PNAS

www.pnas.org

Supplementary Information for

Efficient integration of transmembrane domains depends on the folding properties of the upstream sequences

Marco Janoschke, Mirjam Zimmermann, Anna Brunauer, Raffael Humbel, Tina Junne, and Martin Spiess

Corresponding author: Martin Spiess
Email: martin.spiess@unibas.ch

This PDF file includes:

Tables S1 to S2

Table S1. Amino acid sequences of the model proteins analyzed.

```
ST-DP128:
MEGGEEEVERIPDELFDTKKKHLLDKLIRVGIILVLLIWGTVLLLKSIPHHSNTPDYQEPNSNYTNDGKL
KVSFSVVRNNTFQPKYHELQWISDNKIESNDLGLYVTFMNDSYVVKSVYDDSYNSVLLEGKTFIHNGQNL
TVESITASPDLKRLLIRTNSVQNWRHSTFGGPGAAAALALAALAALALAAAAGPGGWWEITHNTLFIPAN
ETFDRPHNGYVDILPIGGYNHLAYFENSNSSHYKTLTEGKWEVVNGPLAFDSMENRLYFISTRKSSTERH
VYYIDLRSPNEIIEVTDTSEDGVYDVSFSSGRRFGLLTYKGPKVPYQKIVDFHSRKAEKCDKGNVLGKSL
YHLEKNEVLTKILEDYAVPRKSFRELNLGKDEFGKDILVNSYEILPNDFDETLSDHYPVFFFAYGGPNSQ
QVVKTFSVGFNEVVASQLNAIVVVVDGRGTGFKGQDFRSLVRDRLGDYEARDQISAASLYGSLTFVDPQK
ISLFGWSYGGYLTLKTLEKDGGRHFKYGMSVAPVTDWRFYDSVYTERYMHTPQENFDGYVESSVHNVTAL
AQANRFLLMHGTGDDNVHFQNSLKFLDLLDLNGVENYDVHVFPDSDHSIRYHNANVIVFDKLLDWAKRAF
DGQFVKACYPYDVPDYAGYPYDVPDYAYPYDVPDYA●
```

```
RI-DP128:
MEGGEEEVERIPDELFDTKKKHLLDKLIRVGIILVLLIWGTVLLLKSIPHHSNTPDYQEPNSNYTNDGKL
KVSFSVVRNNTFQPKYHELQWISDNKIESNDLGLYVTFMNDSYVVKSVYDDSYNSVLLEGKTFIHNGQNL
TVESITRSMEGGEEEVERIPDELFDTKKKHLLDKLIRVGIILVLLIWGTVLLLKSIPHHSNTPDYQEPNS
NYTNDGKLKVSFSVVRNNTFQPKYHELQWISDNKIESNDLGLYVTFMNDSYVVKSVYDDSYNSVLLEGKT
FIHNGQNLTVESITASPDLKRLLIRTNSVQNWRHSTFGGPGAAAALALAALAALALAAAAGPGGWWEITH
NTLFIPANETFDRPHNGYVDILPIGGYNHLAYFENSNSSHYKTLTEGKWEVVNGPLAFDSMENRLYFIST
RKSSTERHVYYIDLRSPNEIIEVTDTSEDGVYDVSFSSGRRFGLLTYKGPKVPYQKIVDFHSRKAEKCDK
GNVLGKSLYHLEKNEVLTKILEDYAVPRKSFRELNLGKDEFGKDILVNSYEILPNDFDETLSDHYPVFFF
AYGGPNSQQVVKTFSVGFNEVVASQLNAIVVVVDGRGTGFKGQDFRSLVRDRLGDYEARDQISAASLYGS
LTFVDPQKISLFGWSYGGYLTLKTLEKDGGRHFKYGMSVAPVTDWRFYDSVYTERYMHTPQENFDGYVES
SVHNVTALAQANRFLLMHGTGDDNVHFQNSLKFLDLLDLNGVENYDVHVFPDSDHSIRYHNANVIVFDKL
LDWAKRAFDGQFVKACYPYDVPDYAGYPYDVPDYAYPYDVPDYA●
```

The sequences are color-coded to indicate their origin: black, DPAPB; purple, H-segment (as an example L5 is shown); gray, C-terminal triple HA-tag and linker residues. The function of transmembrane segments as signal-anchor, stop-transfer, or re-integration sequences is indicated by a red, blue, or green background, respectively. The relevant loop sequences preceding the H-segments are underlined. Potential glycosylation sites are shown in red.

Table S2. Amino acid sequences of the luminal spacer preceding the stop-transfer H-segments.

ST-DP50: KSIPHHSNTPDYQEPNS**NYT**NDGKLKVLKRLLIRTNSVQNWRHSTFGGPG
ST-Spwt161: KSIPHHSNTPDYQEPNS**NYT**NDGKLKVSFSVVR**NNT**SSSSKL**NES**HRLHQFFRDMDDEESWIKEKKLLVS SEDYGRDLTGVQNLRKKHKRLEAELAAHEPAIQGVLDTGKKLSDDNTIGKEEIQQRLAQFVDHWKELKQL AAARGQRLEEHRLGSTSGGPG
ST-Spmut161: KSIPHHSNTPDYQEPNS**NYT**NDGKLKVSFSVVR**NNT**SSSSKL**NES**HRLHQFFRDMDDEESWIKEKKLLVS SEDYGRDLTGVQNLRKKHKRGPGPGPGPGPAIQGVLDTGKKLSDDNTIGKEEIQQRLAQFVDHWKELKQL AAARGQRLEEHRLGSTSGGPG
ST-GS144: KSIPHHSNTPDYQEPNS**NYT**NDGKLKVSFSVVR**NNT**STQESGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGS QESGSGSGSGSGSGSGSGSGSGSGSGSSGSGSGSGSQESGSGSGSGSGSGSGSGSGSGSGSGSGSSGSGSGSGSQESGSSGGPG
ST-Chp144: KSIPHHSNTPDYQEPNS**NYT**NDGKLKVSFSVVR**NNT**STQESGKSAHWDFAWPWASRSGDSAFWGLWPWEA QESGKSAHWDFAWPWASRSGDSAFWGLWPWEAQESGKSAHWDFAWPWASRSGDSAFWGLWPWEAQESGSSGGPG
ST-Scr144: KSIPHHSNTPDYQEPNS**NYT**NDGKLKVSFSVVR**NNT**STQESGWKWSAHSWRFAPALSWSDGAFEWAPDWG QESGWKWSAHSWRFAPALSWSDGAFEWAPDWGQESGWKWSAHSWRFAPALSWSDGAFEWAPDWGQESGSSGGPG
ST-Non144: KSIPHHSNTPDYQEPNS**NYT**NDGKLKVSFSVVR**NNT**STQESGKSAAGEYYAALASRSGDSAAGEYYAALA QESGKSAAGEYYAALASRSGDSAAGEYYAALAQESGKSAAGEYYAALASRSGDSAAGEYYAALAQESGSSGGPG

The sequences are color-coded to indicate their origin: black, DPAPB (with glycosylation sites in bold); purple, the separator sequence GGPG of the H-segment; green, spectrin; light blue, disrupting GP repeat sequence; orange, GS repeat sequence; red, blue, and gray for chaperone-binding, non-binding, and linker sequences, respectively, as in Fig. 4A.