

---

*Supplementary Data*

## Self-Interaction of Transmembrane Helices Representing Pre-Clusters from the Human Single-Span Membrane Proteins

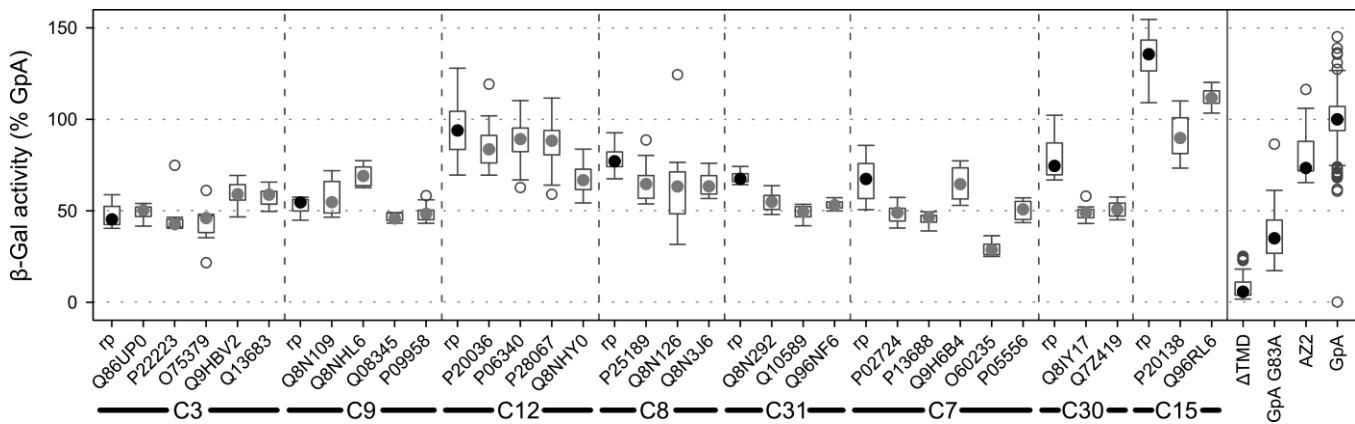
Jan Kirrbach<sup>1</sup>, Miriam Krugliak<sup>2</sup>, Christian L. Ried<sup>1</sup>, Philipp Pagel<sup>3</sup>, Isaiah T. Arkin<sup>2</sup>, and Dieter Langosch<sup>1</sup>

<sup>1</sup>Lehrstuhl für Chemie der Biopolymere, Technische Universität München, Freising, and Munich Center for Integrated Protein Science (CIPS<sup>M</sup>), Munich, Germany

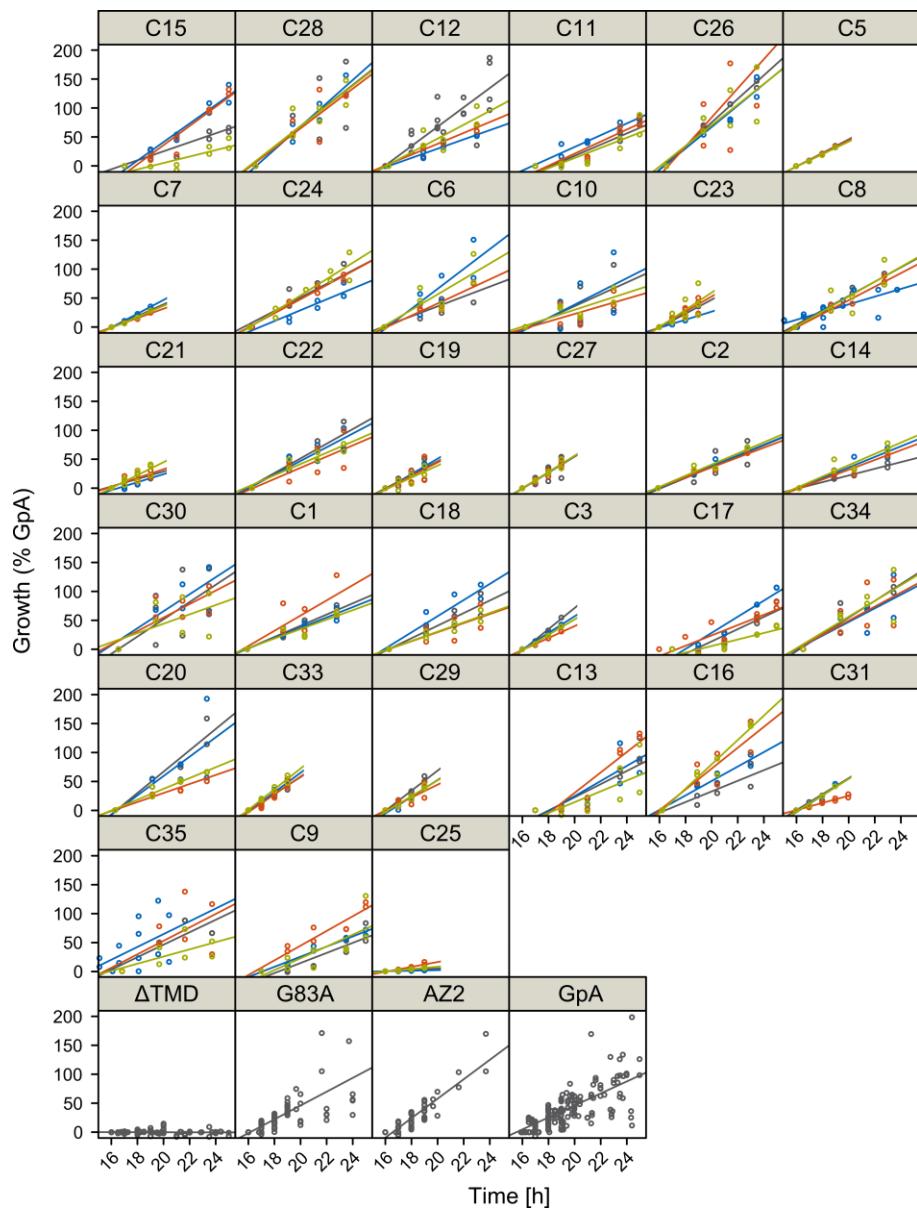
<sup>2</sup>Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel

<sup>3</sup>Lehrstuhl für Genomorientierte Bioinformatik, Technische Universität München, Freising, Germany; Current address: Lip-oFIT Analytic GmbH, Regensburg, Germany

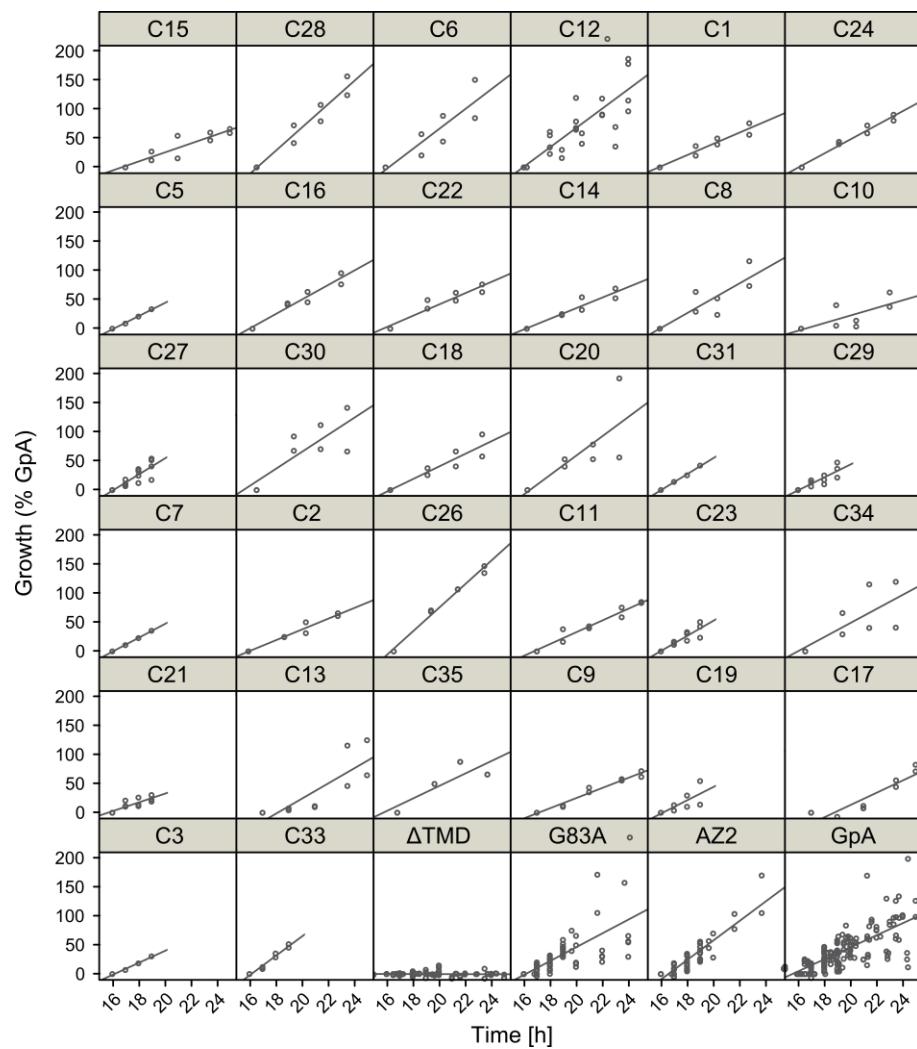
---



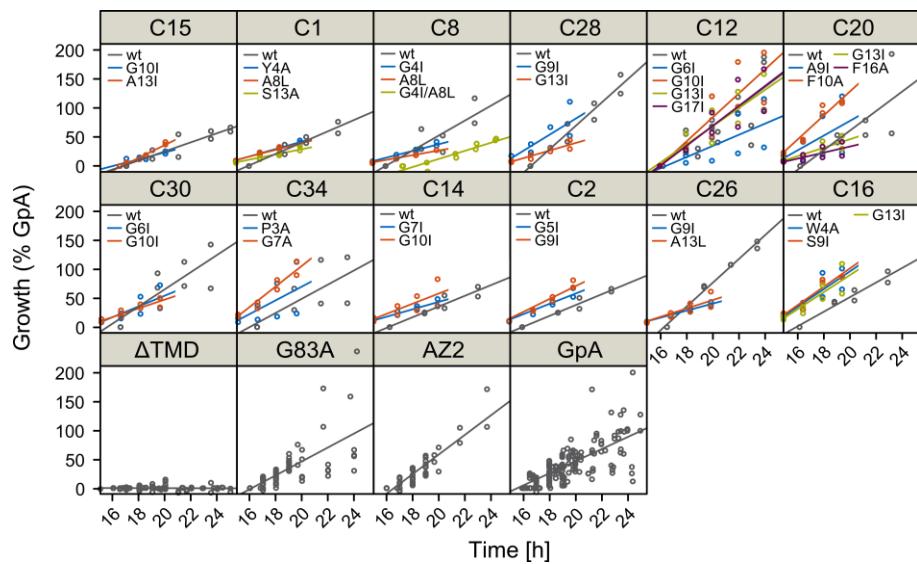
**Fig. S1.** Conservation of self-interaction within exemplary clusters. Data represent relative  $\beta$ -Gal activities (GpA = 100%) as measured with the ToxR system (dot: median, box: interquartile range (IQR), whiskers: upper/lower quartile with max. 1.5 x IQR). Clusters are sorted in descending order of the average conservation of their member's self-interaction. TMDs within each cluster are ordered according to descending similarity to their representative TMD (rp) from left to right. TMDs used for reference are explained in the text. The results of the PD28 assay that controls for membrane insertion are shown in Fig. S5.



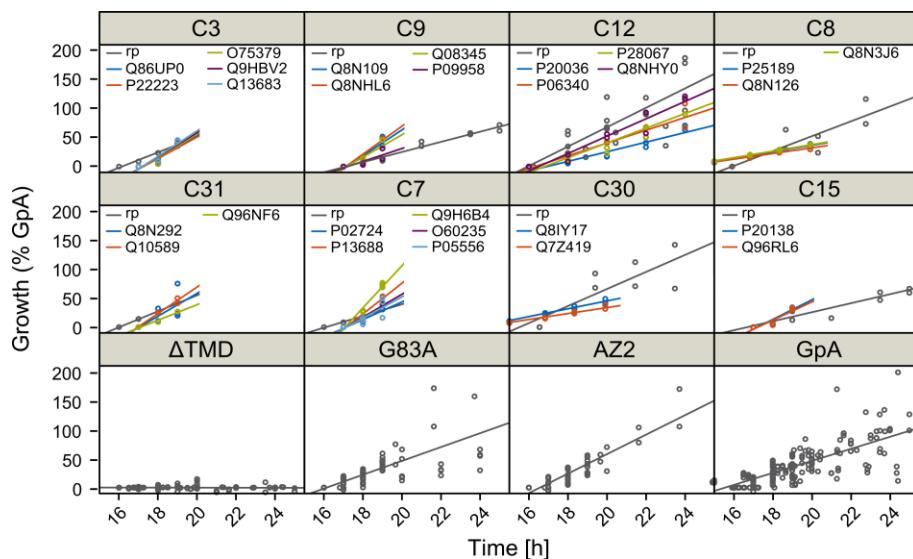
**Fig. S2.** Control for membrane insertion of the ToxR-TMD-MalE fusion proteins tested for self-interaction in Fig 4 A. After incubation for at least 16 h in minimal medium containing 0.4% maltose as sole carbon source, the growth kinetics were obtained by measuring the OD<sub>600</sub> for further 4-8 h and compared to that of the GpA construct (= 100%). For each representative sequence of the top clusters, the four different orientations were measured (-0 grey, -1 blue, -2 orange, -3 green). All clones except C25 were considered to express correctly membrane-integrated ToxR proteins since the slope of their growth curves is at least 50% of GpA.



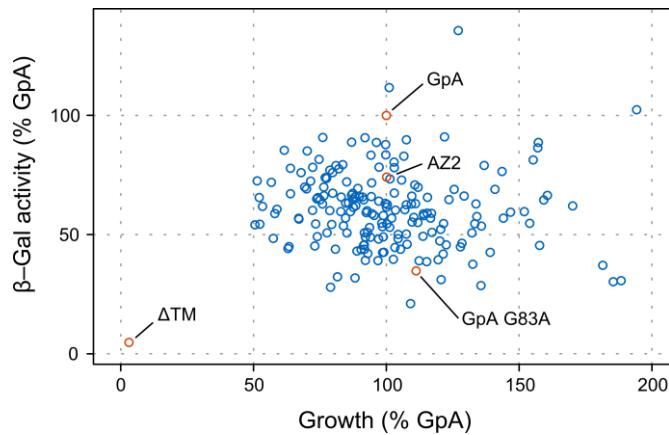
**Fig S3.** Control for membrane insertion of the ToxR-TMD-MalE fusion proteins tested for self-interaction in Fig 4 B. All clones were considered to express correctly membrane-integrated ToxR proteins since the slope of their growth curves is at least 50% of GpA.



**Fig. S4.** Control for membrane insertion of the ToxR-TMD-MalE fusion proteins tested for self-interaction in Fig 5. Each mutated construct was tested. All clones were considered to express correctly membrane-integrated ToxR proteins since the slope of their growth curves is at least 50% of GpA.



**Fig. S5.** Control for membrane insertion of the ToxR-TMD-MalE fusion proteins tested for self-interaction in Fig S1. All clones were considered to express correctly membrane-integrated ToxR proteins since the slope of their growth curves is at least 50% of GpA.



**Fig. S6.** Correlating TMD-TMD interaction to membrane insertion. The measured median  $\beta$ -Gal activities (taken from Figs. 3, 4, and S1) are plotted against the respective PD28 growth kinetics reflecting membrane insertion (taken from Figs. S2 - S5). Note that TMD affinity is not dependent on the slightly different levels of membrane integration observed with the ToxR/TMD/MalE hybrid proteins used in this study.

Table S1. Experimental parameters of TMD-TMD interaction

Cluster	Most prevalent functional annotation <sup>a</sup>	ToxR median <sup>b</sup>	Orientation-dependence <sup>c</sup>	Maximal impact of mutations <sup>d</sup>	Known interaction <sup>e</sup>	Functionally relevant HO <sub>TMD</sub> <sup>g</sup>
C15	Sialic-acid-binding Ig	135.6	0.71	0.79	HO <sub>Ex</sub> [1]	Unknown
C28	Armadillo-repeat containing	<u>102.4</u>	<u>0.71</u>	<u>0.44</u>	Unknown	Unknown
C12	HLA class II $\alpha$ chain	<u>93.9</u>	<u>0.70</u>	<u>0.37</u>	HE <sub>Ex</sub> [2, 3]	Unknown
C6	No prevalent annotation	<u>91.0</u>	0.34	-	-	-
C1	Protocadherin	<u>85.1</u>	0.14	<u>0.63</u>	HO <sub>TMD/Ex</sub> [4]	Yes [4]
C24	Transm. protein 132 family	<u>82.9</u>	0.35	-	Unknown	Unknown
C5	Integrin $\alpha$	<u>81.6</u>	<u>0.41</u>	-	HO <sub>TMD</sub> [5-9], HE <sub>TMD</sub> [5, 9-11]	Controversial [8, 12-14]
C16	Integrin $\alpha$	<u>81.5</u>	0.09	0.09	HO <sub>TMD</sub> [5-9], HE <sub>TMD</sub> [5, 9-11]	Controversial [8, 12-14]
C22	SLIT and NTRK like protein	<u>78.3</u>	0.21	-	Unknown	Unknown
C14	Leukocyte Ig	<u>78.2</u>	0.16	0.16	HO <sub>Ex</sub> [15]	Unknown
C8	Contactin-associated protein-like	<u>77.1</u>	0.27	<u>0.60</u>	HE <sub>Ex</sub> [16, 17]	Unknown
C30	Kin of IRRE-like protein 3	<u>74.5</u>	0.15	0.24	Unknown	Unknown
C10	UDP guanosyltransferase	71.9	0.33	-	HO <sub>Ex</sub> [18]	Unknown
C27	Fibronectin domain containing	70.9	0.19	-	HO <sub>Ex</sub> , HE <sub>Ex</sub>	Unknown
C18	Ion transport regulator	69.4	0.14	-	HE <sub>Ex</sub>	Unknown
C20	HLA class II $\beta$ chain	69.0	0.13	<u>0.34</u>	HE <sub>Ex</sub> [2, 3]	Unknown
C31	No prevalent annotation	67.4	0.08	-	-	-
C29	Desmoglein	67.4	0.12	-	Unknown	Unknown
C7	Integrin $\beta$	67.3	0.36	-	HO <sub>TMD</sub> [5-9], HE <sub>TMD</sub> [5, 9-11]	Controversial [8, 12-14]
C2	HLA class I $\alpha$ chain	67.2	0.19	0.14	HE <sub>Ex</sub> [2, 3]	Unknown
C34	Lysosome associated protein	66.8	0.13	0.18	HE <sub>Ex</sub>	Unknown
C26	VAMP/Synaptobrevin	66.5	<u>0.54</u>	0.11	HO <sub>TMD</sub> [19-21], HE <sub>TMD</sub> [22-24]	Yes [25, 26]
C11	Phospholipase	61.4	<u>0.66</u>	-	HE <sub>Ex</sub>	Unknown
C23	CMRF35-like molecule	60.5	0.30	-	Unknown	Unknown
C13	Syntaxin	56.9	0.11	-	HO <sub>TMD</sub> [22], HE <sub>TMD</sub> [23]	Yes [27]
C21	Nesprin	56.5	0.26	-	HE <sub>Ex</sub>	Unknown
C9	Ig-like receptor	54.6	0.07	-	Unknown	Unknown
C35	Tyrosine-protein phosphatase	54.5	0.08	-	HE <sub>Ex</sub>	Unknown
C19	Leucine-rich repeat containing	47.3	0.21	-	Unknown	Unknown
C17	Neuropilin	45.5	0.13	-	Unknown	Unknown
C3	Cadherin	45.2	0.14	-	HO <sub>TMD/Ex</sub> [28, 29]	Yes [28]
C33	GRAM domain containing	43.8	0.12	-	Unknown	Unknown

<sup>a</sup> Most prevalent function of cluster members as annotated in UniProtKB.<sup>b</sup> The median  $\beta$ -Gal activity measured for the representative TMD in % of GpA. In case of strong self-interaction ( $> AZ2$ ), the value is underlined. Clusters are ordered in descending order of  $\beta$ -Gal activity.<sup>c</sup> A value describing the orientation-dependence of interaction (subtracting the minimal-to-maximal-ratio of median  $\beta$ -Gal activity of the four orientations of the representative TMD from unity yields a value between 0 and 1;  $1 - (\text{activity}_{\min}/\text{activity}_{\max})$ ). Values signifying strong orientation-dependence ( $> 0.4$ ) are underlined.<sup>d</sup> Maximal impact of point mutations of the representative TMD (calculated from the modulus of the mutated-to-wild-type-ratio of median  $\beta$ -Gal activity minus one for the mutation with the most impact on  $\beta$ -Gal activity;  $|(\text{activity}_{\text{mutant}}/\text{activity}_{\text{wild-type}}) - 1|$ ). In case the TMD was denoted as mutation-sensitive ( $> 0.3$ ) the value is underlined. - = not determined.

<sup>e</sup> Homotypic (HO) and/or heterotypic (HE) interaction mediated by either the TMD or extramembranous domains (Ex) as described in the literature.

<sup>f</sup> Functional relevance of homotypic interaction mediated by the TMD as described in the literature.

Table S2. Selected clusters for the comparison of relative affinities within clusters.

Cluster	Conserved motif <sup>a</sup>	Functional diversity <sup>b</sup>	ToxR median <sup>c</sup>	Orientation-dependence <sup>d</sup>	Members <sup>e</sup>	N <sup>f</sup>	Average variance <sup>g</sup>
C3	GxxxxA	16	45.2	0.14	19	5	7.1
C9	-	18	54.6	0.07	11	4	7.4
C12	GxxxxG	22	93.9	0.70	9	4	12.0
C8	GxxxxA	55	77.1	0.27	11	3	13.3
C31	SxxxxG	80	67.4	0.08	5	3	15.1
C7	GxxxxG	83	67.3	0.36	12	5	19.5
C30	GxxxxG	60	74.5	0.15	5	2	24.7
C15	-	13	135.6	0.71	8	2	34.9

<sup>a</sup> Putative interaction motif conserved in the TMD alignment.

<sup>b</sup> The percentage of proteins in the cluster which differ from the main functional annotation in UniProtKB.

<sup>c</sup> The median β-Gal activity measured for the representative TMD in % of GpA.

<sup>d</sup> A value describing the orientation-dependence of interaction (subtracting the minimal-to-maximal-ratio of median β-Gal activity of the four orientations of the representative TMD from unity yields a value between 0 and 1; 1-(activity<sub>min</sub>/activity<sub>max</sub>)).

<sup>e</sup> The number of unique TMDs in the cluster.

<sup>f</sup> The number of TMDs whose relative affinity was compared to the representative TMD of the cluster.

<sup>g</sup> Average variance of median β-Gal activity of TMDs measured for this cluster as percentage of the median of β-Gal activity of the representative TMD.

Table S3. Homologous proteins within top clusters

Cluster	Most representative <sup>a</sup>	Homologs <sup>a</sup>
C1	Q9UN71	Q9Y5H1; Q9Y5I1; Q9Y5H8; Q9Y5F3; Q9Y5H9; Q9UN66; Q96TA0; Q9Y5G0; Q9Y5E1; Q9Y5H2; Q9Y5G2; Q9Y5E8; Q9Y5I3; Q9Y5G1; Q9Y5E2; Q9Y5G3; Q9Y5G6; Q9UN75; O60330; Q9Y5G8; Q9Y5H0; Q9Y5G9; Q9Y5H5; Q9Y5H4; Q9H158; Q9Y5I4; Q9Y5G4; Q9Y5E7
C2	P01892	Q29865; P30504; P01893; P18462; P13747; Q09160; P30443; Q31612; Q9TNN7; P30499; P28827; P10321; P17693; P10314; P30511; P01889; P30453; Q15223; P30464; P30512; P30485
C3	Q9ULB5	Q13634; P55283; P55285; P12830; P22223; P33151; Q9HBV2; Q9Y6N8; Q9ULB4; Q9UJ99; P55291; O75379; P55287; P55289; Q86UP0; P19022; Q13683; Q9HBT6
C5	P78310	P29323; Q13349; P20702; Q9BQS2-2; O75578; P20701; P17301; Q9UIW2; P11215; P51805; Q9UKX5; P38570; O60939; O60830
C6	Q6UWB1	Q96A28; Q07954; Q14761; Q9BUF7; Q5VSK2; O14638; O43557; P00533; P58335; Q9BY79; Q8N423; P50281; O75325; Q9HC10
C7	Q8N967	O60235; Q9BYF1; Q9H6B4; P13688; Q8IVY1; P26010; P02724; O60245; P05556; P15144; Q08174
C8	Q9BZ76	Q96NU0; P25189; Q9UHC6; Q8WYK1; Q8N3J6; Q8N126; Q8NFZ8; Q8IWT1; Q9C0A0; Q9BY67
C9	P43629	O75022; Q8NHJ6; P43628; Q08345; Q6GTX8; Q8NHL6; O75023; P43630; Q8N109; P09958
C10	O75310	Q6UWM9; O75795; P06133; Q9HAW8; P16662; P54855; Q9BY64; Q9Y4X1; P36537
C11	Q6ZV29	Q9HD43; Q6ZV29-2; Q9NWW9; Q86YL7; Q9HBB8; P40198; P40198-3; Q99795; Q86YL7-2
C12	P01908	Q8NHY0; P28067; P01907; P06340; P20036; Q8NHY0-2; P01903; P01909
C13	Q9H1U4	Q19T08; O43451; O14662; P32856; Q9UBN6; Q13277; P16234
C14	Q8IYS5-2	P59901; A6NI73; Q8N149; O75019; O76036; P24071; Q6PI73
C15	Q9Y286	Q9NYZ4; Q96RL6; O15389; Q96LC7; P20138; Q96PQ1; Q9Y336
C16	P56199	P08648; P23229; P06756; Q58EX2; P53708; P08514; P26006
C17	Q8NC67-2	Q8NC67; Q8TDF5; Q9P0K1; P57087; Q30201; P32004; Q9H3R2
C18	P54710	Q14802-2; O00168; Q9H0Q3; P58549; Q14118; Q96DB9; Q14802
C19	P34810	Q96QE4; A6NMS7; A6NN04; Q5BVD1; Q147U4; O60309
C20	P13765	P04229; P13762; P01919; P04440; P01918
C21	Q8NF91	Q6ZMZ3; Q8WXH0; Q6P9G4; Q8WXH0-4
C22	Q8IW52	Q9H5Y7; Q96PX8; O94991; Q9H156; O94933
C23	Q9UGN4	Q13444; Q16671; A2A2V5; Q8TDQ1; Q9UBQ6
C24	Q14DG7	Q14C87; Q24JP5; Q8N3T6; Q6IEE7
C25	Q14954	Q8N743; Q14943; Q14952; P43626
C26	P23763-3	P23763-2; P23763; Q9HCL0; P63027
C27	Q6PJG9	Q96NI6; Q9ULH4; Q9P244; Q9BTN0
C28	Q7L4S7	Q9UH62; Q8N2F6; Q7L311; Q5H9R4
C29	Q14126	Q86SJ6; P16581; P32926; Q02413
C30	Q8IZU9	Q8IZU9-2; Q7Z419; Q8IY17; Q8IY17-2
C31	Q8IUN9	Q96NF6; Q8N292; P07306; Q10589
C33	P32856-2	Q3KR37-3; Q3KR37; Q8IYS0; Q96CP6-2
C34	Q6UXC1	P13473; Q6V0I7; P11279; P23471
C35	Q15262	Q92729; P52799; P55285-2; O14522

<sup>a</sup> UniProt KB identifiers

## REFERENCES

1. Cornish, A.L., et al., Characterization of siglec-5, a novel glycoprotein expressed on myeloid cells related to CD33. *Blood*, 1998. **92**(6): p. 2123-2132.
2. Germain, R.N., Binding domain regulation of MHC class II molecule assembly, trafficking, fate, and function. *Semin Immunol*, 1995. **7**(6): p. 361-372.
3. Schafer, P.H., S.K. Pierce, and T.S. Jardetzky, The structure of MHC class II: a role for dimer of dimers. *Semin Immunol*, 1995. **7**(6): p. 389-398.
4. Chen, X., et al., Structural elements necessary for oligomerization, trafficking, and cell sorting function of paraxial protocadherin. *J Biol Chem*, 2007. **282**(44): p. 32128-32137.
5. Berger, B.W., et al., Consensus motif for integrin transmembrane helix association. *Proc Natl Acad Sci U S A*, 2010. **107**(2): p. 703-708.
6. Li, R., et al., Oligomerization of the integrin alphaIIbbeta3: roles of the transmembrane and cytoplasmic domains. *Proc Natl Acad Sci U S A*, 2001. **98**(22): p. 12462-12467.
7. Li, R., et al., Dimerization of the transmembrane domain of Integrin alphaiIbb subunit in cell membranes. *J Biol Chem*, 2004. **279**(25): p. 26666-26673.
8. Li, R., et al., Activation of integrin alphaIIbbeta3 by modulation of transmembrane helix associations. *Science*, 2003. **300**(5620): p. 795-798.
9. Schneider, D. and D.M. Engelman, GALLEX, a measurement of heterologous association of transmembrane helices in a biological membrane. *J Biol Chem*, 2003. **278**(5): p. 3105-3111.
10. Kim, C., et al., Interactions of platelet integrin alphaIIbb and beta3 transmembrane domains in mammalian cell membranes and their role in integrin activation. *Blood*, 2009. **113**(19): p. 4747-4753.
11. Lau, T.-L., et al., The structure of the integrin alphaIIbbeta3 transmembrane complex explains integrin transmembrane signalling. *EMBO J*, 2009. **28**(9): p. 1351-1361.
12. Parthasarathy, K., et al., Transmembrane helices that form two opposite homodimeric interactions: an asparagine scan study of alphaM and beta2 integrins. *Protein Sci*, 2008. **17**(5): p. 930-938.
13. Schneider, D. and D.M. Engelman, Involvement of transmembrane domain interactions in signal transduction by alpha/beta integrins. *J Biol Chem*, 2004. **279**(11): p. 9840-9846.
14. Wang, W., et al., Tests of integrin transmembrane domain homo-oligomerization during integrin ligand binding and signaling. *J Biol Chem*, 2011. **286**(3): p. 1860-1867.
15. Chen, Y., et al., Crystal structure of myeloid cell activating receptor leukocyte Ig-like receptor A2 (LILRA2/ILT1/LIR-7) domain swapped dimer: molecular basis for its non-binding to MHC complexes. *J Mol Biol*, 2009. **386**(3): p. 841-853.
16. Krämer, E.M., et al., Compartmentation of Fyn kinase with glycosylphosphatidylinositol-anchored molecules in oligodendrocytes facilitates kinase activation during myelination. *J Biol Chem*, 1999. **274**(41): p. 29042-29049.
17. Peles, E., et al., Identification of a novel contactin-associated transmembrane receptor with multiple domains implicated in protein-protein interactions. *EMBO J*, 1997. **16**(5): p. 978-988.
18. Lewis, B.C., P.I. Mackenzie, and J.O. Miners, Homodimerization of UDP-glucuronosyltransferase 2B7 (UGT2B7) and identification of a putative dimerization domain by protein homology modeling. *Biochem Pharmacol*, 2011. **82**(12): p. 2016-2023.
19. Fleming, K.G. and D.M. Engelman, Computation and mutagenesis suggest a right-handed structure for the synaptobrevin transmembrane dimer. *Proteins*, 2001. **45**(4): p. 313-317.
20. Laage, R. and D. Langosch, Dimerization of the synaptic vesicle protein synaptobrevin (vesicle-associated membrane protein) II depends on specific residues within the transmembrane segment. *Eur J Biochem*, 1997. **249**(2): p. 540-546.
21. Roy, R., R. Laage, and D. Langosch, Synaptobrevin transmembrane domain dimerization-revisited. *Biochemistry*, 2004. **43**(17): p. 4964-4970.
22. Laage, R., et al., A conserved membrane-spanning amino acid motif drives homomeric and supports heteromeric assembly of presynaptic SNARE proteins. *J Biol Chem*, 2000. **275**(23): p. 17481-17487.
23. Stein, A., et al., Helical extension of the neuronal SNARE complex into the membrane. *Nature*, 2009. **460**(7254): p. 525-528.
24. Washbourne, P., G. Schiavo, and C. Montecucco, Vesicle-associated membrane protein-2 (synaptobrevin-2) forms a complex with synaptophysin. *Biochem J*, 1995. **305** ( Pt 3): p. 721-724.
25. Kweon, D.-H., C.S. Kim, and Y.-K. Shin, Regulation of neuronal SNARE assembly by the membrane. *Nat Struct Biol*, 2003. **10**(6): p. 440-447.
26. Tong, J., et al., A scissors mechanism for stimulation of SNARE-mediated lipid mixing by cholesterol. *Proc Natl Acad Sci U S A*, 2009. **106**(13): p. 5141-5146.
27. Hofmann, M.W., et al., Self-interaction of a SNARE transmembrane domain promotes the hemifusion-to-fusion transition. *J Mol Biol*, 2006. **364**(5): p. 1048-1060.
28. Huber, O., R. Kemler, and D. Langosch, Mutations affecting transmembrane segment interactions impair adhesiveness of E-cadherin. *J Cell Sci*, 1999. **112** ( Pt 23): p. 4415-4423.
29. Iino, R., I. Koyama, and A. Kusumi, Single molecule imaging of green fluorescent proteins in living cells: E-cadherin forms oligomers on the free cell surface. *Biophys J*, 2001. **80**(6): p. 2667-2677.