

- 0.28 $\pm$ 0.03 (19). The measurements of α using cometary nuclei need to be viewed with caution, because they are observationally challenging, and so the measurements could have problems. For example, measurements of individual comets could have been performed when there was still some small amount of cometary activity—too small to resolve but large enough to affect the photometry. In addition, all objects studied in these surveys are bright active comets when observed near perihelion, which could introduce biases in the way they were selected. What is clearly needed is a population of dormant objects whose members were discovered in a systematic way. The data set studied here supplies us with such a population.
24. Levison et al. assume that active comets fade in brightness over many orbits after they reach $q < 2.5$ AU. However, they do not distinguish between close (say, $q \sim 0.5$ AU) and distant (say, $q \sim 2.5$ AU) passages. We believe, however, that closer perihelion passages lead to increased damage to comets. For this reason, we have added an additional disruption law to Levison et al.'s models. This simple law is designed to mimic the q dependence of fading, while making the fewest changes to the original Levison et al.'s models. If an object evolves onto an orbit with $q < 1$ AU, we assume that it has a 96% chance of disrupting before its next perihelion passage (see SOM). If the comet disrupts, it will not appear in either the active comet or dormant comet populations.
 25. Fortunately, the $a-i$ distribution (i.e., semi-major axis and inclination) of our models is not sensitive to α , so we can perform the absolute magnitude fitting procedure on our best-fit model alone. In particular, we do not need to perform a calculation that varies the four parameters from (13) and α , while fitting to a , i , and H simultaneously.
 26. To do this comparison, we use arguments developed in (13) for relating the observed number of dynamically new comets per year to the total number of HTCs.
 27. There are 22 known active HTCs with $q < 1.3$ AU (a limit set by observational biases), although not all HTCs with $q < 1.3$ AU have yet been discovered. We can predict the total number of active HTCs with our dynamical models. Our best-fit model for the HTCs predicts 84 active HTCs with $q < 1.3$ AU. This value, in turn, implies that there are 194 active HTCs with $q < 3$ AU.
 28. JFCs are comets with Tisserand parameters $2 < T \leq 3$ [see (4, 14)]. These objects are believed to arise in the Kuiper belt (32, 33) or the scattered disk (34)—not in the Oort cloud.
 29. W. F. Bottke Jr. et al., *Icarus* **156**, 399 (2002).
 30. From results given in (48), we estimate that there are 30_{-10}^{+10} active JFCs with $H < 18$ and $q < 1.3$ AU. From (33), for each active comet, there should be between 2.0 and 6.7 dormant comets, with a best-fit value of 3.5. This assumes no disruption. Thus, we estimate that the total number of dormant JFCs with $H < 18$ and $q < 1.3$ AU is $30 \times 3.5 = 105$, with a possible range of 50 to 268, if all inactive comets become dormant. Given Bottke's (29) estimate of dormant JFCs from the NEO surveys (61 with $H < 18$), our best estimate is that 61/105, or $\sim 60\%$, of the JFCs become dormant, as opposed to $\leq 1\%$ of the NICs. Formally, the fraction of JFCs that become dormant could be as small as 7% or as large as 100%. Even the lower limit is about an order of magnitude larger than our upper limit for the fraction of NICs that become dormant.
 31. M. Duncan, T. Quinn, S. Tremaine, *Astron. J.* **94**, 1330 (1987).
 32. ———, *Astrophys. J. Lett.* **328**, 69 (1988).
 33. H. F. Levison, M. J. Duncan, *Icarus* **127**, 13 (1997).
 34. M. J. Duncan, H. F. Levison, *Science* **276**, 1670 (1997).
 35. L. Dones, H. F. Levison, M. J. Duncan, P. R. Weissman, *Am. Astron. Soc. Div. Dynamical Astron. Meeting* **32**, 1103 (2001).
 36. The Kuiper belt is currently collisionally active (50), whereas the Oort cloud is not.
 37. M. Bailey, *Mon. Not. R. Astron. Soc.* **221**, 247 (1984).
 38. J. G. Hills, *Astron. J.* **86**, 1730 (1981).
 39. F. L. Whipple, *Astron. J.* **67**, 1 (1962).

40. G. D. Brin, D. A. Mendis, *Astrophys. J.* **229**, 402 (1979).
41. D. Priyalnik, A. Bar-Nun, *Icarus* **74**, 272 (1988).
42. H. U. Keller et al., *Nature* **321**, 320 (1986).
43. L. Soderblom et al., *Science* **296**, 1087 (2002).
44. H. Boehnhardt, *Science* **292**, 1307 (2001).
45. J. Chambers, *Icarus* **125**, 32 (1997).
46. Y. R. Fernández, D. C. Jewitt, S. S. Sheppard, *Astrophys. J. Lett.* **553**, L197 (2001).
47. E. M. Shoemaker, R. F. Wolfe, in *Satellites of Jupiter*, D. Morrison, Ed. (Univ. Arizona Press, Tucson, AZ, 1982), pp. 277–339.
48. J. A. Fernández, G. Tancredi, H. Rickman, J. Licandro, *Astron. Astrophys.* **352**, 327 (1999).
49. S. C. Lowry, thesis, Queen's University, Belfast, UK (2001).
50. S. A. Stern, J. E. Colwell, *Astrophys. J.* **490**, 879 (1997).
51. We thank M. Duncan, S. A. Stern, and P. Weissman for useful discussions and comments on an early draft of this paper; and M. Bailey for useful comments on this manuscript. H.L., W.B., and L.D. acknowledge support from NASA's Planetary Geology and Geophysics

(NAG5-9479, NAG5-10331, NAG5-9141), Origins of Solar Systems (NAG5-10661), NEO-Observations (NAG5-9951), and Mars-Data Analysis (NAG5-10603) programs. A.M. was subsidized by European Space Agency contract 14018/2000/F/TB. Travel support provided by grants from NATO and NSF/Centre National de la Recherche Scientifique. Observations by Spacewatch and research by R.J. supported by grants from NASA (NAG5-7854, NAG5-10447, NAG5-7533), the Air Force Office of Scientific Research (F49620-00-1-0126), the David and Lucile Packard Foundation, the Steven and Michele Kirsch Foundation, the Paul G. Allen Charitable Foundation, and other contributors.

Supporting Online Material

www.sciencemag.org/cgi/content/full/296/5576/2212/DC1

Materials and Methods
Figs. S1 and S2

References and Notes

25 January 2002; accepted 3 April 2002

Identification of Signal Peptide Peptidase, a Presenilin-Type Aspartic Protease

Andreas Weihofen,¹ Kathleen Binns,² Marius K. Lemberg,¹ Keith Ashman,² Bruno Martoglio^{1*}

Signal peptide peptidase (SPP) catalyzes intramembrane proteolysis of some signal peptides after they have been cleaved from a preprotein. In humans, SPP activity is required to generate signal sequence–derived human lymphocyte antigen–E epitopes that are recognized by the immune system, and to process hepatitis C virus core protein. We have identified human SPP as a polytopic membrane protein with sequence motifs characteristic of the presenilin-type aspartic proteases. SPP and potential eukaryotic homologs may represent another family of aspartic proteases that promote intramembrane proteolysis to release biologically important peptides.

The discovery of intramembrane proteolysis has revealed alternative pathways in cell signaling, cell regulation, and protein processing (1). Dormant, membrane-bound transcription factors, like sterol regulatory element-binding protein (1), activating transcription factor-6 (2), and NOTCH (3), or the growth factor Spitz in *Drosophila* (4), are activated and liberated in regulated processes that culminate in proteolytic cleavage within their membrane anchor. Similarly, β -amyloid (A β) peptides, which are believed to be the main toxic component in Alzheimer's disease, are generated from membrane-anchored β -amyloid precursor protein (β -APP) (5). The critical cleavage in the membrane anchor of β -APP is thought to be catalyzed by the aspartic protease presenilin (6).

Processing of signal peptides by an SPP is related to protein cleavage by presenilin. Both proteases cleave their substrates within the center of a transmembrane region (6, 7). The discovery of posttargeting functions of signal peptides, which are required primarily for the biosynthesis of secretory and membrane proteins, has pointed to a central role for SPP activity (8). Generation of cell surface histocompatibility antigen (HLA)–E epitopes in humans requires processing of signal peptides by SPP (9). HLA-E epitopes originate from the signal sequence of polymorphic major histocompatibility complex (MHC) class I molecules and report biosynthesis of these molecules to the immune system (10). SPP activity is also required for processing hepatitis C virus polyprotein and hence is exploited by the pathogen to produce viral components (11). It is thought that SPP promotes the liberation of functional signal peptide fragments from the endoplasmic reticulum (ER) membrane (8).

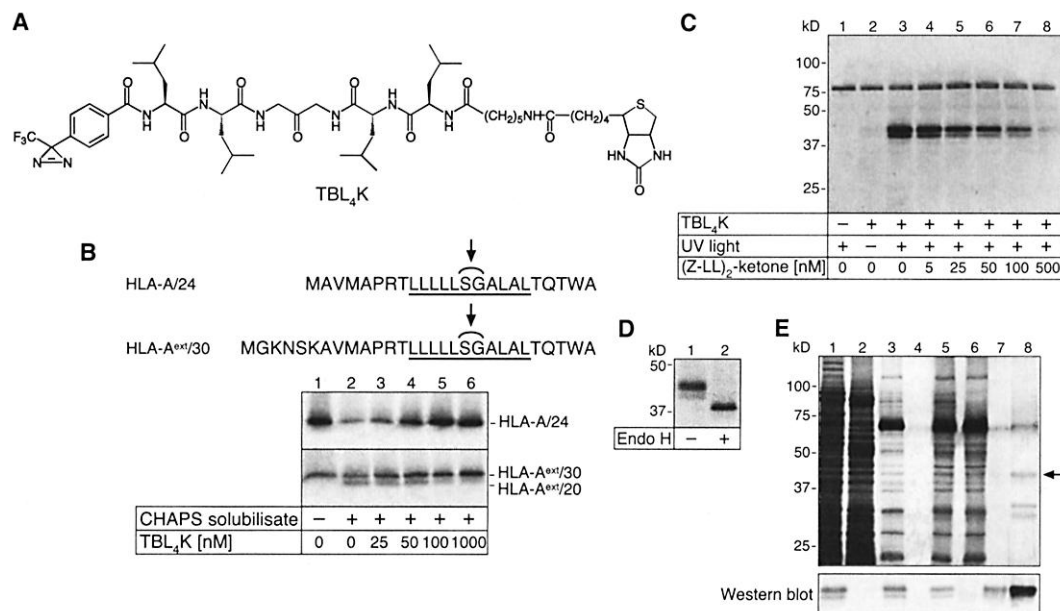
To identify human SPP, we synthesized a ligand affinity probe based on the SPP inhibitor (Z-LL)₂–ketone, which is thought to re-

¹Institute of Biochemistry, Swiss Federal Institute of Technology (ETH), ETH-Hoenggerberg, 8093 Zürich, Switzerland. ²Samuel Lunenfeld Institute, Proteomics, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada.

*To whom correspondence should be addressed. E-mail: bruno.martoglio@bc.biol.ethz.ch

REPORTS

Fig. 1. Inhibition of SPP activity and photolabeling with TBL₄K. (A) Diagram of TBL₄K. (B) TBL₄K inhibits signal peptide processing. Amino acid sequences of substrates HLA-A/24 and HLA-A^{ext}/30. Arrows indicate approximate SPP cleavage sites in the transmembrane region (underlined). HLA-A^{ext}/20, cleavage product of HLA-A^{ext}/30. Substrates were incubated in the presence of detergent (CHAPS)-solubilized SPP activity (lanes 2 to 6) or buffer only (lane 1). TBL₄K was added at indicated concentrations. (C) TBL₄K labels a 42-kD protein. CHAPS-solubilized ER membrane proteins were incubated with 50 nM TBL₄K (lanes 2 to 8) and (Z-LL)₂-ketone at indicated concentrations (lanes 4 to 8). Samples were irradiated with UV light (lanes 1 and 3 to 8) to activate TBL₄K. (D) TBL₄K-labeled protein is N-glycosylated. After labeling with TBL₄K, samples were treated with endoglycosidase H (lane 2) or buffer only (lane 1). (E) Monitoring of purification by SDS-polyacrylamide gel electrophoresis–silver staining and Western blot analysis with biotin-specific antibody. Lane 1, CHAPS-solubilized ER membrane proteins;



lanes 2 and 3, Concanavalin A-Sepharose flow-through and eluate pool; lanes 4 and 5, hydroxyapatite flow-through and eluate pool; lanes 6 and 7, reversed-phase flow-through and peak fraction; lane 8, reversed-phase peak fraction, but five times the equivalents loaded in lane 7. The arrow indicates the position of the TBL₄K-labeled protein.

versibly attack an active-site residue of the protease with its central ketone moiety (7). We synthesized a diazirine-containing derivative of (Z-LL)₂-ketone, TBL₄K (Fig. 1A), because the photo-reactive compound might also be directed to the active site of the protease but irreversibly bind on activation with ultraviolet (UV) light (12). A biotin moiety was also incorporated into the probe to facilitate identification of the adduct (13).

The effect of TBL₄K on SPP activity was examined in vitro with detergent-solubilized SPP activity isolated from canine pancreas ER membranes; the in vitro-synthesized substrates HLA-A/24 corresponding to the signal peptide of HLA-A*0301; and HLA-A^{ext}/30, a mutant signal peptide containing an extended NH₂-region that allowed fixation of the cleavage product (HLA-A^{ext}/20) on the gel (9, 13) (Fig. 1B). Addition of TBL₄K to reactions inhibited cleavage of the peptides in a dose-dependent manner. The apparent median inhibitory concentration (IC₅₀) value (~50 nM) and the concentration that abolished SPP activity (1 μM) were within the range for inhibition of SPP by (Z-LL)₂-ketone (7), indicating that modifications of TBL₄K did not affect targeting to SPP.

To identify protein(s) to which TBL₄K binds, we mixed detergent-solubilized ER membrane proteins with the inhibitor and then exposed the mixture to UV light (13). Western blot analysis with a biotin-specific antibody revealed a major and a minor product of about 42 and 40 kD, respectively (Fig. 1C). Specificity of TBL₄K for these species was confirmed by competition with (Z-LL)₂-

ketone. Endoglycosidase H treatment shifted both TBL₄K-labeled proteins into a single band of ~38 kD, suggesting that they are differentially glycosylated forms of the same protein (Fig. 1D).

To identify the 42-kD species, we performed photolabeling with TBL₄K on a preparative scale and purified the labeled protein by chromatography (Fig. 1E) (13). At each purification step, the presence of TBL₄K-labeled protein was confirmed by Western blot analysis with a biotin-specific antibody. Analysis of the final pooled fractions revealed five proteins (Fig. 1E, lane 8). The TBL₄K-labeled protein was identified by Western blot analysis, excised from a preparative gel, and analyzed by mass spectrometry. Sequences were obtained for six peptides that were compared with predicted translated products in the National Center for Biotechnology Information non-redundant and Expressed Sequence Tags databases (13). With the exception of two amino acid residues, all of the peptide sequences matched a hypothetical human protein of unknown function.

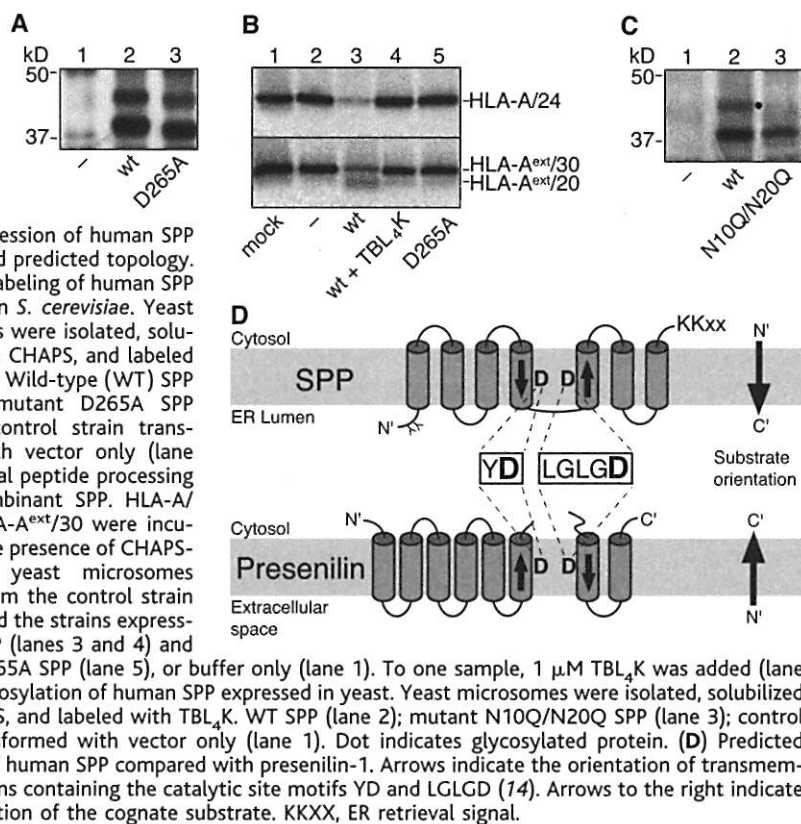
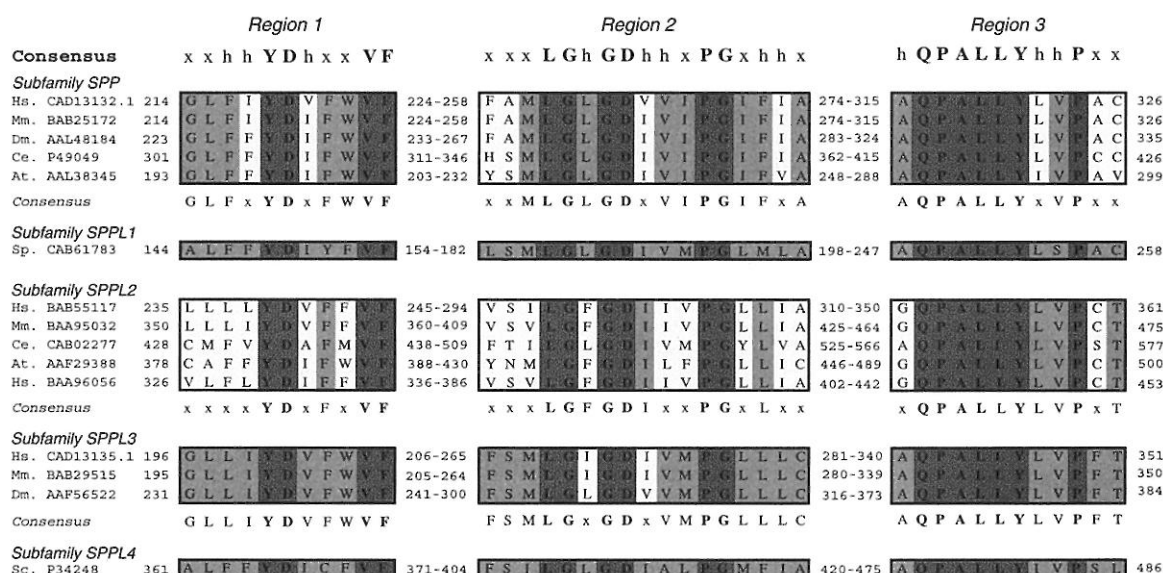
In the absence of homology to other proteins, we screened sequence databases for homologous genes in other species. Potential orthologs of the identified protein were found in higher eukaryotes (Fig. 2) (13), but no functions have yet been described for any of these hypothetical proteins. The most conserved regions contain the motifs YD and LGLGD (14) within two putative transmembrane segments (Fig. 2, regions 1 and 2). Such motifs are characteristic of the presenilin type of aspartic proteases (15, 16). In

presenilin, the two aspartic acid residues within these motifs are required for proteolytic activity and are thought to reside in the protease active site. Apart from these motifs, there is no particular homology between presenilins and the putative SPP. Also, no other potentially conserved catalytic residues such as serine, histidine, or cysteine were found.

Predicted amino acid sequences of more than 15 proteins of unknown function were homologous to human SPP and could be subdivided into at least five subfamilies on the basis of phylogenetic tree analysis (ClustalW 1.4) (Fig. 2). The first subfamily comprises proteins with a COOH-terminal ER-retrieval signal, KKXX (14), and includes SPP. Members of this subfamily are potential orthologs and share homology (up to 94% identity) with SPP throughout the whole amino acid sequence. They are found only in higher eukaryotes. The other four subfamilies comprise SPP homologues without obvious intracellular localization signals. Also, they are homologous to SPP only in the COOH-terminal half of the protein and show substantial variation in the NH₂-terminal regions. This result suggests that the COOH-terminal portions containing the YD and LGLGD motifs and a highly conserved sequence [QPALLYhhP (14); Fig. 2, region 3] form the proteolytic (sub)domain, whereas the NH₂-terminal parts define the specific function of the respective proteins.

Potential orthologs of SPP were not found in the genomes of *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, suggesting that SPP function was acquired late in evo-

Fig. 2. Amino acid comparisons of conserved motifs in members of the SPP clan of putative aspartic proteases (14). Multiple alignment of predicted amino acid sequences of human SPP and its homologs across species. Most conserved regions are shown. Shading indicates residues highly conserved in all homologs (dark gray) and those conserved in the subfamilies (light gray). Subfamily SPP, potential SPP orthologs; subfamilies SPPL1–4, SPP-like proteases; h, hydrophobic amino acid; Hs., *Homo sapiens*; Mm., *Mus musculus*; Dm., *Drosophila melanogaster*; Ce., *Caenorhabditis elegans*; At., *Arabidopsis thaliana*; Sp., *S. pombe*; Sc., *S. cerevisiae*. Accession numbers refer to the EMBL/GenBank/DNA Data Bank of Japan database (www.ncbi.nlm.nih.gov/Genbank/index.html).



lution, possibly to gain signal peptide-derived peptides for signaling or regulatory events, because it seems typical for processes that include intramembrane proteolysis (17). To confirm that the identified protein targeted by TBL₄K is an SPP, we expressed the protein in yeast *S. cerevisiae* (13). We cloned and sequenced the human cDNA from a HeLa cell cDNA library with partial ex-

pressed sequence tag sequences from public databases, and expressed the protein encoded by the recovered cDNA [submitted to the European Molecular Biology Laboratory (EMBL) database; accession number AJ420895]. Yeast microsomal membrane proteins were isolated and solubilized with detergent, and associated proteins were either labeled with TBL₄K or tested for protease

activity (13). The TBL₄K-labeled protein appeared as a glycosylated and a nonglycosylated protein, which were not present in control yeast membranes (Fig. 3A). Furthermore, solubilized membrane proteins containing the human protein showed SPP activity that was sensitive to TBL₄K (Fig. 3B). Mutation of the conserved aspartic acid residue at position 265 of the LGLGD motif to alanine did not affect its labeling with TBL₄K, indicating that substrate binding was not disrupted, but catalytic activity was abolished in the mutant (18) (Fig. 3, A and B). These results suggest that the human TBL₄K-binding protein is SPP and likely an aspartic protease.

Analysis of the amino acid sequence of the predicted protein revealed seven putative transmembrane regions, according to the TMHMM 2.0 prediction program (19): four potential N-glycosylation sites, YD and LGLGD aspartic protease motifs, and the ER retrieval signal KKXX. To assess the topology of SPP, we determined its sites of glycosylation. At Asn⁶², which lies in the loop between putative transmembrane regions one and two, glycosylation is unlikely. This residue is part of a peptide identified by mass spectrometry and hence not modified and presumably facing the cytosol. To test the two N-glycosylation sites Asn¹⁰ and Asn²⁰, we mutated these residues to glutamines (N10Q/N20Q). When wild-type (WT) SPP was expressed in yeast, the TBL₄K-labeled protein appeared as two bands corresponding to glycosylated and nonglycosylated forms of the protein. In contrast, expression of the N10Q/N20Q mutant resulted in a single protein with the expected molecular size of nonglycosylated WT SPP, indicating that the NH₂-terminus of SPP had been translocated

into the ER lumen (Fig. 3C). Glycosylation of Asn¹⁰ and Asn²⁰ was also confirmed in vitro (13).

On the basis of the analysis of glycosylation sites and the prediction of transmembrane regions, we propose a seven-transmembrane topology for SPP with the NH₂-terminus in the ER lumen, the COOH-terminus containing the ER retrieval signal in the cytosol, and the active-site motifs YD and LGLGD in the center of adjacent transmembrane regions (Fig. 3D). Such motifs are also present in adjacent transmembrane regions of presenilins (3, 5, 6). However, the predicted orientation of the transmembrane regions containing the YD and LGLGD motifs is opposite in presenilins compared with that of SPP, in accordance with the opposite orientation of the substrates. The substrates of presenilins, NOTCH-1 and β -APP, are type I membrane proteins, whereas SPP substrates have a type II orientation (Fig. 3D).

Genetic evidence supports the explanation that presenilins are γ -secretases, which catalyze cleavage of β -APP in its transmembrane region and liberate A β peptides (6). Recent observations, however, question presenilins as the proteolytic components that cleave β -APP and NOTCH and suggest alternative functions, such as that of a molecular chaperone for membrane proteins (20). The identification of SPP as a presenilin-type aspartic protease favors the former view that presenilins are proteases. The identification of potential SPP homologs may in fact expand the number of potential proteases, which may account for γ -secretase activity in systems that exclude the action of presenilins (20). Identification of functional human SPP may allow elucidation of the mechanism of intramembrane proteolysis and address the still-unsolved question of how the cleavage of peptide bonds can be achieved in an environment that is thought to preclude hydrolysis.

References and Notes

1. M. S. Brown, J. Ye, R. B. Rawson, J. L. Goldstein, *Cell* **100**, 391 (2000).
2. K. Haze, H. Yoshida, H. Yanagi, T. Yura, K. Mori, *Mol. Biol. Cell* **10**, 3787 (1999).
3. Y. M. Chan, Y. N. Jan, *Neuron* **23**, 201 (1999).
4. J. R. Lee, S. Urban, C. F. Garvey, M. Freeman, *Cell* **107**, 161 (2001).
5. D. J. Selkoe, *Nature* **399**, A23 (1999).
6. H. Steiner, C. Haass, *Nature Rev. Mol. Cell Biol.* **1**, 217 (2000).
7. A. Weihofen, M. K. Lemberg, H. L. Ploegh, M. Bogoy, B. Martoglio, *J. Biol. Chem.* **275**, 30951 (2000).
8. B. Martoglio, B. Dobberstein, *Trends Cell Biol.* **8**, 410 (1998).
9. M. K. Lemberg, F. A. Bland, A. Weihofen, V. M. Braud, B. Martoglio, *J. Immunol.* **167**, 6441 (2001).
10. V. M. Braud et al., *Nature* **391**, 795 (1998).
11. J. McLauchlan, M. K. Lemberg, R. G. Hope, B. Martoglio, in preparation.
12. J. Brunner, *Methods Enzymol.* **172**, 628 (1989).
13. Materials and methods are available as supporting

material on Science online at www.sciencemag.org/cgi/content/full/296/5576/2215/DC1.

14. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. X, any amino acid; h, hydrophobic amino acid.
15. M. S. Wolfe et al., *Nature* **398**, 513 (1999).
16. H. Steiner et al., *Nature Cell Biol.* **2**, 848 (2000).
17. S. Huppert, R. Kopan, *Dev. Cell* **1**, 590 (2001).
18. M. Szelke, in *Aspartic Proteinases and Their Inhibitors*, V. Kostka, Ed. (Walter de Gruyter, Berlin, 1985), pp. 421–441.
19. S. Moller, M. D. Croning, R. Apweiler, *Bioinformatics* **17**, 646 (2001).

20. S. S. Sisodia, W. Annaert, S. H. Kim, B. De Strooper, *Trends Neurosci.* **24**, S2 (2001).

21. We thank J. Brunner, B. Dobberstein, A. Helenius, and J. McLauchlan for discussions; C. Jakob for p426gal; and U. Kutay for the cDNA library. Supported by grants from Natural Sciences and Engineering Research Council of Canada Industrial Postgraduate Scholarship with MDS Sciex (K.B.), Eidgenössische Technische Hochschule Zürich, the National Competence Center for Research on Neuronal Plasticity and Repair, and the Swiss National Science Foundation (B.M.).

15 February 2002; accepted 11 April 2002

Visualization and Functional Analysis of RNA-Dependent RNA Polymerase Lattices

John M. Lyle,^{1*} Esther Bullitt,^{2*} Kurt Bienz,³ Karla Kirkegaard^{1†}

Positive-strand RNA viruses such as poliovirus replicate their genomes on intracellular membranes of their eukaryotic hosts. Electron microscopy has revealed that purified poliovirus RNA-dependent RNA polymerase forms planar and tubular oligomeric arrays. The structural integrity of these arrays correlates with cooperative RNA binding and RNA elongation and is sensitive to mutations that disrupt intermolecular contacts predicted by the polymerase structure. Membranous vesicles isolated from poliovirus-infected cells contain structures consistent with the presence of two-dimensional polymerase arrays on their surfaces during infection. Therefore, host cytoplasmic membranes may function as physical foundations for two-dimensional polymerase arrays, conferring the advantages of surface catalysis to viral RNA replication.

Nucleic acid synthesis is often associated with large, static “factories” (1). For positive-strand RNA viruses such as poliovirus, foot-and-mouth disease virus, hepatitis C virus, and many others, the RNA replication complexes form on the cytosolic surface of cytoplasmic membranes (2–4). The role of this membrane association is not known, but it may be to (i) concentrate and compartmentalize viral products by targeting to a common structure, (ii) provide key lipid constituents to the viral RNA replication complexes, or (iii) physically support the RNA replication complex.

Poliovirus RNA synthesis is catalyzed by a virally encoded RNA-dependent RNA polymerase, termed 3D (5). The 3D polymerase is a soluble enzyme, but it is targeted to membranes by binding to another viral protein, 3AB (6–8), also part of the RNA replication complex. Poliovirus polymerase forms homo-oligomers, as demonstrated by its cooperative elongation and

RNA-binding activity with respect to protein concentration, facile cross-linking with glutaraldehyde (9, 10), and interactions in a yeast two-hybrid system (7, 11). The three-dimensional crystal structure of the poliovirus polymerase determined by x-ray crystallography (12) revealed two interfaces between polymerase molecules (Fig. 1A). Based on a right-hand metaphor for polymerase structure, Interface I involves residues on the side of the thumb domain and residues on the back of the palm of an adjacent polymerase, forming a head-to-tail oligomeric fiber through the crystal (12). Amino acid substitutions predicted to disrupt interface I are lethal to the virus (13, 14) and reduce RNA-binding affinity (14). Interface II involves intermolecular donation of the NH₂-terminal domain of one polymerase to a region of the thumb near the active site of its neighboring polymerase, forming a head-to-tail fiber through the crystal that intersects the fibers formed by interface I at a 90° angle (Fig. 1A) (12). However, the region of 3D polymerase around interface II is not completely resolved in the x-ray structure (12), and alternative crystal forms reveal that alternative packing conformations around interface II allow polymerase-polymerase interactions in this region, even in the absence of the NH₂-terminal donation (15). For wild-type polymerase, the existence of the intermolecular NH₂-terminal donation shown

¹Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA. ²Department of Physiology and Biophysics, Boston University School of Medicine, Boston, MA 02118, USA. ³Institute for Medical Microbiology, University of Basel, CH-4003 Basel, Switzerland.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: karlak@leland.stanford.edu