Human DYSF protein sequence and domain structure

The DYSF protein sequence (accession number NM_003494) is colored according to its known structural domains, identified by sequence homology to other proteins. The domain structure shown here is a compilation of the domain structures found by different algorithms (see References below). The exact boundaries of the domains may vary by a few amino acids. Regions that were classified as C2 domains by only some algorithms are listed as "possible C2 domains" (pink).

Exon boundaries are shown underneath the sequence: for example, 1|2 marks the boundary between exons 1 and 2.

References:
ExPASy
http://www.expasy.org/cgi-bin/niceprot.pl?DYSF_HUMAN
UniProt
http://www.pir.uniprot.org/cgi-bin/upEntry?id=DYSF_HUMAN

Legend:
C2 domain
Possible C2 domain
Ferlin family conserved domain
Region with repeated DYSF domains of unknown function
Transmembrane domain

```
MLRVFILYAE NVHTPDTDIS DAYCSAVFAG VKKRTKVIKN SVNPVWNEGF EWDLKGIPLD     60
                             1|2                       2|3

QGSELHVVVK DHETMGRNRF LGEAKVPLRE VLATPSLSAS FNAPLLDTKK QPTGASLVLQ    120
              3|4                                          4|5

VSYTPLPGAV PLFPPPTPLE PSPTLPDLDV VADTGGEEDT EDQGLTGDEA EPFLDQSGGP    180
                                 5|6

GAPTTPRKLP SRPPPHYPGI KRKRSAPTSR KLLSDKPQDF QIRVQVIEGR QLPGVNIKPV    240
                                            6|7

VKVTAAGQTK RTRIHKGNSP LFNETLFFNL FDSPGELFDE PIFITVVDSR SLRTDALLGE    300
                      7|8                    8|9

FRMDVGTIYR EPRHAYLRKW LLLSDPDDFS AGARGYLKTS LCVLGPGDEA PLERKDPSED    360
 9|10        10|11                                      11|12

KEDIESNLLR PTGVALRGAH FCLKVFRAED LPQMDDAVMD NVKQIFGFES NKKNLVDPFV    420
                      12|13

EVSFAGKMLC SKILEKTANP QWNQNITLPA MFPSMCEKMR IRIIDWDRLT HNDIVATTYL    480
    13|14                        14|15                  15|16

SMSKISAPGG EIEEEPAGAV KPSKASDLDD YLGFLPTFGP CYINLYGSPR EFTGFPDPYT    540
            16|17                 17|18
```

```
ELNTGKGEGV AYRGRLLLSL ETKLVEHSEQ KVEDLPADDI LRVEKYLRRR KYSLFAAFYS   600
     18|19                                       19|20

ATMLQDVDDA IQFEVSIGNY GNKFDMTCLP LASTTQYSRA VFDGCHYYYL PWGNVKPVVV   660
                                    20|21

LSSYWEDISH RIETQNQLLG IADRLEAGLE QVHLALKAQC STEDVDSLVA QLTDELIAGC   720
                                     21|22                         22

SQPLGDIHET PSATHLDQYL YQLRTHHLSQ ITEAALALKL GHSELPAALE QAEDWLLRLR   780
 |23

ALAEEPQNSL PDIVIWMLQG DKRVAYQRVP AHQVLFSRRG ANYCGKNCGK LQTIFLKYPM   840
    23|24                                                    24|25

EKVPGARMPV QIRVKLWFGL SVDEKEFNQF AEGKLSVFAE TYENETKLAL VGNWGTTGLT   900
                                        25|26

YPKFSDVTGK IKLPKDSFRP SAGWTWAGDW FVCPEKTLLH DMDAGHLSFV EEVFENQTRL   960
                                    26|27

PGGQWIYMSD NYTDVNGEKV LPKDDIECPL GWKWEDEEWS TDLNRAVDEQ GWEYSITIPP  1020
         27|28                                        28|29

ERKPKHWVPA EKMYYTHRRR RWVRLRRRDL SQMEALKRHR QAEAEGEGWE YASLFGWKFH  1080
                                     29|30

LEYRKTDAFR RRRWRRRMEP LEKTGPAAVF ALEGALGGVM DDKSEDSMSV STLSFGVNRP  1140
                                     30|31

TISCIFDYGN RYHLRCYMYQ ARDLAAMDKD SFSDPYAIVS FLHQSQKTVV VKNTLNPTWD  1200
     31|32                             32|33

QTLIFYEIEI FGEPATVAEQ PPSIVVELYD HDTYGADEFM GRCICQPSLE RMPRLAWFPL  1260
                                     33|34

TRGSQPSGEL LASFELIQRE KPAIHHIPGF EVQETSRILD ESEDTDLPYP PPQREANIYM  1320
               34|35        35|36        36|37

VPQNIKPALQ RTAIEILAWG LRNMKSYQLA NISSPSLVVE CGGQTVQSCV IRNLRKNPNF  1380
               37|38

DICTLFMEVM LPREELYCPP ITVKVIDNRQ FGRRPVVGQC TIRSLESFLC DPYSAESPSP  1440
      38|39

QGGPDDVSLL SPGEDVLIDI DDKEPLIPIQ EEEFIDWWSK FFASIGEREK CGSYLEKDFD  1500
    39|40                           40|41

TLKVYDTQLE NVEAFEGLSD FCNTFKLYRG KTQEETEDPS VIGEFKGLFK IYPLPEDPAI  1560
   41|42                                              42|43

PMPPRQFHQL AAQGPQECLV RIYIVRAFGL QPKDPNGKCD PYIKISIGKK SVSDQDNYIP  1620
                                        43|44

CTLEPVFGKM FELTCTLPLE KDLKITLYDY DLLSKDEKIG ETVVDLENRL LSKFGARCGL  1680
      44|45

PQTYCVSGPN QWRDQLRPSQ LLHLFCQQHR VKAPVYRTDR VMFQDKEYSI EEIEAGRIPN  1740
    45|46                                               46|47
```

```
PHLGPVEERL ALHVLQQQGL VPEHVESRPL YSPLQPDIEQ GKLQMWVDLF PKALGRPGPP   1800
                                    47|48

FNITPRRARR FFLRCIIWNT RDVILDDLSL TGEKMSDIYV KGWMIGFEEH KQKTDVHYRS   1860
         48|49                                      49|50

LGGEGNFNWR FIFPFDYLPA EQVCTIAKKD AFWRLDKTES KIPARVVFQI WDNDKFSFDD   1920
                                    50|51

FLGSLQLDLN RMPKPAKTAK KCSLDQLDDA FHPEWFVSLF EQKTVKGWWP CVAEEGEKKI   1980
  51|52

LAGKLEMTLE IVAESEHEER PAGQGRDEPN MNPKLEDPRR PDTSFLWFTS PYKTMKFILW   2040
52|53                                          53|54

RRFRWAIILF IILFILLLFL AIFIYAFPNY AAMKLVKPFS   2080
                  54|55
```