SLGIAR LABORATORY MEMBRANE SCAFFOLD PROTEINS
PROTEIN AND NUCLEIC ACID SEQUENCES
File updated March 1, 2008

In this document we compile the protein and nucleic acid sequences for some of the various membrane scaffold protein (MSP) constructs we have made. Note that currently we primarily use two versions, MSP1D1 (earlier called MSP1T2) and MSP1E3D1 and are the only ones for which we have stockpiled protein. The original MSP1 series generates the ~9.7 nm diameter Nanodiscs and had a Factor X protease site. However Factor X also cleaves MSP non-specifically which makes it difficult to remove the His-tag easily. The current versions of scaffold proteins have a TEV protease site (which was engineered by us to give complete and specific cleavage) and also has an additional Δ(1-11) truncation from MSP1 to give more stable discs, since we showed that not all of the first helix was involved in lipid binding (JACS *126*, 3477). To make larger discs we most often use the MSP1E3D1 construct which has an additional three helices (Helix 4, 5, and 6 – see below) inserted into the sequence between Helix 3 and Helix 4 which generates the ~12.1 nm bilayer discs. For our most recent constructs, we have used gene synthesis from Integrated DNA Technologies (http://www.idtdna.com/). Note that all sequences have been published, often in the supplementary material. See for instance JACS *126*, 3477 and NanoLetters *2*, 853. Our laboratory web site provides a lists all our Nanodisc and Membrane Scaffold Protein related publications.

MSP1D1 (MSP1T2):

Protein:

GHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQEM
SKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHELQE
KLSPLGEEMRDRARAHVDALRTHLAPYSDELRQRLAARLEALKENGGARLAEY
HAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQ

DNA:

ACCATGGGTCATCATCATCATCATCATCACGATTATGATATTCCTACTACTGA
GAATTTGTATTTTCAGGGTTCTACCTTCAGTAAACTTCGCGAACAACTGGGCC
CCGTGACGCAGGAATTCTGGGACAACCTGGAAAAAGAAACCGAGGGACTGC
GTCAGGAAATGTCCAAAGATTTAGAAGAGGTGAAGGCCAAGGTTCAGCCATA
TCTCGATGACTTTCAGAAAAAATGGCAGGAAGAGATGGAATTATATCGTCAA
AAGGTGGAACCGCTGCGTGCGGAACTGCAAGAGGGGGCACGCCAAAAACTC
CATGAGCTCCAAGAGAAGCTCAGCCCATTAGGCGAAGAAATGCGCGATCGCG
CCCGTGCACATGTTGATGCACTCCGGACTCATTTGGCGCCGTATTCGGATGAA

CTTCGCCAGCGTTTGGCCGCACGTCTCGAGGCGCTGAAAGAAAACGGGGGTG
CCCGCTTGGCTGAGTACCACGCGAAAGCgACAGAACACCTGAGCACCTTGAG
CGAAAAAGCGAAACCGGCGCTGGAAGATCTACGCCAGGGCTTATTGCCTGTT
CTTGAGAGCTTTAAAGTCAGTTTTCTGTCAGCTCTGGAAGAATATACTAAAAA
GCTGAATACCCAGTAAGCTT

MSP1E3D1:

Protein:

MGHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEKETEGLRQE
MSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLRAELQEGARQKLHEL
QEKLSPLGEEMRDRARAHVDALRTHLAPYLDDFQKKWQEEMELYRQKVEPLRA
ELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDELRQRLAAR
LEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFLSA
LEEYTKKLNTQ

DNA:

ACCATGGGTCATCATCATCATcATCATCACGATTATGATATTCCTACTACTGA
GAATTTGTATTTTCAGGGTTCTACCTTCAGTAAACTTCGCGAACAACTGGGCC
CCGTGACGCAGGAATTCTGGGACAACCTGGAAAAAGAAACCGAGGGACTGC
GTCAGGAAATGTCCAAAGATTTAGAAGAGGTGAAGGCCAAGGTTCAGCCATA
TCTCGATGACTTTCAGAAAAAATGGCAGGAAGAGATGGAATTATATCGTCAA
AAGGTGGAACCGCTGCGTGCGGAACTGCAAGAGGGGGCACGCCAAAAACTC
CATGAGCTCCAAGAGAAGCTCAGCCCATTAGGCGAAGAAATGCGCGATCGCG
CCCGTGCACATGTTGATGCACTCCGGACTCATTTGGCGCCATATCTCGATGAC
TTTCAGAAAAAATGGCAGGAAGAGATGGAATTATATCGTCAAAAGGTGGAA
CCGCTGCGTGCGGAACTGCAAGAGGGGGCACGCCAAAAACTCCATGAGCTCC
AAGAGAAGCTCAGCCCATTAGGCGAAGAAATGCGCGATCGCGCCCGTGCAC
ATGTTGATGCACTCCGGACTCATTTGGCGCCGTATTCGGATGAACTTCGCCAG
CGTTTGGCCGCACGTCTCGAGGCGCTGAAAGAAAACGGGGGTGCCCGCTTGG
CTGAGTACCACGCGAAAGCGACAGAACACCTGAGCACCTTGAGCGAAAAAG
CGAAACCGGCGCTGGAAGATCTACGCCAGGGCTTATTGCCTGTTCTTGAGAG
CTTTAAAGTCAGTTTTCTGTCAGCTCTGGAAGAATATACTAAAAAGCTGAATA
CCCAGTAAGCTT

Human Apo-AI:
For reference we are including the sequence for Human ApoA1 using the assignment of helices from C. G. Brouillette *et al.*, BBA *1531* (2001) 4-46.

```
1_____________________________________________43
DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLN

44------Helix 1----------!66---------Helix 2----------!
LKLLDNWDSVTSTFSKLREQLGPVTQEFWDNLEKETEGLRQEMS

88---Helix 3---!99------Helix 4-------------!
KDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVE

121-------Helix 5------------!143-----Helix 6--------------!
PLRAELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLA

165-------Helix 7------------!188-----Helix 8--------------!
PYSDELRQRLAARLEALKENGGARLAEYHAKATEHLSTLSEKAK

209--Helix 9---!220--------Helix 10------------243!
PALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQ
```

MODULAR PICTURE OF MEMBRANE SCAFFOLD PROTEINS
It may be useful to have a modular picture of the construction of various membrane scaffold proteins. To that end, consider the following schematic:

Building blocks:

GLOB	DEPPQSPWDRVKDLATVYVDVLKDSGRDYVSQFEGSALGKQLN
HisX	MGHHHHHHIEGR
HisTEV	MGHHHHHHHDYDIPTTENLYFQG
Helix 1 (H1):	LKLLDNWDSVTSTFSKLREQLG
Helix 2 (H2):	PVTQEFWDNLEKETEGLRQEMS
Helix 3 (H3):	KDLEEVKAKVQ
Helix 4 (H4):	PYLDDFQKKWQEEMELYRQKVE
Helix 5 (H5):	PLRAELQEGARQKLHELQEKLS
Helix 6 (H6):	PLGEEMRDRARAHVDALRTHLA

Helix 7 (H7): PYSDELRQRLAARLEALKENGG

Helix 8 (H8): ARLAEYHAKATEHLSTLSEKAK

Helix 9 (H9): PALEDLRQGLL

Helix 10(H10): PVLESFKVSFLSALEEYTKKLNTQ

Truncated Helices:

Helix 0.5 (H0.5): STFSKLREQLG

Helix 10.5(H10.5): SALEEYTKKLNTQ

Helix 2S (H2): PVTQEFWDNLEKETEGLRQEMS

Then:

Apo A-I GLOB-H1-H2-H3-H4-H5-H6-H7-H8-H9-H10

MSP1D1 HisTev-H0.5 -H2-H3-H4-H5-H6-H7-H8-H9-H10

MSP1E3D1: HisTev-H0.5-H2-H3-H4-H5-H6-H4-H5-H6-H7-H8-H9-H10

Our roughly 25 other constructs can be described easily in these terms.