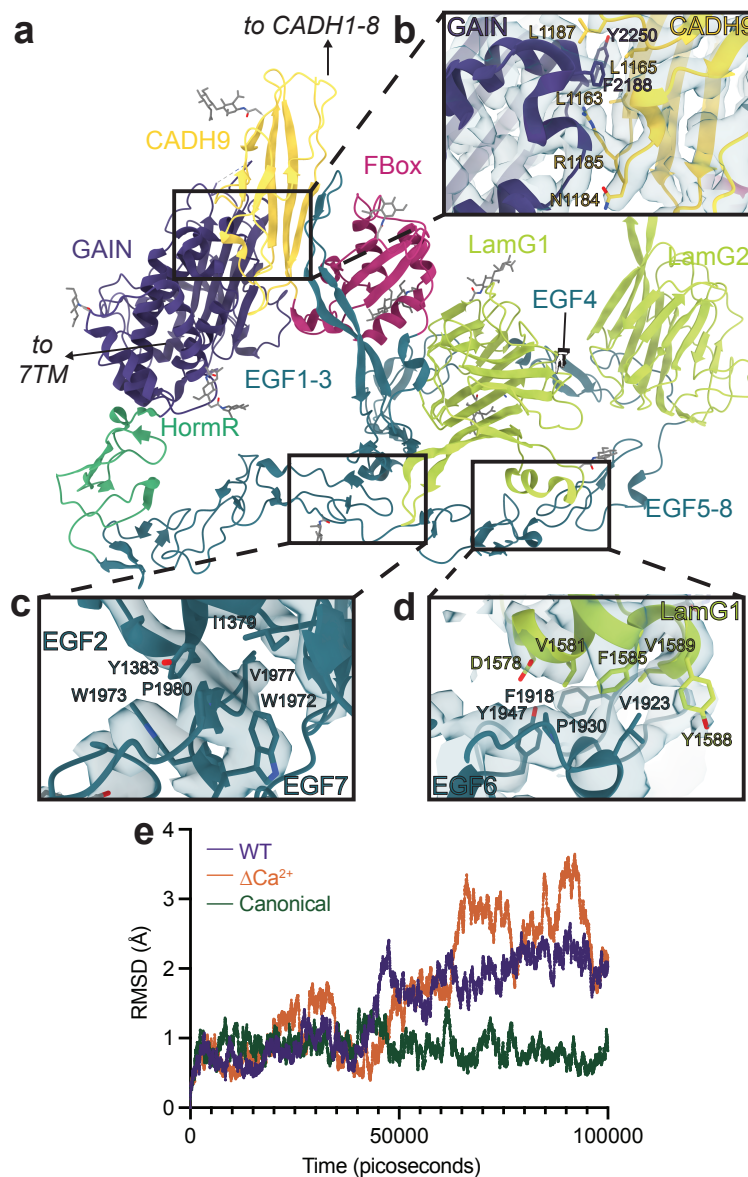


**Supplementary Fig. 3. Local resolution maps of the CELSR1 cryo-EM reconstruction.** **a,b** Density maps of the CELSR1 reconstruction are shown in two views which are colored by local resolution in Å. Maps were generated using the local resolution tools in cryoSPARC and colored using ChimeraX using the viridis color palette. Purple, highest resolution regions; teal and green, intermediate resolution regions; yellow, lowest resolution regions. Domains are labeled and the N- and C-termini are labeled. **c,d** The PDB model for CELSR1 is presented in aligned views to the maps in **a** and **b** for ease of viewing. CADH: cadherin repeat (yellow); Fbox: Flamingo Box domain (pink); EGF: epidermal growth factor repeat (teal); LamG: laminin G repeat (lime green); HormR: hormone receptor domain (green); GAIN: GPCR autoproteolysis-inducing domain (purple); 7TM: seven transmembrane region.



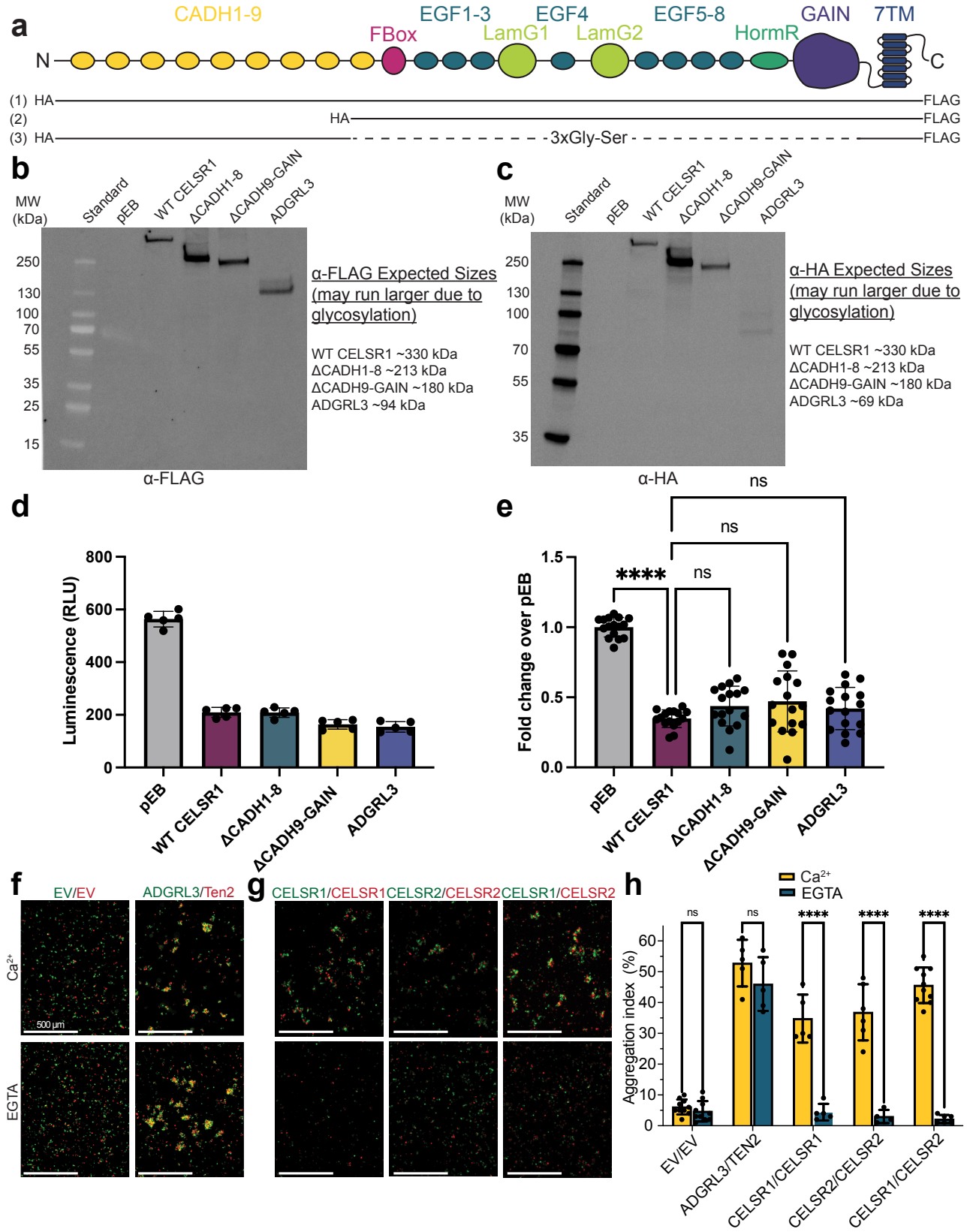
**Supplementary Fig. 4. Analysis of the CELSR1 structure reveals the experimental structure of the Fbox domain and a hypothesis for membrane orientation.** **a** The experimental CELSR1 Fbox structure (pink) is shown in a similar orientation to structural homologs identified using the DALI webserver. The Fbox domain is structurally similar to MAD (neon yellow), SEA (blue, green, turquoise), and Ferridoxin-like (teal) domains. The all-atom RMSD is shown below. **b** Sequence alignments of Fbox domain with zebrafish and human ADGRG6 reveal that the CELSR1 Fbox domain does not have a furin cleavage site where some other SEA domains do. The furin cleavage site in human ADGRG6 is shown in magenta and is underlined. **c** The pCDH15 CADH11-MAD dimerization site is shown (monomers are colored purple or yellow) and **d** the CADH9-Fbox (colored pink) orientation is shown to sterically clash with the pCDH15 mode of dimerization. **e,f** AlphaFold2 predictions for the GAIN-7TM region of CELSR1 overlaid with the CADH9-GAIN CMM suggests a possible membrane orientation for the CADH9-GAIN CMM. CADH: cadherin repeat (yellow); Fbox: Flamingo Box domain (pink); EGF: epidermal growth factor repeat (teal); LamG: laminin G repeat (lime green); HormR: hormone receptor domain (green); GAIN: GPCR autoproteolysis-inducing domain (purple); 7TM: seven transmembrane region (blue).





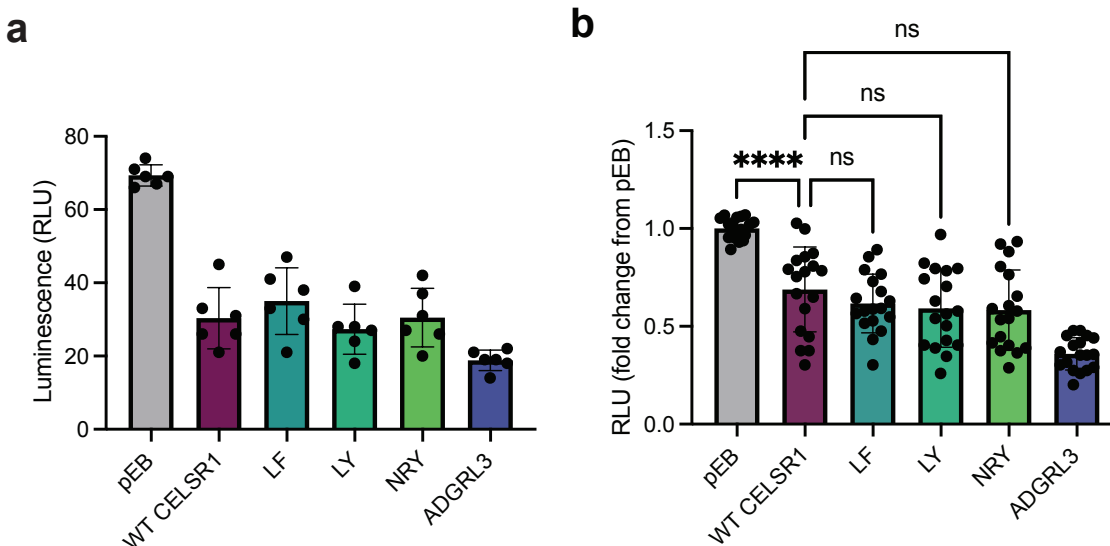
**Supplementary Fig. 5: Interdomain interfaces stabilizing the CELSR1 ECR compact bundle rendered using ChimeraX, with density shown. a** Cartoon representation of CELSR1 ECR structure with key interfaces between **b** CADH9 and GAIN domains, **c** EGF2 and EGF7, and **d** EGF6 and LamG1 with key side chains shown in sticks and density shown using blue surface representation. CADH: cadherin repeat (yellow); Fbox: Flamingo Box domain (pink); EGF: epidermal growth factor repeat (teal); LamG: laminin G repeat (lime green); HormR: hormone receptor domain (green); GAIN: GPCR

autoproteolysis-inducing domain (purple); 7TM: seven transmembrane region. **e** Plots of root mean square deviation (RMSD) for  $\alpha$ -Carbons vs. time for the WT (purple),  $\Delta\text{Ca}^{2+}$  (burnt orange), and Canonical (dark green) simulations from Fig. 4.



**Supplementary Fig. 6. Functional analysis of CELSR ECR constructs.** **a** Domain boundaries of deletion constructs used in this study. CADH: cadherin repeat (yellow oval); Fbox: Flamingo Box domain (pink oval); EGF: epidermal growth factor repeat (teal oval); LamG: laminin G repeat (lime green oval); HormR: hormone receptor domain (green oval); GAIN: GPCR autoproteolysis-inducing domain (purple shape); 7TM: seven transmembrane region (blue cylinders); ICR: intracellular region. Black solid lines show domain boundaries with dotted lines indicating regions that are connected with linkers. **b** Experimental constructs were probed using western blotting against the C-terminal FLAG tag or **c** N-terminal HA tag to confirm the construct size. Expected sizes are noted, some constructs may run larger than expected due to glycosylation. This experiment was repeated three times using independent transfections with similar results. **d** Signaling assay for CELSR function. Experimental constructs are co-transfected with  $\beta_2$ AR and CELSR1 constructs result in lower cAMP relative to empty vector control. ADGRL3 is used as a positive control as previously described<sup>1-3</sup>. Representative raw experiment shown. Each point represents a technical replicate, and data are represented as mean  $\pm$  standard deviation. pEB, grey; WT CELSR1, purple;  $\Delta$ CADH1-8, teal;  $\Delta$ CADH9-GAIN, yellow; ADGRL3, blue. **e** Compiled results of N=3 independent biological replicates performed each with 5 technical replicates presented as fold change over empty vector, each point shown represents a technical replicate. Data are presented as mean  $\pm$  standard deviation. Bars are colored as in **d**. **f,g** Representative cell aggregation assay images showing effect of EGTA on cell aggregation mediated by ADGRL3/Ten2 and CELSR isoforms. Green lettering indicates constructs that were co-transfected with eGFP and red lettering indicates constructs that were co-transfected with dsRed. EGTA does

not affect ADGRL3/Ten2-mediated aggregation but disrupts CELSR-mediated aggregation. CELSR1/CELSR2 can aggregate heterophilically. The experiment was repeated twice with similar results. **h** Quantification of **f**, **g** as aggregation index. Each data point represents a technical replicate from the same biological experiment, and the data is presented as mean +/- standard deviation. Yellow bars represent the presence of calcium and blue bars represent the presence of EGTA. one-way ANOVA with Tukey's correction for multiple comparisons was performed to assess statistical significance between CELSR constructs in panels **e** and **h**. \*\*\*\* corresponds to  $p < 0.0001$ . Each construct was compared to each other construct for the statistical testing, but only certain comparisons are shown for clarity. Source data are provided as a Source Data file.



**Supplementary Fig. 7. Functional analysis of CADH9/GAIN interfacial point mutations.** Experimental constructs are co-transfected with  $\beta$ 2AR and CELSR1 constructs result in lower cAMP relative to empty vector control. LF, L1163A/F2188A; LY, L1163A/Y2250A; NRY, N1184A/R1185A/Y2250A. ADGRL3 is used as a positive control as previously described<sup>1-3</sup>. pEB, grey; WT CELSR1, purple; LF, teal; LY, brilliant green; NRY, light green; ADGRL3, blue. **a** Representative raw experiment shown. Each point represents a technical replicate, and data are represented as mean  $\pm$  standard deviation. **b** Compiled results of N=3 independent biological replicates performed each with 6 technical replicates presented as fold change over empty vector, each point shown represents a technical replicate. Data are presented as mean  $\pm$  standard deviation. one-way ANOVA with Tukey's correction for multiple comparisons was performed to assess statistical significance between CELSR constructs in **b**. \*\*\*\* corresponds to  $p < 0.0001$ . Each construct was compared to each other construct for the statistical testing, but only certain comparisons are shown for clarity. Source data are provided as a Source Data file.



Residue Name	Buried Surface Area (Å <sup>2</sup> )	Contacting residues on opposing domain (Residues <6 Å in distance)
<b>CADH9/GAIN interface</b>		
Q1159	16.6	P2348
G1160	8.1	Y2250
E1162	31.0	Y2250
L1163	99.7	G2187, F2188, R2248, T2249
L1165	31.3	F2188, R2248, Y2250
D1180	40.4	F2188, T2193
D1182	3.7	T2193, R2194
N1183	42.5	F2188, T2193, R2194, E2195
N1184	53.6	R2194, E2195, A2196
R1185	121.9	F2188, R2194, E2195, A2196, T2249, Y2250, L2251, R2252
P1186	32.4	R2252, Y2250
L1187	34.5	T2249, Y2250, L2251, R2252
E1188	16.3	Y2250, R2252
A1189	15.2	Y2250
M1191	7.5	Y2250
G2187	12.5	L1163
F2188	107.2	L1163, S1164, L1165, D1180, L1181, N1183, R1185
T2193	41.7	D1180, D1182, N1183, R1185
R2194	30.2	D1182, N1183, N1184, R1185
E2195	26.8	N1183, N1184, R1185
A2196	31.5	N1184, R1185
R2248	55.7	E1162, L1163, L1165
T2249	24.6	L1163, R1185, L1187
Y2250	110.7	G1160, E1162, L1165, R1185, P1186, L1187, E1188, A1189, M1191
L2251	22.4	R1185, L1187
R2252	73.0	R1185, L1187, E1188
P2348	14.9	Q1159
<b>EGF2/EGF7 interface</b>		
T1371	23.2	W1972
I1379	31.4	W1972, C1978, G1979, P1980
D1380	37.7	V1977, C1978, P1980
Y1383	123.7	V1977, C1978, P1980, W1973, G1974
S1384	7.2	W1973
N1385	16.3	W1973, N1975
W1972	44.1	T1371, I1379
W1973	47.3	Y1383, S1384, N1385
G1974	4.2	Y1383
N1975	4.7	N1385
V1977	57.3	D1380, Y1383
C1978	9.7	I1379, D1380, Y1383
G1979	15.1	I1379, D1380, Y1383
P1980	42.5	I1379, Y1383
<b>LamG1/EGF6 interface</b>		
D1578	41.2	F1918, R1945, Y1947, Y1959
A1580	33.3	P1930, Y1959
V1581	49.4	F1918, A1925, P1930, Y1947
H1584	32.8	K1932
F1585	54.3	A1925, L1928, P1930
Y1588	28.2	K1921, L1928
V1589	76.8	F1918, G1919, K1920, K1921, V1923, L1928
G1590	12.0	G1919, K1920, K1921
N1591	19.7	G1919, K1920
Y1592	34.3	F1918, G1919
F1918	67.0	D1578, V1581, V1589, Y1592
G1919	42.0	V1589, G1590, N1591, Y1592
K1920	28.8	V1589, G1590, N1591
K1921	25.4	Y1588, V1589, G1590
V1923	8.8	V1589
A1925	20.0	V1581, F1585
L1928	71.3	F1585, Y1588, V1589
P1930	38.6	A1580, V1581, F1585
K1932	32.2	H1584
R1945	6.7	D1578
Y1947	25.5	D1578, V1581
Y1959	26.6	D1578, A1580

**Supplementary Table 1: List of interactions in the CELSR1 CADH9-GAIN structure.**

Condition	Simulation Length	Simulation box Dimensions	Total number of atoms (number of water molecules)	Salt Concentration
WT	100 ns	8216.1 nm <sup>3</sup> rhombal dodecahedron with side length 13.9 nm	816681 (270039)	1 mM
$\Delta\text{Ca}^{2+}$	100 ns	8216.1 nm <sup>3</sup> rhombal dodecahedron with side length 13.9 nm	816682 (270039)	1 mM
K533E/T564D (Canonical)	100 ns	8216.1 nm <sup>3</sup> rhombal dodecahedron with side length 13.9 nm	816682 (270042)	1 mM

**Supplementary Table 2: Simulated Systems.**

## Supplementary Methods

### Western blotting

Western blotting was performed as previously described<sup>4,5</sup>. Briefly, HEK293T cells were transfected with 2 micrograms of CELSR constructs using LipoD293 (SignaGen Laboratories SL100668). 48 hours later, cells were washed and resuspended in 500  $\mu$ L of solubilization buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 2 mM MgCl<sub>2</sub> 0.1mM EDTA, 2 mM CaCl<sub>2</sub>, 1% (v/v) Triton X-100). After clarification by high-speed centrifugation, cell lysates were run on SDS-PAGE and transferred to nitrocellulose using a wet transfer system. After blocking, membranes were incubated with either anti-FLAG or anti-HA antibody conjugated to iFluor 488 or iFluor 647, respectively. The next day, membranes were washed and imaged at respective wavelengths using the ChemiDoc imaging system.

### Supplementary References

1. Kordon, S. P. *et al.* Isoform- and ligand-specific modulation of the adhesion GPCR ADGRL3/Latrophilin3 by a synthetic binder. *Nat Commun* **14**, (2023).
2. Kordon, S. P. *et al.* Structural analysis and conformational dynamics of a holo-adhesion GPCR reveal interplay between extracellular and transmembrane domains. *bioRxiv* (2024) doi:10.1101/2024.02.25.581807.
3. Nazarko, O. *et al.* A Comprehensive Mutagenesis Screen of the Adhesion GPCR Latrophilin-1/ADGRL1. *iScience* **3**, 264–278 (2018).
4. Bui, D. L. H. *et al.* The adhesion GPCRs CELSR1-3 and LPHN3 engage G proteins via distinct activation mechanisms. *Cell Rep* **42**, 112552 (2023).
5. Araç, D. *et al.* A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis. *EMBO Journal* **31**, 1364–1378 (2012).
6. Leon, K. *et al.* Structural basis for adhesion G protein-coupled receptor Gpr126 function. *Nat Commun* **11**, (2020).