

# A novel transmembrane topology of presenilin based on reconciling experimental and computational evidence

Anna Henricson, Lukas Käll and Erik L. L. Sonnhammer

Center for Genomics and Bioinformatics, Karolinska Institutet, Stockholm, Sweden

## Keywords

presenilin;  $\gamma$ -secretase; presenilin-like protein; topology prediction; transmembrane topology

## Correspondence

E. Sonnhammer, Center for Genomics and Bioinformatics, Karolinska Institutet, S-171 77 Stockholm, Sweden  
Fax: +46 8337983  
Tel: +46 852486395  
E-mail: erik.sonnhammer@cgb.ki.se

(Received 25 January 2005, revised 22 March 2005, accepted 31 March 2005)

doi:10.1111/j.1742-4658.2005.04691.x

The transmembrane topology of presenilins is still the subject of debate despite many experimental topology studies using antibodies or gene fusions. The results from these studies are partly contradictory and consequently several topology models have been proposed. Studies of presenilin-interacting proteins have produced further contradiction, primarily regarding the location of the C-terminus. It is thus impossible to produce a topology model that agrees with all published data on presenilin. We have analyzed the presenilin topology through computational sequence analysis of the presenilin family and the homologous presenilin-like protein family. Members of these families are intramembrane-cleaving aspartyl proteases. Although the overall sequence homology between the two families is low, they share the conserved putative active site residues and the conserved 'PAL' motif. Therefore, the topology model for the presenilin-like proteins can give some clues about the presenilin topology. Here we propose a novel nine-transmembrane topology with the C-terminus in the extracytosolic space. This model has strong support from published data on  $\gamma$ -secretase function and presenilin topology. Contrary to most presenilin topology models, we show that hydrophobic region X is probably a transmembrane segment. Consequently, the C-terminus would be located in the extracytosolic space. However, the last C-terminal amino acids are relatively hydrophobic and in conjunction with existing experimental data we cannot exclude the possibility that the extreme C-terminus could be buried within the  $\gamma$ -secretase complex. This might explain the difficulties in obtaining consistent experimental evidence regarding the location of the C-terminal region of presenilin.

Presenilins (PSs) are transmembrane (TM) proteins that are highly conserved throughout evolution. Elucidating the TM topology of PSs is vital for understanding their function. Several different models have been suggested based on different experimental findings. It is not possible to produce a membrane topology model that agrees with all published data. In humans there are two PSs (PS1 and PS2) whose protein sequences are > 65% identical to each other. Mutations in these genes cause the majority of early onset familial Alzheimer's disease

cases. PSs are also involved in the proteolytic processing of the Notch receptor, which is responsible for critical signaling events during development. Other substrates have also been identified and there are probably still more to discover [1,2].

PSs are part of  $\gamma$ -secretase, a multisubunit protease that also entails nicastrin, aph-1 and pen-2 [3,4]. This complex is responsible for the intramembrane proteolysis of type I membrane proteins, such as amyloid- $\beta$  precursor protein (APP) and Notch. PS has 10

## Abbreviations

APP, amyloid- $\beta$  precursor protein; CTF, C-terminal fragment; ER, endoplasmic reticulum; GPCR, G protein-coupled receptor; HR, hydrophobic region; IMPAS, intramembrane proteases; NTF, N-terminal fragment; PS, presenilin; PSH, presenilin-homolog; PSL, presenilin-like; SPP, signal peptide peptidase; TLN, telencephalin; TM, transmembrane.

hydrophobic regions (HRs), which may or may not be true TM regions. In the  $\gamma$ -secretase complex, PS exists in its active form as an N- and C-terminal fragment (NTF and CTF, respectively) heterodimer. Endoproteolysis of PS by an unknown 'presenilinase' results in these two fragments [5–7]. The major site of endoproteolytic cleavage in PS1 is between T291 and M292 [5], which is located in HRVII. The first evidence that the NTF–CTF heterodimers are the biologically active form of PS and part of the  $\gamma$ -secretase complex came from inhibition studies using transition state analogues designed to target the diaspartyl putative active site of the protease [8,9]. The two conserved aspartate residues are located in HRVI and HRVIII, respectively (D257 and D385 in PS1). They have been shown to be required for  $\gamma$ -secretase activity [10]. It has also been shown that the PS C-terminus is involved in  $\gamma$ -secretase complex assembly and activity [11–16] and is required for endoplasmic reticulum (ER) retention [13]. It is the conserved 'PAL' sequence in the C-terminal part of the protein that is believed to be essential for the activity. In addition, a wide range of PS interacting proteins have been identified [17].

All previously proposed PS topology models agree that HRI–VI are true TM regions. The debate questions which of HRVII–X are true TM regions, and also the location of the N- and C-termini. It is impossible to produce a membrane topology model that concurs with all previously published models. Most studies conducted are gene fusion experiments using different reporters to indicate cytosolic or noncytosolic location. Such analyses are known to suffer frequently from artifacts induced by the truncation of the protein studied and the nature of the reporter gene used, and therefore these studies must always be interpreted with caution. Antibody experiments on native proteins have been performed by Doan *et al.* [18] and Dewji *et al.* [19,20]. Although antibodies are generally considered more reliable, these two studies produced contradictory results.

The first and most widely accepted model for PS topology is the eight-TM model (excluding HRVII and HRX) with both the N- and C-termini located in the cytosol [21,22]. The experimental results from one antibody study supported this model, while at the same time the authors also speculated about several different possible topology models [18]. A seven-TM model with an extracytosolic N-terminus has also been proposed [19,20]. Other models that have been suggested include a six-TM model with cytosolic N- and C-termini [23], and a 'seven-TM and one membrane-embedded' model with an extracytosolic C-terminus [24]. The eight-TM model is primarily based on studies on the *Caenorhabditis*

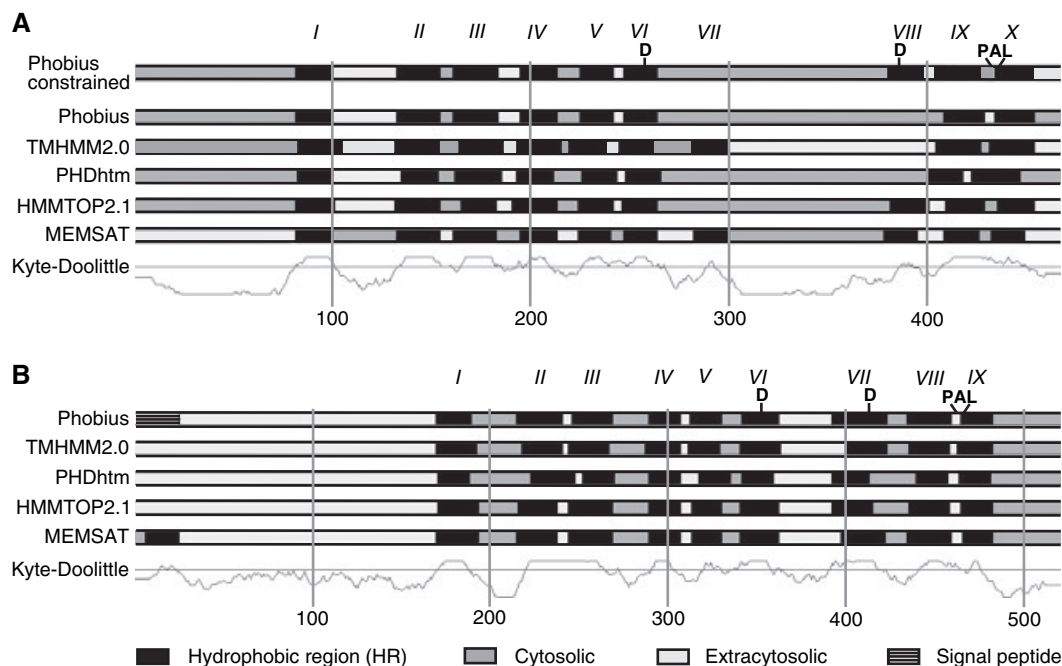
*elegans* ortholog of human PS, Sel-12. The other studies have mainly used human PS1, and to some extent human PS2. However, the topology of both human and worm PSs is considered to be the same.

Recently, a new family of presenilin-like (PSL) proteases (also called presenilin-homologs, PSH; intramembrane proteases, IMPAS) has emerged [25–27]. In humans there are five members of this family, the best known of which is the SPP (signal peptide peptidase). These proteins were discovered due to their sequence homology with PSs. Although the overall sequence homology between the PSs and PSLs is low, they share the conserved aspartate residues that are presumed to be the active site, and the conserved PAL sequence in the C-terminus. Indeed, SPP has been identified as an aspartyl protease [27]. The topology of the PSL proteins has been predicted to be a nine-TM topology with the N-terminus in the extracytosolic space and the C-terminus in the cytosol [28,29]. The aspartate residues are located in TM regions six and seven. This topology has recently been verified experimentally [30].

When comparing the performance of topology prediction methods and low-resolution experiments (such as gene fusion studies and antibody experiments), it has been shown that most predictors are not significantly less accurate than low-resolution experiments [31]. The conflicting results from published studies on PS topology illustrate the limitations of low-resolution experiments. In this study we have used computational prediction methods to analyze PS topology, and have combined these results with previous data from low-resolution experiments and functional data on  $\gamma$ -secretase to produce a novel topology model that is well supported by both prediction methods and experimental results. We propose a novel nine-TM topology model with the C-terminus in the extracytosolic space and the two putative active site aspartate residues in TM regions six and seven, respectively.

## Results and Discussion

All of the experimentally inferred topology models agree that HRI–VI are indeed TM regions [18–24]. A majority of models agree on the localization of the N-terminus and the loops between HRI and HRVI [18,21–24]. Only one model suggests that the N-terminus is located in the extracytosolic space [19,20], leading to opposite localization of the loops between HRI and HRVI compared to the other models. Accumulations of contradictory results are found in the C-terminal part, and consequently several models have been proposed.

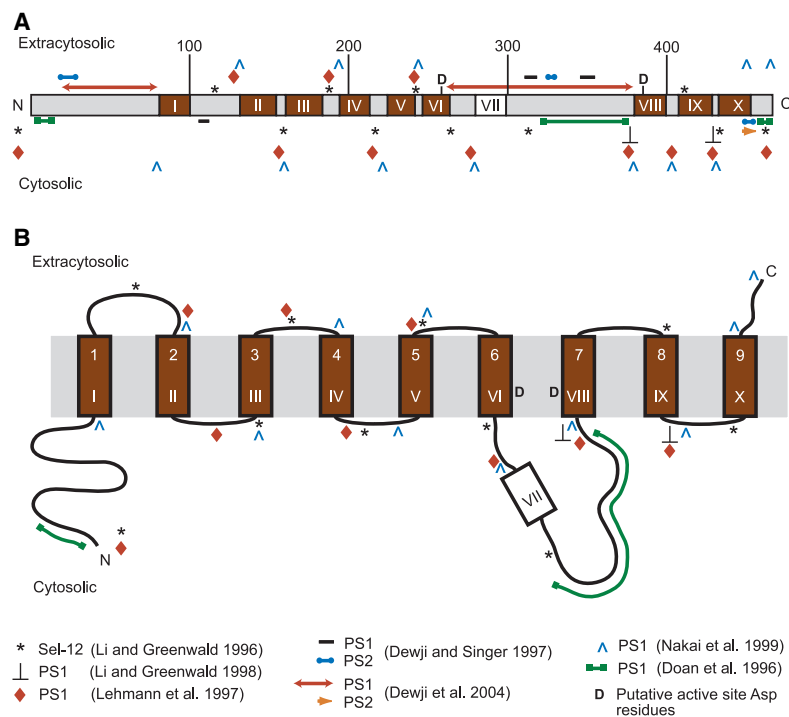


**Fig. 1.** Output from *sFINX* for human presenilin 1 and presenilin-like protein 2. (A) Presenilin 1 (PS1) (UniProt number P49768) has 10 hydrophobic regions (HRs) (indicated in Roman numerals). All topology predictors predict that HRX is a transmembrane (TM) region. A majority of the predictors do not predict HRVIII to be a TM region, because its hydrophobicity is relatively weak. However, when using Phobius constrained prediction with the two putative active site aspartate residues constrained to be in a TM segment, Phobius predicts a nine-TM model with the C-terminus in the extracytosolic space. The aspartate residues are located in TM regions six and seven (D257 and D385 in PS1, indicated with 'D'). This topology is also predicted by HMMTOP2.1 without using constrained prediction. (B) Presenilin-like protein 2 (PSL2) (UniProt number Q8TCT8) has nine HRs (indicated in Roman numerals), and a putative N-terminal signal peptide. All topology predictors predict the same topology, nine TM regions with the C-terminus in the cytosol. The putative active site aspartate residues are located in TM regions six and seven (D351 and D412 in PSL2, indicated with 'D'). Phobius is the only membrane topology predictor that also can predict signal peptides. Other predictors can mistake the signal peptide for a TM region, as MEMSAT does in this example. The conserved PAL motif is indicated with 'PAL' (residues 433–435 in PS1, 463–465 in PSL2).

We analyzed the PS sequences with five different predictors using the *sFINX* tool [32]. Figure 1A shows the output for human PS1. The results for the other PS sequences were essentially the same (not shown). The major insight is that HRX is probably a TM region, which is not the case in most previously published models, including the most widely accepted eight-TM model [21,22]. A minority of the methods predicts that HRVIII is a TM region, when used with default settings. HMMTOP2.1 considers HRVIII to be a TM region, whereas MEMSAT predicts both HRVII and HRVIII to be TM regions. However, the hypothesis that PS is an aspartyl protease performing intramembrane proteolysis of its substrates strongly suggests that the two putative active site aspartate residues are probably located in TM regions. Therefore, we carried out Phobius' [33] constrained prediction with both aspartate residues constrained to be in TM segments. The result was a nine-TM topology with the C-terminus in the extracytosolic space (Figs 1A and 2B). This

is identical to the topology predicted by HMMTOP2.1. This topology was essentially consistent for all PS sequences analyzed.

Our model is supported by 70% of the experimentally determined loop locations (Fig. 2B). The main conflicts between our model and those proposed previously concern HRX and to some extent HRVIII. There are several grounds for HRVIII being a TM region. The most compelling evidence comes from studies showing that the two conserved aspartate residues in HRVI and HRVIII indeed seem to be the active site [8–10]. The proteolysis that these residues perform is intramembrane, implying that they probably are located in TM regions. The relatively weak hydrophobicity of HRVIII can be explained by this embedded catalytic moiety. Furthermore, the membrane topology of the PSL family supports that the putative active site aspartate residues are located in adjacent TM regions (Fig. 1B) [28–30]. Finally, one of the previously published PS topology models supports



**Fig. 2.** Presenilin topology models. (A) Summary of experimental data from low-resolution studies on presenilin (PS) topology. Different regions of PS or single residues have been experimentally determined to be either in the extracytosolic or cytosolic space. The protein sequence shown is human PS1 (UniProt number P49768). When human PS2 or Sel-12 (UniProt numbers P49810 and P52166, respectively) had been analyzed, the positions were correlated to human PS1. (B) Our proposed nine-transmembrane (TM) topology. The TM regions correspond to hydrophobic region (HR) I–VI and HRVIII–X. The N-terminus and the long loop are in the cytosol, whereas the C-terminus is located in the extracytosolic space. The putative active site aspartate residues are located in TM regions six and seven (indicated with 'D'). HRs are indicated in Roman numerals.

HRVIII as a TM region [21,22]. In our model, as well as in the latter, the orientation of the PS active site is inverted compared to the PSL. A rationale for this is that the PS substrates are of inverse orientation compared to PSL substrates; they are type I and type II membrane proteins, respectively [1,2].

There are several reasons why we believe that HRX is a TM region. Undoubtedly, the membrane topology model for the PSL proteins [28–30] supports HRX as a TM region (Fig. 1B). The PAL motif that is conserved between the PS and PSL is located in the proximity of, or within, HRX. In addition, HRX is a strongly hydrophobic region and all the predictors consider it to be a TM region. The PAL sequence in the C-terminus has been shown to be essential for  $\gamma$ -secretase activity [11–13,15,16]. In fact, it may constitute part of the active site, as mutations in these residues exhibit the same phenotypes as mutations in the aspartate residues. This implies that the PAL sequence might be located in a TM region. The C-terminus has also been shown to interact with a number of other proteins, which could give an indication as to whether it is located in the cytosol or extracytosolic space. One of the members of the  $\gamma$ -secretase complex, nicastrin, has been shown to bind at the C-terminus of PS [13]. The PS binding site in nicastrin is likely to be located in its TM region, suggesting that the PS C-terminus dives into or penetrates the membrane. Telencephalin

(TLN) and APP have also been shown to interact with PS. Once again, PS binds with its C-terminus to the TM domain of TLN and APP [34]. Furthermore, the topology model proposed by Nakai *et al.* supports HRX as a TM region [24].

In the model proposed here, only approximately 14 amino acid residues in the C-terminus are located in the extracytosolic space. Considering the available experimental data and the fact that these residues are relatively hydrophobic, it is possible that the extreme C-terminus might be buried in the  $\gamma$ -secretase complex. The PS C-terminus has also been shown to be required for ER retention [13]. The sequence identified by Kaether *et al.* consists of 22 amino acids that include the PAL motif and HRX, and it is well conserved between species. As noted by Kaether *et al.* this sequence is not similar to any known cytoplasmic ER retention signal. However, subunits of multi-protein complexes possess retention signals in their TM domains, which do not have the classical three-to-four amino acid motif. These retention signals prevent the export of unassembled subunits from the ER. Once the complex is assembled, the retention signal is masked and the protein complex can be exported. In another study the PS C-terminus was also found to be required for efficient transport [35], however, no details regarding the experiment were given.

Altogether, there are experimental data that support HRX as a TM region. At the same time, there are experimental studies that contradict HRX as a TM region. Only the topology model by Nakai *et al.* supports HRX as a TM region, although Doan *et al.* [18] speculate that HRX might be a TM region without presenting any experimental evidence. A reason why HRX has not been considered a TM segment is that in most gene fusion studies the loop before HRX and the C-terminus were determined to be located in the cytosol. However, as pointed out by Nakai *et al.* this could be due to the nature of the reporter gene used. The C-terminus of PS is relatively hydrophobic and fusion to a reporter gene with a relatively hydrophobic N-terminus could create an artificial TM region at the fusion point, leading to incorrect localization of the reporter gene. The contradicting results regarding the C-terminal location from the antibody studies [18–20] might be an artifact due to the overexpression of PS.

There are also protein interaction experiments that indicate a cytosolic location of the C-terminus. PS has a potential PDZ domain recognition sequence in its extreme C-terminus, and PDZ domain-containing proteins that recognize PS through this motif have been reported [36–39]. Taken together, the experimental data regarding the location of the C-terminus is conflicting. We believe that the topology predictions for PS and the confirmed topology for the PSL proteins, combined with published PS experimental studies, suggest that the C-terminus is located in the extracytosolic space. A possible explanation for the conflicting results is that the C-terminus could be buried in the  $\gamma$ -secretase complex, as discussed previously. Another possibility, although less likely, is that there are two molecular species of PS coexisting in the cell differing in the location of the C-terminus. This would implicate that the C-terminal region could interact with factors in both the cytosol and the extracytosolic space.

The antibody studies by Dewji *et al.* [19,20] have the most conflicts with our model. They propose a seven-TM topology model with the N-terminus and the long loop between HRVII and HRVIII located in the extracytosolic space, and consequently the C-terminus in the cytosol. The localization of the N-terminus and the long loop not only conflicts with our model, but also with all other experimental studies. The proposition by Dewji *et al.* [19] that PSs are G protein-coupled receptors (GPCRs) and have a seven-TM topology with an extracytosolic N-terminus is unlikely due to several reasons. First, there are eight very hydrophobic regions in the PS sequence (HRI–VI and HRIX–X). Second, the N-terminus is probably located in the cytosol as shown by four experimental studies [18,21–24] and our

analysis. Third, the PSs do not exhibit topological characteristics commonly found in different subfamilies of known GPCRs. For example, loop six (extracytosolic loop three) in the Dewji *et al.* model is very long (approximately 140 residues), which has not been found in GPCRs [40]. Also, the first cytosolic loop is longer than those usually found in GPCRs.

In summary, we propose a novel nine-TM topology for the PSs with the C-terminus located in the extracytosolic space (Fig. 2B). The nine TM regions correspond to HRI–VI and HRVIII–X. The putative active site aspartate residues are located in TM regions six and seven in our model. It is impossible to decide on a single topology model that agrees with all published experimental data on PSs. Determining their topology is critical for understanding their normal and pathogenic functions. However, the ultimate solution to the topology of PSs can only be achieved through atomic resolution studies of the whole  $\gamma$ -secretase complex. Meanwhile, we offer an alternative topology model of PSs, reconciling the previously published experimental data with results from our topology prediction analysis.

## Experimental procedures

The PS sequences in the full alignment from the Pfam database, version 16.0, were analyzed, excluding sequence fragments. Membrane topology for each sequence was predicted with five different methods, and the sfinx tool [32] (<http://sfinx.cgb.ki.se/>) was used to display the results. The predictors were Phobius [33], TMHMM2.0 [41], PHDhtm [42], HMMTOP2.1 [43] and MEMSAT [44]. In addition, a Kyte–Doolittle hydrophobicity curve [45] was constructed for each PS sequence. PHOBIUS also predicts N-terminal signal peptides. Each program was used with default settings. The sequences were also analyzed using PHOBIUS constrained predictions with the two putative active site aspartate residues constrained to be in TM segments.

## Acknowledgements

This work was supported by a grant from the Swedish Knowledge Foundation and Pfizer Corporation to A.H., and L.K., and by grants from Pfizer Corporation to E.S.

## References

- 1 Xia W & Wolfe MS (2003) Intramembrane proteolysis by presenilin and presenilin-like proteases. *J Cell Sci* **116**, 2839–2844.
- 2 Martoglio B & Golde TE (2003) Intramembrane-cleaving aspartic proteases and disease: presenilins, signal

- peptide peptidase and their homologs. *Hum Mol Genet* **12 Spec 2**, R201–R206.
- 3 Kimberly WT & Wolfe MS (2003) Identity and function of gamma-secretase. *J Neurosci Res* **74**, 353–360.
  - 4 Fraering PC, Ye W, Strub JM, Dolios G, LaVoie MJ, Ostaszewski BL, van Dorsselaer A, Wang R, Selkoe DJ & Wolfe MS (2004) Purification and characterization of the human gamma-secretase complex. *Biochemistry* **43**, 9774–9789.
  - 5 Steiner H, Romig H, Pesold B, Philipp U, Baader M, Citron M, Loetscher H, Jacobsen H & Haass C (1999) Amyloidogenic function of the Alzheimer's disease-associated presenilin 1 in the absence of endoproteolysis. *Biochemistry* **38**, 14600–14605.
  - 6 Thinakaran G, Borchelt DR, Lee MK, Slunt HH, Spitzer L, Kim G, Ratovitsky T, Davenport F, Nordstedt C, Seeger M *et al.* (1996) Endoproteolysis of presenilin 1 and accumulation of processed derivatives *in vivo*. *Neuron* **17**, 181–190.
  - 7 Jacobsen H, Reinhardt D, Brockhaus M, Bur D, Kocyba C, Kurt H, Grim MG, Baumeister R & Loetscher H (1999) The influence of endoproteolytic processing of familial Alzheimer's disease presenilin 2 on abeta42 amyloid peptide formation. *J Biol Chem* **274**, 35233–35239.
  - 8 Li YM, Xu M, Lai MT, Huang Q, Castro JL, DiMuzio-Mower J, Harrison T, Lellis C, Nadin A, Neduvetil JG *et al.* (2000) Photoactivated gamma-secretase inhibitors directed to the active site covalently label presenilin 1. *Nature* **405**, 689–694.
  - 9 Esler WP, Kimberly WT, Ostaszewski BL, Diehl TS, Moore CL, Tsai JY, Rahmati T, Xia W, Selkoe DJ & Wolfe MS (2000) Transition-state analogue inhibitors of gamma-secretase bind directly to presenilin-1. *Nat Cell Biol* **2**, 428–434.
  - 10 Wolfe MS, Xia W, Ostaszewski BL, Diehl TS, Kimberly WT & Selkoe DJ (1999) Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature* **398**, 513–517.
  - 11 Bergman A, Laudon H, Winblad B, Lundkvist J & Naslund J (2004) The extreme C terminus of presenilin 1 is essential for gamma-secretase complex assembly and activity. *J Biol Chem* **279**, 45564–45572.
  - 12 Wang J, Brunkan AL, Hecimovic S, Walker E & Goate A (2004) Conserved 'PAL' sequence in presenilins is essential for gamma-secretase activity, but not required for formation or stabilization of gamma-secretase complexes. *Neurobiol Dis* **15**, 654–666.
  - 13 Kaether C, Capell A, Edbauer D, Winkler E, Novak B, Steiner H & Haass C (2004) The presenilin C-terminus is required for ER-retention, nicastrin-binding and gamma-secretase activity. *Embo J* **23**, 4738–4748.
  - 14 Tomita T, Takikawa R, Koyama A, Morohashi Y, Takasugi N, Saido TC, Maruyama K & Iwatsubo T (1999) C terminus of presenilin is required for overproduction of amyloidogenic Abeta42 through stabilization and endoproteolysis of presenilin. *J Neurosci* **19**, 10627–10634.
  - 15 Tomita T, Watabiki T, Takikawa R, Morohashi Y, Takasugi N, Kopan R, De Strooper B & Iwatsubo T (2001) The first proline of PALP motif at the C terminus of presenilins is obligatory for stabilization, complex formation, and gamma-secretase activities of presenilins. *J Biol Chem* **276**, 33273–33281.
  - 16 Takasugi N, Takahashi Y, Morohashi Y, Tomita T & Iwatsubo T (2002) The mechanism of gamma-secretase activities through high molecular weight complex formation of presenilins is conserved in *Drosophila melanogaster* and mammals. *J Biol Chem* **277**, 50198–50205.
  - 17 Chen Q & Schubert D (2002) Presenilin-interacting proteins. *Expert Rev Mol Med* **2002**, 1–18.
  - 18 Doan A, Thinakaran G, Borchelt DR, Slunt HH, Ratovitsky T, Podlisny M, Selkoe DJ, Seeger M, Gandy SE, Price DL & Sisodia SS (1996) Protein topology of presenilin 1. *Neuron* **17**, 1023–1030.
  - 19 Dewji NN, Valdez D & Singer SJ (2004) The presenilins turned inside out: implications for their structures and functions. *Proc Natl Acad Sci USA* **101**, 1057–1062.
  - 20 Dewji NN & Singer SJ (1997) The seven-transmembrane spanning topography of the Alzheimer disease-related presenilin proteins in the plasma membranes of cultured cells. *Proc Natl Acad Sci USA* **94**, 14025–14030.
  - 21 Li X & Greenwald I (1996) Membrane topology of the *C. elegans* SEL-12 presenilin. *Neuron* **17**, 1015–1021.
  - 22 Li X & Greenwald I (1998) Additional evidence for an eight-transmembrane-domain topology for *Caenorhabditis elegans* and human presenilins. *Proc Natl Acad Sci USA* **95**, 7109–7114.
  - 23 Lehmann S, Chiesa R & Harris DA (1997) Evidence for a six-transmembrane domain structure of presenilin 1. *J Biol Chem* **272**, 12047–12051.
  - 24 Nakai T, Yamasaki A, Sakaguchi M, Kosaka K, Miura K, Amaya Y & Miura S (1999) Membrane topology of Alzheimer's disease-related presenilin 1. Evidence for the existence of a molecular species with a seven membrane-spanning and one membrane-embedded structure. *J Biol Chem* **274**, 23647–23658.
  - 25 Grigorenko AP, Moliaka YK, Korovaitseva GI & Rogae EI (2002) Novel class of polytopic proteins with domains associated with putative protease activity. *Biochemistry (Mosc)* **67**, 826–835.
  - 26 Ponting CP, Hutton M, Nyborg A, Baker M, Jansen K & Golde TE (2002) Identification of a novel family of presenilin homologues. *Hum Mol Genet* **11**, 1037–1044.
  - 27 Weihofen A, Binns K, Lemberg MK, Ashman K & Martoglio B (2002) Identification of signal peptide peptidase, a presenilin-type aspartic protease. *Science* **296**, 2215–2218.

- 28 Grigorenko AP, Moliaka YK, Soto MC, Mello CC & Rogae EI (2004) The *Caenorhabditis elegans* IMPAS gene, *imp-2*, is essential for development and is functionally distinct from related presenilins. *Proc Natl Acad Sci USA* **101**, 14955–14960.
- 29 Henricson A, Sonnhammer EL, Baillie DL & Gomes AV (2004) Functional characterization in *Caenorhabditis elegans* of transmembrane worm-human orthologs. *BMC Genomics* **5**, 85.
- 30 Friedmann E, Lemberg MK, Weihofen A, Dev KK, Dengler U, Rovelli G & Martoglio B (2004) Consensus analysis of signal peptide peptidase and homologous human aspartic proteases reveals opposite topology of catalytic domains compared with presenilins. *J Biol Chem* **279**, 50790–50798.
- 31 Chen CP, Kernysky A & Rost B (2002) Transmembrane helix predictions revisited. *Protein Sci* **11**, 2774–2791.
- 32 Sonnhammer EL & Wootton JC (2001) Integrated graphical analysis of protein sequence features predicted from sequence composition. *Proteins* **45**, 262–273.
- 33 Kall L, Krogh A & Sonnhammer EL (2004) A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* **338**, 1027–1036.
- 34 Annaert WG, Esselens C, Baert V, Boeve C, Snellings G, Cupers P, Craessaerts K & De Strooper B (2001) Interaction with telencephalin and the amyloid precursor protein predicts a ring structure for presenilins. *Neuron* **32**, 579–589.
- 35 Nufer O, Gulbrandsen S, Degen M, Kappeler F, Paccaud JP, Tani K & Hauri HP (2002) Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. *J Cell Sci* **115**, 619–628.
- 36 Gupta S, Singh R, Datta P, Zhang Z, Orr C, Lu Z, Dubois G, Zervos AS, Meisler MH, Srinivasula SM *et al.* (2004) The C-terminal tail of presenilin regulates Omi/HtrA2 protease activity. *J Biol Chem* **279**, 45844–45854.
- 37 Ni CY, Yuan H & Carpenter G (2003) Role of the ErbB-4 carboxyl terminus in gamma-secretase cleavage. *J Biol Chem* **278**, 4561–4565.
- 38 Lau KF, McLoughlin DM, Standen C & Miller CC (2000) X11 alpha and x11 beta interact with presenilin-1 via their PDZ domains. *Mol Cell Neurosci* **16**, 557–565.
- 39 Xu X, Shi Y, Wu X, Gambetti P, Sui D & Cui MZ (1999) Identification of a novel PSD-95/Dlg/ZO-1 (PDZ)-like protein interacting with the C terminus of presenilin-1. *J Biol Chem* **274**, 32543–32546.
- 40 Inoue Y, Ikeda M & Shimizu T (2004) Proteome-wide classification and identification of mammalian-type GPCRs by binary topology pattern. *Comput Biol Chem* **28**, 39–49.
- 41 Krogh A, Larsson B, von Heijne G & Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567–580.
- 42 Rost B, Fariselli P & Casadio R (1996) Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci* **5**, 1704–1718.
- 43 Tusnady GE & Simon I (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**, 849–850.
- 44 Jones DT, Taylor WR & Thornton JM (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33**, 3038–3049.
- 45 Kyte J & Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105–132.

## Supplementary material

The following material is available from <http://www.blackwell-publishing.com/products/journals/suppmat/EJB/EJB4691/EJB4691sm.htm>

**Table S1.** Summary of experimental studies on presenilin topology.

**Table S2.** Results from Phobius constrained prediction.