Dear Patrik.

Hope you are doing well. It was great to discuss with you and I am curious what you will discover. I don’t want to spam you with to much information. But please, don’t hesitate to contact me if you have any questions or if you like to get more information.

I very much appreciate that all the information is between us and will not be shared with others.

Here is how I started: (i) I downloaded the sequence of validated proteins from UNIPROT; (ii) predicted the transmembrane spans; (iii) cut out the transmembrane spans (TMSs) and left 10 amino acids before and after the TMS (if possible), as I thought they might be important for headgroup recognition; (iv) alligned the prolonged TMSs (see below); (v) removed those that did not align well.

As discussed, we need to distinguish the orientation of the TMSs (type I or a type II), which I didn’t do with the transmembrane spans below. For a multi-transmembrane spanning protein, their TMSs would than be assigned to one out of two clusters, depending on their orientation. To simplify things, it would be best to start with single TMS proteins.

I attached one Excel file, with the information spread over different sheets. There, you will find:

1. A list of all the 328 sphingolipid-binding proteins (all SLBP). These are proteins, which were crosslinked to photoactivatable sphingolipids, e.g. sphingosine (Sph), ceramide (Cer), sphingomyelin (SM), glucosylceraide (GlcCer).
2. Subsets of the „all SLBP“ list. Proteins are clustered according whether they showed a reduced („negative“) or increased („positive“) crosslink efficancy.

My simple thinking:

* all SLBP: contains proteins interacting with sphingosine (Sph), ceramide (Cer), glucosylceramide (GlcCer) and sphingomyelin (SM)
* FB1-negative: interactions with Cer, SM and GlcCer
* FB1-positive: interactions with Sph
* PPMP-negative & PDMP-negative: interactions with GlcCer
* PPMP-positive & PDMP-positive: maybe interactions with Cer or SM

I would start with the FB1-negative list, as this list contains various „p24-relaxed motif“-containing proteins and multiple proteins that I validated (e.g. GPAA1, TF, HS2ST, JAM3, SSRA, IMPA3).

Let me know what you think.

All the best, Per

My first try to look for a motif:

SSRA

MRLLPRLLLLLLLVFPATVLFRGGTRGLLAVAQDLTEDEETVEDSIIEDEDDEAEVEEDEPTDLVEDKEEEDVSGEPEASPSADTTILFVKGEDFPANNIVKFLVGFTNKGTEDFIVESLDASFRYPQDYQFYIQNFTALPLNTVVPPQRQATFEYSFIPAEPMGGRPFGLVINLNYKDLNGNVFQDAVFNQTVTVIEREDGLDGETIFMYMFLAGLGLLVIVGLHQLLESRKRKRPIQKVEMGTSSQNDVDMSWIPQETLNQINKASPRRLPRKRAQKRSVGSDEAANDILDYKDDDDKV\*

MRLLPRLLLLLLLVFPATVLFRGGTRGLLAV

EREDGLDGETIFMYMFLAGLGLLVIVGLHQLLESRKRKRPIQ

GPAA1

MGLLSDPVRRRALARLVLRLNAPLCVLSYVAGIAWFLALVFPPLTQRTYMSENAMGSTMVEEQFAGGDRARAFARDFAAHRKKSGALPVAWLERTMRSVGLEVYTQSFSRKLPFPDETHERYMVSGTNVYGILRAPRAASTESLVLTVPCGSDSTNSQAVGLLLALAAHFRGQIYWAKDIVFLVTEHDLLGTEAWLEAYHDVNVTGMQSSPLQGRAGAIQAAVALELSSDVVTSLDVAVEGLNGQLPNLDLLNLFQTFCQKGGLLCTLQGKLQPEDWTSLDGPLQGLQTLLLMVLRQASGRPHGSHGLFLRYRVEALTLRGINSFRQYKYDLVAVGKALEGMFRKLNHLLERLHQSFFLYLLPGLSRFVSIGLYMPAVGFLLLVLGLKALELWMQLHEAGMGLEEPGGAPGPSVPLPPSQGVGLASLVAPLLISQAMGLALYVLPVLGQHVATQHFPVAEAEAVVLTLLAIYAAGLALPHNTHRVVSTQAPDRGWMALKLVALIYLALQLGCIALTNFSLGFLLATTMVPTAALAKPHGPRTLYAALLVLTSPAATLLGSLFLWRELQEAPLSLAEGWQLFLAALAQGVLEHHTYGALLFPLLSLGLYPCWLLFWNVLFWK

HS2ST

MGLLRIMMPPKLQLLAVVAFAVAMLFLENQIQKLEESRSKLERAIARHEVREIEQRHTMDGPRQDATLDEEEDMVIIYNRVPKTASTSFTNIAYDLCAKNKYHVLHINTTKNNPVMSLQDQVRFVKNITSWKEMKPGFYHGHVSYLDFAKFGVKKKPIYINVIRDPIERLVSYYYFLRFGDDYRPGLRRRKQGDKKTFDECVAEGGSDCAPEKLWLQIPFFCGHSSECWNVGSRWAMDQAKYNLINEYFLVGVTEELEDFIMLLEAALPRFFRGATELYRTGKKSHLRKTTEKKLPTKQTIAKLQQSDIWKMENEFYEFALEQFQFIRAHAVREKDGDLYILAQNFFYEKIYPKSN

GLLRIMMPPKLQLLAVVAFAVAMLFLENQIQKLEESRS

JAM3

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTSDPRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLKIWNVTRRDSALYRCEVVARNDRKEIDEIVIELTVQVKPVTPVCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRANPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVYDLNIGGIIGGVLVVLAVLALITLGICCAYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVIAANDILDYKDDDDKV

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQ

QEMEVYDLNIGGIIGGVLVVLAVLALITLGICCAYRRGYFINN

P24

MVTLAELLVLLAALLATVSGYFVSIDAHAEECFFERVTSGTKMGLIFEVAEGGFLDIDVEITGPDNKGIYKGDRESSGKYTFAAHMDGTYKFCFSNRMSTMTPKIVMFTIDIGEAPKGQDMETEAHQNKLEEMINELAVAMTAVKHEQEYMEVRERIHRAINDNTNSRVVLWSFFEALVLVAMTLGQIYYLKRFFEVRRVV

MVTLAELLVLLAALLATVSGYFVSIDAHAEECF

Contreras

V V L W S F F E A L V L V A M T L G Q I Y

TM1 ---------MRLLPRLLLLLLLVFPATVLFRGGTR--- 26 SSRA

TM3 ----------MMPPKLQLLAVVAFAVAMLFLENQIQKL 28 HS2ST

TM6 ---------VVLWSFFEALVLVAMTLGQIY-------- 21 p24

TM5 YDLNIGGIIGGVLVVLAVLALITLGICCAYRRG----- 33 JAM3

TM4 --PPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNR—34 JAM3

TM2 ----LDGETIFMYMFLAGLGLLVIVGLHQLLESRKR--- 32 SSRA

: : \* .