



Identification of residues critical for topology inversion of the transmembrane protein TM4SF20 through regulated alternative translocation

Received for publication, January 22, 2019, and in revised form, February 5, 2019. Published, Papers in Press, February 26, 2019, DOI 10.1074/jbc.RA119.007681

Jingcheng Wang[‡], Lisa N. Kinch^{§¶}, Bray Denard[‡], Ching-En Lee[‡], Elina Esmaeilzadeh Gharehdaghi[‡], Nick Grishin^{§¶}, and Jin Ye^{¶1}

From the Departments of [‡]Molecular Genetics and [§]Biophysics and Biochemistry, and [¶]Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390

Edited by Karen G. Fleming

Adopting a proper topology is crucial for transmembrane proteins to perform their functions. We previously reported that ceramide regulates a transmembrane protein called TM4SF20 (transmembrane 4 L six family member 20) through topological inversion by altering the direction through which the protein is translocated across membranes during translation. This regulatory mechanism, denoted regulated alternative translocation (RAT), depends on a GXXXN motif present in the first transmembrane helix of TM4SF20. Here, using site-directed mutagenesis, we show that Asn-26 in the motif is crucial for RAT of TM4SF20, as it cannot be replaced even by Gln. In contrast, Gly-22 in the motif could be substituted by other small residues such as Ala and Ser without affecting RAT of TM4SF20. We further demonstrate that the GXXXN motif alone is insufficient to induce RAT of a transmembrane protein because TM4SF4, a relative of TM4SF20 that also contains the motif in the first transmembrane helix, did not undergo RAT. Using TM4SF40–TM4SF20 chimeras, we identified Pro-29 of TM4SF20 as another important element required for RAT of the protein. Substituting Pro-29 alone did not affect RAT of TM4SF20, whereas replacing Pro-29 together with either Leu-25 or Val-17 of TM4SF20 with the corresponding residues of TM4SF4 abolished RAT of TM4SF20. Because Val-17, Gly-22, Leu-25, Asn-26, and Pro-29 are predicted to reside along the same surface of the transmembrane helix, our results suggest that interactions with other proteins mediated by this surface during translocation may be critical for RAT of TM4SF20.

Unlike cytosolic and nuclear proteins, segments of transmembrane proteins located at either side of membranes are exposed to different environments. Thus, the topology of transmembrane proteins, which depicts orientation of their segments at either side of membranes, is crucial for the function of these proteins. In eukaryotic cells, the topology of polytopic membrane proteins is primarily determined during their trans-

location by the direction through which the first transmembrane helix is translocated across membranes of the endoplasmic reticulum (ER)² (1).

We recently reported that ceramide regulates a polytopic transmembrane protein called TM4SF20 (transmembrane 4 L six family member 20) by inverting the topology of the protein (2). In the absence of ceramide, the N-terminal end of the first transmembrane helix is inserted into ER lumen, producing TM4SF20(A), which inhibits proteolytic activation of a membrane-bound transcription factor called CREB3L1 (cAMP responsive element-binding protein 3-like 1) (2, 3) (Fig. 1A). Accumulation of ceramide or related sphingolipids inverts the direction through which the first transmembrane helix is inserted into the ER, leading to production of TM4SF20(B) in which the N-terminal end of the transmembrane helix is located at cytosol (2) (Fig. 1A). In contrast to TM4SF20(A), TM4SF20(B) stimulates proteolytic activation of CREB3L1 (2), allowing the transcription factor to drive expression of genes required for assembly of collagen-containing matrix and those inhibiting cell proliferation (3–5). Because this regulatory mechanism does not flip TM4SF20 that has already been synthesized but inverts the topology of the newly synthesized protein by changing the direction through which the first transmembrane helix is translocated across membranes, we designated this process as regulated alternative translocation (RAT) (2).

We then determined that RAT of TM4SF20 relies on a GXXXN motif present in the first transmembrane helix of the protein, as mutating either the Gly-22 or Asn-26 in the motif to leucine completely abolished the topological regulation, causing the protein to be locked into TM4SF20(B) regardless of the presence of ceramide (2). In the current study, we performed more detailed analyses on this motif. Our data indicate that the presence of a G/A/SXXXN motif in the first transmembrane helix is required but not sufficient to drive RAT of TM4SF20. The results suggest that the two critical residues in the motif

This work was supported by National Institutes of Health Grants GM-116106 and HL-20948 and Welch Foundation Grant I-1832 (to J.Y.). The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

¹ To whom correspondence should be addressed: 5323 Harry Hines Blvd., Dallas, TX. Tel.: 214-648-3461; Fax: 214-648-8804; E-mail: Jin.Ye@UTSouthwestern.edu.

² The abbreviations used are: ER, endoplasmic reticulum; TM4SF20, transmembrane 4 L six family member 20; RAT, regulated alternative translocation; TM4SF4, transmembrane 4 L six family member 4; CREB3L1, cAMP responsive element-binding protein 3-like 1; PNGase F, peptide N-glycosidase F; TRAM2, translocating chain-associated membrane protein 2; TLC domain, TRAM/LAG1/CLN8 homology domain; TRAM1, translocating chain-associated membrane protein 1; GPCR, G protein-coupled receptor.

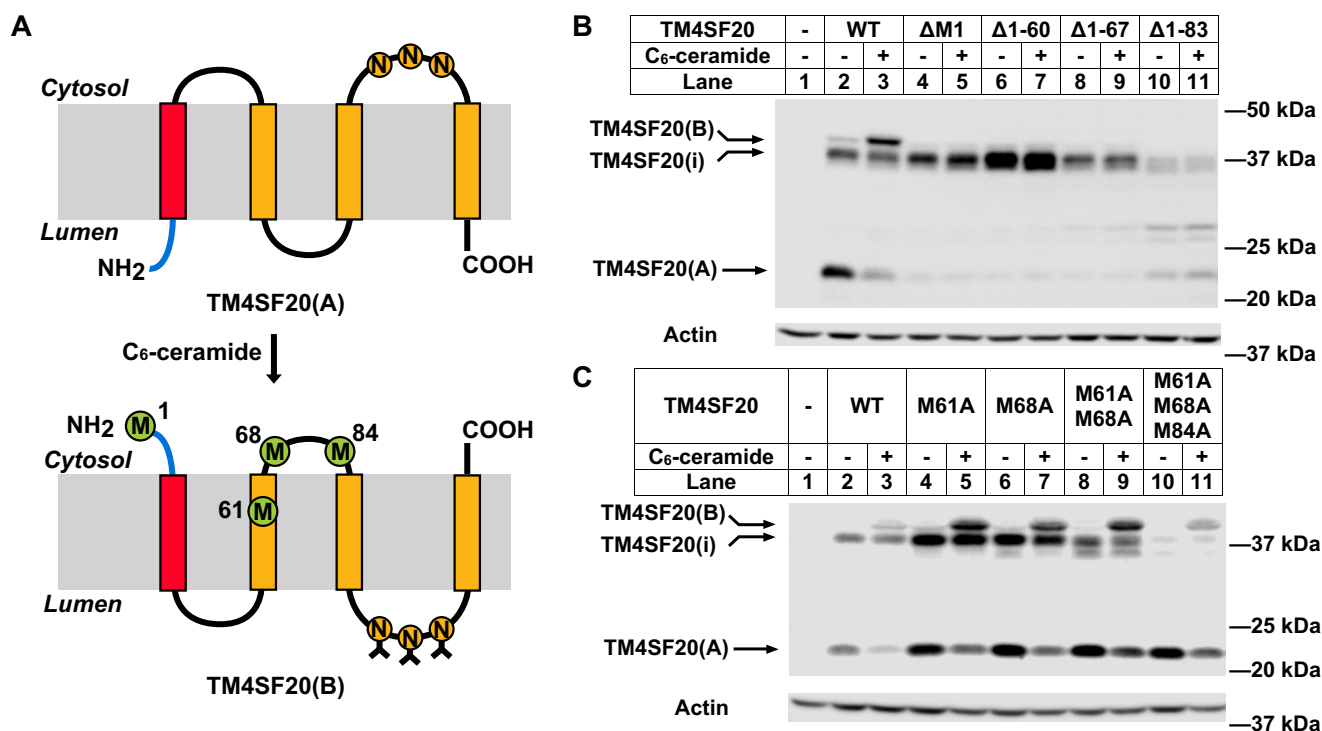


Figure 1. Internal translation start sites contribute to generation of TM4SF20(i). *A*, a model illustrating RAT of TM4SF20. The positions of Met-1, -61, -68, and -84 residues in TM4SF20(B) are indicated. Asparagine residues glycosylated in TM4SF20(B) but not TM4SF20(A) are highlighted. *B* and *C*, on day 0, A549 cells were seeded at 5×10^5 cells per 60-mm dish. On day 1, cells were transfected with 0.1 μ g/dish of plasmid encoding C terminally Myc-tagged TM4SF20 containing the indicated mutation or deletion. 4 h after transfection, cells were treated with or without 5 μ M C₆-ceramide for 16 h. On day 2, cell lysates were subject to immunoblot analysis with anti-Myc to detect TM4SF20.

along with other residues aligned with them at the same surface of the helix may interact with ER translocation machinery, and this interaction may be critical for RAT of TM4SF20.

Results

Characterization of TM4SF20(i), another isoform of the protein produced through alternative translation initiation

TM4SF20 contains three potential *N*-linked glycosylation sites, all located at the longest loop between the third and fourth transmembrane helix (2). These sites are glycosylated in TM4SF20(B), as the loop is within the ER lumen (Fig. 1A) (2). In contrast, these sites are not glycosylated in TM4SF20(A), as the loop is located at the cytosol (Fig. 1A) (2). This difference in glycosylation is one of the reasons why the apparent molecular weight of TM4SF20(B) is higher than that of TM4SF20(A) (2). A typical immunoblot analysis detecting the Myc epitope tagged at the C terminus of TM4SF20 was shown in Fig. 1B: in untreated cells, a band migrated close to the predicted molecular weight of TM4SF20(A) was detected (Fig. 1B lane 2). Treatment with C₆-ceramide, a ceramide analogue that is converted to naturally existing ceramide inside cells (6), led to production of TM4SF20(B) with a higher molecular weight (Fig. 1B, lane 3). The residual TM4SF20(A) remained in these cells (Fig. 1B, lane 3) was synthesized before the ceramide treatment (2). In addition to ceramide-regulated TM4SF20(A) and TM4SF20(B), we also detected another form of the protein with a molecular weight slightly lower than that of TM4SF20(B), the presence of which was not affected by ceramide treatment (Fig. 1B, lanes 2 and 3).

Although fallen short of revealing the molecular identity of the protein, our previous study demonstrated that this form of TM4SF20 was glycosylated, and it can only be detected with epitopes tagged at the C but not N terminus of the protein (2). A plausible explanation for these results is that this unknown isoform could be an N terminally-truncated TM4SF20 synthesized from a downstream internal translation initiation site. After the first methionine, the next possible translation initiation sites could be closely positioned Met-61, Met-68, and Met-84. The protein produced from these internal translation initiation sites should not contain the first transmembrane helix of TM4SF20, which is crucial to make the protein to adopt a topology consistent with TM4SF20(A) (2). Thus, this truncated protein is expected to be locked into the topology of TM4SF20(B) regardless of the presence of ceramide. This hypothesis could explain why the apparent molecular weight of the unknown form of TM4SF20 is smaller than that of TM4SF20(B), why the protein is glycosylated, and why expression of the protein is not regulated by ceramide.

To test this hypothesis, we transfected cells with plasmids encoding various N terminally-truncated mutants of TM4SF20 tagged with Myc at the C terminus, and detected expression of the proteins by immunoblot analysis with anti-Myc. Deleting the first methionine (Δ M1) abolished expression of TM4SF20(A) and TM4SF20(B) but not this unknown form of TM4SF20 (Fig. 1B, lanes 4 and 5), indicating that this protein was not initiated from the first methionine. The same result was obtained for TM4SF20 initiated from M61(Δ 1-60) or M68(Δ 1-67) (Fig. 1B, lanes 6-9). In con-

Residues critical for RAT of TM4SF20

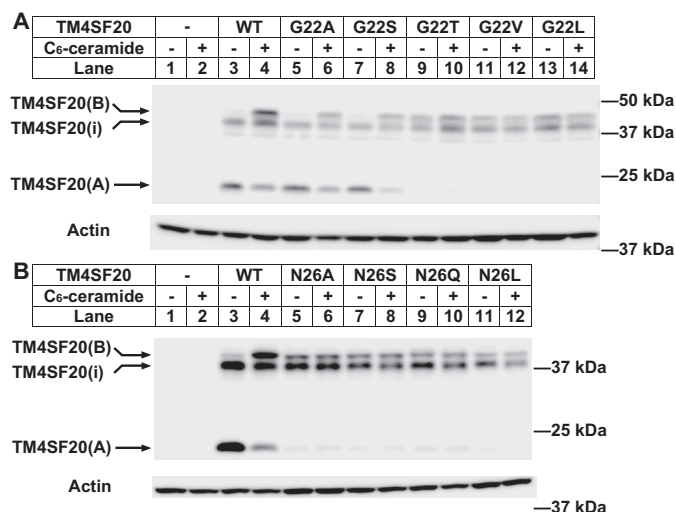


Figure 2. Mutation analysis of the Gly-22 and Asn-26 in the GXXXN motif of TM4SF20. A549 cells were seeded, transfected with C terminally Myc-tagged TM4SF20 mutants as indicated, treated, and analyzed as described in the legend to Fig. 1B.

trast, the molecular weight of the truncated TM4SF20 initiated from M84(Δ 1–83) was slightly smaller than that of the unknown form of the protein (Fig. 1B, lanes 10 and 11). These results suggest that the unknown form of TM4SF20 was initiated from an internal methionine N-terminal to Met-84. We thus designate this form of the protein as TM4SF20(i) to reflect its initiation from an internal start codon.

To further characterize TM4SF20(i), we mutated the two potential internal initiation codons at Met-61 and Met-68 to alanine. Although mutating either methionine did not affect migration of TM4SF20(i) on immunoblot analysis (Fig. 1C, lanes 4–7), mutating both residues reduced the molecular weight of TM4SF20(i) to that of TM4SF20(Δ 1–83) (Fig. 1C, lanes 8 and 9). TM4SF20(i) almost disappeared when Met-61, -68, and -84 were all mutated to alanine (Fig. 1C, lanes 10 and 11). These results suggest that TM4SF20(i) is a N terminally-truncated protein, the translation of which is initiated from a cluster of methionines between the second and third transmembrane domain, namely Met-61, Met-68, and Met-84.

Mutagenesis analysis of the GXXXN motif

We previously reported that a GXXXN motif present in the first transmembrane helix of TM4SF20 is crucial for the topological regulation of TM4SF20, as mutating either one of the two critical residues in the motif to leucine (G22L or N26L) completely blocked RAT of the protein, locking the topology of the protein into TM4SF20(B) regardless of the presence of ceramide (2). To gain more insights into this motif, we mutated Gly-22 to other smaller residues such as Ala and Ser did not affect RAT of TM4SF20 (Fig. 2A, lanes 3–8), replacing it with a larger residue such as Thr or Val completely inhibited RAT of TM4SF20 just like G22L, causing production of only B but not A form of the protein regardless of ceramide treatment (Fig. 2A, lanes 9–14). In contrast to Gly-22, the requirement of Asn-26 for RAT of TM4SF20 appears to be more stringent, as replacing

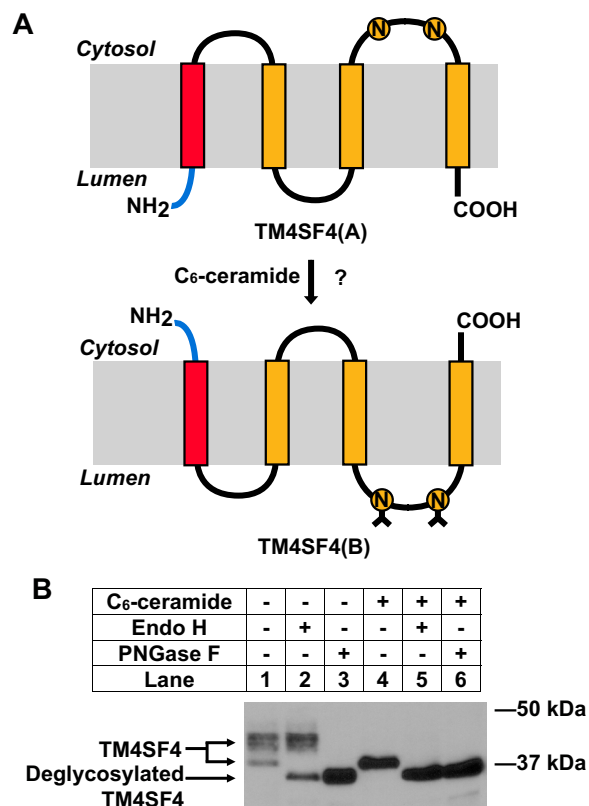


Figure 3. TM4SF4 does not go through RAT. A, a hypothetical model of RAT of TM4SF4. The only two potential N-linked glycosylation sites, Asn-124 and Asn-156, are predicted to be glycosylated only in TM4SF4(B) but not in TM4SF4(A). B, A549 cells were seeded, transfected with C terminally Myc-tagged TM4SF4, and treated with C₆-ceramide as described in the legend to Fig. 1B. On day 2, equal amounts of lysate treated with or without the indicated endoglycosidase were subject to immunoblot analysis with anti-Myc to detect TM4SF4.

Asn to other polar or nonpolar residues such as Ala, Leu, Ser, or Gln all abolished RAT of the protein (Fig. 2B).

The GXXXN motif is not sufficient to induce RAT

TM4SF20 belongs to a family of proteins that contain four transmembrane domains (7). Within this family, TM4SF4 is another protein that contains a GXXXN motif in the first transmembrane domain. We thus decided to determine whether TM4SF4 is subjected to ceramide-induced topological inversion through RAT. Similar to TM4SF20, TM4SF4 has two predicted N-linked glycosylation sites in the loop between the third and fourth transmembrane helix. Thus, if TM4SF4 behaves similarly to TM4SF20 undergoing ceramide-induced RAT, we expect the protein to be unglycosylated in untreated cells, but become glycosylated in cells treated with ceramide (Fig. 3A). Surprisingly, TM4SF4 was glycosylated regardless of ceramide treatment, as treatment with PNGase F, an endoglycosidase that removes N-linked glycan, reduced the apparent molecular weight of the protein in either condition (Fig. 3B, lanes 1, 3, 4, and 6). In contrast to PNGase F, endoglycosidase H, another endoglycosidase that trims N-linked sugars from proteins localized in the ER but not those in the post-Golgi compartment (8), was only active for deglycosylation of TM4SF4 in ceramide-treated but not untreated cells (Fig. 3B, lanes 1, 2, 4, and 5). These results suggest that trafficking but not topology of

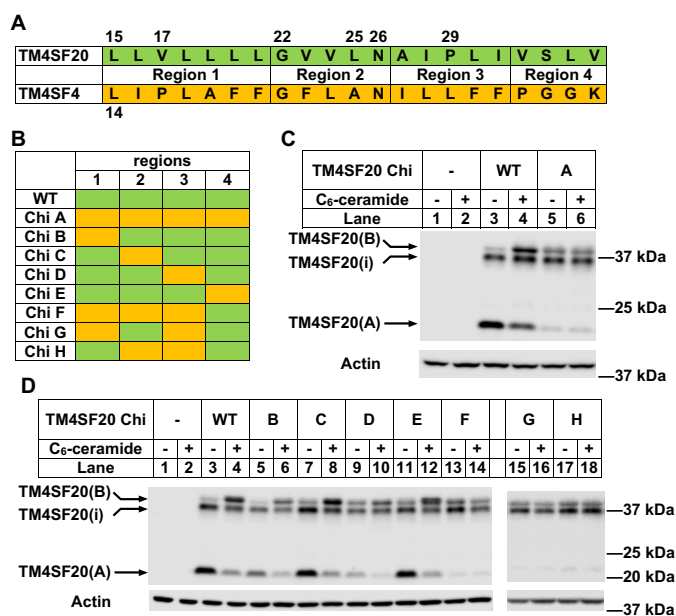


Figure 4. Elements other than the GXXXN motif in the first transmembrane domain contributes to RAT of TM4SF20. A, alignment of the first transmembrane domain of TM4SF20 with that of TM4SF4. The domains are arbitrarily split into four regions as indicated. B, graphic illustration of chimeras in which the indicated regions in the first transmembrane domain of TM4SF20 shown in A (green) were replaced by the corresponding sequence of TM4SF4 (orange). Chi, chimera. C and D, A549 cells were seeded, transfected with C terminally Myc-tagged TM4SF20/TM4SF4 chimeras as indicated, treated, and analyzed as described in the legend to Fig. 1B.

TM4SF4 is affected by ceramide, and TM4SF4 always adopts a topology consistent with that of TM4SF20(B) regardless of the presence of ceramide.

We previously reported that the first transmembrane helix of TM4SF20 is required and sufficient to drive topological inversion of the protein in response to ceramide accumulation (2). We thus hypothesized that even though TM4SF4 contains a GXXXN motif, other elements in the transmembrane helix may prevent RAT of the protein. To test this hypothesis, we replaced the first transmembrane helix of TM4SF20 with the corresponding sequence of TM4SF4 (Fig. 4A). Unlike TM4SF20, the chimeric protein (chimera A in Fig. 4B) did not undergo RAT and was locked into the configuration of TM4SF20(B) regardless of ceramide treatment (Fig. 4C).

To further narrow down the regions of TM4SF4 that block RAT, we divided the first transmembrane domain of the protein into four regions (Fig. 4A). Sequence alignment revealed that region 4 located at the C-terminal end of the transmembrane helix shared the least sequence homology between TM4SF4 and TM4SF20 (Fig. 4A). However, substituting residues of TM4SF20 in this region with that of TM4SF4 (chimera E, Fig. 4B) did not affect ceramide-induced RAT of TM4SF20 (Fig. 4D, lanes 11 and 12). Likewise, placing back residues of TM4SF20 in this region to the chimera in which the transmembrane helix of TM4SF20 was replaced by the corresponding sequence of TM4SF4 (chimera F, Fig. 4B) did not restore RAT of the protein (Fig. 4D, lanes 13 and 14). These results suggest that this region is not important for RAT of TM4SF20.

Further chimeric analyses revealed that substituting residues of TM4SF20 in regions 1 or 2 by that of TM4SF4 (chimera B or

C, Fig. 4B) barely affected RAT of TM4SF20 (Fig. 4D, lanes 5–8). Although replacing residues in region 3 of TM4SF20 with that of TM4SF4 alone (chimera D, Fig. 4B) partially impaired RAT of the protein (Fig. 4D, lanes 9 and 10), combining this substitution with that switching the residues in regions 1 or 2 (chimera G or H, Fig. 4B) completely abolished RAT of the protein (Fig. 4D, lanes 15–18).

The results shown in Fig. 4 suggest that combining mutations in region 3 with that in regions 1 or 2 results in complete inhibition of RAT of TM4SF20. To identify the residues in region 3 that are important for RAT, we fixed the substitution in region 1 (chimera B, Fig. 4B), and then mutated each residue of TM4SF20 in region 3 to the corresponding residue of TM4SF4. This analysis revealed that only the P29L mutation under this circumstance completely inhibited RAT of the chimeric protein (Fig. 5A, lanes 8 and 9). These results suggest that Pro-29 in region 3 of TM4SF20 may also be important for RAT of the protein.

We then sought to identify residues in region 1 supporting RAT of TM4SF20. For this purpose, we substituted each residue of TM4SF20 in region 1 with that of TM4SF4 on TM4SF20(P29L). Although P29L by itself did not affect RAT of TM4SF20 (Fig. 5B, lanes 5 and 6), combining this mutation with V17P but no other substitutions in region 1 abolished RAT of TM4SF20 (Fig. 5B, lanes 7–14). Under this circumstance, it appears that it was gaining a Pro but not losing a Val that resulted in inhibition of RAT, as mutating Val-17 to other residues such as Ala or Val did not affect RAT of TM4SF20 (Fig. 5C). Using the same strategy, we identified that it was the substitution of L25A in region 2 combined with the P29L mutation in region 3 resulted in inhibition of RAT of TM4SF20 (Fig. 5D).

Discussion

We previously reported that ceramide triggers topological inversion of TM4SF20 through RAT, and this process depends on a GXXXN motif present in the first transmembrane helix of the protein (2). Here we present more evidence that the first transmembrane helix is crucial for RAT. TM4SF20(i), another isoform of the protein, does not undergo RAT owing to the lack of the first transmembrane helix because its translation is initiated from a cluster of methionines between the second and third transmembrane domain. Interestingly, expression of TM4SF20(i) was significantly inhibited only when all three start codons were mutated in the mRNA encoding this section of the protein. This observation suggests that the mRNA surrounding these start codons lies within a strong internal translation initiation site so that translation of TM4SF20(i) can be started from either one of the in-frame start codons. This scenario may also explain why mutating all three start codons in this region markedly reduced but did not completely eliminate production of TM4SF20(i), as multiple non-AUG, noncanonical start codons (9) present in this strong translation initiation site may still allow residual amounts of the protein to be produced.

We also performed more mutational analyses on the GXXXN motif. Although the Gly-22 position can tolerate other small residues such as Ala or Ser without losing the topological regulation, Asn-26 appears to be very critical for RAT as it cannot even be replaced by Gln. These observations are supported by

Residues critical for RAT of TM4SF20

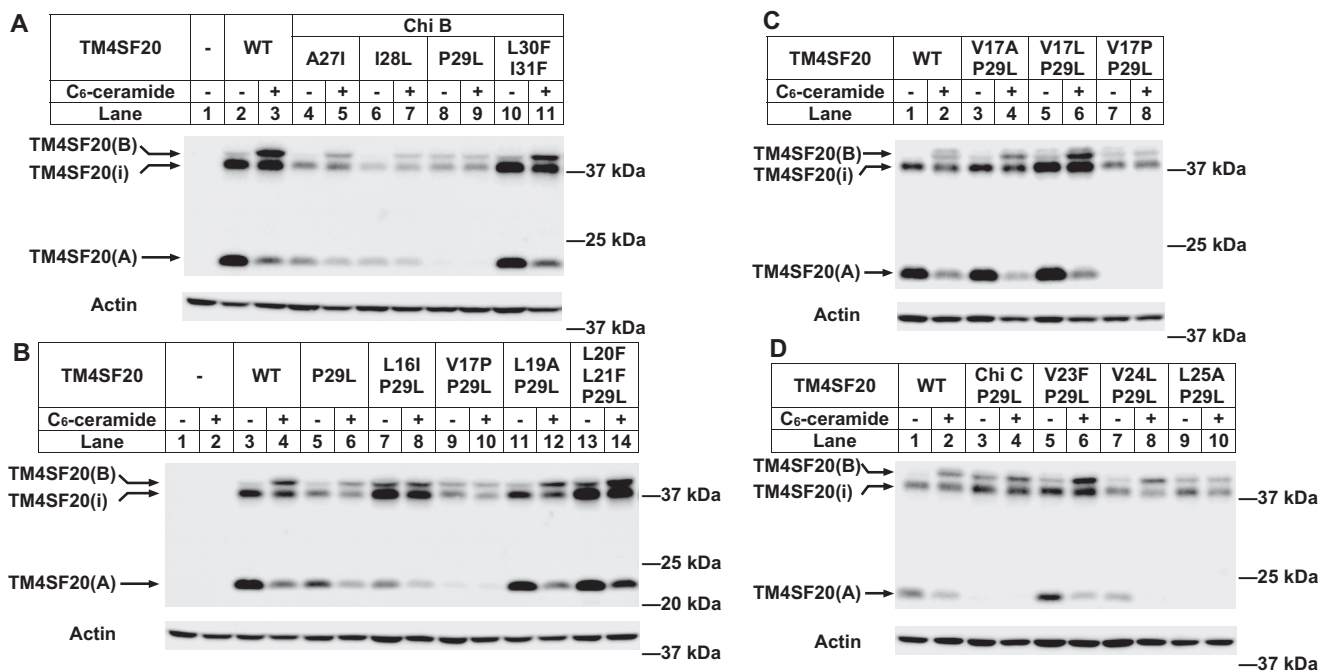


Figure 5. Pro-29 in combination with Val-17 or Leu-25 is critical for RAT of TM4SF20. A549 cells were seeded, transfected with C terminally Myc-tagged TM4SF20/TM4SF4 chimeras and/or mutations as indicated, treated, and analyzed as described in the legend to Fig. 1B.

sequence alignment of TM4SF20 from various species, which demonstrates that Asn-26 is conserved across species, whereas Gly, Ala, or Ser can be found at position 22 of the human TM4SF20 (Fig. 6A). To gain structural insights into this motif, we used the structural data of the first transmembrane domain of the P2Y₁₂ receptor (10, 11) as a template to model the first transmembrane helix of TM4SF20 (Fig. 6B), as both transmembrane helices contain a GXXXN motif and share some sequence similarity. This structural model predicts that Gly-22 is adjacent to Asn-26 at the same side of the transmembrane α -helix. A small side chain at position 22 may favor formation of a hydrogen bond between the backbone C = O group at this location and the Asn-26 side chain NH₂ group (10, 12) (Fig. 6B). Formation of this hydrogen bond may orientate the side chain of Asn-26 for interaction with another protein during its translocation across ER membranes, and this interaction may be critical for RAT of TM4SF20. The energetic tendency for protein-protein interaction mediated by the asparagine residue in the GXXXN motif present in the center of a transmembrane helix has been reported previously (13).

In addition to Gly-22 and Asn-26, we identified Pro-29 as a residue important for RAT of TM4SF20. This finding is supported by the sequence alignment, as Pro-29 is also conserved in TM4SF20 across species (Fig. 6A). Unlike Gly-22 and Asn-26, replacing the proline residue at this position with leucine by itself did not affect RAT of TM4SF20. This mutation has to combine with other amino acid substitution within the transmembrane helix to disrupt the topological regulation. One of the mutations acting together with P29L to disrupt RAT is L25A. Given the conservation of Pro-29 within the TM4SF20 family, the presence of Pro-29 may induce a kink in the α -helix by disrupting the backbone hydrogen bond to be formed by Pro-29 and the neighboring Leu-25 in a transmembrane helix

structure model (14) (Fig. 6B, left). Although formation of this hydrogen bond may become possible with the P29L mutation, the influence of the neighboring Leu-25 side chain may still favor the native kinked helix conformation (15) (Fig. 6B, middle). The additional L25A mutation, in which the bulky side chain on Leu-25 was replaced by the smaller one of alanine, might allow straightening of the helix (Fig. 6B, right). If this explanation is correct, then the presence of this kink may also be a structural requirement for RAT. Another mutation acting together with P29L to disrupt RAT is V17P. Interestingly, Val-17, Gly-22, Leu-25, Asn-26, and Pro-29 are all located along the same surface of the transmembrane helix revealed by the net projection (Fig. 6C). We currently do not understand why combining the P29L mutation with V17P disrupts RAT of TM4SF20. Because it is gaining a proline rather than losing a valine at this position that blocked RAT of TM4SF20, it is likely that introducing a proline at this position may disrupt the helical surface formed by the residues described above. Taken together, our results suggest that a protein-interacting surface containing Val-17, Gly-22, Leu-25, Asn-26, and Pro-29 may bind to a component of ER translocon, and this interaction may be critical for RAT of TM4SF20.

A candidate protein that may interact with the first transmembrane helix of TM4SF20 during its translocation is translocating chain-associated membrane protein 2 (TRAM2). We previously reported that knockdown of TRAM2 stimulated production of TM4SF20(B) even in the absence of ceramide (2). TRAM2 is highly homologous to TRAM1, a component of ER translocon responsible for transporting nascent peptide chains across ER membranes during translation (16, 17). TRAM2 contains a TLC domain postulated to bind ceramide or related sphingolipids (18). A possible explanation of these results is that in the absence of ceramide, TRAM2 interacts with the first

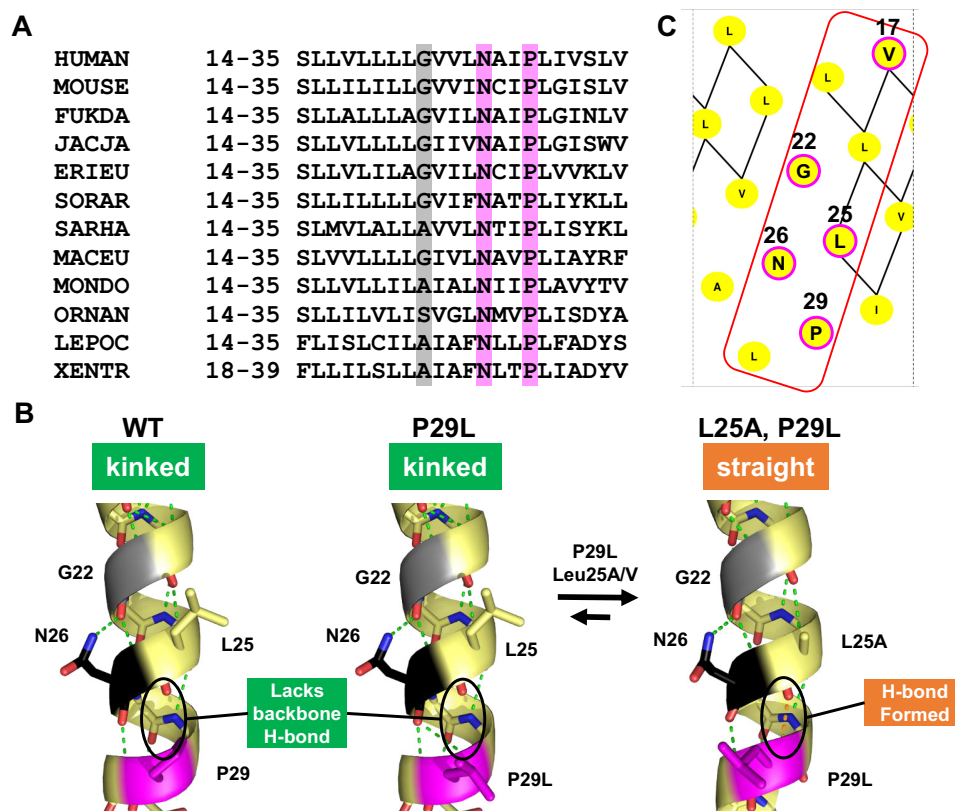


Figure 6. Structural modeling of the first transmembrane helix of TM4SF20. *A*, alignment of the first transmembrane domains of TM4SF20 in select species. Corresponding residues at the position of Gly-22, Asn-26, and Pro-29 of human TM4SF20 are highlighted. *B*, speculated structural model to demonstrate the effect of L25A and P29L mutations on a kink in the first transmembrane helix of TM4SF20. Side chains of indicated residues are shown. The hydrogen bond formed between the helix backbone at positions 26 and 29 of TM4SF20(L25A,P29L) is highlighted in orange, whereas all other hydrogen bonds are highlighted in green. *C*, projection of the first transmembrane helix of TM4SF20 created by NetWheels to show the surface containing Val-17, Gly-22, Leu-25, Asn-26, and Pro-29.

transmembrane helix portion of the nascent peptide chain of TM4SF20, and this interaction enables an unusual translocation process by driving the sequence N-terminal to the transmembrane helix into the ER lumen to produce TM4SF20(A). Ceramide, through its potential binding with TRAM2, may block this interaction. In the absence of this interaction, the translocation goes through the default pathway by pushing the sequence C-terminal to the transmembrane helix into the ER lumen (19) to produce TM4SF20(B). We have tried to disrupt RAT of TM4SF20 by substituting polar residues within transmembrane helices of TRAM2 that might interact with Asn-26 of TM4SF20 with leucine. Unfortunately, these efforts so far have been unsuccessful. More sophisticated analyses may be needed to uncover the biochemical interaction between the first transmembrane helix of TM4SF20 and ER translocon machinery that is critical for RAT of TM4SF20.

Notably, all mutations we made in the first transmembrane helix of TM4SF20 that inhibited expression of TM4SF20(A) in the absence of ceramide increased synthesis of glycosylated TM4SF20(B) but not any other forms of the protein with a lower molecular weight. These results suggest that these mutations did not disrupt the first transmembrane helix, as in the absence of this transmembrane helix, the loop containing the glycosylation sites would have been exposed to cytosol, thereby producing an unglycosylated protein with a molecular weight

lower than that of TM4SF20(B). Thus, the first transmembrane helix appeared to remain intact in these mutant proteins.

Lipid-induced topological inversion has emerged as a novel mechanism to regulate transmembrane proteins both in prokaryotic and eukaryotic cells (2, 20–24). To search for more mammalian proteins subjected to this regulation, we used bioinformatics analysis to look for proteins containing a GXXXN motif in the first transmembrane helix. This analysis revealed that this motif is enriched in the first transmembrane domain of G protein-coupled receptors (GPCRs). We then identified one of these receptors, namely C-C chemokine receptor 5, which indeed undergoes ceramide-induced RAT.³ The current study demonstrates that TM4SF4, which also contains a GXXXN motif in the first transmembrane helix, is not subjected to topological inversion through RAT. Through various chimera between TM4SF20 and TM4SF4, we determined that no single amino acid substitution could explain the difference in topological regulation between the two proteins. Instead, it was the combined replacements of Pro-29 with Leu-25 or Val-17 of TM4SF20 with the corresponding residues of TM4SF4 that led to inactivation of RAT. However, placing Pro-29 and Val-17 back still failed to restore the topological regulation of the TM4SF20 chimera in which the rest of the

³ B. Denard, S. Han, J. Kim, E. M. Ross, and J. Ye, unpublished data.

Residues critical for RAT of TM4SF20

residues in the first transmembrane helix were replaced by TM4SF4.⁴ Thus, in addition to the residues we have identified, combined substitutions of other residues of TM4SF20 with that of TM4SF4 may also block RAT of the fusion protein. Thus, sequence analysis alone may not be sufficient to predict transmembrane proteins subjected to RAT. Detailed biochemical analysis is required to demonstrate whether a transmembrane protein is subjected to topological regulation through RAT.

Experimental procedures

Materials and methods

Hybridoma cells producing IgG-9E10, a mouse mAb against Myc tag, were obtained from the American Type Culture Collection. Rabbit anti-actin antibody was obtained from Sigma. Horseradish peroxidase-conjugated donkey anti-mouse and donkey anti-rabbit secondary antibodies were obtained from Jackson ImmunoResearch Inc. C₆-ceramide was obtained from Sigma.

Plasmids

pCMV-TM4SF20-Myc (2) and pCMV-TM4SF4-Myc encodes the indicated full-length human protein followed by five tandem repeats of the Myc epitope tag and a His₆ tag. Mutations based on pCMV-TM4SF20-Myc were performed using QuikChange Lightning multisite-directed mutagenesis kit (Agilent Technologies). Truncated protein expression plasmids are made with TaKaRa Ex Taq polymerase and New England Biolabs cloning enzymes. The open reading frames of TM4SF20 in all plasmids were confirmed by Sanger sequencing.

Cell culture and transfection

A549 cells were maintained at 37 °C, 8% CO₂, in medium A (1:1 mixture of Ham's F-12 medium and Dulbecco's modified Eagle's medium containing 100 units/ml of penicillin, 100 µg/ml of streptomycin sulfate, and 5% (v/v) fetal calf serum). Transfection was carried out in 60-mm dishes containing 5 × 10⁵ cells, using X-tremeGENE HP DNA Transfection Reagent (Roche Applied Science) according to the manufacturer's suggestion. The total amount of DNA was adjusted to 2 µg/dish with empty vector.

Immunoblot

Cells were harvested for whole cell lysate as described (25) and analyzed by SDS-PAGE followed by immunoblot assays with the indicated antibodies (1:2,000 dilution for mouse anti-Myc, 1:10,000 dilution for rabbit anti-actin, and 1:5,000 dilution for secondary antibodies). Horseradish peroxidase-conjugated secondary antibodies were visualized using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Scientific number 34580). The results were captured by films (Phenix Research Products) (only Fig. 3B) or the LI-COR Odyssey Imaging System.

De-glycosylation

Cell lysate was split into three equal fractions, treated with glycoprotein denaturing buffer (New England Biolabs), and

then treated with either endoglycosidase H (New England Biolabs P0702) or PNGase F (New England Biolabs P0704) according to the manufacturer's recommendations.

Structural modeling of the first transmembrane helix of TM4SF20

Sequence orthologs of TM4SF20 from various species that share more than 80% identity with the human protein were downloaded from the OMA (Orthologous Matrix) browser (26) and aligned using the multiple alignment using fast Fourier transform (MAFFT) server (27). For Fig. 6C, the sequence LVLLLLGVVNAIPLI within the first transmembrane of human TM4SF20 was submitted to the NetWheels web server (28) to generate a net projection (default parameters).

For Fig. 6B, the sequence corresponding to the first transmembrane helix of human TM4SF20 (GFSLLVLLLLGVVNAIPLIVSLV) was submitted to HHPRED (29) to identify potential structure templates from the Protein Data Bank. The top scoring templates corresponded to the first transmembrane helix of GPCRs containing a similar GXXXN motif. We used P2Y12 as the template to build the structural model not only because of the presence of the GXXXN motif but also because the first transmembrane helix of the protein could be either straight (Protein Data Bank 4ntj) or kinked (Protein Data Bank 4pxz) dependent on ligand binding (10, 11). The kinked helical template of the P2Y12 receptor was used to build the model for WT TM4SF20 and the P29L mutant, whereas the straight helical template was used to construct the model for TM4SF20(P29L,L25A).

Author contributions—J. W., L. N. K., B. D., N. G., and J. Y. conceptualization; J. W., L. N. K., B. D., C.-E. L., and E. E. G. data curation; J. W., L. N. K., C.-E. L., N. G., and J. Y. formal analysis; J. W. validation; J. W., B. D., and J. Y. investigation; J. W., L. N. K., B. D., and J. Y. methodology; J. W., L. N. K., and J. Y. writing-original draft; J. W., L. N. K., and J. Y. writing-review and editing; L. N. K. visualization; J. Y. resources; J. Y. supervision; J. Y. funding acquisition; J. Y. project administration.

Acknowledgments—We thank Jungyeon Kim for technical assistance, Lisa Beatty and Camille Harry for cell culture, and Nancy Heard for graphic illustration.

References

1. Zimmermann, R., Eyrich, S., Ahmad, M., and Helms, V. (2011) Protein translocation across the ER membrane. *Biochim. Biophys. Acta* **1808**, 912–924 [CrossRef Medline](#)
2. Chen, Q., Denard, B., Lee, C.-E., Han, S., Ye, J. S., and Ye, J. (2016) Inverting the topology of a transmembrane protein by regulating the translocation of the first transmembrane helix. *Mol. Cell* **63**, 567–578 [CrossRef Medline](#)
3. Chen, Q., Lee, C.-E., Denard, B., and Ye, J. (2014) Sustained induction of collagen synthesis by TGF-β requires regulated intramembrane proteolysis of CREB3L1. *PLoS ONE* **9**, e108528 [CrossRef Medline](#)
4. Denard, B., Seemann, J., Chen, Q., Gay, A., Huang, H., Chen, Y., and Ye, J. (2011) The membrane-bound transcription factor CREB3L1 is activated in response to virus infection to inhibit proliferation of virus-infected cells. *Cell Host Microbe* **10**, 65–74 [CrossRef Medline](#)
5. Denard, B., Lee, C., and Ye, J. (2012) Doxorubicin blocks proliferation of cancer cells through proteolytic activation of CREB3L1. *eLife Sci.* **1**, e00090 [CrossRef](#)

⁴ J. Wang, unpublished data.

6. Kitatani, K., Idkowiak-Baldys, J., and Hannun, Y. A. (2008) The sphingolipid salvage pathway in ceramide metabolism and signaling. *Cell Signal.* **20**, 1010–1018 [CrossRef Medline](#)
7. Wright, M. D., Ni, J., and Rudy, G. B. (2000) The L6 membrane proteins: a new four-transmembrane superfamily. *Protein Sci.* **9**, 1594–1600 [CrossRef Medline](#)
8. Breitling, J., and Aebi, M. (2013) N-Linked protein glycosylation in the endoplasmic reticulum. *Cold Spring Harb. Perspect. Biol.* **5**, a013359 [Medline](#)
9. Kearse, M. G., and Wilusz, J. E. (2017) Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.* **31**, 1717–1731 [CrossRef Medline](#)
10. Zhang, J., Zhang, K., Gao, Z.-G., Paoletta, S., Zhang, D., Han, G. W., Li, T., Ma, L., Zhang, W., Müller, C. E., Yang, H., Jiang, H., Cherezov, V., Katritch, V., Jacobson, K. A., *et al.* (2014) Agonist-bound structure of the human P2Y12 receptor. *Nature* **509**, 119–122 [CrossRef Medline](#)
11. Zhang, K., Zhang, J., Gao, Z.-G., Zhang, D., Zhu, L., Han, G. W., Moss, S. M., Paoletta, S., Kiselev, E., Lu, W., Fenalti, G., Zhang, W., Müller, C. E., Yang, H., Jiang, H., *et al.* (2014) Structure of the human P2Y12 receptor in complex with an antithrombotic drug. *Nature* **509**, 115–118 [CrossRef Medline](#)
12. Van Rhee, A. M., Fischer, B., Van Galen, P. J., and Jacobson, K. A. (1995) Modelling the P2Y purinoceptor using rhodopsin as template. *Drug Des. Discov.* **13**, 133–154 [Medline](#)
13. Lear, J. D., Gratkowski, H., Adamian, L., Liang, J., and DeGrado, W. F. (2003) Position-dependence of stabilizing polar interactions of asparagine in transmembrane helical bundles. *Biochemistry* **42**, 6400–6407 [CrossRef Medline](#)
14. Huang, Y. H., and Chen, C. M. (2012) Statistical analyses and computational prediction of helical kinks in membrane proteins. *J. Comput. Aided Mol. Des.* **26**, 1171–1185 [CrossRef Medline](#)
15. Langelaan, D. N., Wiczorek, M., Blouin, C., and Rainey, J. K. (2010) Improved helix and kink characterization in membrane proteins allows evaluation of kink sequence predictors. *J. Chem. Inf. Model* **50**, 2213–2220 [CrossRef Medline](#)
16. Görlich, D., Hartmann, E., Prehn, S., and Rapoport, T. A. (1992) A protein of the endoplasmic reticulum involved early in polypeptide translocation. *Nature* **357**, 47–52 [CrossRef Medline](#)
17. Voigt, S., Jungnickel, B., Hartmann, E., and Rapoport, T. A. (1996) Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *J. Cell Biol.* **134**, 25–35 [CrossRef Medline](#)
18. Winter, E., and Ponting, C. P. (2002) TRAM, LAG1 and CLN8: members of a novel family of lipid-sensing domains? *Trends Biochem. Sci.* **27**, 381–383 [CrossRef Medline](#)
19. Rapoport, T. A., Li, L., and Park, E. (2017) Structural and mechanistic insights into protein translocation. *Annu. Rev. Cell Dev. Biol.* **33**, 369–390 [CrossRef Medline](#)
20. Zhang, W., Bogdanov, M., Pi, J., Pittard, A. J., and Dowhan, W. (2003) Reversible topological organization within a polytopic membrane protein is governed by a change in membrane phospholipid composition. *J. Biol. Chem.* **278**, 50128–50135 [CrossRef Medline](#)
21. Zhang, W., Campbell, H. A., King, S. C., and Dowhan, W. (2005) Phospholipids as determinant of membrane protein topology: Phosphatidylethanolamine is required for the proper topological organization of the γ -aminobutyric acid permease (GabP) of *Escherichia coli*. *J. Biol. Chem.* **280**, 26032–26038 [CrossRef Medline](#)
22. Bogdanov, M., and Dowhan, W. (2012) Lipid-dependent generation of dual topology for a membrane protein. *J. Biol. Chem.* **287**, 37939–37948 [CrossRef Medline](#)
23. Bogdanov, M., Dowhan, W., and Vitrac, H. (2013) Lipids and topological rules governing membrane protein assembly. *Biochim. Biophys. Acta* **1843**, 1475–1488 [CrossRef Medline](#)
24. Vitrac, H., MacLean, D. M., Karlstaedt, A., Taegtmeier, H., Jayaraman, V., Bogdanov, M., and Dowhan, W. (2017) Dynamic lipid-dependent modulation of protein topology by post-translational phosphorylation. *J. Biol. Chem.* **292**, 1613–1624 [CrossRef Medline](#)
25. DeCaprio, J., and Kohl, T. O. (2017) Detergent lysis of tissue culture cells for immunoprecipitation. *Cold Spring Harb. Protoc.* **2017**, pdb prot098558 [Medline](#)
26. Altenhoff, A. M., Glover, N. M., Train, C. M., Kaleb, K., Warwick Vesztrocy, A., Dylus, D., de Farias, T. M., Zile, K., Stevenson, C., Long, J., Redestig, H., Gonnet, G. H., and Dessimoz, C. (2018) The OMA orthology database in 2018: retrieving evolutionary relationships among all domains of life through richer web and programmatic interfaces. *Nucleic Acids Res.* **46**, D477–D485 [CrossRef Medline](#)
27. Katoh, K., and Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform.* **9**, 286–298 [CrossRef Medline](#)
28. Mol, A. R., Castro, M. S., and Fontes, W. (2018) NetWheels: a web application to create high quality peptide helical wheel and net projections. *bioRxiv* 416347
29. Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A. N., and Alva, V. (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 [CrossRef Medline](#)