BIOCHEMICAL AND STRUCTURAL CHARACTERIZATION OF DI-ASPARTYL INTRAMEMBRANE PROTEASES

by

Celia Mirtha Torres Arancivia

A dissertation submitted to the Graduate Faculty in Biochemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy,

The City University of New York

2011

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Dr. Iban Ubarretxena

_____ _____

Date Chair of Examining Committee

Dr. Edward J. Kennelly

_____ _____

Date Executive Officer

Dr. K.V. Lakshmi

Dr. Ranajeet Ghose

Dr. Horst Schulz

Dr. David Stokes

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

BIOCHEMICAL AND STRUCTURAL CHARACTERIZATION OF DI-ASPARTYL INTRAMEMBRANE PROTEASES

by

Celia Mirtha Torres Arancivia

Adviser: Professor Iban Ubarretxena

A new era in the study of proteases was initiated with the discovery of intramembrane-cleaving proteases (I-CLiPs). In an unusual mechanism, I-CLiPs are able to catalyze peptide bond cleavage of substrates within the membrane bilayer, an environment poorly permeable to water. Here we will focus on γ-secretase and presenilin(PS)-like proteases, membrane proteases that use two aspartic acids as catalytic residues and constitute the most relevant family of I-CLiPs. We first describe the purification of γ-secretase from a mammalian expression system. We used this preparation for negative-stain single-particle electron microscopy and determined the structure of a native-like catalytically active γ-secretase complex at a resolution of 25 Å. Antibody labeling of the extracellular domain of Nicastrin was employed to ascertain the topology of the reconstruction. In addition, active site labeling with a gold-coupled inhibitor demonstrated that γ-secretase contains a single active site facing a large conical internal cavity.

Secondly, we used the yeast *P. pastoris* as an expression system to isolate milligram quantities of active and fully mature γ-secretase. The higher yield of γ-secretase obtained

from this system allowed the addition of a size exclusion chromatography step to the purification procedure to reduce the structural heterogeneity of the preparation and thereby improve the structural analyses.

Finally, we provide novel proof for the existence of a presenilin-like protease in the euryarchaeota *M. marisnigri* (MCMJR1). We tested *in vitro* protease activity on chimeric protein substrates containing transmembrane domain (TMD) regions and found that MCMJR1 was indeed able to generate a proteolytic product. Mutagenesis confirmed that two aspartic acid residues were essential for the activity of the protease. Also, the activity of MCMJR1 was blocked when an inhibitor of γ-secretase was incubated with the enzyme. Mass spectrometry analysis revealed that MCMJR1 cleaves the chimeric substrates close to the middle of the TMD with promiscuous peptide bond selectivity.

In conclusion, our study has enhanced the understanding of PS/γ-secretase and provided new routes for functional and structural studies on di-aspartyl intramembrane proteases, both of which are pivotal in unraveling the mechanistic details of intramembrane proteolysis in biology and disease.

ACKNOWLEDGMENTS

I am infinitely grateful to my mentor, Dr. Iban Ubarretxena. The culmination of my studies and of this dissertation would have not been possible without his support, understanding and constant encouragement.

I also thank friends and colleagues at The City College of New York, New York Structural Biology Center and Mount Sinai School of Medicine for their help and advice. I feel fortunate for having met such an amazing group of people. I am especially thankful to my friend Jose Chavez for sharing his experience through enlightening discussions and technical assistance.

Finally, I would like to express my gratitude to my parents and husband for their infinite patience, moral support and words of encouragement during this long experience at the Academy. To them I express my love and appreciation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

Aβ amyloid β-peptides

AD Alzheimer's disease

AICD APP intracellular domain

APH-1 anterior pharynx-defective 1

APP amyloid precursor protein

βCTF β-secretase-cleaved APP C-terminal fragment

βME β-mercaptoethanol

BSA bovine serum albumin

CBP calmodulin binding peptide

CHAPSO 3-[(3-cholamidopropyl) dimethylammonio]-2-hydroxy-1-propanesulfonate

CHO Chinese hamster ovary

CTF C-terminal fragment

DDM n-dodecyl-ß-D-maltoside

E-64 trans-epoxysuccinyl-L-leucylamido-(4-guanidino)butane

ECL electrochemiluminesce

EGFR epidermal growth factor receptor

EM electron microscopy

ER endoplasmic reticulum

FC-12 fos choline 12

FEG field-emission gun

FSC fourier shell correlation

HEK human embryonic kidney

I-CLiPs intramembrane-cleaving proteases

IPTG isopropyl-β-D-thio-galactoside

MBP maltose binding protein

MRA multi-reference alignment

MS mass spectrometry

MSA multivariate statistical analysis

NCT nicastrin

NTF N-terminal fragment

PBS phosphate buffered saline

PEN-2 presenilin enhancer 2

PMSF phenylmethylsulfonyl fluoride

Prot A IgG binding domains of *S. aureus* protein A

PS presenilin

RIP regulated intramembrane proteolysis

SP signal peptidase

SPEM single particle electron microcopy

SPP signal peptide peptidase

SPPLs SPP-like proteases

SREBP sterol regulatory element binding protein

SUMO small ubiquitin modifying protein

TAP tandem affinity purification

TBS tris buffered saline

TCA trichloroacetic acid

TMD transmembrane domain

TX-100 Triton X-100

Chapter I

GENERAL INTRODUCTION

Biological importance of proteases

Early studies on proteases were mainly focused on their nonspecific degradative functions during protein catabolism (food digestion) and generation of amino acids (intracellular protein turnover). Later, the involvement of proteases in the catalysis of highly specific proteolytic processing reactions was proposed [1]. During the years that follow, a solid understanding on the proteolytic activation of precursors of enzymes (zymogens) and hormones (prohormones), identified in biological processes such as digestion, haemostasis (coagulation), complement activation, development and regulation of blood pressure, was achieved [2].

Nowadays, in accordance with the large number of entries (belonging mostly to serine and metalloproteases) in MEROPS [3], a comprehensive database of proteases and inhibitors, proteases have been shown to be involved in a larger number of biological processes such as cell-cycle progression, cell proliferation and cell death, DNA replication, tissue remodeling and immune response. For instance, in immunity a serine protease, granzyme, induces apoptosis within virus-infected cells [4]. For other serine proteases with selectivity for small peptides, an involvement in cell signaling has been suggested [5]. The most studied metalloprotease, matrix metalloproteinase (MMPs),

controls the degradation of the extracellular matrix in physiological functions that involve tissue growth and remodeling during development [6, 7].

Several proteases are also key modulators of processes during the pathological events of diseases that constitute leading causes of death in the U.S. [8]. The zinc-endopeptidase MMPs have been shown to be involved in the pathology of heart failure (MMP-1 and MMP-14) as well as embracing cancer growth and invasion [7]. Upregulation of cysteine cathepsins has also been demonstrated in many human tumors, including breast, lung, brain, gastrointestinal, head and neck cancer and melanoma [9]. Serine proteases, more specifically matriptases and kallikreins, are also associated with a variety of cancer types as indicated by their high gene expression levels in these conditions [10, 11]. Acetylcholinesterase (ACE) is another important serine protease that catalyzes hydrolysis of acetylcholine required for proper function of cholinergic neurons [5]. Inhibition of ACE is key in the treatment of neuromuscular and neurological diseases such as Alzheimer's disease [12] in which the aspartic proteases β- and γ-secretases have also been shown to play a pivotal role in the development of its pathology (for review, see refs [13] and [14]). Aspartic proteases are also closely associated with diseases that are a major cause of death in both the developed and developing worlds such as AIDS, peptic ulcer disease, malaria and cancer [15].

Membranes and lipids

Proteases are not only found in aqueous environments within and outside the cell, but they are also found associated with the lipid bilayer. These proteases display different levels of association with the membrane and for some of them strong hydrophobic interactions keep the protein domains embedded within the lipid bilayer where they are capable of cleaving their membrane associated protein susbtrates. Identified membrane proteases include serine, metallo and aspartyl proteases, with the last catalytic type being the main subject of this research. Studies carried out on aspartyl intramembrane proteases will be summarized following a brief introduction on biological membranes.

Biological membranes define the external boundaries of cells and regulate transport processes and enzymatic activities [16] (Fig. 1). All biological membranes are composed mainly of lipids and proteins. There are three major kinds of membrane lipids, all amphiphilic molecules: phospholipids, cholesterol and glycolipids. Phospholipids, the major class of membrane lipids, are constructed from two hydrocarbon tails, usually fatty acids (hydrophobic moiety), a platform to which fatty acids are attached (e.g. glycerol), and a polar head group (hydrophilic moiety). Due to their cylindrical and amphiphatic nature, the energetically most-favored structure for most phospholipids is to form bilayers in aqueous environment [17]. These bilayers are composed of 2 lipid sheets (each a monolayer) with the hydrophobic tails of each individual sheet interacting with one another forming a hydrophobic interior ~30 Å thick. This leaves the

hydrophilic head group layers, each 10-15 Å thick, to interact with the aqueous media on each side of the bilayer [18].

In membranes, a large fraction of membrane proteins are found in fluid disordered regions. In other regions, phospholipids associate with membrane proteins forming submicroscopic assemblies called lipid rafts [19]. Lipids rafts constitute specialized membrane microdomains engaged in signal transduction, membrane trafficking, cytoskeletal organization and motility, polarization and pathogen entry [20, 21]. The tightly packed lipids rafts are proposed to be highly dynamic assemblies that float freely in the plane of a surrounding membrane that is more fluid [20].

The fluidity of the membrane depends on the degree of unsaturation (kinked fatty acyl chains) and length of the hydrocarbon tail. The latter defines a fundamental property of synthetic bilayers made from a single type of phospholipid, the phase transition temperature (T_c). Close to the T_c, the hydrocarbon chains are roughly parallel and have low mobility and close proximity (crystalline type state). Above the T_c, a change occurs from the rigid crystalline state to a disordered state (liquid type state) characterized by an increased mobility of the hydrocarbon chains, which in turn thins the bilayer [22].

Cholesterol, essential in the assembly of lipid rafts, also moderates the fluidity of the membrane. Cholesterol inserts into the bilayer with its long axis perpendicular to the plane of the membrane forming specific complexes with certain classes of phospholipids. When present in large amounts in eukaryotic plasma membranes,

cholesterol tends to make lipid bilayer less fluid decreasing the permeability of the bilayer to small water-soluble molecules [23]. Cholesterol also inhibits phase transitions by preventing hydrocarbon chains from coming together and crystallizing [24].

The plasma membranes of many mammalian cells are composed predominantly of four types of phospholipids: phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine and sphingomyelin, each type carrying head groups that differ en size, shape and charge. The proportion of each phospholipid varies not only according to the type of membrane (i.e. species, cell type, organelle) but also between the two monolayers of the lipid bilayer. Added to this asymmetry is the presence of glycolipids, found exclusively in the noncytosolic monolayer of the lipid bilayer. The asymmetrical distribution observed in the lipid bilayer is functionally important. The hydrophilic head groups of glycolipids are sugar groups that act as specific sites for recognition by carbohydrate-binding proteins. In a similar way, binding and functionality of some cytosolic enzymes and peripheral membrane proteins depends on the presence of specific phospholipid head groups.

Membrane proteins

The protein content in the plasma membrane varies from 18% to up to 75% in energy transduction membranes [25]. Many properties of membrane proteins, acting as pumps, channels, receptors and enzymes, are greatly dependent on the surrounding lipids. Thus, a positive match between proteins and lipids allows for correctly folded and functional

membrane proteins. The most common motif in membrane proteins is the α-helical conformation, although other structures such as β-barrels also have similar energetic advantages [26]. The stability of individual helical structures is achieved through (a) hydrophobic side chains that establish contacts with the hydrophobic region of lipids, and (b) strong main-chain hydrogen bonding [27]. In folded proteins, helix-helix interactions are more favorable than helix-lipid interactions. The association of individual helices to constitute folded proteins are driven or favored by the presence of helix recognition motifs (GXXXG, leucine zipper or polar residues), the thickness of the bilayer or the helix backbone dipole [28].

The main categorization classifies membrane proteins in integral or peripheral proteins [17] (Fig. 2). Peripheral proteins are associated with the membrane through electrostatic interactions and hydrogen bonding with hydrophilic domains of integral proteins or with the polar heads of membrane lipids. Changes in the pH or the addition of salts are enough to disrupt the interactions between peripheral proteins and the membrane. In contrast, integral proteins are mostly membrane-spanning proteins that are strongly associated with the membrane by hydrophobic interactions with the hydrocarbon chains of membrane lipids. Only agents that interfere with these interactions are capable of removing them. Based on the spatial relationship of protein domains to the lipid bilayer, integral membrane proteins fall into 6 categories. Those with a single transmembrane helix are either type I, with the N-terminal domain oriented towards the outside or type II, with the N-terminal domain oriented towards the inside. Sometimes classified as type III and IV, are those that have multiple TMDs in a single or several different

polypeptides, respectively. Type V integral membrane proteins are covalently anchored to the membrane by fatty acid (e.g. palmitoyl groups) or lipid chains (e.g. farnesyl groups), while in type VI, the anchor is glycosil phosphatidyl inositol (GPI) [16].

Signal Transduction

The interaction between membrane proteins and lipids allow four types of transmembrane α-helices motions in the membrane: translation, piston, rotation parallel to the membrane (pivot) and rotation perpendicular to the membrane [29]. These proposed motions, as well as function-related conformational changes occurring as a result of the presence of α-helix destabilizing residues [30], have been suggested to be part of the mechanism of signal transduction. During signal transduction, information is transmitted from the surroundings to the cell's interior [31] through membrane proteins that act as cell-surface receptors and convert extracellular ligand-binding events into intracellular signals. These signals are amplified and integrated to finally activate or inactivate target proteins that alter the metabolism of the cell, gene expression, cell movement or cell shape. There are three major types of cell-surface receptors [32] (Fig. 3). The ion-channel linked receptor allows inorganic ions, primarily Na^+, K^+, Ca^{++}, Cl^-, to diffuse passively and selectively down the electrochemical gradient across the lipid bilayer. These receptors open or close in response to a specific stimulus (e.g. neurotransmitter binding) and are responsible for mediating the electrical signaling in muscle cells and the nervous system. G-protein-linked receptors act indirectly through the trimeric G-binding protein to regulate the activity of a separate target membrane

protein. Binding of a signal molecule to the receptor causes the receptor change conformation, which in turn alters the conformation of the G-protein bound to the receptor. Active G-protein subunits that can regulate the activity of target proteins (enzymes or ion channels) are then generated. Smell and vision depend on G-protein-linked receptors. Finally, enzyme-linked receptors activate their intrinsic enzyme activity or activate other enzymes upon binding of signal molecules that are often collectively called growth factors. They promote growth, proliferation, differentiation or survival of cells in animal tissue.

Additional signaling pathways depend on the regulated peptide bond cleavage of membrane bound proteins within their transmembrane regions. This regulated intramembrane proteolysis (RIP) results in the liberation of active cytoplasmic or extracellular/luminal domains that serve as biological effectors in other intracellular locations [33]. RIP plays a prominent role in biological processes such as cell differentiation, unfolded protein response, cholesterol metabolism, immune surveillance and intercellular communication [34]. The proteases responsible for this RIP are an evolutionary conserved group of integral membrane proteases called intramembrane-cleaving proteases (I-CLiPs).

Intramembrane-cleaving proteases (I-CLiPs)

Three major groups of I-CLiPs have been identified: site-2 protease zinc metalloproteases, rhomboid serine proteases and the two related di-aspartyl proteases,

presenilin (PS)/γ-secretase complex and signal peptide peptidase (SPP) (Fig. 4). All I-CLiPs share at least four common features: (a) they are involved in the regulation of cellular physiology; (b) each protease is a polytopic integral membrane protein with their catalytic residues buried in the membrane bilayer; (c) they hydrolyze substrates within their TMD regions; and (d) they have short sequence motifs surrounding their catalytic residues that are archetypal of the water-soluble proteases of the same mechanistic class.

The subject of this thesis focuses on GXGD-type di-aspartyl intramembrane proteases, a family of membrane protein enzymes that has one of their two catalytic aspartic acid residues contained within a GXGD motif. For clarity, however, we will first introduce proteases from a general perspective before focusing on intramembrane proteases.

Proteases

Proteases are enzymes that catalyze the hydrolysis of peptide bonds of proteins [35]. They are involved in the gross hydrolysis of proteins during protein catabolism and generation of amino acids as well as in vital processes that require specific proteolytic reactions such as cell death [36]. Owing to the increasing number of genes encoding proteases (more than 4% in the human genome and about 2% in other genomes), there is currently a rising interest in the functional and structural characterization of the many proteases present from archaea to man [3].

Protease terminology

The term protease was introduced in the late nineteenth century and remains nowadays the term most commonly used. Considering that proteolytic enzymes are more specifically "peptide bond hydrolases", a more logical term is peptidases. Indeed, peptidase is the word that the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) recommends as the general term for all proteolytic enzymes. In terms of the reaction catalyzed, proteases are primarily divided into two classes: (1) proteinases or endopeptidases, which are enzymes that cut the polypeptide chain internally to produce large fragments and; (2) exopeptidases, which are enzymes that act near the end of the polypeptide chain to liberate products of just a few amino acid residues long. The latter, when acting close to the N-terminus or C-terminus are termed aminopeptidases or carboxipeptidases respectively. In addition, six types of proteases are recognized on the basis of the chemical nature of the catalytic amino acid residue: serine, threonine, cysteine (grouped under the catalytic type protein nucleophile), aspartic, glutamic and metallo (each, subdivision of the catalytic type water nucleophile) proteases, as well as some of unknown catalytic type. The NC-IUBMB took both classification criteria and assigned a unique number to each peptidase, the EC (from Enzyme Commission) number, with a major contribution in nomenclature (Table I).

A new form of classification that reflects structural features and evolutionary relationships, criteria not considered by the NC-IUBMB, has been proposed by the

MEROPS system [3]. *MEROPS* uses a hierarchical, structure-based classification of peptidases. In this system, each peptidase is assigned to a "Family" on the basis of statistically significant similarities in amino acid sequence (common ancestry). Families that are thought to be homologous (catalytic mechanism) are grouped together in a "Clan". Thus, a three-layer system that combines the *MEROPS* system of peptidase clans and families, and the EC recommendations on enzyme nomenclature, has been developed. This three-layer system classifies peptidases by (a) catalytic type – level 1, (b) molecular structure (clans and families) – level 2, and (c) individual peptidases – level 3 (Table I).

Mechanisms of catalysis

As mentioned above (and exemplified in table I), two major catalytic types are now recognized: protein nucleophile and water nucleophile. The main distinction relies on having strongly nucleophilic amino acids at the catalytic site of the first type while those proteases from the second type catalyze the hydrolysis of peptide bonds without nucleophilic attack by the functional group of the side chain [37]. Of all proteases, perhaps the best studied are serine proteases. The generally accepted chemical mechanism of serine protease catalysis involves acylation and deacylation reactions [38] (Fig. 5A). In the acylation reaction, an ester is formed between the oxygen of Ser (acting as a nucleophile) and the acyl portion of the peptide substrate yielding a tetrahedral intermediate. The oxyanion of the tetrahedral intermediate is stabilized by hydrogen bonding with the backbone NH of the oxyanion hole (Gly and Ser) before it collapses

with the expulsion of the amino portion of the substrate, generating the acylenzyme intermediate. In the deacylation reaction, a water molecule attacks the acylenzyme intermediate yielding a second tetrahedral intermediate that later collapses regenerating the active site and expelling a carboxylic acid product.

The mechanism of catalysis of cysteine proteases is highly similar to that of serine proteases (Fig. 5B). In this instance, a covalent intermediate is also formed after the nucleophilic attack by the sulfur atom of the catalytic cysteine side chain. As in the mechanism of serine proteases, a nearby histidine side-chain acts as hydrogen acceptor/shuttle with two stabilizing -NH- groups also present in the oxyanion hole. Similarly, a threonine protease mechanism has been proposed from studies of the proteasome catalytic subunit and mutational strategies have shown that an N-terminal threonine residue is indeed responsible for the nucleophilic attack in peptide hydrolysis [39, 40].

In aspartic proteases, a nucleophilic attack also takes place, but contrary to serine, threonine and cysteine proteases, a water molecule acts as the nucleophile (Fig. 5C). The most widely accepted mechanism is an acid-based system involving two aspartic acid residues in the active site with the nucleophilic water molecule residing between them [41]. These two aspartic acid residues act as proton donor and acceptor coordinating the nucleophilic water molecule. After activation of the water molecule by the negative aspartate side chain, a specific carbonyl carbon in the substrate scissile bond is attacked generating an oxyanion tetrahedral intermediate. Protonation of the scissile amide N

atom and rearrangement result in the breakdown of the tetrahedral intermediate to the acyl product complex or an amino product complex [42]. A covalently bound intermediate is not involved during catalysis by aspartic proteases [43]. In metalloproteases, hydrogen bonding via an oxyanion hole is replaced by coordination of typically two or three histidines to a metal ion that assists in the attack by a water molecule [44] (Fig. 5D). This water molecule is also hydrogen-bonded to a glutamic acid, which plays an analogous role to the histidine in serine and cysteine proteases, and to one of the catalytic groups in aspartic proteases.

In the active site of glutamic proteases, the nucleophilic water is activated and the tetrahedral intermediate stabilized by an unusual dyad consisting of glutamine and glutamate residues [3, 45]

Substrate specificity and peptide bond selectivity in soluble proteases

In soluble proteases it is common to observe that the specificity and cleavage site in the substrate differ from one member of a family to another (Table II). This is observed, for example in chymotrypsin-like serine proteases for which efficiently processed substrates are usually categorized in terms of the P1/S1 interaction. Briefly, based on the protease specificity terminology [46] the active site of a given enzyme is composed of subsites S1, S2, … Sn and S1', S2', …Sn'. In the substrate, amino acid residues at positions numbered P1, P2, …Pn and P1', P2', …Pn' (counted from the cleavage site in both directions, towards the N and C-terminus respectively) occupy the subsites with the same numbers (Fig. 6). Thus, in chymotrypsin, the specificity is determined by Ser189,

Gly216 and Gly226 (hydrophobic pocket S1) and correlates with the hydrophobicity of the P1-Phe preferred residue. With trypsin, a negatively charged S1 pocket that includes Asp189, Gly216 and Gly226 accounts for the specificity for substrates containing Arg or Lys at P1. In another example, the smaller S1 site of elastase (Val216 and Thr226) grants the protease a preference for substrates with small aliphatic residues at P1. Interactions at the S2-Sn sites of chymotrypsin display little specificity contrary to enterokinase, which recognizes the P5-P1 sequence D-D-D-D-K with high precision [38, 47].

The fairly broad specificity possessed by cysteine proteases is a consequence of having an extended active site consisting of seven subsites (S1-S4 and S'1-S'3), each able to accommodate one amino acid residue of a substrate (P1-P4 and P'1-P'3). A defined residue selectivity is nevertheless observed at the S2 subsite, with the P2/S2 interaction acting as the major specificity determinant in many cases [48]. Many cysteine proteases show a strong preference for hydrophobic residues at P2 [49]. More specific requirements have been established for caspases and calpains. In caspases, the P1 residue is Asp, with the single exception of *Drosophila* caspase Dronc, and the P1' residue is small and uncharged [50]. In calpains, preferred residues in the P2 position are Leu, Thr and Val and in P1 Lys, Tyr and Arg. In position P1' small residues such as Ser, Thr and Ala occur most often [51].

Extensive reviews on aspartic proteases include the pepsin-like protease family and the retroviral peptidases. HIV-1 retroviral peptidase is capable of recognizing many different

sequences as substrates yet certain determinants of its protease specificity have been clarified. The crystal structure of HIV-1 bound to synthetic peptides as substrates revealed a set of hydrogen-bond interactions involving the substrate main chain. Each substrate side chain can then orient into a subsite within the enzyme giving the main chain of the substrate an extended β-sheet configuration. The β-sheet main chain interactions start with the P4 residue having a preference for small amino acids found in β-turns or known to be α-helix breakers [52]. A preference for aromatic residues or larger polar residues has been observed at the P3 position. Fewer amino acid residues are allowed in the P2 position (either Asn or smaller hydrophobic and β-branched residues) depending on the type of subsites [53, 54]. P1 and P1' show a preference for large hydrophobic or aromatic residues with hydrophobic S1 and S1' subsites. Preferences for glutamic acid and hydrophobic amino acids have been observed at P2' and P3' positions respectively [53]. Finally, side-chain interactions between P4' residue and the external surface of HIV-1 has been proposed [52].

The preference of pepsin for hydrolysis has been analyzed based on the frequency of appearance of amino acid residues in several proteins of known sequences. At P1 and P1' positions, phenylalanine appears at very high frequencies followed by leucine and tyrosine, with the residues at the P1 position influencing cleavage of the substrate to a somewhat higher degree than those on P1' [55]. For pepsin it was then determined that a hydrophobic interaction constitutes an essential step of the enzymatic reaction [56]. Nevertheless, fungal peptidases are unique within the pepsin family in the sense that they have the ability to cleave a bond with Lys in the P1 position [57].

Mechanisms of inhibition

Inhibitors designed to inactivate protein nucleophile catalytic type proteases are based on electrophilic groups. For instance, diisopropylfluorophosphate (DFP) and the less toxic phenylmethylsulfonyl fluoride (PMSF) inactivate virtually all serine proteases by reacting irreversibly with serine to form a stable tetrahedral adduct that mimics the tetrahedral intermediate/transition state [38]. Other compounds inactivate serine proteases by forming stable acylenzyme intermediates such as 3-benzyl-6-chloro-2-pyrone [58].

Inhibitors designed for *water nucleophile* proteases require a greater necessity for secondary binding interactions along the active site cleft. In this regard, the main design for aspartic protease inhibitors is based on the mimicry of the tetrahedral transition state of the proteolytic reaction (Fig. 7) [15]. Basically, the core for the development of these inhibitors is a good non-cleavable isostere such as statin, phosphinate or the hydroxyethylamine moiety in place of the normal hydrolyzable amide bond [42]. For example, when a hydroxymethylcarbonyl isostere is introduced in a substrate inhibitor the critical hydroxyl group forms hydrogen bond interactions with the carboxylic acid moieties of the two catalytic aspartic acid residues in the protease. The carbonyl oxygen in the inhibitor is also involved with one of the aspartic acids to form a hydrogen bond network that ultimately suspends the protease's biochemical actions [41]. In the case of metalloproteases, introduction of compounds such as EDTA, EGTA and phenanthroline lead to nearly irreversible metal ion chelation that impairs nucleophilic attack by the

water molecule [59].

Di-aspartyl intramembrane proteases

PS and SPP are involved in signaling events as well as in the degradation of membrane-bound proteins with no apparent biological functions [60, 61]. They span the membrane nine times and their catalytic motifs YD and GLGD are located in the middle of adjacent TMD regions [62, 63]. Bacterial type IV prepilin peptidases (TFPPs) [64, 65] and archaeal TFPP-like [66] are also GXGD-type polytopic di-aspartyl membrane proteases (Fig. 8). In TFPPs two highly conserved aspartic acid residues, one contained within the GXGD motif, are responsible for processing leader peptides of flagellins and other secretory proteins [64, 67]. However, different from PS and SPP, in TFPPs and TFPP-like, the two active-site aspartate residues are found at either the cytoplasm-membrane interface or within a cytoplasmic loop [67]. TFPPs then remove leader peptides of selected substrates cleaving at a position between hydrophobic and hydrophilic domains close to the cytoplasmic side of the membrane [64, 68, 69]. The latter sets TFPPs apart from the family of intramembrane proteases.

A third motif, common to both PS and SPP, contains a C-terminal PAL sequence that contributes to active site conformation [70-72]. These di-aspartyl intramembrane proteases also share equivalent inhibitor binding pharmacology for transition state analogs, non-transition state inhibitors and modulator compounds [73]. However the similarities end there and important structural and functional aspects differentiate

between the two. Whereas PS is the catalytic component of a multiprotein complex, SPP acts alone. PS and SPP also display inverted overall topologies [74] that may account for substrate specificity of SPP preferentially towards type II transmembrane proteins and of PS towards type I transmembrane proteins [75, 76]. The following paragraphs cover specific features of these two di-aspartyl intramembrane proteases in more detail.

Presenilin (PS)

Presenilin or PS derives its name from its role in the development of Alzheimer's disease (AD). Specifically, in the brain of an AD patient, a chronic and substantial decline in cognitive function occurs due to the deposition of amyloid or "senile" plaques [76]. The major constituent of these plaque depositions is insoluble amyloid β-peptides (Aβ; molecular mass ~4 kDa) that are produced by proteolysis of a large glycosylated type I membrane protein, the amyloid precursor protein (APP) by γ-secretase [77, 78].

Genetic approaches backed up with *in vivo* reconstitution studies have demonstrated that PS is part of the multisubunit γ-secretase complex that includes three additional integral membrane proteins: nicastrin (NCT), anterior pharynx-defective 1 (APH-1) and presenilin enhancer 2 (PEN-2) [79] (Fig. 9). Several studies have hypothesized on the biochemical functions and individual roles of NCT, APH-1 and PEN-2 in the assembly and activation of the γ-secretase complex. Mutational studies and RNA interference techniques have shown that APH-1 may contribute to the initial assembly of the γ-secretase complex by playing a scaffold role and that the APH-1•NCT subcomplex

constitutes a prerequisite in assembly and stabilization of full-length PS [80, 81]. In embryonic fibroblast cells lacking endogenous NCT the role of this subunit was also studied. It was shown that, unlike wild type cells in which the γ-secretase complex is located in the trans-Golgi network, PEN-2, APH-1 and full-length PS remained associated with each other in the endoplasmic reticulum (ER). A putative role of NCT in the stability and proper intracellular trafficking of the complex was then assigned [82]. Moreover, Shah et al. [83] demonstrated that NCT assists in the recognition of the newly generated N-terminus after ectodomain shedding of γ-secretase substrates. Specifically, the carboxylate side chain of NCT, residue E333, interacts with the substrate positioning it into the initial substrate-binding site. In addition, PEN-2 seems to elicit the final maturation of the γ-secretase complex by inducing the endoproteolysis of PS into N-terminal fragment (NTF) and C-terminal fragment (CTF) and by stabilizing the NTF/CTF heterodimer [84, 85]. Finally, in PS, the catalytic component of the γ-secretase complex, two conserved aspartate residues in TMDs 6 (D257) and 7 (D385) were shown to be critical for γ-secretase activity [86] and thus PS was hypothesized to be a di-aspartyl intramembrane protease.

Notch, ErbB-4 and APP are well-studied targets of intramembrane proteolysis by PS/γ-secretase. Notch is a type I membrane-bound transcription factor that determines cell fate during development and in the adult [87]. In response to ligand binding, the Notch receptor is cleaved by PS releasing a cytoplasmic domain that translocates to the nucleus to activate transcription [88]. Similarly, ligand stimulation of ErbB-4, a member of the epidermal growth factor receptor (EGFR) family, promotes its cleavage by PS. The

intracellular domain of ErbB-4 is then released into the cytosol and upon translocation to the nucleus regulates gene transcription and cell fate in the developing brain [89]. APP is initially cleaved under normal physiological conditions within the Aβ peptide sequence by α-secretase and then by γ-secretase, resulting in the generation of soluble amino-terminal derivatives, which might have biological functions in growth regulation and neuroprotection [90] (Fig. 10). In early-onset AD, the presence of a double mutation in APP (Swedish mutation; APPSwe) switches the preference of APP as a substrate for α-secretase and instead β-secretase is the enzyme that cleaves APP at the N-terminus of the Aβ peptide sequence to produce a soluble ectodomain and a membrane-anchored C-terminal fragment, βCTF [91]. Successive cleavage events have been described for the processing of βCTF by PS/γ-secretase [92]. An initial intramembrane cleavage occurs near the inner leaflet at the ε position generating the APP intracellular domain (AICD) which upon translocation to the nucleus, initiates transcriptional events. The longer counterparts of AICD, Aβ49 or Aβ48, remain in the active site of γ-secretase where a trimming process starts successively cleaving every three amino acid residues upstream at the ς- and then at the γ-cleavage site. This model proposes that Aβ49 is processed to Aβ46, Aβ43 and finally to Aβ40, whereas Aβ48 is trimmed to Aβ45, Aβ42 and Aβ39. AD-causing mutations lead to an increase in the Aβ42/Aβ40 ratio implying a processing route of APP that initiates at a ε-cleavage site that preferentially generates Aβ48. A sequential proteolysis of this type has also been proposed for Notch and CD44 by PS/γ-secretase. The Notch receptor is cleaved twice by PS, first at a S3/ε-cleavage site and then at a S4/γ position, an event that is dependent on the S3/ε cleavage [93]. Similarly, CD44, a cell surface adhesion molecule, is also cut in two positions. One cleavage (close

to the cytoplasmic border) liberates the intracellular domain of CD44 for putative nuclear signaling, and the other cut splits the TMD resulting in the secretion of a peptide similar to Aβ [94]. Initially, it was suggested that two pairs of catalytic aspartates, each contained in a single PS molecule, were necessary to carry out the described multiple intramembrane cleavages [95]. Later, cysteine mutagenesis and disulphide cross linking experiments showed no evidence of a PS dimer in the core of the γ-secretase complex [96]. For other γ-secretase substrates, such as the receptor tyrosine kinase, ErbB-4 [97] and E-cadherin [98], only a single γ-secretase cleavage site has been identified.

Given the minimal sequence conservation among the nearly 60 identified γ-secretase substrates [99] and that several mutations in substrates are well tolerated by γ-secretase [100, 101], factors at higher level of protein structure are more likely to provide insights into this unusual mechanism of broad cleavage specificity. Lichtenthaler et al. [101] propose that the active site of γ-secretase interacts with βCTF by binding to one side of the α-helical TMD of βCTF and mutations interfering with this interaction alter cleavage site specificity. In a different approach, a secondary structural prediction on all γ-secretase substrates with known primary cleavage sites was performed [99], finding that helix destabilizing residues may be important for PS-mediated proteolysis. Finally, more than 25% of γ-secretase substrates contain the transmembrane dimerization motif GxxxG in their TMD [99] and the pairwise replacement of these glycine residues in APP led to drastic changes in the generation of Aβ species [102, 103] implying that substrate dimerization alters γ-secretase cleavage precision.

Remarkably, the participation of PS has also been confirmed in regulatory events such as apoptosis, protein trafficking and cell adhesion that occur independently of its activity as a protease [104].

Signal peptide peptidase (SPP)

The clearance of signal peptides by proteolysis within the ER by SPP was initially related to protein cleavage by PS and for a period of time SPP was considered to be a PS homolog [105]. The limited overall sequence homology between SPP and PS (~30% identical) in addition to the attribution of biological roles to SPP-mediated intramembrane proteolysis made SPP a new member of the di-aspartyl intramembrane protease family [106]. A central role for SPP activity has been recognized in the generation of human histocompatibility leukocyte antigen (HLA-E) epitopes during virus infection or tumor growth [107] and in the maturation of the membrane-anchored core proteins of hepatitis C [108] and GB viruses in the ER of infected cells [109]. The liberation of bioactive peptides from substrates such as IgSF1 (an Ig protein domain) [110], the eosinophil cationic protein (biomarker for asthma and allergic inflammatory diseases) [111] and preprocalcitonin by SPP [112] have also been described.

The abolishment of catalytic activity due to mutations of the conserved aspartic acid residues D219 and D265 (within the YD and GLGD motifs respectively) to alanine reinforced the role of SPP as an aspartic protease [106, 113]. In addition, as described for PS, an initial requirement for SPP intramembrane cleavage to occur is the release of

part of the extracellular domain from the substrate. Analogous to α- and β-secretase, a signal peptidase (SP) liberates the signal peptide from the preprotein before SPP can cleave within the transmembrane region [106, 114]. Moreover, an important requirement for intramembrane cleavage by SPP seems to be the presence of helix-breaking residues providing flexibility to the transmembrane region and exposing the scissile peptide bond to the protease [115]. Nevertheless, different from PS, the need for protein cofactors for SPP activation or an endoproteolytic event as that described for PS into a CTF and NTF has not been reported to date.

Sequence database screening studies [105] revealed another 4 human genes and several others in animals, plants, yeast and archaea with sequence similarities to both PS and SPP. Since these proteins are predicted to share the orientation of SPP, the nomenclature of SPP-like proteases (SPPLs) was given. These human SPPLs are SPPL3 and the SPPL2 family members (SPPL2a/b/c). SPPL3 is only ~25% identical to SPP and both are predominantly localized to the ER [116]. In a cell-based reporter assay for intramembrane cleavage, SPPL3 was able to process an SPP cleavable type II membrane protein domain, suggesting that they may share similar biological functions in the ER [117]. In comparison to SPPL2a and SPPL2b, which are mainly located in late endosomal compartments, studies in SPPL2c are scarce. BRI2, a membrane protein that may be involved in cell signaling events by acting as a receptor, and which when mutated causes familial British and familial Danish dementias [118], has been identified as being processed by SPPL2a/b. Tumor necrosis factor α (TNFα) and Fas ligand (FasL), both involved in intracellular signaling, are also processed by SPPL2a/b. More

specifically, processing of TNFα by SPPL2b utilizes multiple intramembrane cleavages to liberate the intracellular domain of TNFα into the cytosol and the carboxy-terminal into the extracellular space. The similarity observed in the cleavage pattern of γ-secretase and SPPL2b suggests common principles for regulated intramembrane proteolysis by the members of aspartyl intramembrane proteases [119].

What questions regarding GXGD di-aspartyl intramembrane proteases will be addressed in this thesis?

Despite remarkable effort devoted in the past few years to understanding how presenilin and SPP function, much remains unknown. In particular, is still unclear how GXGD di-aspartyl intramembrane proteases recognize their substrates and how they select which peptide bond to cleave. In addition, the role of conserved amino acid residues and their mechanism of action is not well understood. Finally, despite the fact that we now have representative prokaryotic structures of serine and metallo intramembrane proteases [120, 121], we still lack a structural model of a GXGD di-aspartyl intramembrane protease of similar resolution. In this thesis we will focus on addressing these relevant questions. We believe that the insights derived from this study will contribute to a better understanding of GXGD di-aspartyl intramembrane proteases.

How have we addressed these questions?

The pace in our mechanistic and structural understanding of GXGD di-aspartyl intramembrane proteases in general, and of PS in particular, has been hindered by difficulties to purify functional, highly homogeneous γ-secretase complex on one hand, and by the lack of more tractable systems from prokaryotic sources on the other. Initially, from a homogeneous preparation of γ-secretase purified from mammalian cells, we were able to perform an EM analysis and obtain a 3D map of the complex from a negatively stained sample. Although relevant conditions and features (choice of detergent and active site location) were unique to our map, the aggressive competition in the field soon made our 3D reconstruction obsolete with the publication of γ-secretase maps at 15 and 12 Å by other groups [122, 123].

More recently, we did succeed: (1) in purifying active γ-secretase expressed in the yeast *Pichia pastoris* to levels suitable for improved structural studies, and (2) in identifying and characterizing MCMJR1, a GXGD di-aspartyl intramembrane protease from archaea with significant biochemical similarities to PS. Our strategy is to capitalize on these two breakthroughs and perform mechanistic and structural studies. First, our innovative purification procedure from *P. pastoris* represents a breakthrough relative to the microgram quantities previously reported and paves the way for in depth biochemical characterization and for future structural studies. Second, working under the hypothesis that MCMJR1 constitutes a simpler model system to study GXGD di-aspartyl intramembrane proteases, as a general strategy for biochemical experiments, we

formulated and tested hypotheses on MCMJR1, to then selectively validate them on recombinant γ-secretase. Further, to gain structural insights, in the future we will prioritize crystallization and structure determination of MCMJR1.

Chapter II

STRUCTURAL ANALYSIS OF γ-SECRETASE

Introduction

X-ray crystallography is in general an established technique to determine atomic structures of large individual proteins, although it can run into several limitations when it comes to macromolecular assemblies [124, 125]. Large assemblies often resist attempts to form crystals, and when crystals are formed, packing may interfere with the study of macromolecular interactions. In the case of large multi-component membrane proteins of eukaryotic origin, their low expression levels and inherent instability when extracted with detergents from their native membrane, which is often concomitant with loss of biological activity, have become major obstacles in the production of large, well ordered 2D and 3D crystals [126]. Therefore, crystallographic approaches to elucidate the structure of γ-secretase are challenging. Single-particle electron microscopy (SPEM), briefly introduced in the following paragraphs, has thus become a valuable structural method complementary to X-ray for demanding macromolecular complexes such as γ-secretase.

Here, we employed a tandem affinity purification (TAP) protocol based on the original method by Puig et al. [127] to purify γ-secretase from human embryonic kidney (HEK) cells stably expressing TAP-PS, NCT, APH-1, and PEN-2. The TAP tag linked to PS

allowed the purification of fully assembled γ-secretase from detergent-solubilized membranes after two consecutive affinity purification steps. Detergent-purified γ-secretase was able to process C100Flag into Aβ40 and Aβ42 in ratios similar to those found under physiological conditions. This preparation yielded an optimal sample for the EM studies presented below in more detail.

Electron microscopy

Electron microscopy (EM) techniques are capable of providing structural information of a large variety of biological samples, ranging from small proteins to large macromolecular assemblies [128-130]. Substantial structural information can be qualitatively obtained from EM through direct analysis of micrographs (photographs obtained from the electron microscope), identifying the arrangements or quaternary state of the particle under study. Nevertheless, more precise information can be extracted upon image processing to ultimately obtain a three dimensional structure of the original object under study. The latter, named 3D electron microscopy (3D EM), allows the three dimensional study of proteins to high resolutions for rigid structures such as the ribosome [131], to up to near atomic resolution (3.8 Å) for specimens capable of forming bi-dimensional crystals or that possess internal symmetry, such as viruses [128, 129].

Transmission electron microscope

In the transmission electron microscope (TEM), the source of illumination consists of a filament that emits electrons that are accelerated due to a high voltage-electromagnetic field (100kV – 300kV) from the top of the microscope column. Inside, the electron path is conducted and focalized through the use of magnetic lenses and apertures. In addition, high vacuum conditions are necessary to avoid the collision of electrons with molecules of air, reducing their dispersion and ensuring in this way, an optimal incidence of electrons on the sample. The electron path then transverses the sample and creates a projection image that is recorded on a fluorescent screen or a charge-coupled device (CCD; digital camera). The interaction between the path of electrons and the atoms of the sample generates different types of dispersions that are responsible for the appearance of contrast in the image. The process of image formation is yet inevitably affected by instrumental imperfections (i.e. lenses aberrations, astigmatism) that generate an altered version of the object under study. The final resolution achieved is highly dependant on the appropriate treatment of these detrimental imperfections.

Sample preparation for electron microscopy

Inside the TEM, the sample is subjected to deleterious conditions mainly provoked by the electron radiation and high vacuum. To protect the sample from chemical instability and dehydration as a result of the conditions mentioned above, two successful techniques have been developed. The first sample preparation involves the adsorption of the

purified sample to a continuous carbon film (Fig. 11 steps 1 and 2) deposited on a grid. The sample is finally embedded in a thin layer of an electron dense heavy metal stain (Fig. 11 step 3 (a)) not only to protect the particle under study from radiation but also to increase the contrast of the final image (negative staining). The main drawback of this technique consists on the limiting penetration power of heavy metals, truncating the attainable resolution to typically 20 Å [130]. Negative staining can also introduce artifacts attributable to incomplete stain embedding and flattening upon drying of the grid [132].

The second technique proposes a better sample preservation at the expense of a more elaborated preparation. In more detail, the sample is quickly frozen in liquid ethane at -180°C (cryo-EM) promoting the formation of ice of amorphous structure (vitreous/amorphous ice). The speed of freezing avoids the formation of ice of crystalline structure, protecting in that way the particles and preserving a hydration state similar to that of the native sample (Fig. 11 step 3 (b)). The absence of stain allows the acquisition of an image that closely represents the density of the particle, sometimes along with internal features. Nevertheless, due to the low difference between the density of water (in the form of vitreous ice) and that of the sample protein along with the low levels of electron dose used [133], the particle under study has to be of sufficient size to make up for the poor contrast acquire from this type of sample preparation [134].

Processing of digital images

As mentioned above, the images collected from a TEM (Fig. 11 step 4) gather destructive features during their generation. Then, to maximize the preservation of structural detail is recommended to reduce the electron dosage (10-20e/Å^2), which unfortunately results in noisy images (poor signal-to-noise ratio). In addition, particles from the same sample are not identical, with a heterogeneity based mainly on structural and orientational differences. The first one refers to different conformational and/or dissociation states that a protein or macromolecular complex may adopt, with each state treated as a different object. The heterogeneity based on orientation arises when the particle is adsorbed or deposited on the carbon film at a certain angle.

During the analysis and processing of a large population of 2D images of single particles (single particle electron microcopy; SPEM) (Fig. 11 step 5), rotational and translational alignment and averaging techniques are use to enhance the structural information (signal) contained within these images. Classification algorithms finally separate sets of images into homogeneous classes representing different views (projection images) of the same particle (Fig. 11 step 6).

3D reconstruction

Once a set of homogenous data has been obtained is then possible to combine the information of the different orientations to reconstruct the three-dimensional structure of

the object (Fig. 11 step 7). In an ideal situation, the adsorption of a particle on the surface of the grid support occurred in such a way that a high number of different orientations were adopted and are sufficient to cover the whole angular space [135]. In this specific situation, reconstruction methods are based on a reference volume [136], from which projections from different orientations are generated and to which the experimental images are matched. From this projection matching, angles corresponding to the direction of orientation of the reference volume are assigned to the experimental images that can then be back-projected to directly create a 3-D reconstruction. In this method, a successful reconstruction is highly dependent on the initial volume used as a reference, being necessary a thorough analysis before its selection.

In the "common lines" technique, based on the Central-Section Theorem [137], a reference volume is not required. The theorem states that, in Fourier space, two projections from the same object share a common line. Once the experimental images have been classified and the signal enhanced, this method can be applied with reliable results.

In several situations, the particle under study adopts only a few orientations on the carbon film that would normally be insufficient to cover the whole angular space. In this situation, the sample is tilted inside the microscope and images of additional orientations are obtained [138]. In this method, an initial reference is also not needed since the orientation is taken from the angle of the particles in the tilted images and the rotation from the same particle in the image acquired at 0°. Normally, reconstructions obtained

through this method lack resolution in a certain direction (missing cone) due to instrumental limitations (tilt angle ≤60°).

Materials and Methods

Constructs and cell lines

The TAP tag was N-terminally inserted in frame with the coding region of human PS sequence (TAP-PS). Using standard DNA cloning procedures, TAP-PS construct was cloned into expression vector pAG3. HEK 293S cells [139] were stably transfected with pAG3-TAP-PS construct alone or along with pAG3 constructs of NCT, APH1 and PEN-2 together with 0.1 mg of pIREShyg plasmid (Clontech) using the calcium phosphate method [140]. Positive clones were selected with 400 ug/ml of Hygromycin B (Invitrogen) and stably transfected cell lines cultured in Dulbecco's modified Eagle's medium containing 10% (v/v) fetal bovine serum (FBS), penicillin-streptomycin and 200 ug/ml of Hygromycin B. For suspension culture, HEK293S cells were grown in IS-GRO medium (Irvine Scientific) at 37°C without CO_2 and supplemented with 10% FBS, 10 mM L-glutamine, penicillin-streptomycin and 200 µg/ml of Hygromycin B.

Expression and detergent solubilization of γ-secretase from HEK293S cells

HEK293S cells over expressing TAP-PS, NCT, APH-1, and PEN-2 were harvested by centrifugation at 500xg for 5 minutes at 4°C and then washed once with ice-cold

phosphate buffered saline (PBS) and centrifuged again. The resulting cell pellet was weighed and resuspended in 5 volumes of homogenization buffer (10 mM MOPS, 10 mM KCl, 1 mM EDTA, 1x protease inhibitor cocktail) and homogenized with a glass homogenizer. Following homogenization, the lysate was centrifuged at 1000xg for 10 minutes to collect the post nuclear supernatant fraction and the pellet was resuspended in 5 volumes of homogenization buffer and homogenized for a second time. After centrifugation at 1000xg, the supernatant was combined and centrifuged at 100,000xg for 1 hour at 4°C to pellet the membranes.

The membrane pellet was weighed and then solubilized on ice for 2 hours in 5 volumes of solubilization buffer (50 mM sodium citrate, pH 6.4, 10% glycerol, 1 mM EDTA, 1x protease inhibitor) containing 1% of the detergent 3-[(3-cholamidopropyl) dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO; CMC=0.5%, Anatrace). After detergent solubilization, the sample was centrifuged at 100,000xg for 1 hour at 4°C and the supernatant, containing solubilized membrane proteins, were saved for γ-secretase purification. Total protein concentration of the detergent solubilizate was determined using the BCA protein assay kit (PIERCE).

Purification of detergent solubilized γ-secretase.

The two sequential affinity steps employed during purification are based on the TAP tag linked to PS consisting of a duplicate of IgG binding domains of *S. aureus* protein A (Prot A) and a calmodulin binding peptide (CBP). The CHAPSO-solubilized fraction

was diluted twice in 20 mM Tris HCl pH 8.0 containing 150 mM NaCl and incubated for 2 hours in the cold room with IgG-sepharose (Pharmacia) beads pre-equilibrated in binding buffer A (10 mM Tris pH 7.4, 150 mM NaCl, 0.5% CHAPSO). Following incubation, the beads were extensively washed with binding buffer A followed by one wash with TEV protease buffer (10 mM Tris pH 8.0, 150 mM NaCl, 0.5% CHAPSO, 0.5 mM EDTA, 1 mM DTT). TEV protease digestion was then performed at 16°C to release TAP-PS and its associated proteins from the beads. The TAP-PS eluate was adjusted to a final concentration of 1 mM CaCl2 and incubated for 1 hour at 4°C with calmodulin beads (Stratagene) that were pre-equilibrated with binding buffer B (10 mM Tris pH 7.4, 150 mM NaCl, 1 mM Mg-acetate, 2 mM $CaCl_2$, 0.5% CHAPSO, 10 mM β-mercaptoethanol [βME]). Following incubation, the beads were washed extensively with binding buffer B, and bound TAP-PS and other γ-secretase components were eluted in elution buffer (10 mM Tris pH 7.4, 150 mM NaCl, 1 mM Mg-acetate, 0.5% CHAPSO, 5 mM EGTA). The eluate was then concentrated using a 3K centrifugal filter device (Millipore). Purified γ-secretase was resolved on a 12% SDS-PAGE and visualized by silver staining. Pure bovine serum albumin (BSA) at different concentrations was also run on the same gel and used as standards to determine the concentration of γ-secretase.

Antibodies and western blotting

Purified γ-secretase was fractionated on either 10% or 12% SDS PAGE and transferred to a PVDF membrane. γ-Secretase components were probed with PS-NT (recognizes residues 1-65 of PS), aPSLoop (recognizes residues 320-375 of PS), NCT54 (recognizes

residues 242-546 of NCT ectodomain) and PNT-2 (recognizes 1-26 residues of PEN-2, a kind gift from Dr. Gopal Thinakaran) antibodies.

γ-Secretase *in vitro* activity assay

0.5 to 4 ul of purified γ-secretase complex was used to examine γ-secretase activity by incubating the preparation with purified chimeric C100-Flag substrate [141]. Adding radio-immuno precipitation assay (RIPA) buffer and boiling the mixtures for 5 minutes terminated the reactions. The samples were centrifuged and the Aβ peptides generated quantified using an electrochemiluminesce (ECL)-based assay.

EM analysis

Sample preparation

NCT54 antibody labeling

Active detergent-purified γ-secretase (15-20 nM) in 10 mM Tris pH 7.4, 150 mM NaCl, 1 mM Mg-acetate, 0.25% CHAPSO and 5 mM EGTA was incubated for 1 hour at room temperature with NCT54 (1:1000), a polyclonal antibody generated against residues 242–546 of the extracellular domain of human NCT [142]. Two microliters of γ-secretase complexed with NCT54 was then applied on 300-mesh carbon-coated copper/rhodium grids that were previously glow-discharged. After one minute, the

excess solution applied to the grid was blotted and immediately rinsed with a solution of 6% ammonium molybdate, 2% threalose pH 8.3 for 30 seconds. The excess of ammonium molybdate staining solution was finally blotted from the side and air-dried before visualization in the electron microscope.

Streptavidin-gold labeling

Nanogold®-Streptavidin (Nanoprobes), consisting of affinity-purified streptavidin covalently conjugated to an extremely uniform 1.4 nm-diameter Nanogold® particle, was used to label a modified version of the Merck transition state analogue L-685,458 [143]. Modified L-685,458 (2 uM), carrying a 39.4Å-length spacer arm between the inhibitor warhead and a biotin moiety, was incubated with an aliquot of γ-secretase (70 nM) for 1 hour at room temperature. After inhibition of γ-secretase, Nanogold®-Streptavidin was added at a final concentration of 0.1 mM and allowed to interact with the biotin moiety of the modified L-685,458 for 1 more hour at room temperature. Finally, the mixture was incubated for another hour at room temperature with antibody NCT54 (in a 1:10 dilution). Two microliters of this reaction mixture were then applied to carbon-coated copper/rhodium grids for 1 minute. The specimen was partially blotted, and the grid was applied onto the surface of a 40 ul droplet of gold enhancing reagent (Nanoprobes) for 20 seconds to generate a better contrast signal in the electron microscope. Sample staining for visualization in the electron microscope under negative stain conditions was then carried out as explained above.

Electron microscopy

Negatively stained specimens were imaged with a Tecnai G2 F20 field-emission gun (FEG) transmission electron microscope operating at an accelerating voltage of 200 kV. Micrographs were collected in low-dose (<20 e/Å2) mode with flood beam illumination at a nominal magnification of 88 250x at which the effective pixel size was 1.7 Å. Micrographs were then recorded on a Tietz 4kx4k CCD camera (TVIPS, Germany) with underfocus ranging from 11,000 to 28,000Å.

Image processing

Particle selection, windowing, floating, filtering and determination of the defocus and amplitude contrast used for correcting the effects of the contrast transfer function were performed with the software package EMAN [144]. Subsequent, particle intensity normalization, alignment, classification and model building steps were carried out with the software package SPIDER [145].

To build an initial 3D model of γ-secretase complexed with NCT54, a subset of 1040 particles was subjected to reference-free classification by multivariate statistical analysis (MSA). Hierarchical ascendant classification produced 13 reference-free class averages and the relative orientations of their characteristic views were determined using the cross-common lines technique. The resulting averages were combined to build a 3D model. This model was then used to generate 83 2D reference-projections (15-degree

intervals) that were then subjected to multi-reference alignment (MRA) to evaluate angular space coverage. Upon inspection of cross-correlation coefficients, the bottom 7% of the particles were eliminated and the rest was used to build the initial 3D model.

A new set of projections was generated from the initial 3D model and used as references for MRA of 15,040 particle images divided into six defocus groups. Particles were then subjected to MSA and the classes were aligned through a model-based iterative refinement process until the orientation of the images stabilized. The final 3D model was built out of 551 class averages for a total of 15,040 particle images and its resolution determined by Fourier Shell Correlation (FSC) analysis. The map was rendered into surface views using CHIMERA [146].

Results

TAP-tagged PS is subject to endoproteolytic activation.

The endoproteolysis of full length PS into PS-CTF and PS-NTF constitutes the final step in the activation and maturation of γ-secretase. The formation of the PS-CTF/NTF heterodimer was then examined in HEK293S cells transiently expressing TAP-PS alone, or together with NCT, APH-1 and PEN-2. Transient expression of untagged wild-type PS (wt-PS) in HEK293S cells led to the accumulation of ~43kDa PS holoprotein (Fig. 12, lane 1), while as expected, coexpression of wt-PS with NCT, APH-1 and PEN-2 lead to high levels of PS-NTF and PS-CTF (Fig. 12, lane 2). Similar to wt-PS, transient

expression of N-terminally TAP-tagged PS lead to the accumulation of ~72kDa TAP-PS holoprotein (Fig. 12, lane 3) that was subjected to endoproteolytic processing upon coexpression of APH-1, NCT and PEN-2 generating a ~55kDa TAP-PS-NTF and ~20kDa PS-CTF (Fig. 12, lane 5). Finally, we generated a stable HEK293S cell line that coexpressed TAP-PS, together with NCT, APH-1 and PEN-2 and demonstrated the high level of accumulation of TAP-PS-NTF and PS-CTF derivatives (Fig. 12, lane 5).

Collectively, these data reveal that in transfected cells, TAP-PS is subject to endoproteolytic processing indicative of a fully assembled complex.

Purification of γ-secretase.

At all stages during purification, fractions were monitored for γ-secretase components by western blotting. Since the input amount of CHAPSO-solubilized γ-secretase was adjusted so that the bulk of TAP-PS bound to the IgG-Sepharose beads, only low levels of γ-secretase components were detected from trichloroacetic acid (TCA)-precipitated fractions corresponding to the flow through and wash steps (Fig. 13A).

We also observed that TEV protease digestion effectively released CBP:PS-NTF and other associated γ-secretase components from the IgG-Sepharose beads. This fraction was supplemented with 1 mM CaCl2 and used as input for the next purification step using calmodulin beads. Again, we detected very low levels of γ-secretase components in the flow through and wash steps after TCA precipitation of these fractions. The final

elution of γ-secretase from the calmodulin beads included 5 mM EGTA in the buffer solution. While evaluation of the final eluate by western blot analysis revealed the presence of all components of the complex, it was remarkable that the vast majority of NCT in the preparation were the mature, highly glycosylated species (Fig. 13A, lane 8; compare to input in lane 1).

The concentrated final eluate was loaded on SDS-PAGE and visualized by silver staining revealing that the preparation is >95% pure (Fig. 13B). By comparison with BSA standards ran in parallel on SDS-PAGE, we calculated that from 5g of pelleted cells corresponding to 1 L of HEK293S cells, we typically obtained 5 to 10 ug of γ-secretase.

Detergent-purified γ-secretase is enzymatically active

To assess the γ-secretase activity of the purified complex, we performed an *in vitro* assay using the chimeric C100Flag substrate and quantified the production of Aβ40 and Aβ42 using an ECL-based assay. The production of Aβ40 was greatly inhibited by L-685,458 [143], a potent and selective transition state isosteric inhibitor of γ-secretase activity (Fig. 13C), as shown by the significant decrease in the ECL signal compared to that of DMSO alone (Fig. 13C).

Moreover, the ratio of Aβ40 to Aβ42 generated by the purified γ-secretase preparation very closely resembled the ratio observed in the conditioned medium of mammalian

cells and *in vivo* settings [147] (Fig. 13D). The calculated specific activity of our γ-secretase preparation for production of Aβ40 and Aβ42 are 22.5 $pM \cdot min^{-1}$ and 4.9 $pM \cdot min^{-1}$, respectively. Under the described experimental conditions a single γ-secretase complex catalyzed the production of 28 to 56 molecules of Aβ40 per min and 6 to 12 molecules of Aβ42 per min.

Electron microscopy of single particles of γ-secretase complexes

We imaged γ-secretase complexes in 0.25% CHAPSO in complex with the polyclonal antibody NCT54. γ-Secretase/NCT54 complexes were applied to glow discharged carbon-coated copper grids and stained with a solution containing 6% ammonium molybdate and 2% trehalose. This stain allowed us to enhance the contrast in the micrographs collected (Fig. 14A) revealing highly homogeneous globular particles of approximately 150 Å in diameter. A closer inspection revealed that a fraction of raw images displayed γ-secretase complexes decorated with the characteristic trilobed shape of a single NCT54 antibody (Fig 14A, inset), indicating the presence of a single NCT molecule per γ-secretase complex. Raw particles (a total of 15,040) were manually selected with EMAN [144] and the 2D images processed with SPIDER [145] from which a stable refined structure was generated after 26 iterations of particle parameter refinement. Back projections of the 3D reconstruction were obtained and found to closely correspond with the experimental class averages generated (Fig 14B). The consistency between the reconstructed structure and the particle dataset was also confirmed when found that the final distribution of Euler angles covered all the angular

space (Fig. 14C). The final structure was determined to a resolution of 25 Å based on a Fourier shell correlation (FSC) of 0.5 between two half-datasets of unmasked particles (Fig 14D).

Description of the 3D reconstruction of the negatively stained γ-secretase complex

The 3D reconstruction for the γ-secretase complex was rendered into surface views using CHIMERA. For clarity, figure 15A displays a front view of the map contoured at voxel thresholds of 3 (transparent blue) and 3.5 (solid green) times the standard deviation (σ) above the mean voxel intensity. At 3σ the complex is 130Å wide (x-axis), 110Å deep (y-axis) and 150Å tall (z-axis), and using an average protein density of 1.41g/cm^3 [148], a mass of ~ 500kDa was calculated. In addition, we were able to observe that in a plane perpendicular to the z-axis, a strong EM density of ~50Å high surrounded the entire map as a continuous belt (Fig. 15A-E), a feature that likely corresponds to the transmembrane region of the complex. Moreover, is notorious the asymmetric distribution of mass outside of the plane of the membrane (z-axis) with most of it located on one side (top side in Fig. 15A). Firstly, a comparison between 2D images of γ-secretase particles that are decorated with NCT54 (Fig 14A, inset) and the EM map of γ-secretase, suggested that NCT54 binds to this side. Secondly, a large domain was found to cover most of the membrane surface with a maximum height of ~50Å and a calculated mass of ~150kDa, which is consistent with the mass of a fully glycosylated ectodomain of NCT. On the cytoplasmic side of the complex, only a relatively small amount of mass was observed. Taken together, we can conclude that the distribution of

mass in our 3D reconstruction is consistent with the overall topology predicted for γ-secretase.

The surface of the map is continuous except in three regions where low-density openings can be observed (Fig. 15, B-G, white arrows). These openings lead to a large internal cavity (comprising a volume of ~150,000Å^3) revealed when the map is sliced at different heights along an axis parallel to the membrane plane (Fig. 16, A-F). The cavity extends from the extracellular surface of the membrane where it has maximum dimensions (70 x 50Å in the membrane plane) (Fig. 16D), to past the membrane centre where it starts narrowing (Fig. 16E), to finally close in the cytoplasmic side (Fig. 16F). On the extracellular side, the large opening appears to be blocked by a region attributed to the ectodomain of NCT (Fig. 16, B-C).

Localization of the active site of γ-secretase with a gold-labeled transition state analog inhibitor

To address the issue of the location of the active site, we used a transition state analog known to target the active site in presenilin, L-685,458, [143] to which a newly designed spacer arm conjugated to biotin was attached [149]. Since gold labeling is often used to mark specific protein sites for EM [150, 151], we argued that after inhibition, binding of Nanogold®-Streptavidin to the biotin moiety of the inhibitor would allow us to label the active site and by default PS.

Initially, in an attempt to determine the optimum streptavidin-inhibitor distance for active site labeling, inhibitors with different spacer lengths where immobilized to streptavidin beads and incubated with detergent-solubilized γ-secretase. Only when the length of the spacer arm was >30.5Å could the complex be captured (Yueming Li, not shown), suggesting that the active site of γ-secretase lies in the interior of the complex and is distant >30.5Å from the outside surface of the complex. Therefore, a 39.4Å spacer arm between the inhibitor and the biotin moiety was used for gold-labeling experiments.

Figure 17A shows a gallery of projection images of γ-secretase marked with a gold-labeled transition state analog inhibitor. Only one gold label (~5 nm in diameter) per complex can be clearly identified by their strong positive contrast. This one-to-one stoichiometry between the inhibitor and γ-secretase suggests that γ-secretase contains a single presenilin subunit, which is in agreement with recent biochemical studies that revealed a 1:1:1:1 stoichiometry of the four individual components in active γ-secretase [152]. Single-particle averaging of labeled complexes (100 particles) and cross-correlation against back projections of the EM map of γ-secretase was carried out to determine the location of the gold-streptavidin relative to the 3D reconstruction. As shown schematically in Figure 17B and 17C, the gold label is located in the periphery of the complex in an area corresponding to the interface between the transmembrane and extracellular region of the 3D reconstruction (Fig. 17C). A range of positions of the gold labels fit to an arc (Fig. 17B), presumably because of the flexibility conferred by the long (~39.4Å) spacer arm between the inhibitor and the label. Such a distribution is consistent with a location of the γ-secretase active site at a point (~39.4Å away) near the inside surface facing the internal cavity (Fig. 17D).

Discussion

The progress in our understanding of the structural and functional biology of membrane proteins is not as remarkable as that achieved in water-soluble proteins. The difficulty in expressing and purifying membrane proteins in high yields and to homogeneity constitutes the main challenges encountered in this field. Typical limitations towards a successful purification include the inherent propensity of membrane proteins to aggregate or to be degraded by proteases outside the lipid bilayer in addition to interferences of the detergent during separation methods. In the case of PS/γ-secretase, the most prominent member of the emerging family of I-CLiPs, certain expression approaches and purification strategies have somehow alleviated these difficulties. The choice of adequate eukaryotic expression systems (i.e. drosophila, yeast, mammalian cell culture) along with the incorporation of tags, the use of affinity resins and compatible detergents has greatly improved the yield of homogenous preparations of active γ-secretase. Given the progress achieved from these inventive purification strategies, γ-secretase has now evolved into a good candidate for structural studies using EM. In comparison to crystallographic studies for which highly purified protein in milligram amounts is necessary, very little sample is required for EM studies. For EM studies, the molecule under study must fulfill two requirements: (1) has to be large enough to be seen in images; and (2) must exist in many identical copies [153]. A mature, fully active γ-secretase, with all four components present in a biological stoichiometric ratio then constitutes an ideal sample for EM.

From HEK cells we were able to express and purify fully assembled and catalytically active γ-secretase complexes. Membranes were solubilized and the complex eluted in CHAPSO, a detergent in which the enzyme displays native-like activity [141]. We also labeled the active site of the enzyme with a transition state analog inhibitor coupled to gold. The latter allowed us to identify the location of the active site, and by default PS, and to verify that the imaged complexes indeed corresponded to catalytically active species. This strategy yielded a 3D reconstruction of ~500kDa with a protein mass distribution in strong agreement with the predicted γ-secretase topology as determined by antibody labeling of the extracellular domain of NCT.

Remarkably, imaging of γ-secretase complexes bound with a gold-coupled transition state analog inhibitor indicates that the active site in γ-secretase faces a large conical internal cavity. This cavity is surrounded by a ~35Å thick transmembrane protein wall and extends from the extracellular side, where it is covered by the large extracellular domain of NCT, to past the membrane centre, where it narrows to finally close at the cytoplasmic side. The presence of uniform stain surrounding the internal cavity suggests that it might be exposed to water, thus raising the possibility that the active site of γ-secretase is water-accessible. Such an arrangement would allow access to water needed for hydrolysis of a hydrophobic transmembrane helix substrate. Indeed, this hypothesis is supported by biochemical data, suggesting that the region around the two catalytic aspartates in TMD6 and TMD7 of PS is water exposed [96, 154]. Moreover, structural support comes from recently solved X-ray structures of Rhomboid [121] and Site-2 intramembrane proteases [120]. Rhomboid has a water-filled cavity that converges on

the catalytic Ser and opens to the extracellular side, whereas Site-2 protease contains a polar channel that allows water entry to the catalytic zinc atom. In either example, a wall of TMDs is used to create hydrophilic regions that are sequestered from the surrounding membrane lipids [155, 156].

Based on existing biochemical evidence and structural insight from our 3D reconstruction we suggest a model for the hydrolysis of a transmembrane substrate by γ-secretase (Fig. 18). Ectodomain removal of type I membrane proteins produce γ-secretase substrates containing newly generated short N-termini that are specifically recognized by the ectodomain of NCT [83] during the initial steps of substrate interaction with γ-secretase. In our map the ectodomain of NCT is located at a position covering most of the extracellular region of the γ-secretase consisting with its role as a substrate receptor (Fig. 18). The described interaction between NCT and the substrate may be concomitant with docking of the TMD of the substrate to a site on the external lipid-facing surface of γ-secretase. The existence of a docking site that is distinct, but in proximity to the active site has been previously suggested by data showing that helical peptide inhibitors of γ-secretase bind at a different site compared to transition-state analog inhibitors [157, 158]. In an ensuing step the substrate is transferred from the membrane exposed docking site to the active site where it is hydrolyzed by catalytic aspartyl moieties at the interface of the water-accessible internal cavity. Partial unwinding of helical amino acid residues around the scissile peptide bond has to probably precede cleavage. A tantalizing possibility is that exposure to the hydrophilic environment of the active site contributes to local disorder in the transmembrane helix.

The first structural glimpse into γ-secretase was nevertheless published during the last stages of our structural study. Lazarov et al. [122] obtained a negatively stained EM 3D reconstruction at a resolution of 15 Å based on ~300 kDa (measured by blue-native PAGE) γ-secretase complexes solubilized in 0.1% of the detergent digitonin under conditions in which, as previously reported by the authors [159], γ-secretase does not produce measurable physiological ratios of Aβ40 and Aβ42. In addition, the majority of the protein mass in the EM map was attributed to the transmembrane region, in contradiction with topology studies from several groups demonstrating that ~54% of the protein mass of γ-secretase is contained in the extracellular domain [160-162] (Fig. 19A). For the reconstruction of a more relevant EM map, Osenkowski et al. [123] collected cryo-EM data on γ-secretase complexes of ~230 kDa (measured by scanning transmission electro microscopy) from an enzymatic preparation eluted also in 0.1% digitonin. In this 3D reconstruction generated at a resolution of 12 Å, NCT tagged with GST was localized and with that the orientation of the complex in the membrane determined. In comparison to the first EM structure, a larger domain at the extracellular side was found (Fig. 19B). Although our reconstruction was obtained from γ-secretase complexes extracted in CHAPSO, a detergent in which the enzyme displays native-like activity, in addition to propose the location of the active site through gold labeling, our EM map was not competitive enough in the field. More aggressive approaches such as structure determination in vitreous ice to obtain an improved cryoEM map or even crystallographic experiments will be attempted in the future. To fulfill these expectations, we rely on our innovative expression and purification protocol from *P. pastoris* explained in the next chapter.

Chapter III

PURIFICATION OF γ-SECRETASE COMPONENTS FOR STRUCTURAL STUDIES

Introduction

Several groups have reported the purification of γ-secretase expressed in eukaryotic hosts. Fraering et al. reported the purification of the enzyme over expressed in Chinese hamster ovary (CHO) cells [159]. More recently, Cacquevel et al. obtained a γ-secretase preparation with higher specific activity from a more efficient line of CHO cells [163]. Despite the high merit of these efforts, the amounts of purified enzyme obtained are substantially inferior to what is generally needed for high-resolution structural studies. In contrast, we have developed a procedure to purify milligram amounts of catalytically functional and fully assembled human γ-secretase complex from the yeast *P. pastoris* in a simple and inexpensive way. The higher yields obtained allowed the addition of a size exclusion chromatography step, from which we expect to achieve a remarkable improvement in the homogeneity of the preparation for future structural studies.

Also included in this chapter is the expression and purification tested for each individual component of the γ-secretase complex from *E.coli*. Based on the fact that SPP is active on its own, *in vitro* activity assays were performed with detergent-purified PS.

***Pichia pastoris* as an expression system**

Pichia pastoris is a methylotrophic yeast species widely used for the production of recombinant proteins [164]. The use of *P. pastoris* as a host system presents several advantages relative to bacteria and other fungi. As a unicellular microorganism like bacteria, *P. pastoris* grows relatively fast in an economical manner. However, yeasts are eukaryotes with the same intracellular environment and capabilities for post-translational modifications as higher eukaryotes [165], with a high likelihood of correctly fold, process and assemble the heterologous protein into functional molecules [166].

Relative to other yeast species, *P. pastoris* can grow at very high cell densities in the presence of methanol as the sole carbon and energy source upon induction of transcription of the metabolic enzyme alcohol oxidase (AOX) [167, 168]. AOX expression is tightly regulated at the level of transcription of its gene, *AOX1*, which also serves as a promoter to control foreign gene expression [169]. In principle, the gene of interest can be stably integrated into the yeast's genome via homologous recombination at a position downstream of *AOX1* [170]. Plasmids design for this purpose harbor the wild-type allele of biosynthetic pathway genes (i.e. *HIS4*; histidinol dehydrogenase gene) as markers for selection of *his4* auxotrophic mutant hosts of *P. pastoris*. In the past, the limited number of existing markers for manipulation of *P. pastoris* genome rendered almost impossible the construction of *P. pastoris* strains with more than a few recombinant modifications. The isolation and characterization of new biosynthetic marker genes and with that, the development of new auxotrophic host strains [171] now

allow researchers to produce foreign proteins that are hetero-oligomeric or that require an additional protein not present in yeast [172]. For all these reasons, we therefore consider *P. pastoris* to be an expression system with high potential to improve the yield of fully assembled γ-secretase compared to that obtained from more elaborated approaches that use mammalian expression systems.

Materials and Methods

Generation of transgenic *Pichia pastoris*.

PCR amplified products corresponding to PEN-2 (pubmed accession number NP_758844), PS (pubmed accession number P49768), NCT (pubmed accession number NP_056146) and his-myc-tagged APH-1 (pubmed accession number AAH01230), were each inserted into EcoRI-digested intracellular expression vectors. The marker genes used for each expression vector corresponded to HIS4 for PS, URA3 for APH-1, ADE1 for NCT and ARG4 for PEN-2 resulting in plasmids pIB4-PS WT, pBLURAIX-APH-1, pBLADEIX-NCT and pBLARGIX-PEN2, respectively. Expression vectors were integrated into the genome of *P. pastoris* strain JC308 (genotype *ade1 arg4 his4 and ura3*) [171] into their respective loci after linearization and homologous recombination. All four genes corresponding to the γ-secretase complex were inserted under control of the strictly methanol-regulated AOX promoter.

Expression and purification of γ-secretase complex

One colony was picked from YPD (1% yeast extract, 2% peptone, 2% glucose, 2% agar, 20 mg/L adenine sulfate, 20 mg/L uracil) plates to inoculate 50 ml of YPG (1% yeast extract, 1% peptone, 2% glycerol) medium. Cells were grown at 30°C for 48 hours under constant agitation. Cells were then collected by centrifugation at 3000 x g at 4°C and resuspended in 200 ml of BMMY (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% yeast nitrogen base, 0.4 mg/L biotin, and 0.5% methanol) induction medium. Induction occurred at 30°C for 72 hours with a 0.5% methanol replenishment every 24. Cells were then harvested (OD_{600} >25) by centrifugation at 3000 x g for 10 minutes at 4°C and washed with lysis buffer (50 mM Tris HCl pH 7.0, 1 mM EDTA, 10% glycerol). The cell pellet was collected again by centrifugation at 10,000xg for 10 minutes at 4°C and resuspended in lysis buffer containing 1x protease inhibitor cocktail for yeast consisting of 10 ug/ml E-64, 100 uM PMSF and 1 mM phenanthroline and lysed with a microfluidizer operating at 40000 psi. The lysate was centrifuged at 10,000xg for 20 minutes at 4 °C to eliminate unbroken cells and the supernatant collected and ultracentrifuged for 1 hour at 100,000 x g at 4°C to collect the microsomal membrane pellet. The membranes were washed in 5 volumes of sodium bicarbonate buffer (0.1M $NaHCO_3$ pH 11.3) and ultracentrifuged again. The membrane pellet was washed twice in solubilization buffer (50 mM Tris HCl pH 7.0, 400 mM NaCl, 10% glycerol) before resuspension in 30 ml of solubilization buffer containing 1x protease inhibitor cocktail for yeast and CHAPSO to a final concentration of 24 mM (1.5% v/v; 3 times the CMC). Membrane solubilization was allowed to occur for 4 hours at 4°C

followed by ultracentrifugation at 100,000xg for 1 hour to separate the insoluble material. The membrane solubilizate was then incubated for 4 hours with 1 ml of pre-equilibrated Ni-NTA resin (Qiagen). The resin was packed into a disposable column, washed twice using 10 column volumes of washing buffer (50 mM Tris HCl pH 7.0, 400 mM NaCl, 10 mM imidazole, 10% glycerol and CHAPSO to a final concentration of 4 mM) and the protein eluted in 10ml of elution buffer (50 mM Tris HCl pH 7.0, 400 mM NaCl, 250 mM Imidazole, 10% glycerol and 4 mM CHAPSO). The fractions eluted from the Ni-NTA resin were combined and concentrated to 500 ul using a 30K centrifugal filter device (Millipore) and injected into a high performance size exclusion chromatography (SEC) column (TSK-G4000SWxl, TOSOH Bioscience) with a separation range of 10^3 – 7 x 10^6 Da (globular proteins). Fractions corresponding to γ-secretase were finally collected in SEC buffer (50 mM Tris HCl pH 7.0, 400 mM NaCl, 10% glycerol and 4 mM CHAPSO) in the presence of 1x protease inhibitor cocktail for yeast.

Antibodies and western blotting

The presence of γ-secretase components during purification was confirmed by immunoblot analysis. Samples were fractionated on either 12% or 15% SDS PAGE and transferred to a PVDF membrane. The antibodies used included mouse monoclonal anti-Nicastrin (NCT C-terminus; Millipore), polyclonal rabbit anti-PEN-2 (PEN-2 C-terminus, SIGMA), mouse monoclonal anti-Myc (myc tagged Aph1; Millipore) and mouse monoclonal anti-Presenilin 1 (full-length PS and PS CTF; ABR).

In vitro γ-secretase activity

Samples at a final concentration of 1 ug/ml corresponding to the membrane solubilizate in 24 mM CHAPSO, the elution from the Ni-NTA column and to the fractions collected from the SEC were incubated with C100Flag at a final concentration of 1uM in the presence or absence of the γ-secretase inhibitor L685,458. γ-Secretase-mediated cleavage to generate the Aβ40 peptide was measured by electrochemiluminescence (ECL) assays using ruthenylated G2–10 murine monoclonal anti-Aβ40 and detected with biotinylated murine monoclonal antibody 4G8.

Results

Purification of mature γ-secretase from the cell membrane of *Pichia pastoris*

Pichia pastoris cells that stably coexpress NCT, PEN-2, PS and his-myc-tagged APH-1 under control of the strict inducible AOX1 promoter were generated. From samples of lysed cells taken after prolonged induction of protein production with methanol for 72 hours, we were able to observe cross-reactivity after anti-PS C-terminus immunoblot analysis (Fig. 20A). The generation of a band corresponding to PS CTF (Fig. 20A), was indicative of the endoproteolytic processing of PS holoprotein dependant on the assembly of all four components and therefore, of the heterologous expression of fully assembled γ-secretase complex from *P. pastoris*.

The purification of fully assembled γ-secretase complex was greatly facilitated by the insertion of a histidine tag in frame with the sequence corresponding to APH-1. In addition, APH-1 bearing a myc tag improved the detection of this component by immunoblot analysis with anti-myc monoclonal antibodies. Thus, detection of a predominant band corresponding to his-myc APH after elution from the Ni-NTA column confirmed that a higher yield of complex was obtained when the elution buffer contained CHAPSO at a final concentration of 24 mM (Fig. 20B). We further confirmed that the majority of the γ-secretase complex bound to the Ni-NTA affinity resin and was selectively eluted in the presence of 250 mM of imidazole as shown by the detection of bands at the corresponding molecular weight for each γ-secretase individual subunit (Fig. 20C).

In addition, different from other methods reported for the purification of γ-secretase complex, a SEC step was included (Fig. 20D). After evaluation of the fractions by immunoblot analysis, the presence of all components of the complex was confirmed. Since the vast majority of NCT in the preparation was the mature, highly glycosylated species (unglycosylated NCT is 78 kDa), we argue that the bulk of γ-secretase isolated using this purification strategy selected for fully matured γ-secretase complexes that reside in post ER compartments. The yield of mature γ-secretase complex was approximately, 0.2 mg/liter of *P. pastoris* culture.

Detergent-purified γ-secretase complex is active

Ectodomain shedding of APP precedes cleavage by γ-secretase. Hence, to assay the activity of the γ-secretase fractions collected throughout the purification procedure we incubated the enzymatic preparation with C100Flag, a substrate consisting of the βCTF portion of APP generated after processing by β-secretase. Cleavage of C100Flag measured by ECL using derivatized anti-Aβ antibodies was detected in low levels for γ-secretase present in the CHAPSO solubilizate fraction. Instead, a higher production of Aβ40 was detected after the incubation of C100Flag with eluted γ-secretase from the Ni-NTA affinity resin and with that collected from the SEC. Finally, full inhibition of the generation of Aβ40 by the potent transition state analog inhibitor L-685,458 confirmed that the activity displayed was γ-secretase specific (Fig. 21).

***E. coli* as an expression system**

We initially established as our goal to produce milligram quantities of homogeneous and functional human PS, APH-1, PEN-2 and NCT separately from *E. coli* to be used for biochemical, biophysical and structural studies. The task proved to be challenging and we channeled our efforts to express and purify PSdelta9, an AD-associated mutation within the PS gene that deletes exon 9 due to a splicing error and results in the accumulation of the uncleaved full-length protein [173]. Based on the fact that SPP, different from PS, does not require the presence of accessory proteins to be active, we then decided to test a similar cleavage capacity with PSdelta9 in *in vitro* activity assays.

Materials and Methods

Cloning, plasmids and strains

E coli strain DH5α (Invitrogen) was used for all the cloning steps. *E. coli* competent cell strains Rosetta(DE3)pLysS and BL21(DE3) (Novagen) were employed for over expression of individual γ-secretase components and the C100Flag substrate, respectively.

Standard PCR techniques were use to amplify the human genes corresponding to PEN-2, APH-1, NCT, PS and PSΔexon9 from template plasmids. Amplified genes were then cloned into a T7 promoter-based, IPTG-inducible, pET system (Novagen) developed for expression of recombinant proteins in *E. coli*. Translation of the constructs was engineered to express a recombinant protein bearing a hexa histidine-tag at either the N- (pET15b) or C-terminus (pET29b). The linker between the gene of interest and the hexa histidine tag has 27 and 10 amino acid residues when the tag is located at the C- and at the N-terminus respectively, in both cases including a thrombin cleavage site. After confirming the identity of the cloned fragments by DNA sequencing, the recombinant plasmids were used to transform *E. coli* host strain Rosetta(DE3)pLysS, a BL21 derivative is designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*.

Construction of chimeric PSΔexon9 involved the amplification of the gene

corresponding to PSΔexon9 along with the fragment corresponding to the flag tag (DYKDDDDK octapeptide) at the C terminus. The amplified product was cloned into pET29 expression vector and the recombinant plasmid used to transform *E. coli* competent cell strain Rosetta(DE3)pLysS.

The recombinant C100Flag substrate (108 amino acid residues; MW=12.3 kDa) derives from the APP C-terminal fragment (βCTF) corresponding to amino acids Met596-Asn695 of the 695-amino acid long isoform of APP (pubmed accession number CAA68374). An amplified DNA fragment encoding 100 amino acids corresponding to the βCTF in addition to an octapeptide corresponding to the flag tag at the C terminus was cloned into pET21 expression vector (Novagen) [141]. The recombinant plasmid was used to transform *E. coli* host strain BL21 (DE3).

Standard purification procedure

For a standard purification experiment, an aliquot from a starting culture grown at 37°C in Luria-Bertani (LB) broth was taken to inoculate 1 lt of LB broth. Cells were allowed to grow to mid-log phase at 37°C before induction with isopropyl-β-D-thio-galactoside (IPTG, Gold Biotechnology) to a final concentration of 0.4 mM. After 4 hours, cells were harvested by centrifugation at 6 000 x g and the pellet resuspended in lysis buffer (50 mM Tris-HCl pH 8.00, 300 mM NaCl, 10 mM Imidazole) in the presence of 1x protease inhibitor cocktail for bacteria consisting of 10 ug/ml for E-64, 14 ug/ml for Bestatin and 100 uM for PMSF. After passing the suspension twice through a French

press (35mL pressure cell) operating at 11,000 *psi*, unbroken cells were separated by centrifugation at 43 000 x g for 15 minutes at 4°C and the soluble fraction was ultracentrifuged at 148 000 x g for 30 minutes. Inner membranes were then collected and resuspended in lysis buffer in the presence of 1x protease inhibitor cocktail for bacteria. This membrane suspension was subjected to solubilization for 4 hours at 4°C by adding the detergent Triton X-100 (TX-100, CMC = 0.013-0.056%; Anatrace) to a final concentration of 2% w/v. Insoluble material was then removed by ultracentrifugation (30 minutes, 148 000 x g at 4°C) and the supernatant containing the detergent-solubilized membranes was incubated with 1 ml of pre-equilibrated Ni-NTA resin. After 2 hours at 4°C under slow rotation, the resin was poured onto a disposable column and first washed with 15 ml of lysis buffer containing TX-100 to a final concentration of 0.2% w/v and secondly with 15 ml of lysis buffer containing n-dodecyl-ß-D-maltoside (DDM, CMC=0.0077%; Anatrace) to a final concentration of 0.1% w/v. Finally, a third wash was carried out to elute unspecific bound proteins using wash buffer (50 mM Tris-HCl, 300 mM NaCl, 20 mM Imidazole) plus DDM to a final concentration of 0.1% w/v. Histidine-tagged proteins were eluted in elution buffer (50 mM Tris-HCl, 300 mM NaCl, 250 mM Imidazole and 0.1% DDM) in the presence of 1x protease inhibitor cocktail for bacteria.

Purification of PSΔexon9

An aliquot from a starting culture was taken and use to inoculate 4 lts of 2XYT media. After cells reached an A_{600} ~0.7 shaking at 37°C, induction of PSΔexon9 production

with IPTG (0.4 mM final concentration) was carried out at 18°C for 16 hours. The cell pellet after harvesting (of approx. 20 ml) was resuspended in 60 ml of resuspension buffer (20 mM NaHepes pH 7.5, 250 mM NaCl, 1 mM EDTA, 10% glycerol, DNaseI) in the presence of 1x protease inhibitor cocktail for bacteria. After breaking the cells using a French press operating at 11,000 *psi*, unbroken material was centrifuged and inner membranes collected and resuspended in 50 ml of solubilization buffer (Tris-buffered saline; TBS pH 7.4, 10% glycerol) containing 1% DDM in the presence of 1x protease inhibitor cocktail for bacteria. After 3 hrs of solubilization, the insoluble material was ultracentrifuged and the supernantant incubated with 300 ul of pre-equilibrated anti-flag M2 affinity resin (Sigma). Binding was allowed to occur for 2 hrs at 4°C, under constant, slow rotation. The resin was then packed into a disposable column and washed with 30 column volume (cv) of solubilization buffer containing 0.1% DDM. PSΔexon9 was eluted by competition with 3x flag peptide (Sigma) diluted in TBS pH 7.4, 0.1% DDM in the presence of 1x protease inhibitor cocktail for bacteria.

Purification of recombinant C100Flag substrate

An aliquot from a starting culture was taken and used to inoculate 2 lts of LB broth. The cell culture was allowed to grow at 37°C until an $A_{600} \sim 0.7$ was reached before inducing protein over expression by adding IPTG to a final concentration equal to 0.4 mM. After 4 hours, cells were harvested by centrifugation at 6 000 x g and the pellet resuspended in ~20ml of lysis buffer (10 mM Hepes pH 7.40, 150 mM NaCl, 1 mM EDTA, DNaseI and 10% glycerol) in the presence of 1x protease inhibitor cocktail for bacteria. The

suspension was passed through a French press operating at 11,000 *psi* and the unbroken material spun down at 43 000 x g for 15 minutes. The supernatant was collected and the inner membranes isolated after ultracentrifugation at 148 000 x g for 45 minutes. Membranes were resuspended in PBS, pH 7.4 in the presence of 1x protease inhibitor cocktail for bacteria and membrane solubilization was allowed to occur for 4 hrs at 4°C after addition of TX-100 to a final concentration of 2% in a final volume of 25 ml. The insoluble material was separated by ultracentrifugation and the supernatant incubated with 300 ul of pre-equilibrated anti-flag M2 affinity resin (Sigma). After 1 hour under slow rotation at 4°C, the M2 resin was first washed with 3 x 5 ml PBS plus 0.2% TX-100 followed by 3 x 5ml PBS plus 0.1% DDM. Chimeric C100Flag substrate was finally eluted by competition with 3x flag peptide diluted in TBS pH 7.4 in the presence of 1x protease inhibitor cocktail for bacteria plus 0.1% DDM.

PSΔexon9 identification by mass spectrometric analysis

Eluted protein from the M2 resin was separated by SDS-PAGE and visualized by coomassie staining. A protein band presumably corresponding to PSΔexon9 was excised from the gel and washed with acetonitrile/100 mM ammonium bicarbonate (ABC, 45:55, v/v). The gel slice was reduced (10 mM TCEP), alkylated (50 mM iodoacetamide), and then digested *in situ* with trypsin (200 ng per band in 50 mM ABC). The tryptic peptides were extracted using POROS 20 R2 beads (Applied Biosystems) in the presence of 5% formic acid and 0.2% trifluoracetic acid and then dried with a SpeedVac evaporator. The resulting peptides were dissolved in 8 ul of HPLC sample solvents containing

water:methanol:acetic acid (70:30:1 v/v/v). Micro-HPLC-MS/MS analysis was conducted on an LCQ electrospray ionization ion trap mass spectrometer (ThermoFinnigan) coupled with an online MicroPro-HPLC system (Eldex Laboratories). Two microliters of the tryptic peptide solution was injected into a Magic C18 column (5um, 200 Å, 0.2 × 50 mm, Michrom BioResources) pre-equilibrated with 70% solvent A [0.5% acetic acid and 0.01% TFA in water/methanol (95:5, v/v)] and 30% solvent B [0.5% acetic acid and 0.01% TFA in methanol/water (95:5, v/v)]. Peptides were separated and eluted from the HPLC column with a linear gradient from 30 to 95% of solvent B in 15 minutes at a flow rate of 3 ul/min. The eluted peptides were sprayed directly into the LCQ mass spectrometer. The LCQ mass spectrometer was operated in a data-dependent mode for measuring the molecular masses of peptides (parent peptides) and collecting MS/MS peptide fragmentation spectra. The measured molecular masses of parent peptides and their MS/MS data were used to search the National Center for Biotechnology Information DNA/protein non-redundant sequence databases using the program KNEXUS (Genomic Solutions).

Results

Initial efforts were mainly focused on expressing and purifying PEN-2, APH-1 and PS. From small-scale expression trials, the level of expression of constructs in pET15b were very low compared to those detected upon expression of the human genes in pET29b constructs (Fig. 22), carrying out further experiment with the latter.

Large-scale purification of chimeric PEN-2 (MW=15.5 kDa) yielded protein in sufficient quantities to add two more steps to the standard purification procedure, an anion exchange chromatography (High-Q) and a size-exclusion (superdex HR200) chromatography (Fig. 23 A, B). On the other hand, modifications in the standard procedure had to be introduced to achieve improvements during the purification of chimeric APH-1 (MW=32.5 kDa). From small-scale expression trials, anti-His immunoblot analysis revealed a band likely to correspond to a degradation product derived from chimeric APH-1 (<25 kDa) in addition to an equally relevant band at the corresponding molecular weight for full-length APH-1 (Fig. 24A). We speculated that protein induction at 20°C rather than at 37°C could avoid proteolytic degradation of the full-length chimeric APH-1. Indeed, by lowering the temperature from 37°C to 18°C after addition of IPTG, inactivation of this proteolytic event was achieved. Furthermore, higher yields of chimeric APH-1 were detected when DDM (at a final concentration of 2%) was used to solubilize inner membranes instead of TX-100 (FIG. 24B). Binding of the C-terminally tagged protein to Ni-NTA resin was nevertheless inefficient under the conditions tested severely compromising the purification of APH-1.

Drawbacks during cloning of NCT and during the purification of PEN-2, PS and APH-1 were later encountered. The individual purification of all four membrane proteins to accomplish *in vitro* reconstitution of active γ-secretase complex therefore became challenging. We then focused our attention on PSΔexon9, an AD-associated mutation lacking exon 9, site of endoproteolysis of PS. Although the endoproteolytic activation of PS is a necessary step during the maturation of the γ-secretase complex, conveniently,

PSΔexon9 results in active uncleaved full-length protein as seen in K293 cells stably transfected with APP containing the Swedish mutation [173].

Purification of chimeric PSΔexon9 (MW=53.7kDa) was carried out with several modifications introduced into the standard protocol. The addition of an affinity M2 resin step greatly improved the quality of purified PSΔexon9 in sufficient quantities for protein identification from an excised band after separation from a coomassie-stained SDS-PAGE (Fig. 25). Based on the literature demonstrating that SPP is a di-aspartyl protease capable of intramembrane proteolysis in the absence of other proteins acting as cofactors, we tested whether PSΔexon9 also displayed activity on its own. Samples corresponding to the membrane solubilizate in 1% DDM and to the elution from the M2 affinity resin were incubated with C100Flag as described in the *in vitro* activity assay for γ-secretase purified from *P. pastoris*. An ECL signal was nevertheless undetectable suggesting that under the purification conditions employed PSΔexon9 fails to generate the Aβs peptide from C100Flag.

Discussion

Two studies are of special interest in regards to PS/γ-secretase expression and purification of active γ-secretase in amounts adequate for structural studies. In the first one, Fraering *et al.* [159] described a multi-step procedure for high-grade purification of intact, proteolytically active γ-secretase complexes in 0.1% digitonin from CHO cells. Briefly, membranes were prepared from CHO cells expressing human PS, N-terminally

flag-tagged PEN-2 and C-terminally HA-tagged APH-1. Isolation of the complex was achieved by performing two sequential affinity steps changing detergents among CHAPSO, digitonin and Brij35 to finally elute the complex in 1% Brij35-HEPES buffer. A γ-secretase transition state analog (31C) affinity resin followed by an M2 anti-flag affinity resin separated the complex from other proteins based on the affinity tag linked to PEN-2. The cleavage efficiency of the eluted γ-secretase complex adjusted to 0.1% digitonin by dilution was assessed in terms of the *in vitro* production of Aβ40 and Aβ42 after incubation with C100Flag [141] and by mass spectroscopy. This study also identified certain conditions that are critical for cleavage efficiency such as cell membrane lipid composition.

In the second study, the γ-secretase complex was isolated from a more efficient line of CHO cells expressing human PS, N-terminally flag-tagged PEN-2, C-terminally HA-tagged APH-1 and C-terminally V5-tagged NCT, this time using a single affinity step during a shorter purification procedure [163]. For this preparation, membranes were initially solubilized and processed in 1% CHAPSO to finally dilute the final eluate in buffer containing 0.1% digitonin. This protocol allowed the production of high levels of mature, active, and homogeneous γ-secretase particles with ~4-5-fold higher specific activity compared to the cell line used in the protocol described above.

A strategy involving the use of the yeast *Pichia pastoris* as an expression system seems to constitute a very promising and more advantageous tool to achieve higher yields. Although the expression of exogenous PS fully replaced endogenous levels in HEK cells

as seen before in neuroblastoma cell lines [174], advantageously, in *P. pastoris* endogenous γ-secretase activity is not present [175]. In the described procedure, the cell mass of *P. pastoris* reached high densities from which a resulting higher yield of active fully assembled γ-secretase complex was obtained. Different from other purification procedures reported to date, this strategy allowed the incorporation of a separation step involving SEC in order to improve the homogeneity of the complex. This preparation has now become a preliminary sample for EM studies from which we expect to obtain a cryo-EM map of improved resolution.

From a more challenging approach, expression of each individual γ-secretase component from *E. coli* for *in vitro* reconstitution following their purification was attempted. Expression of tagged APH-1, PEN-2, PS and PSdelta9 was achievable although further improvements in purification were attained only for the last. Based on evidences that another di-aspartyl I-CLiP, SPP, is capable of processing membrane protein substrates in the absence of accessory proteins, we then explored whether our DDM-purified PSdelta9 is capable of the same. Conveniently, PSdelta9 mimics an endoproteolyzed and biologically active PS heterodimer [173, 176-178], increasing the probability of success of our *in vitro* cleavage assay in the absence of other members of the γ-secretase complex. Under the conditions tested, PSdelta9 nevertheless failed to display activity when incubated with C100Flag. In the future, additional purification and incubation conditions are worth of being tested for PSdelta9 however due to the inherent complexity of eukaryotic proteins we believe that a more tractable prokaryotic system embraces more promising outcomes.

Chapter IV

IDENTIFICATION OF A GXGD TYPE DI-ASPARTYL INTRAMEMBRANE PROTEASE IN EURYARCHAEOTA *METHANOCULLEUS MARISNIGRI* STRAIN JR1

Introduction

Sequence analysis of PS and SPP raised the hypothesis that archaea may possess GXGD type di-aspartyl intramembrane proteases [105]. Based on sequence alignment we were able to identify a putative homolog from euryarchaeota *Methanoculleus marisnigri* JR1 (MCMJR1) with the following biochemical characteristics: (1) MCMJR1 is an integral membrane protein with eight predicted TMDs; (2) The protein has YD and GXGD motifs, which contain catalytic aspartates on adjacent TMDs; (3) The enzyme cleaves model intramembrane protease substrates at multiple sites within their TMD; (4) MCMJR1 also cleaves a *bona fide* PS substrate; (5) MCMJR1 activity is inhibited by a PS transition state analogue inhibitor; and (6) The mutation G219A in MCMJR1, adjacent to the catalytic D220, slows down cleavage, as it does in both PS and SPP. To our knowledge MCMJR1 represents the first archaeal GXGD di-aspartyl intramembrane protease with significant biochemical similarities to PS and with high potential to be employed as a model system for biochemical and structural studies of γ-secretase.

Materials and methods

Cloning, plasmids and strains

E coli strain DH5α (Invitrogen) was used for all the cloning steps. *E. coli* competent cell strains Rosetta(DE3)pLysS and BL21 (DE3) (Novagen) were employed for expression of proteases and protein substrates, respectively.

Based on sequence alignment to published PS homologs [105] using BLAST (Basic Local Alignment Search Tool), twelve PS- and nine TFPP-like genes were identified in commercially available archaeal genomes. Standard PCR techniques were used to amplify the genes of interest that were subsequently cloned into a T7 promoter-based, IPTG-inducible, pET system (Novagen). Translation of the constructs based on PS- and TFPP-like genes predict a recombinant protein tagged with 10 histidines at either the N- or C-terminus (10xHis*gene*). Constructs corresponding to PS-like genes were also engineered to express as N-terminal fusions to small ubiquitin modifying protein (SUMO) following a hexa histidine-tag (6xHisSUMO*gene*).

The chimeric protein substrates contain the predicted TMD regions of human APP (pubmed accession number CAA68374; Gly625-Met647) and that of ligands of the EGFR from *Drosophila melanogaster*, spitz (pubmed accession number NP_476909; Ala142-Leu164), keren (pubmed accession number NP_524129; Ala122-Leu144) and gurken (pubmed accession number NP_476568; Ile248-Leu271). Using standard PCR

techniques these sequences were amplified and cloned into our modified PET29b plasmid (MBPpET29b). Translation of the chimeric substrates constructs predict either a 458 (MBPSpitz/Keren/Gurken; MW=50.3 kDa) or a 436 amino acid-residue protein (MBPAPP; MW=47.5 kDa), both including a C-terminal hexa histidine-tag and an N-terminal MBP (MW=42 kDa) tag separated from the TMD region by a thrombin site.

Expression and large-scale purification of proteases

Constructs corresponding to PS-like genes, MCMJR1 (from *Methanoculleus marisnigri* JR1; pubmed accession number YP_001047832; 6xHisSUMOMCMJR1) and MHJF (from *Methanospirillum hungatei* JF-1; pubmed accession number YP_503938; 6xHisSUMOMHJF) and TFPP-like genes, MCV (from *Methanococcus voltae*; pubmed accession number AAM34242; 10xHisMCV) and MMS2 (from *Methanococcus maripaludis* S2; pubmed accession number NP_987675; 10xHisMMS2) were chosen for scaling up. Recombinant plasmids were transformed into Rosetta(DE3)pLysS and cells allowed to grow overnight in 10 ml of LB media. The cell pellet was collected from the starting culture and resuspended in 2XYT (yeast extract tryptone) media from which 800 ul were taken to inoculate 800 ml of 2XYT. Cells reached mid-log phase after approximately 3 hours shaking at 37°C. Protein production was induced with IPTG to a final concentration of 0.1 mM and protein over expressed at 18°C for 16 hours. The cell pellet of approximately 5gr was collected after centrifugation at 6000 x g and resuspended in 20 ml of buffer A (20 mM NaHepes pH 7.5, 250 mM NaCl, 1 mM $MgSO_4$, 1mM βME) containing DNaseI and 1x protease inhibitor cocktail for bacteria.

This suspension was passed twice through a French Press operating at 11,000 psi and the debris cleared by centrifugation at 43,000 x g for 15 minutes. Membranes were isolated by ultracentrifugation at 125,000 x g for 1 hour, resuspended in a final volume of 50 ml of buffer A and solubilized with 1% (w/v) of either DDM or Fos-choline-12 (FC-12, Anatrace).

Insoluble material was removed by ultracentrifugation (30 minutes, 148 000 x g in a Type 50.2 Ti rotor, at 4°C) and the supernatant containing the detergent-solubilized membranes was incubated with 1.5ml of pre-equilibrated Ni-NTA resin. The detergent concentration was maintained at either 0.1% for DDM or 0.2% for FC-12 during all the steps following the initial solubilization.

After 2 hours at 4°C under slow rotation, the resin was poured onto a column and washed with 10 ml of Buffer B (20 mM NaHepes pH 7.5, 250 mM NaCl, 50 mM imidazole and 1 mM βME). Histidine-tagged proteins were eluted in Buffer C (20 mM NaHepes pH 7.0, 250 mM NaCl, 200 mM Imidazole and 1 mM βME) in the presence of 1x protease inhibitor cocktail for bacteria and fractions containing recombinant proteins 6xHisSUMOMCMJR1 or 6xHisSUMOMHJF were pooled in a final volume of 2 mL. Histidine-tagged SUMO protease was added (approximately 35 ug) to allow the release of the 6xHisSUMO tag while dialyzing over night against buffer D (20 mM NaHepes pH 7.0 containing, 250 mM NaCl, 2 mM βME). After re-passing the mixture through pre-equilibrated Ni-NTA resin and retaining both, the histidine-tagged SUMO protease and the released 6xHisSUMO tag, MCMJR1 and MHJF were collected in the flow-

through.

Similarly, eluted 10xHisMCV and 10xHisMMS2 from Ni-NTA resin were pooled and dialyze against buffer D.

Expression and large-scale purification of substrates

Constructs corresponding to chimeric substrates were transformed into BL21(DE3) cells, positive clones picked and used to inoculate an overnight starting culture from which an aliquot was used to inoculate 1lt of enriched media (0.85% Na_2HPO_4, 0.03% KH_2PO_4, 0.5 g/L NaCl, 0.01% NH_4Cl, 1% tryptone, 0.5% yeast extract, 2 mM $MgSO_4.7H_2O$, 1 mM $CaCl_2$, 0.6% Glucose, 10 mg/ml Thiamine, 40 μg/ml Kanamycin). The cell culture was allowed to grow at 37°C until an A_{600}~0.7 was reached before inducing protein production by adding 0.4 mM IPTG. After 4 hours the cell pellet was collected by centrifugation and resuspended in 20 ml of lysis buffer (50 mM Tris-HCl pH 7.4, 300 mM NaCl, 10 mM Imidazole, 10% glycerol and 0.1 μg/ml DnaseI) in the presence of 1x protease inhibitor cocktail for bacteria. The suspension was passed through a French press operating at 11,000 psi and the unbroken material spun down at 43 000 x g for 15 minutes. The supernatant was collected and bacterial membranes were isolated by ultracentrifugation at 125 000 *x g* for 45 minutes. Inner membranes were solubilized in a total volume of 25 ml of lysis buffer plus 2% TX-100 for 4 hours. The insoluble material was separated by ultracentrifugation and the supernatant incubated with 0.5 ml of pre-equilibrated Ni-NTA resin. After 4 hours at 4°C, the resin was packed into a disposable

column and the detergent exchanged to 0.1% DDM while washing the resin with lysis buffer containing 0.2% TX-100 and then with wash buffer (50 mM Tris-HCl pH 7.4, 300 mM NaCl, 20 mM Imidazole and 10% glycerol) containing 0.1% DDM. The MBP chimeric substrate was then eluted in 50 mM Tris-HCl pH 7.4, 300 mM NaCl, 250 mM Imidazole, 10% glycerol in the presence of 1x protease inhibitor cocktail for bacteria and 0.1% DDM. Finally, the protein solution was dialyzed against PBS pH 7.40 containing 10% glycerol and 0.1% DDM.

In vitro activity assay

MBP chimeric substrates were incubated with 10xHisMCV, 10xHisMMS2, MHJF1 and MCMJR1 purified in either DDM or FC-12 and mutant variants of MCMJR1 purified in DDM to test their activity. Protein concentration was determined by mixing 5ul of purified protein with coomassie (Bradford) protein assay reagent (Thermo Scientific) and reading absorbance at 595 nm. Chimeric membrane protein substrates (1.5 uM) were mixed with each of the four detergent-purified archaeal membrane proteins (4 uM) in a final volume of 25 ul in PBS buffer pH 7.4, containing the appropriate detergent in the presence of 1x protease inhibitor cocktail for bacteria. After the reaction mixture was incubated for 8 hours at 37°C, protease activity generated a N-terminal product (~46 kDa) that contains the MBP tag and a C-terminal product that contains the hexa histidine-tag (~4 kDa). Detection of the former was achieved by taking an aliquot of the reaction mixture containing approximately 1.5 ug of chimeric substrate, adding sample loading buffer and boiling for 5 minutes prior to loading the sample onto a 10% tris-

glycine gel for SDS-PAGE. The band corresponding to the N-terminal product was then visualized from the coomassie-stained gel. For confirmation of the identity of the band by anti-MBP immunoblot analysis, an aliquot equivalent to approximately 4 ng of chimeric substrate was also taken from the mixture, loaded onto a 10% SDS- PAGE and transfer to a PVDF membrane. Although uncommon, incubation of the reaction mixture with 10ul of pre-equilibrated Ni-NTA resin was necessary to achieve a better discrimination between the band corresponding to the full-length substrate and the N-terminal product before anti-MBP immunoblot analysis. After 1 hour at 4°C, samples were taken from the flow through.

In addition, DDM-purified MCMJR1 was incubated with γ-secretase recombinant substrate C100Flag. The incubation conditions for the *in vitro* activity assay included larger amounts of C100Flag (15 uM) in order to improve the detection of a flag-tagged C-terminal product by anti-Flag immunoblot analysis.

Processing of the SDS-PAGE gel image with Image J (image processing and analysis in Java, NIH) allowed a comparative quantification of the N-terminal product generated when necessary.

Site-directed mutagenesis

MCMJR1 D5A, D40A, D128A, D162A, D195A, D220A, D236A, D162N, D162E, D220N, D220E and G219A single-point mutant variants were constructed by the

Quickchange (Stratagene) site-directed mutagenesis method. This technique was also employed to mutate chimeric substrate MBPSpitz, MBPKeren and MBPGurken. Mutations were introduced at 6 positions within the TMD region of MBPSpitz involving residues M149, C150, V153, L156, F157 and V158. Based on the high homology between MBPSpitz and MBPKeren, these amino acid residues were each substituted by the corresponding residue present in the sequence of MBPKeren as follows: M149T, C150L, V153L, L156A, F157M and V158C. The designed MBPSpitz mutant variants then included six single-point mutants bearing the substitutions just mentioned and five multiple-point mutants, two carrying the triple mutation M149T/C150L/V153L and L156A/F157M/V158C and three carrying the double substitutions M149T/C150L, M149T/V153L and C150L/V153L.

Moreover, nine single-point mutations to alanine involving residues F252, P253, V254, L255, L256, M257, L258, S259 and S260 were introduced within a stretch of residues upstream and downstream of the identified cleavage sites in MBPGurken.

Cleavage site identification

The identification of MCMJR1 cleavage site involved the incubation of larger amounts of MBP chimeric substrate for a longer period of time compared to the *in vitro* cleavage assay. In a typical assay, we mixed the chimeric substrate (10 uM) with DDM-purified MCMJR1 (2 uM) and the total mixture taken to a final volume of 1 mL with PBS buffer pH 7.4, 0.1% DDM, in the presence of inhibitors. After 24 hours, 500ul of pre-

equilibrated Ni-NTA resin was added to the reaction mixture to capture the undigested chimeric substrate and the C-terminal product. For mass spectrometry analysis, the flow through containing the N-terminal product was concentrated to 200 ul to which 1ml of 10% TCA in acetone was added and kept at -20°C overnight. The precipitated protein was collected after centrifugation at 17 000 x g for 5 minutes and the pellet washed in 1ml of acetone at -20°C. Protein reconstitution was achieved by adding 500 ul of 8 M urea and the suspension passed through a 30K centrifugal filter device (Millipore) while exchanging the urea for ammonium bicarbonate buffer pH 7.4. The protein sample was then concentrated to a final volume of ~40ul. Thrombin (0.1 U) cleavage was carried out on 20ul of the concentrated protein sample to generate a small peptide between the thrombin site and the MCMJR1 cleavage site. The N-terminal product (control sample) and the N-terminal product treated with thrombin (thrombin-treated, TT sample) were analyzed in a MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) mass spectrometer (Voyager DE-STR, Applied Biosystems) operating in linear delayed extraction mode. Each mass spectrum was averaged from 500 measurements and calibrated with myoglobin as an internal calibrant. The spectra were smoothed and further analyzed using the software M-over-Z.

Inhibition treatment

DDM-purified MCMJR1 (~0.8 uM) was pre-incubated overnight at 37°C in the presence of γ-secretase inhibitors in addition to the inhibitors present in our 1x protease inhibitor cocktail for bacteria (E-64 at a final concentration of 10 ug/ml, Bestatin at 14 ug/ml and

PMSF at 100 uM). The reversible inhibitor of aspartic proteases (pepstatin A; Calbiochem), dipeptide known to inhibit γ-secretase activity (DAPT; Calbiochem), transition-state analog inhibitor of γ-secretase (31C; Calbiochem), dipeptide isostere (L-685,458; Calbiochem) and the peptidomimetic nontransition-state analog inhibitor of γ-secretase (Compound E; Calbiochem) used for this assay were dissolved in DMSO. The volume of pepstatin A (235 uM), DAPT (235 uM), 31C (100 uM), L-685,458 (20 uM) and Compound E (235 uM) incubated with MCMJR1 accounted for 2% of DMSO in the final volume. MBPGurken was added the next morning also to a final concentration of ~0.8 uM and samples were taken after 2, 3, 4, 6, 8 and 24 hours at 37°C. Similarly, increasing concentrations of 31C (10, 25, 50, 75 and 100 uM) and L-685,458 (Sigma; 10, 25, 50, 75, 100 and 200 uM) were pre-incubated with MCMJR1 before the addition of MBPGurken. For each incubation mixture, the inhibition level of the formation of the product of digestion was assayed by western blot.

Results

Four putative archaeal di-aspartyl protease-like genes expressed in *E. coli*.

Expression yields of recombinant proteins bearing either an N- or a C- terminal histidine-tag were tested in a 100 ml culture and purified by metal affinity chromatography in one of two detergents, DDM or FC-12. Deca histidine tagged MHJF1 (Fig. 26A, lane 6), MCMJR1 (Fig. 26A, lane 8), MCV (Fig. 26B, lane 3) and MMS2 (Fig. 26B, lane 4), were able to express as shown by the presence of bands migrating

with an apparent molecular weight that corresponded well with the predicted mass. One clear conclusion from these small-scale experiments is the preference for an N-terminal affinity tag and for the detergent FC-12.

The linker between the N-terminal deca-histidine tag and the gene of interest possesses a tobacco etch virus (TEV; [179]) protease site. Removal of the tag with TEV protease was although unsuccessful for MHJF1 and MCMJR1 choosing to re-engineer these constructs to express as N-terminal fusions to SUMO tag following a hexa histidine-tag. Small-scale experiments yielded comparable results to those shown in figure 26A.

Detergent-purified MCMJR1 displays proteolytic activity

In vitro cell-free assays with detergent-solubilized membrane fractions or purified protease constitute reliable tools to monitor protease activity. In a very elegant approach, cleavage by γ-secretase was followed by the generation of Aβ derivatives accurately measured by ECL derivatized anti-Aβantibodies [141]. In a simpler strategy, biochemical demonstration of activity by SPP [180] and TFPP [64, 181, 182] was acquired by incubating purified protease with chimeric proteins based on identified substrates. The processed form of these chimeric substrates was then detected in immunoblots as a faster migrating cross-reactive species.

Here we used a comparable approach to test whether our detergent-purified archaeal proteins possess protease activity. *In vitro* assays using our chosen TFPP-like proteins,

MCV and MMS2, have previously been tested resulting in the processing of the preflagellin FlaB2 into mature flagellin [181, 183]. Contrary, *in vivo* substrates if any, for our chosen PS-like proteins, MCMJR1 and MHJF are unknown. We then tested all four targets in an *in vitro* intramembrane protease activity assay that uses substrates designed to trigger the activity of Rhomboids, the only serine intramembrane proteases known to date. Briefly, in our laboratory we designed chimeric substrates in which the EGFR ligand ectodomain of the Drosophila Rhomboid substrates Spitz, Keren and Gurken, was replaced by MBP (42 kDa), generating chimeric substrates MBPSpitz, MBPKeren and MBPGurken, respectively.

Cleavage of the chimeric substrate by a protease generates a short C-terminal product and a N-terminal product that corresponds to the MBP moiety of the chimeric substrate and that can be distinctively detected from the larger undigested chimera substrate. In more detail, full-length MBPSpitz, MBPKeren and MBPGurken migrate as a single band at ~50kDa on a 10% SDS PAGE. Provided that cleavage of the detergent-purified chimeric substrate occurs, a C-terminal product is then released and a product of ~46 kDa can be detected. After the incubation of the chimeric substrates, each with either DDM- and FC-12-purified 10xHisMCV, 10xHisMMS2, MHJF1 and MCMJR1 (Fig. 27), we monitored the generation of the ~46 kDa-product by loading an aliquot of the incubation mixtures onto a 10% SDS PAGE followed by immunoblot analysis.

A band at ~46 kDa was detected only after the incubation of the chimeric substrates with DDM-purified MCMJR1 (Fig. 28). MHJF1, 10xHisMCV and 10xHisMMS2 purified in

either DDM or FC-12 as well as FC-12-purified MCMJR1 failed to display evidences of protease activity under the conditions tested. The activity displayed by MCMJR1 in the presence of amino-peptidases (bestatin) and cysteine/serine proteases inhibitors (E-64/PMSF) added to the reaction mixture, generated a band representative of the MBP moiety of the chimeric substrate lacking the C-terminus. Interestingly, although Spitz and Keren display high sequence homology, the specificity displayed by MCMJR1 against MBPSpitz was very low compared to that displayed for MBPKeren and MBPGurken as shown by the signal strength of the band below that corresponding to the undigested chimeric substrate (50 kDa). Thus, for the following tests, incubation mixtures included DDM-purified MCMJR1 with either MBPKeren or MBPGurken.

MCMJR1 is a GXGD-type di-aspartyl protease

Sequence alignment suggested that MCMJR1 is an archaeal GXGD-type di-aspartyl protease (Fig. 29A). In addition, from the prediction of membrane-spanning regions and orientation (Fig. 29B), we were able to identify important similarities between MCMJR1 and human PS. MCMJR1 is predicted to span the membrane eight times, a topology that localizes the "catalytic loop", i.e. the hydrophilic region that links the TMDs containing the catalytic motifs, towards the cytoplasm. In addition, amino acid residues D162 and D220 are contained within the conserved motifs YD and GMGD respectively, located in adjacent TMDs (Fig. 29B). Taken together, these similarities led us to hypothesize the involvement of these residues in the catalytic event observed for MCMJR1 in our *in vitro* cleavage assay. To address the contribution of MCMJR1 to the generation of a

product of proteolysis derived from our chimeric substrates, all seven aspartic acid residues in the amino acid sequence of MCMJR1 were individually mutated to alanine to eliminate the side chain beyond the β carbon (Fig. 30) [184]. The yield and purity obtained of the aspartate-to-alanine MCMJR1 mutants after expression and purification were comparable to that of the wild-type protease. As described for the *in vitro* cleavage assay, wild-type MCMJR1 and mutant variants were each incubated with MBPKeren and MBPGurken and the activity assayed. Incubation of the chimeric substrates with D5A, D40A, D195A and D236A MCMJR1 mutants generated the N-terminal product of ~46 kDa indicative of proteolysis (Fig. 31A). The activity displayed by these mutants was comparable to that displayed by wild-type MCMJR1 as judged by the intensity of the band corresponding to the N-terminal product measured from a processed coomassie stained SDS-PAGE gel. Based on the latter, we observed that D128A MCMJR1 mutant displayed an intermediate level of protease activity. Finally, a complete absence of a band corresponding to the N-terminal product was noticed when either, MBPKeren and MBPGurken substrates were incubated with D162A and D220A MCMJR1 mutants (Fig. 31A). Anti-MBP immunoblot analysis confirmed the absence of a N-terminal product after the incubation of the chimeric substrates with D162A and D220A MCMJR1 mutants as shown by the lack of cross reactivity (Fig. 31B). Taken together, our findings indicate that residues D162 within the YD motif and D220 within the GMGD motif represent candidate active site aspartic acids in MCMJR1 and that the activity displayed in our *in vitro* cleavage assay was MCMJR1 specific and not due to some residual *E. coli* protease activity carried over during purification.

To further confirm the contribution of D162 and D220 in peptide bond cleavage, both residues were mutated to asparagine. Since asparagine and aspartic acid are essentially similar in structure (differing on the side chain's functional group), this replacement would likely maintain the wild-type conformation of the protease. After repeating our *in vitro* cleavage assay, we observed that aspartate-to-asparagine MCMJR1 mutants completely lacked activity against MBPKeren and MBPGurken, as expected when an active site residue is mutated to a nonfunctional substitute (Fig. 32). Most importantly, this suggests that alterations in the active site conformation due to the exchange of D162 and D220 for an alanine residue were not responsible for the loss of protease activity of D162A and D220A MCMJR1 mutants. In addition, the direct participation in catalysis of the carboxylic acid functional group in D162 and D220 was tested *in vitro* through aspartate-to-glutamic acid MCMJR1 mutants. We nevertheless observed that the D162E and D220E MCMJR1 mutants were unable to restore protease activity under the conditions tested (Fig. 32) suggesting that aspartic acid residues at these positions are critical.

MCMJR1 is an ICLi-P

Cleavage of MBPKeren and MBPGurken by MCMJR1 produces N- and C-terminal products (Fig. 33A). The C-terminal product contains a significant part of the TMD sequence and thus is bound to a significant amount of detergent. In general, mass spectrometry approaches for characterization of detergent solubilized membrane proteins involve extraction into organic solvents to strip out bound detergent and enable data

collection of unfolded proteins [185]. Since complete detergent removal seldom is achieved [186, 187], accurate membrane protein mass determination by MS remains a challenge [188, 189]. We therefore focused our attention on the N-terminal product, which is essentially MBP followed by a hydrophilic C-terminal linker segment. To determine the substrate cleavage site we then isolated this N-terminal product from the incubation mixture and carried out a MALDI-TOF MS analysis (Fig. 33, A-B).

MS analysis revealed three major intramembrane cleavage sites within MBPKeren. Observed masses for the N-terminal product (control sample) derived from MBPKeren after cleavage by MCMJR1 were: 46 122, 46 338 and 46 450 daltons (corresponding masses as calculated by peptidemass analysis tool on the ExPASy server: 46 106, 46 332 and 46 479 daltons, respectively). These masses correspond to cleavage sites between amino acid residues Ala-131/Leu-132, Leu-133/Phe-134 and Phe-134/Met-135 (numbering according to the full-length sequence of wild-type Keren), after the tenth, twelfth and thirteenth amino acid residue respectively of the predicted TMD region of Keren. However, the determination of a precise mass of the analyzed species is restricted when their molecular weight is high, yielding an error of ±40 Da between the observed and calculated masses. To circumvent this problem, an aliquot of the N-terminal product derived from the digestion of MBPKeren was incubated with thrombin. As explained in the materials and methods section, the MBP chimeric substrates consist of a N-terminal MBP tag separated from the TMD of the *D. melanogaster* EGFR ligand by a thrombin site. Then, the incubation with thrombin of the N-terminal product (thrombin-treated; TT sample) generated after MBPKeren processing produces a small peptide between the thrombin site and the MCMJR1 cleavage site. The mass corresponding to this peptide

can therefore be calculated more accurately, compared to the mass collected for the N-terminal product, yielding an error of only ±5. Peptides with masses of 3615.8, 3842.2 and 3989.5 daltons were observed. The calculated masses for the cleavage positions specified above are 3615, 3841 and 3989 respectively confirming in this way that MCMJR1 cleaves MBPKeren close to the middle of the TM domain between Ala-131/Leu-132, Leu-133/Phe-134 and Phe-134/Met-135 (Table III).

Similarly, for the purified N-terminal product derived from MBPGurken after cleavage by MCMJR1, MALDI-TOF MS analysis revealed two intramembrane cleavage sites. From the control sample, observed masses were: 45 940 and 46 011 daltons (calculated masses: 45 901 and 46 014 daltons respectively) (Fig. 34). These masses corresponded to cleavage sites between amino acid residues Leu-255/Leu-256 and Leu-256/Met-257. Analysis of the TT sample revealed instead three intramembrane cleavage sites with observed masses of 3411.1, 3524.7 and 3656.0 daltons. The calculated masses for the first two cleavage positions are 3410 and 3523 daltons, respectively, confirming in this way that MCMJR1 indeed cleaves MBPGurken within the predicted TMD region between the amino acid residues specified above. Due to the higher accuracy expected from the analysis of the TT sample compared to the control sample, it is likely that MCMJR1 is also able to recognize three positions within the TMD region of Gurken being the third one between amino acid residues Met-257/L-258 (Table III).

MCMJR1 cleaves chimeric γ-secretase substrates

Given the similarities between MCMJR1 and presenilin we questioned whether MCMJR1 could cleave *bona fide* presenilin substrates, unique to this enzyme. We then tested MCMJR1 activity on C100Flag (Fig. 35A), an APP-based substrate widely used to assess the activity of γ-secretase preparations based on their capacity to generate Aβ40 and Aβ42. After incubation of MCMJR1 with C100Flag a very faint band was detected around 8 kDa (Fig. 35, B-C). Detection of that band after anti-flag immunoblot analysis suggested that it may correspond to a C-terminal product originated after cleavage by MCMJR1 at a position close to the middle of the predicted TMD. The corresponding N-terminal product derived from this proteolytic reaction was not detectable after PAGE since it is estimated to be of very small mass, around 4 kDa (Fig. 35B).

To map the cleavage profile by mass spectrometry, we incubated chimeric MBPAPP substrate with DDM-purified MCMJR1 and observed the generation of a band at ~44 kDa representative of the MBP moiety of MBPAPP lacking part of the C-terminal portion (Fig. 36A). From MS analysis, two major intramembrane cleavage sites within MBPAPP were revealed. Observed masses obtained after analysis of the control sample were 43 674 and 43 860 daltons (calculated masses: 43 680 and 43 879 daltons respectively) (Fig. 36B). These masses corresponded to cleavage sites between amino acid residues Gly-709/Val-710 and Val-711/Ile-712 (numbering according to the full-length sequence of APP), after the tenth and twelfth amino acid residue respectively of the predicted TMD of APP (Fig. 36C). Remarkably, these sites corresponded to γ-

secretase cleavage sites that generate Aβ38 and Aβ40. These results suggest that MCMJR1 displays the rather unique property of being able to cleave substrates derived from APP.

A transition-state analog inhibitor of γ-secretase inhibits MCMJR1 activity

Considering that a number of inhibitors affect the activity of both members of the aspartyl I-CLiP family [73], we tested whether γ-secretase inhibitors alter the activity of MCMJR1 (Fig. 37). MBPGurken was chosen for this assay given the larger amount of the N-terminal product generated compared to that from MBPKeren, MBPAPP or C100Flag after incubation with MCMJR1. The formation of the N-terminal product after digestion of MBPGurken can be detected as early as 3 hours after the beginning of the reaction as showed by bands that weakly cross-reacted at the corresponding molecular weight. Samples taken after 8 and 24 hours of incubation thus showed more consistent results (Fig. 38, A-B). We observed that pre-incubation of MCMJR1 with 31C (100 uM final concentration) affected the generation of the N-terminal product as shown by the lack of cross-reaction at the corresponding molecular weight after 8 hours (Fig. 38A). After 24 hours, the effect of 31C on MCMJR1 was weakly reversed and a faint band corresponding to the N-terminal product started to show (Fig. 38B). To further assess whether the lack of cleavage product is specific to an inhibitory effect caused by 31C, we pre-incubated MCMJR1 with increasing concentrations of this transition-state analog. Pre-incubation of MCMJR1 with 31C at a final concentration of 10 uM did not affect the generation of the N-terminal product. At higher concentrations of 31C used,

cleavage of MBPGurken by MCMJR1 was gradually reduced confirming in this way that the diminished activity observed was due to an inhibitory effect specific of MCMJR1 by 31C (Fig. 39A).

On the other hand, the cleavage capacity of MCMJR1 was not affected by the pre-incubation with L-685,458, a γ-secretase inhibitor structurally similar to 31C and of irreversible nature. Since the commercially available stock solution of L-685,458 used in our assay only allowed the incubation of this peptidomimetic inhibitor at a final concentration as high as 20 uM, five times less than that used of 31C, we then assessed whether the lower concentration used was in correlation with the lack of effectiveness in inhibiting MCMJR1 shown in figure 38. From a lyophilized form of L-685,458 we were able to prepare a stock solution of higher concentration. After 24 hours of incubation of pre-inhibited MCMJR1 with MBPGurken we observed that in the presence of a final concentration of 200 uM of L-685,458, MCMJR1 was still capable of generating an N-terminal product (Fig. 39B).

FAD-associated mutation slows intramembrane proteolysis by G219A MCMJR1 mutant variant

Immediately adjacent to the critical aspartate contained in the GLGD motif in PS, a very early-onset (mean onset age 35 years) AD-associated G384A mutation has been identified [190]. This mutation produces substantial amounts of the pathological Aβ42 implying a selective slowed production of Aβ40 in stably transfected HEK293 cells over

expressing APP Swedish mutant [63]. Moreover, mutation to alanine of the most C-terminal glycine residue in the GXGD motif of SPPL2b slowed intramembrane proteolysis of TNFα causing an increase of longer intracellular products, most likely as a result of slowed sequential intramembrane cleavage [191]. To prove whether such amino acid exchange slows the processing of MBPGurken by MCMJR1 in a time based manner, the corresponding conserved glycine residue G219 was mutated to alanine to generate MCMJR1 G219A. As for PS and SPPL2b, the generation of the N-terminal product derived from the proteolysis of MBPGurken by MCMJR1 G219A was indeed slowed compared to that carried out by wild type MCMJR1 (Fig. 40).

Insights into MCMJR1 substrate specificity

From the initial experiments carried out to test the capacity of MCMJR1 to cleave MBP chimeric substrates, MCMJR1 clearly showed to be more specific towards MBPKeren compared to MBPSpitz (Fig. 28). The overall homology displayed between wild-type Keren and Spitz is very high and we believe that by targeting the amino acids that do not contribute to that homology, insights into the substrate specificity of MCMJR1 may be provided. We then performed single, double and triple mutations that involved six amino acid residues within the TMD region of the chimeric MBPSpitz substrate. These amino acids were chosen based on their proximity to the cleavage site positions identified in MBPKeren and extrapolating those positions into the amino acid sequence of MBPSpitz. Each individual residue was exchanged to the corresponding one in the amino acid

sequence of Keren, in this way, making MBPSpitz more homologous to MBPKeren (Fig. 41A).

In vitro activity assays including M149T, C150L, V153L, L156A, F157M and V158C MBPSpitz single-point mutant variants did not contribute to an improvement in the cleavage capacity of MCMJR1 compared to that exhibit towards wild-type MBPSpitz. Instead, a more efficient specificity towards the MBP chimeric substrate seems to occur when the enzyme was incubated with M149T/C150L/V153L MBPSpitz triple mutant variant, albeit not as efficient as when incubated with MBPKeren (Fig. 41B). For a clearer insight into the substrate specificity of MCMJR1 we then focused our attention on this stretch of amino acids. When incubated with M149T/C150L and C150L/V153L MBPSpitz double mutants, MCMJR1 displayed a comparable activity as that observed towards the triple mutant. Contrary, a band corresponding to the N-terminal product was not generated upon incubation of M149T/V153L MBPSpitz double mutant variant with MCMJR1 (Fig. 41C).

From a different approach, to assess whether the substrate specificity and maybe peptide bond selectivity of MCMJR1 are modulated by substitutions for smaller amino acid residues, we exchanged each amino acid residue around the cleavage sites identified within the TMD region of MBPGurken (Fig. 42A). Since MCMJR1 seems to display a higher selectivity for MBPGurken compared to MBPKeren, MBPGurken was chosen for this assay and amino acid residues from positions 252 to 260 (numbering according to the full-length sequence of wild-type Gurken) were substituted for alanine residues each

at a time. After incubation of MCMJR1 with wild-type MBPGurken and with each of the nine chimeric substrates bearing the single-point mutation, the intensity of the band corresponding to the N-terminal product was measured from a processed coomassie stained SDS-PAGE using image J. A significant difference in the intensity of the band corresponding to the product of proteolysis generated from P253A MBPGurken mutant variant was observed (Fig. 42B). Specifically, the capability of MCMJR1 to cleave MBPGurken seems to have largely decreased as a consequence of the substitution of Pro253 for alanine as compared to the intensity measured for the rest of the N-terminal products generated by the other mutant variants.

The insufficient amount of N-terminal product generated from P253A MBPGurken mutant variant also affected its isolation for further analysis by MS. For all other N-terminal products, thrombin treatment and MS analysis were carried out as described before. We found that MS peaks corresponding to the cleavage between amino acid residues Met-257/Leu-258 were absent in both MS spectra corresponding to the N-terminal and TT samples, indicating that the substitution of Met257 for an alanine residue abrogated cleavage at this position (Fig. 42C).

Discussion

Presenilin (PS), the catalytic subunit of the γ-secretase complex, and Signal Peptide Peptidase (SPP) belong to the di-aspartyl intramembrane-cleaving protease (I-CLiPs) family. A distinctive feature of this family is the location of the catalytic aspartic acid

residues within the conserved motifs YD and GXGD each located in two adjacent TMD regions, an unexpected location for a proteolytic reaction that requires water. Due to its role in biology and disease, considerable efforts have been especially devoted into understanding the unusual mechanism adopted by PS to process membrane bound substrates within the lipid bilayer. From the available negatively stained EM and cryo-EM structures, hypotheses can now be made in regards to the mechanism of hydrolysis of PS/γ-secretase, although, compared to other I-CLiPs, much remains unknown. For metallo and serine I-CLiPs, the identification of homologs of eukaryotic Rhomboids and S2P intramembrane proteases in bacteria and archaea [34, 192, 193] somehow aided in overcoming the challenges met during purification of membrane proteins. The expression and purification in sufficient quantities of the *E. coli* homolog of Rhomboid and archaeal homolog of S2P from *E. coli* resulted in successful structural studies from which a legitimate description of the mechanism of hydrolysis by these I-CLiPs was proposed. PS, as the active component of the γ-secretase complex not only requires the presence of three other integral membrane proteins to become active but it also has to go through endoproteolytic activation into a CTF/NTF heterodimer. The higher degree of complexity of PS/γ-secretase is indisputable compared to other intramembrane proteases, thus affecting its expression in high yields and purification as an active, highly homogeneous enzymatic preparation.

Sequence database searches suggest that PS-like proteins are present throughout the eukaryote species, in yeast, plants, molluscs, insects, fish, birds and mammals [105, 194]. Based on sequence homology to TMD regions of PS, PS-like proteins have also

been found in four species of archaea [105, 194]. Evidence of cleavage of the γ-secretase substrate Notch exists only for two of the genes in *C. elegans*, sel-12 and hop-1 [195, 196]. To our knowledge, a plausible di-aspartyl protease activity has not yet been proposed for any other PS-like protein. Based on sequence alignment, here we identified putative membrane bound aspartyl proteases from archaeal origin and purified them in detergent micelles from *E.coli* membranes. We tested protease activity and found that the candidate from *Methanoculleus marisnigri* strain JR1 (MCMJR1) was able to carry out proteolysis *in vitro* of our chimeric transmembrane substrates. In MCMJR1, we found that the two aspartic acid residues D162 and D220 within the motifs YD and GMGD respectively, were responsible for the activity of the protease. Mutation of the catalytic aspartic acid residues to alanine, a replacement expected to maintain the main-chain conformation [197], and to the nonfunctional asparagine residue confirmed the catalytic role of these aspartates as observed by abrogation of proteolysis. On the other hand, a protease activity comparable to that displayed by wild-type di-aspartyl protease MCMJR1 was not recovered when D162 and D220 were mutated to glutamic acid residues. The latter do not necessarily rule out the participation of the carboxylic acid functioning in acid-base catalysis, instead, we believe that the longer side chain of the glutamic acid residue may be drastically affecting the local conformation of a likely small active site pocket confirming in this way the necessity for aspartic acid residues at these positions. Additionally, our MS data revealed that processing by MCMJR1 of our chimeric membrane protein substrates based on ligands of EGFR occurred at several positions, mainly, close to the middle of the predicted TMD region. For PS, a similar cleavage pattern carries considerable interest since the pathology of AD is determined by

the cleavage of APP at a preferential position to generate the longer pathological form of Aβ of 42 amino acid residues in length over the benign form of only two amino acid residues shorter (Aβ40).

Additional features of MCMJR1 such as specificity towards APP based substrates, effect on substrate processing after inhibitor treatment and the impact of the introduction of a FAD-associated mutation in proteolysis proved to be highly relevant to establishing similarities to known di-aspartyl intramembrane proteases, especially PS. The approximate molecular weight of the product released from C100Flag suggests that MCMJR1 cleaved this substrate close to the middle of the TMD corresponding to APP. Confirmation of the latter can be obtained in the future by immunprecipitating the products of proteolysis derived from C100Flag with monoclonal anti-Aβ antibodies followed by MALDI-TOF mass spectrometric analysis [159] or using ECL derivatized anti-Aβ antibodies as described by Li *et al.* [141, 159]. Alternatively, we incubated chimeric MBPAPP with MCMJR1 and found that cleavage occurred close to the middle of the TMD region at two positions that are also recognized by γ-secretase (Fig. 43). Cleavage between amino acid residues Gly-709/Val-710 and Val-711/Ile-712 by PS generates Aβ38 (γ38 site) and Aβ40 (γ40 site), both forms present in a ~1:22 ratio as measured from conditioned medium of HEK293 cells stably co-expressing APP Swedish mutant together with wild-type PS [178]. Notably, from our *in vitro* assay and MS analysis, we observed a preference by MCMJR1 to also cleave at these positions.

We also found that 31C (IC50 ~10 nM in cell-free assay), a compound structurally similar to the hydroxyethylene transition-state analogue L-685,485, inhibited the production of the N-terminal product derived from MBPGurken in our *in vitro* assay at a concentration of 100 uM. Surprisingly, the more potent L-685,485 γ-secretase inhibitor (IC50 ~0.3 nM in cell-free assay) did not show inhibitory effects at concentrations of up to 200 uM. In soluble aspartyl proteases, hydroxyethylene-based inhibitors adopt an extended conformation in the active site cleft binding symmetrically between the catalytic carboxyl groups [15] and in HIV-1 for example, the catalytic aspartates interact with the isostere hydroxy group of these inhibitors [42, 198]. We speculate that an additional interaction between the catalytic aspartates D162 and D220 with the isostere amine of the hydroxyethylamine-based inhibitor 31C provided a higher efficiency compared to L-685,485 to block cleavage by MCMJR1. Such an interaction has been observed between the catalytic aspartates in HIV-1 with the isostere amine of hydroxyethylamine-based inhibitor SQ which belongs to a novel peptidomimetic series resistant to drug induced mutations [199]. Moreover, the structure of L-685,485 bears an additional phenyl group at P3' compared to 31C (Fig. 44). In HIV-1, subsites consisting of 6-8 residues accommodate the phenyl groups of inhibitors [199] allowing us to hypothesize that an additional subsite is not available in MCMJR1 to optimally interact with L-685,485 (Fig. 44).

Finally, we observed that the substitution to alanine of the glycine residue immediately adjacent to the catalytic aspartate contained in the motif $GMGD_{220}$ in MCMJR1 slowed the generation of the N-terminal product. We focused on this substitution since it is

highly conserved in aspartyl membrane proteases of the GXGD type and it has been found that mutagenesis of this glycine residue affects the activity of PS and SPP. For SPPL2b containing the G420A substitution, a slowed intramembrane proteolysis of TNFα was observed [191]. In PS, the substitution G384A is actually a naturally occurring FAD-associated mutation that causes an exceptional increase of the toxic form of Aβ of 42 amino acids in length [63, 200], an increase higher than any other FAD-associated mutation [190]. With partially purified γ-secretase derived from cells expressing the PS G384A mutation, it was also found that Aβ40 generation was slowed compared to wild type γ-secretase with the amount of Aβ42 being considerably higher than Aβ40 [191]. The model of sequential proteolysis of APP proposed for PS by Qi-Takahara and colleagues [92] is supported by these findings and allow conjectures about a similar mechanism for SPPL2b. The data that we have collected on the intramembrane cleavage performed by MCMJR1 is still preliminary in terms of suggesting a model of processing that fits that described for PS and SPPL2b. Whether along with the slowed production of N-terminal product derived from MBPGurken, an accumulation of a longer cleavage product (plausible intermediates during sequential intramembrane proteolysis by MCMJR1) occurs in parallel, will be subject of a thorough time-based analysis involving C-end-specific antibodies along with appropriate electrophoretic systems.

While assessing the stability of MCMJR1 for crystallization and structural determination, an additional similarity with PS was found. According to MCMJR1 TM topology prediction, TMDs 5 and 6, each harboring the conserved catalytic motifs YD_{162}

and $GMGD_{220}$ respectively, are connected by a loop that faces the cytoplasmic side (approximately Ala163 – Ala213). In PS, the large "catalytic loop" located between TMDs 6 and 7, also faces the cytoplasm and regulates γ-secretase processing as observed by the reduced cleavage at the γ40 site when a large portion was deleted [201]. From a parallel research oriented towards obtaining crystallographic data on MCMJR1, speculations were made in regards to an increase in stability of MCMJR1 bearing truncations at the level of the cytoplasmic loop, an approach used in successful structure determination studies [202]. Then, when monitoring the activity of these mutant variants we found that substrate processing by MCMJR1 was severely affected. In specific, deletions larger than ~15 residues or closer to the catalytic aspartic residues completely abrogated the generation of N-terminal product derived from intramembrane cleavage of MBPGurken allowing us to hypothesized on an important role in activity of the cytoplasmic loop of MCMJR1 (Fig. 45).

An additional requirement for intramembrane cleavage by the aspartic I-CLiPs PS and SPP has been described. For γ-secretase catalysis to occur, removal of the luminal domain from full-length type I membrane proteins is required for subsequent intramembrane proteolysis mediated by PS [203]. In a similar way, a cleavage performed by a signal peptidase at the exoplasmic side of the ER membrane liberates the signal peptide from the pre-protein generating a substrate that is more efficiently processed by SPP [61]. In our *in vitro* cleavage assay, we observed that MCMJR1 was capable of cleaving chimeric membrane substrates in which the ectodomain corresponding to the wild-type type I membrane protein was replaced by MBP. A prior

processing to remove the large domain corresponding to MBP was not necessary for MCMJR1 intramembrane proteolysis proving an additional advantage provided by our assay.

Probably the most striking feature described for MCMJR1 is its capability to cleave chimeric membrane substrates in the absence of accessory proteins. In this regard, the activity displayed by MCMJR1 is more closely related to SPP since substantial evidence has now demonstrated that PS alone is incapable of accomplishing protease activity. PS is part of a large molecular weight complex and the presence of three accessory proteins, PEN-2, NCT and APH is necessary to recover γ-secretase activity [175]. Although additional studies are still necessary to identify potential binding partners for MCMJR1, and with that possibly boost its activity, here we have shown that alone is able to carry out detectable cleavage events.

Taking into account multiple lines of evidence pointing towards MCMJR1 being an archaeal GXGD type di-aspartyl intramembrane protease, we focused on examining MCMJR1 requirements for substrate recognition and selection of the scissile bond. We have shown that MCMJR1 displays activity against MBPKeren and MBPGurken, but surprisingly, almost undetectable levels of proteolysis were displayed towards MBPSpitz. Since MBPKeren and MBPSpitz are almost identical, it is plausible that the observed difference in substrate specificity by MCMJR1 is a result of the differences in amino acid residues between them. To acquire initial insights into the substrate specificity of MCMJR1, we then introduced mutations within the TMD region of

MBPSpitz. Based on the sequence alignment of these two very homologous substrates, we identified the non-matching amino acid residues in MBPSpitz and substituted them for the corresponding ones in the sequence of MBPKeren. We were able to observe an increased activity against MBPSpitz variants bearing amino acid substitutions in the immediate vicinity of the corresponding cleavage sites identified in MBPKeren. From among the M149T, C150L and V153L substitutions, the first one is likely to be more relevant compared to the others. The replacement of Met149 for the β-branched threonine, a residue with low helix propensity, is likely to contribute to a local destabilization leading to the exposure of the scissile bond. Although Val153 is also a β-branched amino acid, a helix destabilizing effect derived from its replacement for leucine is not expected since both amino acid residues are equally strong helix stabilizers in systems that mimic the structural features of biological membranes [204]. In the future, improving the product recovery during sample preparation for MS analysis will reveal whether these residues are also in close proximity to cleavage positions in MBPSpitz mutant variants.

Further insights into the requirements of MCMJR1 for intramembrane proteolysis were obtained from an orientative alanine scanning mutagenesis-based assay on MBPGurken, for which MCMJR1 displays the highest specificity in our *in vitro* assay. We found that the replacement of Pro253 for an alanine residue considerably reduced the specificity of MCMJR1 towards the MBPGurken single-point mutant variant. The destabilizing effects of proline residues in α-helical conformations in water-soluble globular proteins are well known [205-208]. In TM α-helices of integral membrane proteins, the presence of

proline residues induces dynamic flexibility (hinges) [209] and / or helix distortion by introducing kinks [210], both with apparent roles in signal transduction [211], in the gating mechanism of ion channels [209] and membrane protein folding [212]. In MBPGurken we believe that the presence of Pro253 provides this substrate with a local distortion or flexibility, a few residues away as suggested by Cordes *et al.* [26] coincidentally, at the positions identified as cleavage sites for MCMJR1. We then speculate that in P253A MBPGurken mutant variant, peptide bonds at these positions failed to be exposed to MCMJR1.

From the MS analyses carried out on the cleavage products derived from the rest of the MBPGurken single-point mutant variants we observed that the substitution of Met257 for alanine modulated the number of cleavages performed by MCMJR1. Hydrolysis of the peptide bond when Met257 is located at the P1 position was abolished in the M257A MBPGurken mutant. In the future, from a systematic approach involving more drastic mutations (i.e. bulky, polar or helix stabilizing residues) we expect to achieve conclusive remarks into MCMJR1 requirements.

Combining these results we demonstrate that MCMJR1 is indeed a GXGD type di-aspartyl intramembrane protease with significant biochemical similarities to PS. We believe that MCMJR1 may likely be a less complex, probably more tractable, prokaryotic system for functional and structural studies on di-aspartyl intramembrane proteases. In the future, we expect to make structure-based functional hypothesis from MCMJR1 that could finally be extended to eukaryotic PS.

Future work

1) High-resolution structural studies in γ-secretase. Given its complexity, structure determination of intact γ-secretase is a challenging task. However, it is important that we built on our ability to produce milligram amounts of the enzyme. To this end, we plan to optimize the expression, purification and detergent-stability of recombinant γ-secretase from *P. pastoris* to yield functional material of sufficient amounts and purity for future crystallization experiments.

2) Structure of MCMJR1 by 2D or 3D crystallization. Prokaryotic membrane proteins are more tractable than their eukaryotic counterparts and their structures tend to be the first to be solved, as the elegant work on channels and transporters so clearly demonstrates [213]. Prioritizing the structure determination of an archaeal GXGD enzyme would, as of now, provide the first snapshot at atomic detail of a GXGD di-aspartyl intramembrane protease and would allow us to make structure-based functional hypotheses bound to extend their relevance to eukaryotic PS.

3) Amino acid substitutions in MCMJR1. We will investigate the role of amino acid residues that are conserved among GXGD di-aspartyl intramembrane proteases. For example, we will carry out systematic mutagenesis of G219 in MCMJR1 to test the idea that it allows rotational freedom of the adjacent catalytic aspartate. We will also systematically mutate G217 and the position X within the GXGD motif. In addition, we will study the effect that other familial Alzheimer's disease-associated PS mutations

[214], within and adjacent to the YD and PAL motifs (i.e. Y256S, V261L, V261F), have on MCMJR1 activity.

As an alternative strategy, we will carry out deletions and we will substitute substantial regions of the TMDs of PS and SPP into analogous regions in MCMJR1 to test their effect on activity. These MCMJR1 chimeras will allow us to cover ground rapidly and zoom in into sequence regions that are of importance to GXGD di-aspartyl intramembrane protease function.

4) MCMJR1 pharmacology. SPP and γ-secretase appear pharmacologically similar in that they are targeted by many of the same small molecules, including transition state analogs, non-transition state inhibitors, and Aβ modulators [73]. Moreover, they both apparently have a substrate-binding site that is distinct from the active site and certain nonsteroidal anti-inflammatory drugs known to shift the site of proteolysis by γ-secretase also affect SPP [180]. Differences in the pharmacology of both enzymes have also been recently reported [215]. Based on our preliminary data we propose to test the effect of γ-secretase and SPP inhibitors and modulators such as (Z-LL)2-ketone [61], BMS-299897 [216], IN973 [217] and BMS-433796 [218] for their effect on the activity and peptide bond selectivity of MCMJR1. Binding assays will also be performed to determine V_{max} and K_d, essentially as described previously [73]. These inhibition studies will also allow us to identify other compounds, like 31C, which can be used in our crystallization experiments. Additionally, nonsteroidal anti-inflammatory drugs known to modulate the site of proteolysis by γ-secretase [219-221] will be tested.

5) Cysteine-scanning mutagenesis in MCMJR1. Since the cleavage sites for substrates of γ-secretase are located within the hydrophobic region of the membrane, the question of how to include the water needed to carry out the proteolytic reaction has been posed. Sato *et al*. [154] and Tolia *et al.* [96] have recently employed substituted cysteine accessibility methods to show that TMDs 6 and 7 of presenilin contribute to the formation of a hydrophilic pore within the membrane. Furthermore, residues in the luminal portion of TMD6 have been predicted to form a sub-site for substrate binding that is exposed to a hydrophilic milieu, whereas those around the second GXGD catalytic motif within TMD7 are highly water accessible [154]. These biochemical data suggest that the active site in γ-secretase is water accessible. We will employ similar approaches to probe the environment and structure of the active site of MCMJR1. We will introduce cysteines in the putative active site region by mutagenesis, and probe their accessibility to both membrane permeable and impermeable reagents.

6) Substrate specificity of MCMJR1. We will carry out experiments aimed at identifying the substrate motifs that trigger MCMJR1 activity. We will initially focus on studying the cleavage determinants of the TMD of APP. First, we will introduce deletions in the TMD to show if the intact hydrophobic domain is required for efficient cleavage. Secondly, we will study the effect of helix breaking residues by systematically substituting positions along the TMD by helix stabilizing (I or L), neutral (A) and destabilizing (N or P) amino acid residues. Thirdly, the motifs identified with this strategy will be introduced into TMDs that are not efficiently processed substrates of MCMJR1, such as Spitz, to evaluate if we can convert them into substrates.

Chapter V

CONCLUDING REMARKS

Introduction

In recent years, the term regulated intramembrane proteolysis (RIP) has been used to describe an atypical hydrolytic event that takes place in the lipid bilayer. Through this unusual chemistry that we haltingly start to comprehend, RIP of membrane bound proteins releases active peptides from their inoperative state allowing them to travel within or outside the cell where they can fulfill their purpose. Indeed, RIP regulates cellular physiology during EGF signaling, activation of cholesterol biosynthesis, cell differentiation, migration and proliferation among several other events [34]

Responsible for RIP are a group of enzymes called intramembrane proteases. Intramembrane proteases are classified in metalloproteases, serine proteases and di-aspartyl proteases according to the mechanistic class described for their water-soluble counterparts. Typifying each group are site-2 protease (S2P) zinc metalloproteases, rhomboid serine proteases and the two related di-aspartyl proteases, presenilin (PS)/γ-secretase complex and signal peptide peptidase (SPP).

Sequence database searches suggest that di-aspartyl intramembrane proteases are well conserved throughout eukaryotes with putative homologues present in yeast, plants,

mollusks, insects, fish, birds and mammals [105]. The existence of presenilin-like proteins in archaea has been suggested [105], but never proven biochemically. For S2P and Rhomboids, homologues are instead found in all three domains of life [192, 222] and a functional characterization has been achieved for several (for reviews [223, 224]). This nearly ubiquitous presence has tremendously contributed to acquiring novel insights into these two large families of intramembrane proteases. Indeed, the availability of homologues of the well studied Drosophila Rhomboid-1 and mammalian S2P in lower organisms has resulted in structural and functional breakthroughs that for di-aspartyl intramembrane proteases we can just expect to achieve in a near future.

In this section, we will summarize the advances achieved in the field of intramembrane proteases, in particular the progress involving S2P and Rhomboid in addition to supplement the current knowledge on di-aspartyl proteases with our discovery of MCMJR1, an archaeal di-aspartyl intramembrane protease that recapitulates key biochemical properties of the elusive eukaryotic PS and SPP.

S2P and Rhomboids: functions and their ubiquitous presence

The role of S2P was first described in cholesterol homeostasis in animal cells as a response during sterol deprivation [225, 226]. In this scenario, proteolysis of the membrane-bound transcription factor, the sterol regulatory element binding protein (SREBP), releases the N-terminal segment that then enters the nucleus activating the transcription of genes encoding 3-hydroxyl-3-methylglutaryl CoA and the low-density

lipoprotein receptor [225]. SREBP is initially cleaved by site-1 protease within the short hydrophilic loop that joins the two transmembrane segments of the transcription factor followed by an intramembrane cleavage carried out by S2P. In the mid nineties, such an unusual and extraordinary cleavage event resembled only that described for the APP to generate Aβ peptides during AD pathogenesis [227]. During the past years, the roles of S2P have been extended to activating fatty acid biosynthesis in Drosophila [228] and during unfolded protein response in mammals [229]. Functional roles for bacterial S2P have also been characterized during sporulation in *B. subtilis* [230], periplasmic stress response in *E. coli* [231], control of cell envelope composition in *M. tuberculosis* [232] among others (reviewed in [224]).

The involvement of Rhomboids is also described in well-studied areas of cell biology. RIP by Rhomboids was first found to regulate the activation of the epidermal growth factor Spitz in Drosophila. During activation as a receptor ligand within the signal-sending cell, trafficking of Spitz from the ER to the Golgi apparatus is assisted by the transport factor Star. In this location, the ligand domain of Spitz is cleaved from its transmembrane anchor by Rhomboid in the first described activation of a growth factor by RIP [233]. Additional biological functions are identified for Rhomboid besides that in intercellular signaling in Drosophila and in the analogous event in *P. stuartii* named quorum sensing. Well-characterized functions include the cleavage of adhesins by rhomboid during parasite invasion in addition to regulate aspects of mitochondrial morphology in yeast, mammals and Drosophila.

As mentioned earlier, homologues of S2P and Rhomboid are present among all forms of life. PSI-BLAST searches identified 334 unique sequences of homologues to S2P when conserved sequence elements were used as seeds [222]. Also, a nearly ubiquitous presence of the Rhomboid family is observed in bacteria, archaea and eukaryotes [192]. Interestingly, for this highly conserved and widespread group of membrane proteins, evidence of Rhomboids being functionally conserved across kingdoms was acquired when the aarA rhomboid of *P. stuartii* was substituted by Drosophila rhomboid-1, and vice versa, and their roles recovered in each organism [234]. Whether the mechanism by which rhomboid-1 is able to activate the signal that mediates quorum sensing in *P. stuartii* or aarA is capable of generating active EGFR ligands in Drosophila is comparable between the substituted proteases, remains an open question.

Here we have introduced MCMJR1 as a novel GXGD-type di-aspartyl intramembrane protease from archaea and in an analogous way, shown that it can recapitulate key biochemical properties of eukaryotic presenilins and SPPs.

MCMJR1 is a functionally conserved archaeal di-aspartyl intramembrane protease

Sufficient evidence has demonstrated that PS constitutes the catalytic core of γ-secretase [86], a multi-protein complex composed of PS, NCT, APH-1 and PEN-2 [79]. As γ-secretase consists of four hydrophobic proteins totaling at least 19 TMDs its structural and functional characterization is particularly demanding [235]. Several are the challenges that are to be surmounted by researches when attempting high-resolution

structural studies in γ-secretase. These include the elaborated maturation pathway of fully assembled and glycosylated enzyme, the requirement for specific lipids for activity, the use of detergents that often lead to complex dissociation, the presence of isoforms of APH and PS in the mature complex and finally low yields of purified and highly homogenous enzyme. In our own study, we initially assayed the individual expression and purification from *E. coli* of each component of the human γ-secretase complex. Although promising results were observed for human Pen-2 and PSdelta9 in terms of their purification, challenging barriers made hard to obtain purified APH and NCT to finally pursue our initial goal of achieving an *in vitro* reconstitution of γ-secretase complex. In the absence of all 4 members, we then tested the *in vitro* activity of PSdelta9, which as discussed earlier, mimics an endoproteolyzed and biologically active PS heterodimer. We found that under the conditions tested, detergent purified PSdelta9 was incapable of displaying *in vitro* activity.

In recent years, the discovery of SPP resulted promising in the field of di-aspartyl intramembrane proteases offered an alternate route to study the evasive PS. Human SPP is a polytopic integral membrane protein with seven predicted TMDs [106]. Like presenilin, the two catalytic aspartates in SPP are found within the conserved YD and GXGD motifs located in adjacent TMDs [106]. In addition, SPP can be photolabeled by a γ-secretase transition state analog inhibitor, suggesting a possible conservation of active-site structure within the two enzymes [236], and even Aβ modulators affect SPP activity [180]. These findings have provided strong evidence that SPP and PS share significant biochemical properties. Different from PS though, SPP does not require neither

complexation with additional proteins for activity nor an endoproteolytic activation, thus promoting its use as a model system to study PS. Early studies employed Chinese hamster ovary (CHO), human embryonic kidney 293 (HEK), or human neuroglioma (H4) cells for expression of SPP from which the enzyme was partially purified [180, 237]. An important study on SPP involved the bacterial expression of the enzyme as we initially tested for the γ-secretase complex. Human and *Drosophila* SPP and that from *S. pombe* were expressed in *E. coli*, followed by their purification and reconstitution in proteoliposomes [238]. The enzyme was proteolytically active ratifying that SPP does not require any additional proteins as cofactors for proteolysis as confirmed previously when SPP activity was reconstituted in *S. cerevisiae,* an organism that does not encode the protease [106]. Despite the availability of relevant functional studies on SPP, protocols for the high-grade purification of this protease for detailed structural analyses have not been reported. The absence of this type of studies is most likely due to the eukaryotic nature of SPP and the inherent complexity that is then added to purification procedures, precluding further insights into the nature of di-aspartyl intramembrane proteases.

Advantageously, for the other two families of intramembrane proteases, the presence of homologues of rhomboids and S2P in lower organisms resulted in the purification in high yields of the *E. coli* rhomboid, GlpG and the archaeal S2P, mjS2P from simple expression systems. This strategy was followed by a crystallographic analysis that provided the first structural and mechanistic insight into intramembrane proteases. For GlpG [121, 239], these studies revealed a catalytic dyad composed of Ser201 and His254 buried within a water accessible cavity that opens towards the extracellular side. The

origin of the cavity is attributed to the shorter nature of TM4, with S201 located at its N-terminus, 10Å below the membrane surface, situating this active site residue at the bottom of the cavity. In both published structures, the hydrophilic cavity is partially surrounded by TM helices with an additional flexible component that acts as a lateral gate for substrate entry (TM5 in [239] and L1 loop in [121]). From the structure of GlpG provided, not only many of the functional and mechanistic hypotheses on Rhomboids can be confirmed but it also provides a plausible route for protein substrates of intramembrane proteases to access an active site that although buried within the membrane, is in contact with water molecules.

A lateral gating mechanism has also been described for mjS2P [120]. In this archaeal homolog of the eukaryote S2P, TM1 and TM5-TM6 together, seem to move away from each other to allow substrate entry and catalysis by His54, His58 (both in TM2) and Asp148 (in TM4). In this open conformation, the funnel-shaped channel originates from the coordinated zinc atom, 14Å below the lipid membrane surface, to the cytosolic side. Although a similar mechanism seems to govern water access to the active site in both GlpG and mjS2P, in the latter, water molecules seem to have constant access.

Despite the fact that we now have representative prokaryotic structures of serine and metallo intramembrane proteases, we still lack a structural model of a GXGD di-aspartyl intramembrane protease of similar resolution as well as a mechanistic understanding on how they work. In our study we have provided two separate novel routes to approach a degree of understanding of di-aspartyl intramembrane proteases comparable to that

achieved for the structurally characterized GlpG and mjS2P. The first one involved the expression and purification in high yields of human γ-secretase using *P. pastoris* as an expression system, an application that provides the same advantages as the expression in mammalian or insect cells, but in a less elaborated manner. Improvements on this procedure are expected in a near future to provide an optimal sample for high-resolution structural studies. Alternatively, we identified the archaeal di-aspartyl intramembrane protease MCMJR1 and probed its biochemical similarity to PS and SPP. Indeed, mutation to alanine of any of the aspartic amino acid residues within the signature motif YD or GXGD abolishes MCMJR1 proteolytic activity, as does the corresponding mutations in presenilins [86] and SPPs [106]. Although we cannot at this point be fully confident about the validity of the predicted topology of MCMJR1, we note that in analogy to PSs and SPPs these two motifs are located towards the C-terminal of MCMJR1, in adjacent TMDs that are joined by a relatively long loop. Employing substituted cysteine accessibility methods may show in future experiments that TMDs 5 and 6 of MCMJR1, containing the YD and GMGD motifs respectively, contribute to the formation of a hydrophilic pore within the membrane. Furthermore, the mutation G219A - adjacent to D220 in the GMGD motif - slows down proteolytic activity considerably, in clear analogy to the effect observed when the corresponding mutation was introduced in both presenilin [63] and SPP [191]. Additional aminoacid exchanges not only within the catalytic motifs YD162 and GMGD220 but also of those residues conserved among other putative archaeal di-aspartyl intramembrane proteases and human PS (Fig. 46) are worthy of being tested in the future in order to characterize the activity of MCMJR1. In addition, in analogy to SPP, neither ectodomain removal nor the presence of accessory proteins

appears to be a requisite for MCMJR1-catalyzed proteolysis of TMD substrates. Moreover, MCMJR1 displays promiscuous peptide bond specificity as processing of these chimeric substrates was shown to occur at multiple positions within the hydrophobic region of the predicted TMD, also a salient feature of presenilin [240]. The identification of the cleavage site entails great potential for accomplishing future kinetic studies on the activity of MCMJR1 by fluorescence resonance energy transfer by introducing a donor/acceptor pair on either side of a short aminoacid sequence that includes the scissile bond.

Perhaps the most remarkable characteristic of MCMJR1 is its ability to cleave the established *in vitro* presenilin substrate C100Flag [141]. Moreover, the chimera containing the TMD of APP was hydrolyzed by MCMJR1 at two transmembrane sites, which correspond to known presenilin cleavage positions. This feature of MCMJR1 deserves further investigation with enzyme and substrates reconstituted in lipid bilayers to best mimic an *in vivo* situation.

Taken together, it can then be suggested, as found for rhomboids, that eukaryote and putative prokaryote di-aspartyl intramembrane proteases may also be functionally conserved thus validating the use of proteases from prokaryotic sources as more tractable and presumably simpler proteases to implement adequate biophysical and structural approaches.

Table I
Clans and families of proteases

protein nucleophile: serine-, threonine- and cysteine peptidase				
Clan	Family	EC	Example	Active site residues
PA	S1	3.4.21.4	Trypsin	Ser; His; Asp
SB	S8	3.4.21.62	Subtilisin E	Ser; His; Asp
SF	S26	3.4.21.89	Signal peptidase I	Ser; Lys/His
ST	S54	3.4.21.105	Rhomboid	Ser; His
PB	T1	3.4.25.1	Proteasome	Ser; Thr
CA	C1	3.4.22.2	Papain	Cys; His; Asn
CA	C2	3.4.22.52	Calpain-1	Cys; His; Asn;Gln
CD	C14	3.4.22.36	Caspase-1	Cys; His
water nucleophile: aspartic-, glutamic- and metallo peptidase				
AA	A1	3.4.23.1	Pepsin A	Asp; Asp; Tyr
AA	A2	3.4.23.16	HIV-1	Asp; Asp
GA	G1	3.4.23.32	Scytalidoglutamic peptidase	Glu; Gln
MA	M10	3.4.24.7	Matrix metallopeptidase-1	His; His; His/Glu
MA	M4	3.4.24.27	Thermolysin	His; His/Glu
MA	M2	3.4.15.1	Angiotensin-converting enzyme peptidase unit 2	His; His/Glu

Table II
Cleavage pattern of soluble proteases

Family	Protease	Cleavage pattern	# of cleavages based on
Serine	Chymotrypsin B	a/A/p/Fw†-/-/-/-	10
	Trypsin 1	-/-/-/KR†-/-/-/-	13760
	Elastase 2	-/-/-/viat†-/-/-/-	400
Cysteine	Cathepsin B	-/-/l/r†-/-/g/-	81
	Caspase 3	D/e/v/D†gs/-/-/-	414
	Calpain 1	-/-/l/-†as/-/-/-	101
Aspartyl	HIV-1 retropepsin	-/-/ve/l†l/eva/-/-	1059
	Pepsin A	P/a/e/F†fl/R/L/-	79

Listed are representative members of serine, cysteine and aspartyl protease families. For each, sequences around cleavage sites (ten or more) were aligned from which the specificity preferences for amino acid residues are summarized. Up to four residues on either side of the scissile bond (cross) are shown. Upper- or lower-case letters are used to give an intuitive feel for the strength of the specificity preference. *Rawlings, N.D., Morton, F.R., Kok, C.Y., Kong, J. & Barret, A.J. (2008) MEROPS: the peptidase database. Nucleic Acids Res 36, D320-D325*

Table III
Observed and calculated masses of MCMJR1 cleavage products

Substrate	*Calculated mass of products (Da)*		*Observed mass of products (Da)*
	Product sequence	Mass	Mass
Gurken-TMD	**MBP**-PRGS--PVL_{255}	45900.9	45940
	GS--PVL_{255}	3409.9	3411.3
	MBP-PRGS--$PVLL_{256}$	46014.1	46011
	GS--$PVLL_{256}$	3523.1	3524.5
	GS--$PVLLM_{257}$	3654.3	3655.6
Keren-TMD	**MBP**-PRGS--TLA_{131}	46106.1	46122
	GS--TLA_{131}	3615.1	3615.8
	MBP-PRGS--$TLALL_{133}$	46332.4	46338
	GS--$TLALL_{133}$	3841.4	3842.2
	MBP-PRGS--$TLALLF_{134}$	46479.6	46450
	GS--$TLALLF_{134}$	3988.6	3989.5

Figure 1. Structure of biological membranes. Schematic depicts the current view of biological membranes, in which the lipid bilayer has variable patchiness and thickness. These membranes are typically crowded with integral membrane proteins and peripheral membrane proteins. Membrane protein oligomers and complexes with their soluble partners are common. *Engelman, D., Membranes are more mosaic than fluid. Nature, 2005. 438(1): p.578-80.*

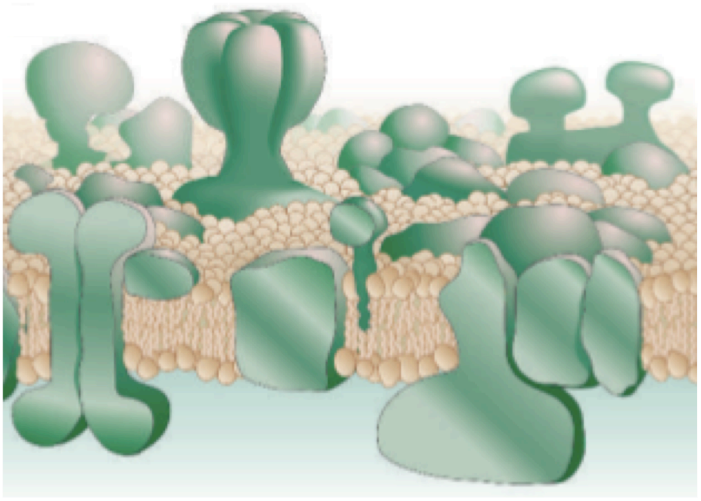


FIGURE 1

Figure 2. Schematic of membrane protein types. Single-pass transmembrane proteins **(A)** type I and **(B)** type II indicating the N- and C-termini, **(C)** multipass transmembrane, **(D)** lipid chain-anchored membrane, **(E)** GPI-anchored membrane, and **(F)** peripheral membrane proteins. TMD regions are in gray. Elongated rectangle depicts the lipid bilayer.

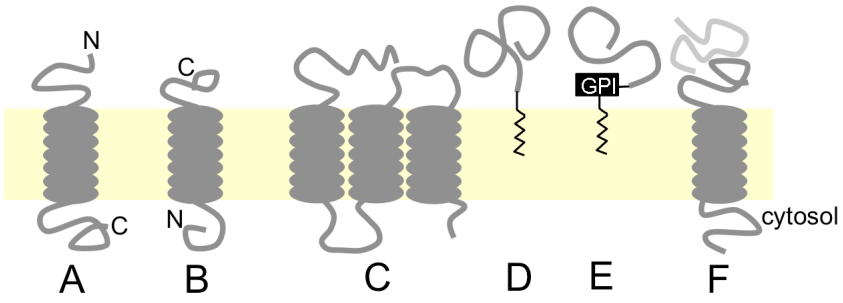


FIGURE 2

Figure 3. Signal transduction across cell membranes. Four classes of cell surface receptors involved in signal transduction: **(A)** Ion-channel-linked receptors, **(B)** G-protein-linked receptors, **(C)** enzyme-linked receptors, and **(D)** signaling dependent on regulated intramembrane proteolysis.

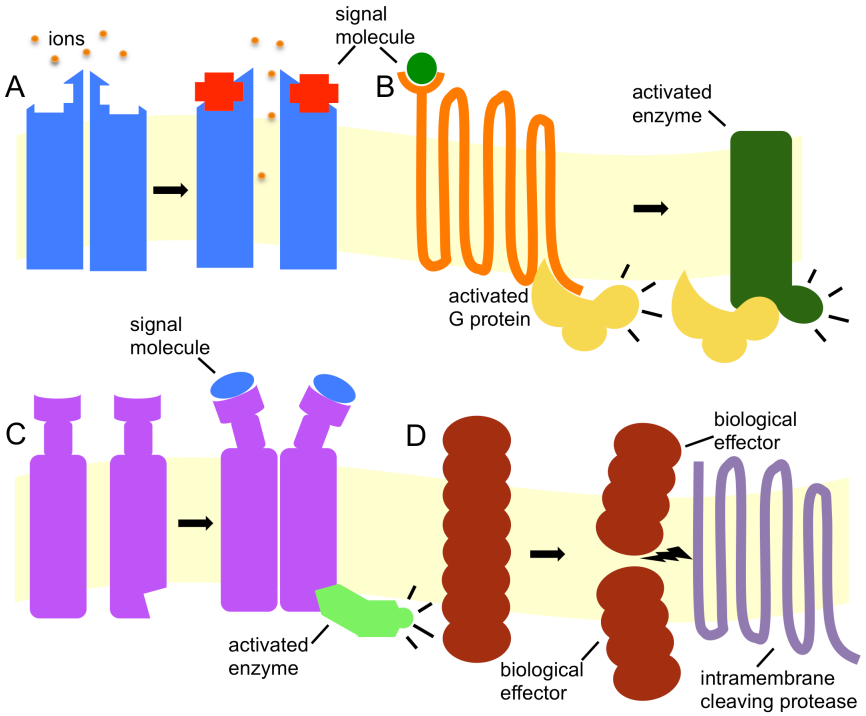

ions
signal molecule
A
B
activated enzyme
activated G protein
signal molecule
C
D
biological effector
activated enzyme
biological effector
intramembrane cleaving protease

FIGURE 3

Figure 4. Intramembrane-cleaving proteases. Dark cylinders across the lipid bilayer (continuous slab) depict the TMD regions of Site-2 protease (S2P) and Rhomboid and those predicted for Presenilin and Signal peptide peptidase. Catalytic residues are assigned with the one-letter code (bold) and the coordinated zinc ion (Zn) in S2P is circled.

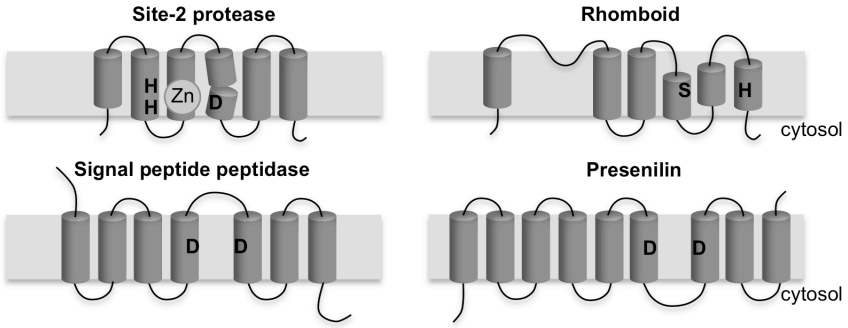

Site-2 protease
H
H
Zn
D
Rhomboid
S
H
cytosol
Signal peptide peptidase
D
D
Presenilin
D
D
cytosol

FIGURE 4

Figure 5. Schematic of the mechanism of catalysis of soluble proteases. (A) The catalytic triad in the active site of serine proteases consists of aspartic acid, histidine and serine. Deprotonation of the serine hydroxyl aided by aspartic acid and histidine enables the nucleophilic attack of the substrate carbonyl carbon. The backbone NH groups of glycine and serine located in the oxyanion hole stabilize the tetrahedral transition state intermediate. **(B)** In cysteine proteases the nucleophilic attack is carried out by the sulphur atom of a cysteine residue. As described for serine proteases, cysteine proteases also feature an oxyanion hole formed by cysteine and glutamine. **(C)** A general acid-base mechanism takes place in aspartyl proteases. A water molecule (red) is activated by the two aspartic acid residues in the active site of the enzyme enabling the polarized water to attack the substrate carbonyl carbon. **(D)** A coordinated metal, often zinc, participates in the catalytic mechanism of metalloproteases. Coordination is accomplished by three or two histidines and an acidic side chain such as glutamate as shown here activating the water molecule (red). The zinc ion stabilizes the oxyanion.

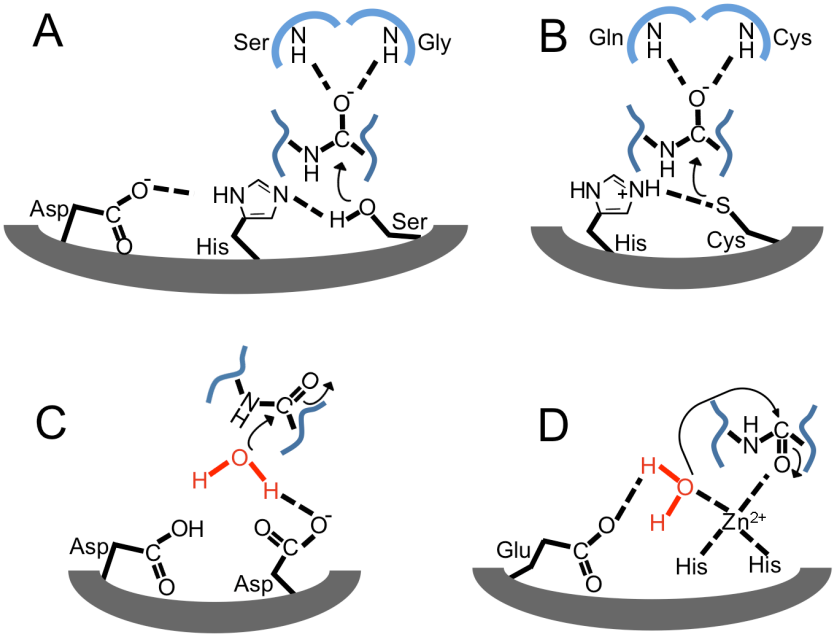


FIGURE 5

Figure 6. Schematic representation of an enzyme-substrate complex. The standard nomenclature P_1 … P_n, P_1' … P_n', designates amino acid residues of peptide substrates that have the same numbering (counted from the point of cleavage) as the subsites they occupy in the active site of the enzyme (S_1 … S_n, S_1' … S_n'). The scissile bond is indicated (arrow).

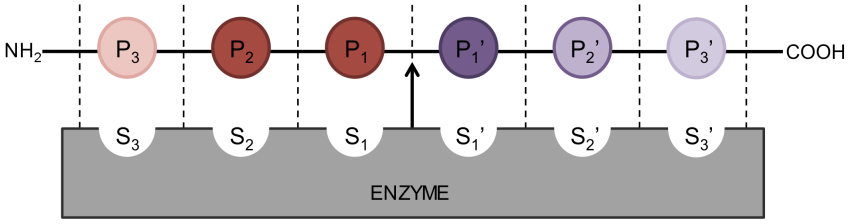


FIGURE 6

Figure 7. Proposed catalytic mechanism for soluble aspartic proteases. (A) In this mechanism, the Asp group that is closer to the nucleophilic water molecule is assigned a negative charge. The nucleophilic water molecule held between the catalytic aspartates attacks the carbonyl group in the substrate scissile bond to generate an oxyanion tetrahedral intermediate (in brackets). Protonation of the scissile amide N atom and rearrangement result in the breakdown of the tetrahedral intermediate to the hydrolysis products. **(B)** Peptidic substrate analogue containing a non-hydrolyzable P_1-P_1' amide bond such as the hydroxymethylcarbonyl isostere transition state inhibitor depicted. *Jeffrey-Tri Nguyen, Yoshio Hamada, Tooru Kimura, and Yoshiaki Kiso. Design of Potent Aspartic Protease Inhibitors to Treat Various Diseases. Arch. Pharm. Chem. Life Sci. 2008, 341: p. 523 – 535*

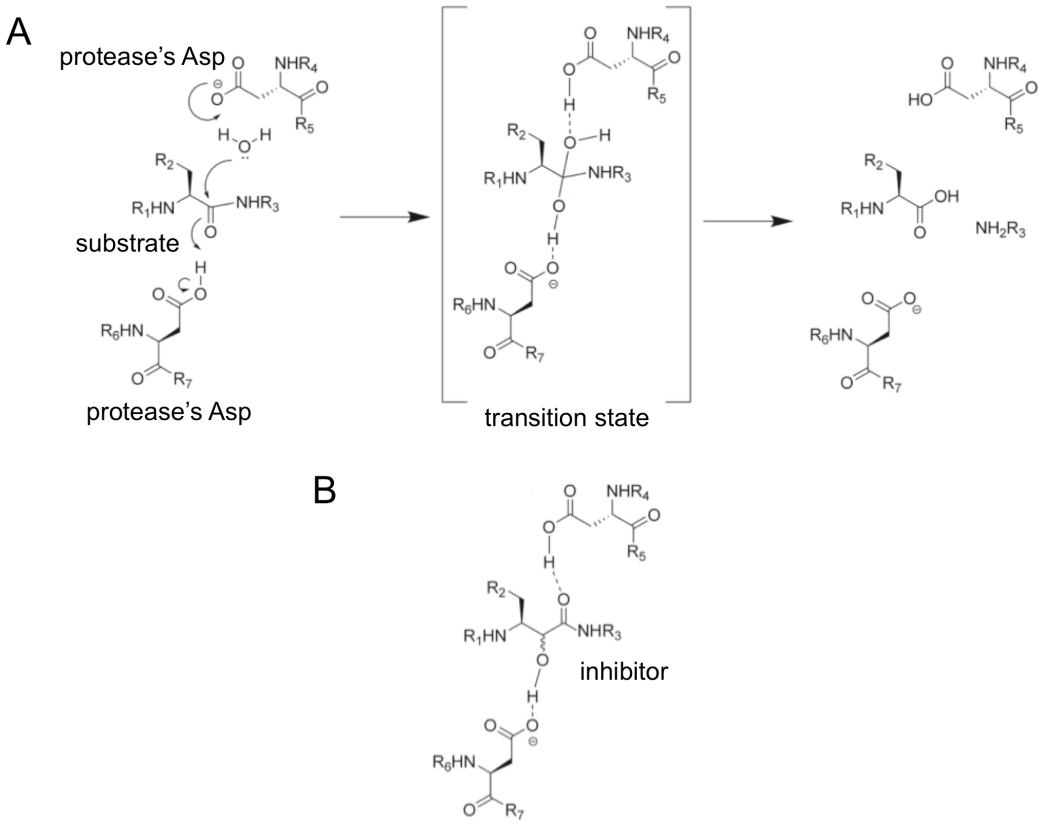

A
protease's Asp
substrate
protease's Asp
transition state
B
inhibitor

FIGURE 7

Figure 8. GXGD-type di-aspartyl intramembrane proteases. **(A)** Cylinders depict the predicted TMD regions of presenilin (PS), signal peptide peptidase (SPP) and the pilin peptidase from *V. cholerae*, TcpJ. Stars show the predicted positions of the catalytic aspartic acid residues and for each the surrounding amino acid residues are also indicated. A continuous slab depicts the membrane. **(B)** Multiple sequence alignment of the amino acid sequences of PS homologs, SPP homologs and TcpJ. Residues around the N-terminal (Asp 1) and C-terminal aspartate residues (Asp 2) and the C-terminal PAL motif, when present, are shown. Numbering of the aspartate residues and that of proline in the PAL motif are given according to the full-length protein.

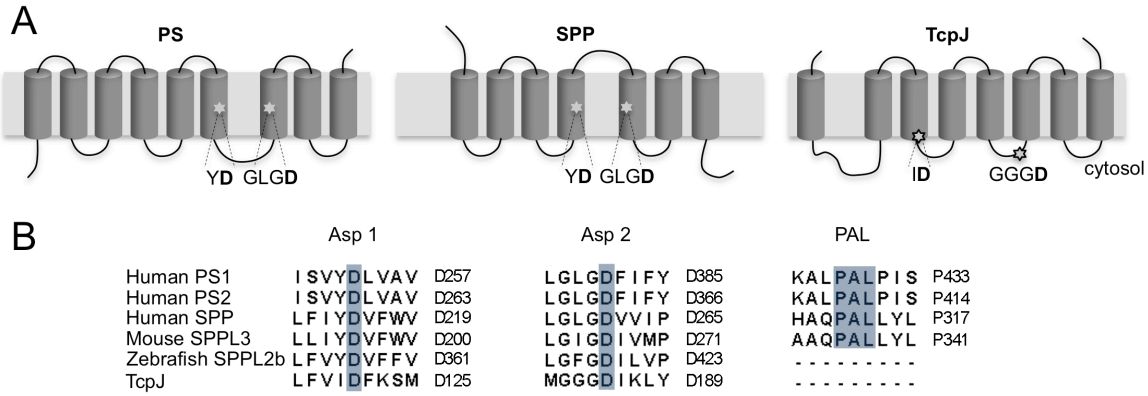

A
PS
SPP
TcpJ
YD GLGD
YD GLGD
ID
GGGD
cytosol
B
Asp 1
Asp 2
PAL
Human PS1 ISVYDLVAV D257 LGLGDFIFY D385 KALPALPIS P433
Human PS2 ISVYDLVAV D263 LGLGDFIFY D366 KALPALPIS P414
Human SPP LFIYDVFWV D219 LGLGDVVIP D265 HAQPALLYL P317
Mouse SPPL3 LLIYDVFWV D200 LGIGDIVMP D271 AAQPALLYL P341
Zebrafish SPPL2b LFVYDVFFV D361 LGFGDILVP D423 ----------
TcpJ LFVIDFKSM D125 MGGGDIKLY D189 ----------

FIGURE 8

Figure 9. Schematic representation of the subunit topology and composition of γ-secretase. The TMDs are depicted as cylinders and the ectodomain of NCT is shown as a sphere with schematic glycosylation sites (Glyc). A larger cylinder depicts APP, located between the PS fragments and in close proximity to the catalytic aspartic acid residues. A continuous slab depicts the membrane.

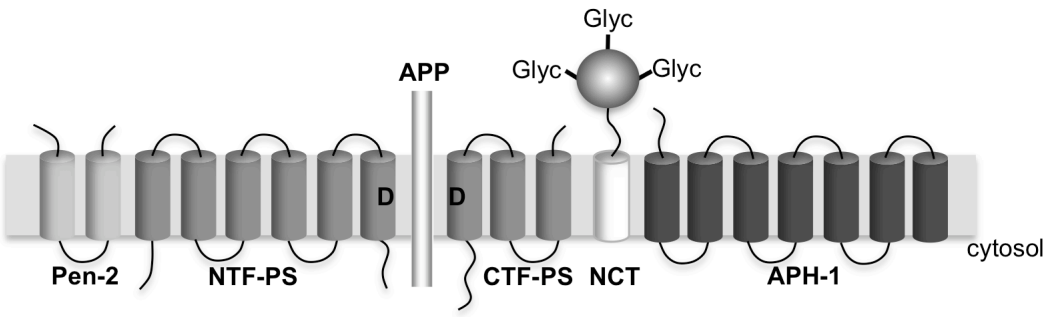


FIGURE 9

Figure 10. Schematic diagram of proteolytic degradation of the amyloid precursor protein. **(A)** Depicted is the 770-amino acid residue long isoform of the amyloid precursor protein (APP), an integral membrane protein containing a Kunitz-type protease inhibitor (KPI) domain and a single TMD region. The ectodomain of APP can be processed either by α-secretase or β-secretase followed by intramembrane proteolysis catalyzed by γ-secretase to produce p3 or amyloid β-peptides (Aβ) and their soluble counterparts (sAPPα and sAPPβ). **(B)** Detail of the amino acid sequence of Aβ domain (in bold) and the TMD (underlined) of APP. The different cleavage sites by α-, β- or γ-secretases are marked. Also marked are the positions where γ-secretase cleaves to generate products of 40 (γ40), 42 (γ42) and 49 (ε49) amino acids in length. The mutations known to cause FAD are shown in italics.

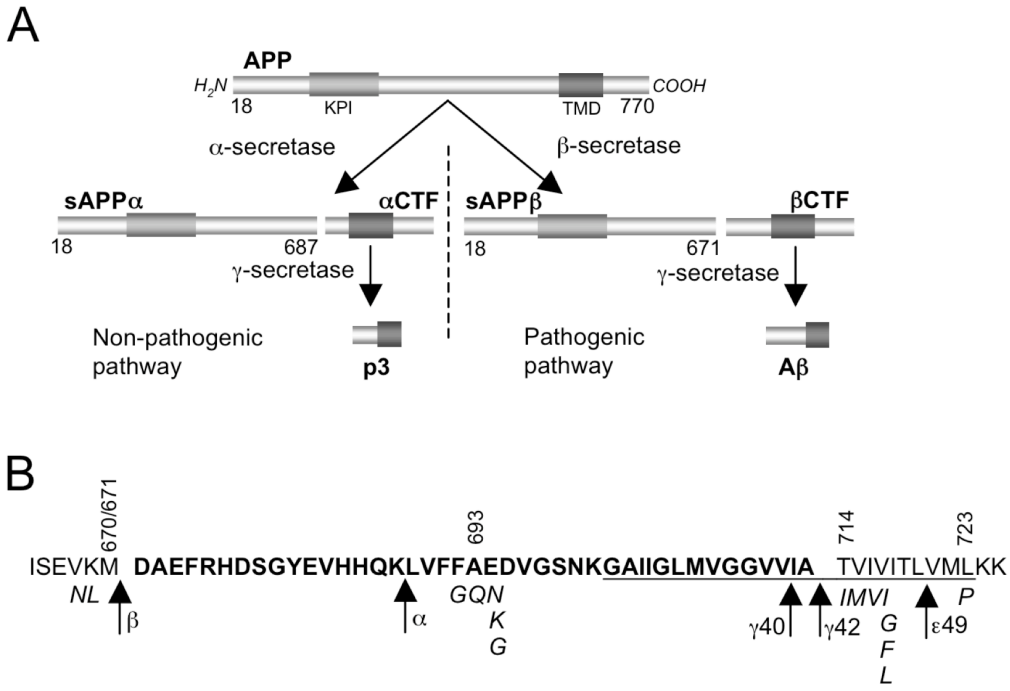

A
APP
H2N
COOH
18
KPI
TMD
770
α-secretase
β-secretase
sAPPα
αCTF
sAPPβ
βCTF
18
687
18
671
γ-secretase
γ-secretase
Non-pathogenic pathway
p3
Pathogenic pathway
Aβ
B
670/671
693
714
723
ISEVKM DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA TVIVITLVMLKK
NL
β
α
GQN
K
G
γ40
γ42
IMVI
G
F
L
ε49
P

FIGURE 10

Figure 11. General scheme of the steps involved in the reconstruction of a 3D map. The macromolecular sample preparation purified to homogeneity (1) is applied for a few seconds to an EM grid of appropriate material and mesh size (2). After blotting the excess solution, the sample can be either embedded in a heavy metal stain (3a) or quickly frozen in liquid ethane at -180°C using a freeze plunger (3b). The grid is then inserted into the TEM and micrographs of the sample generated (4) from which 2D images of individual particles are later collected using computational tools (5). The orientation and position of each 2D image is determined (6) and this information finally combined to reconstruct the 3D structure of the object (7).

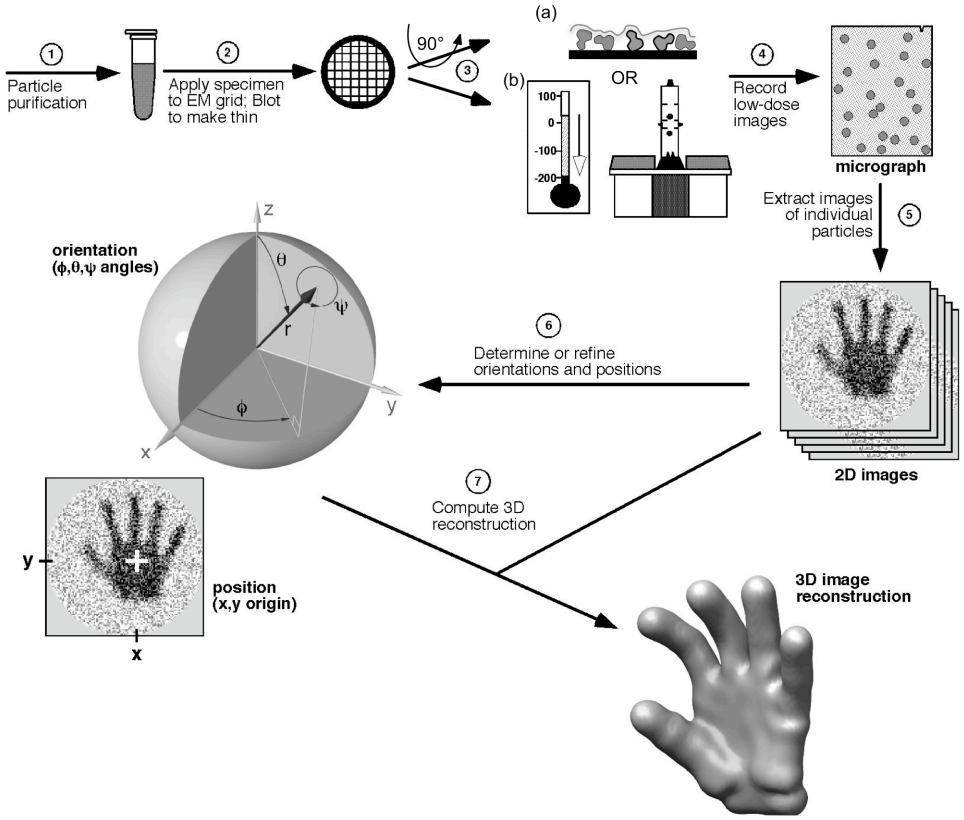

1
Particle purification
2
Apply specimen to EM grid; Blot to make thin
90°
3
(a)
(b)
OR
100
0
-100
-200
4
Record low-dose images
micrograph
Extract images of individual particles
5
z
orientation (ϕ,θ,ψ angles)
θ
ψ
r
ϕ
y
x
6
Determine or refine orientations and positions
2D images
7
Compute 3D reconstruction
y
x
position (x,y origin)
3D image reconstruction

FIGURE 11

Figure 12. TAP-PS undergoes endoproteolysis in HEK293 cells. Cell lysates from HEK293 cells that were transiently transfected with wt-PS (lane 1); wt-PS, NCT, APH-1 and PEN-2 (lane 2); TAP- PS (lane 3); TAP-PS, NCT, APH-1 and PEN-2 (lane 4), and cell lysates from HEK293S cells stably co-expressing TAP-PS, NCT, APH-1 and PEN-2 (lane 5) were resolved by SDS-PAGE. Full length PS (PS-FL) and its derivatives PS-NTF and PS-CTF were probed by western blotting. The migration position of the molecular weight marker is shown on the right.

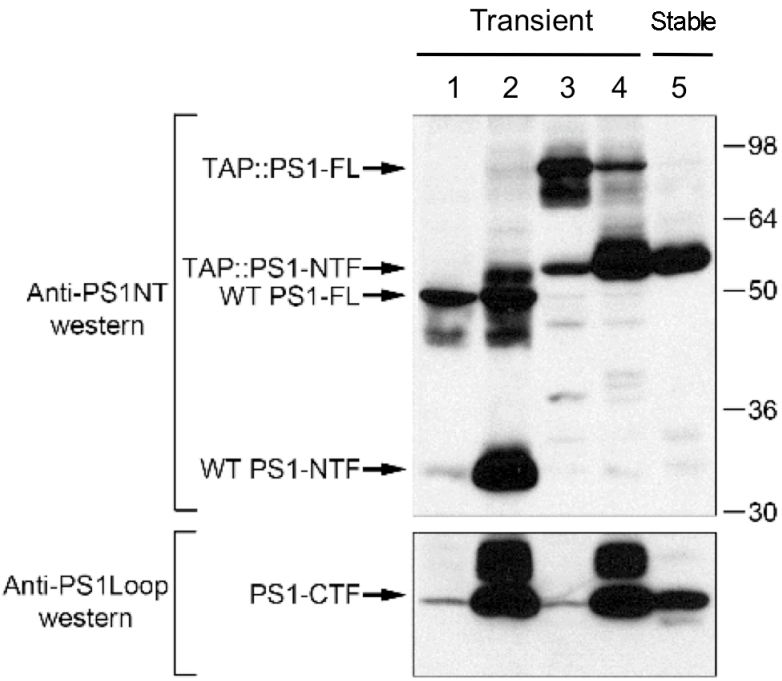

Transient
Stable
1
2
3
4
5
TAP::PS1-FL
TAP::PS1-NTF
WT PS1-FL
WT PS1-NTF
Anti-PS1NT western
Anti-PS1Loop western
PS1-CTF
98
64
50
36
30

FIGURE 12

Figure 13. Purification of mature and enzymatically active γ-secretase in detergent. **(A)** Tandem affinity purification (TAP) of γ-secretase from HEK293 cells that over-express TAP-PS, NCT, APH-1 and PEN-2. 20 ug of solubilized γ-secretase input, an equivalent volume of IgG sepharose flow through, 0.5% of IgG sepharose eluate as well as 0.5% of final eluate were all directly mixed with 3x sample buffer and loaded onto the gel. 10% of IgG wash fraction, 10% of calmodulin flow through and 20% of calmodulin wash fraction were first TCA precipitated, resuspended in 1x sample buffer before loaded onto the gel. Full length TAP-PS and its endoproteolytic derivatives, NCT and PEN-2 were probed by western blotting. Each fraction is indicated on top of the blot image. FT indicates flow through. **(B)** Purified γ-secretase was fractionated by SDS-PAGE and visualized with silver staining. The bands corresponding to the γ-secretase components are labeled. **(C)** The production of Aβ40 by purified γ-secretase can be inhibited by the transition state analog inhibitor L-685,458. **(D)** Purified γ-secretase catalyzes the conversion of the substrate C100Flag to both Aβ40 and Aβ42 products in a ratio analogous to that found *in vivo*.

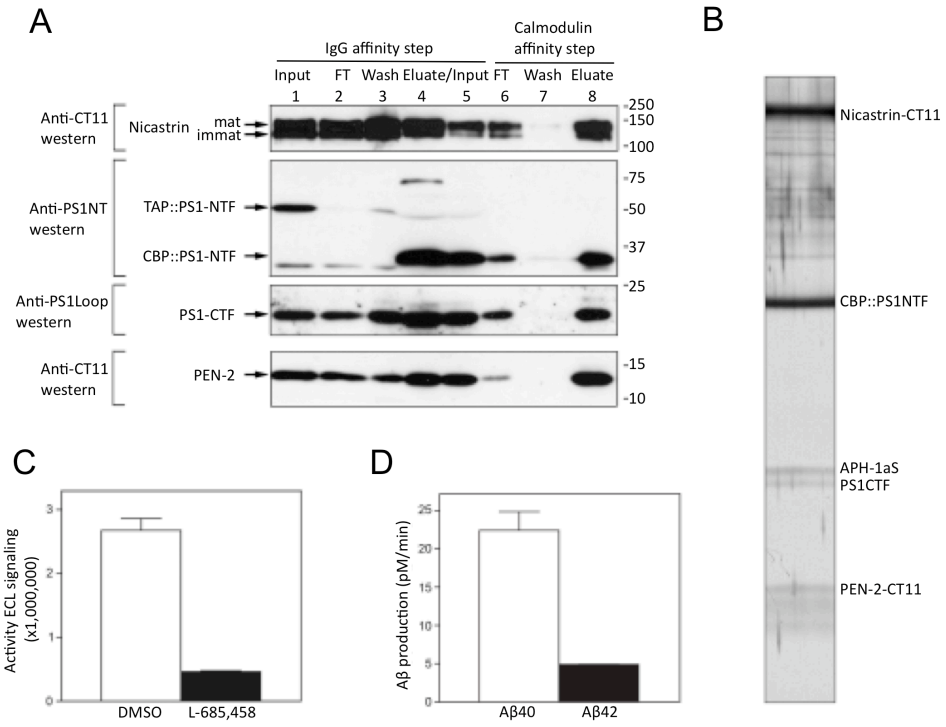

A
IgG affinity step
Calmodulin affinity step
Input FT Wash Eluate/Input FT Wash Eluate
1 2 3 4 5 6 7 8
Anti-CT11 western
Nicastrin
mat
immat
Anti-PS1NT western
TAP::PS1-NTF
CBP::PS1-NTF
Anti-PS1Loop western
PS1-CTF
Anti-CT11 western
PEN-2
250
150
100
75
50
37
25
15
10
B
Nicastrin-CT11
CBP::PS1NTF
APH-1aS
PS1CTF
PEN-2-CT11
C
Activity ECL signaling (x1,000,000)
DMSO
L-685,458
D
Aβ production (pM/min)
Aβ40
Aβ42

FIGURE 13

Figure 14. Electron microscopy of γ-secretase. (A) Low-dose CCD image of γ-secretase after incubation with the antibody NCT54 and negative staining with ammonium molybdate. The inset on the right shows a gallery of γ-secretase particles decorated with a single bound NCT54 molecule (contoured in white). **(B)** Plot of the Euler angle distribution of class averages used to reconstruct the 3D map. The radii of the circles are proportional to the number of particles in each class. **(C)** Gallery of class-averages and their corresponding back projections from the 3D map. **(D)** Fourier shell correlation (FSC) curve. The scale bars represent 150Å

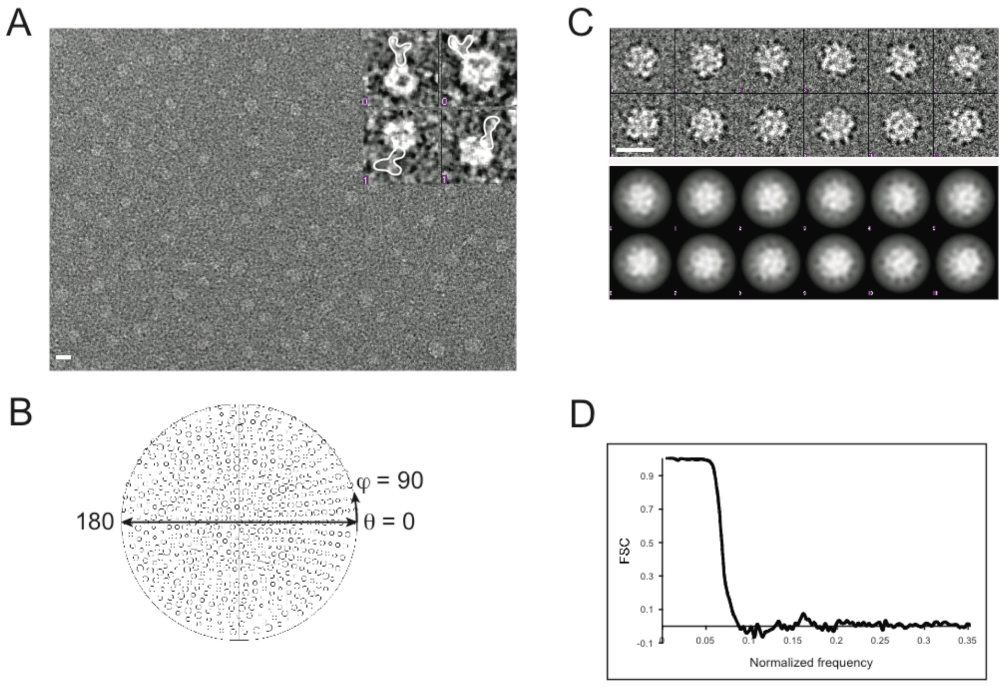

A
C
B
D
φ = 90
180
θ = 0
FSC
0.9
0.7
0.5
0.3
0.1
-0.1
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
Normalized frequency

FIGURE 14

Figure 15. 3D reconstruction of negatively stained γ-secretase. Surface rendition of different views of γ-secretase**. (A)** Front view of the structure contoured at 3σ (blue) and 3.5σ (green). An asterisk marks the position of the Fab fragment of NCT54 and the location of the ectodomain of NCT is indicated. The dimensions of the map in a direction (z-axis) perpendicular to the membrane plane and its topology are given on the left. For size comparison the structure of the Ca-ATPase (PDB ID 1SU4) is shown on the right and its topology has been inverted for clarity. **(B)** Front view (same view as in **A** but countered at 3σ). **(C)** Side view. **(D)** Back view. **(E)** Side view. **(F)** Extracellular view. **(G).** Cytoplasmic view. From **B-G** the map is contoured at 3σ and the axis of rotation and angle relative to **B** is given. White arrows mark regions of low-density on the surface of the map. Scale bars are included.

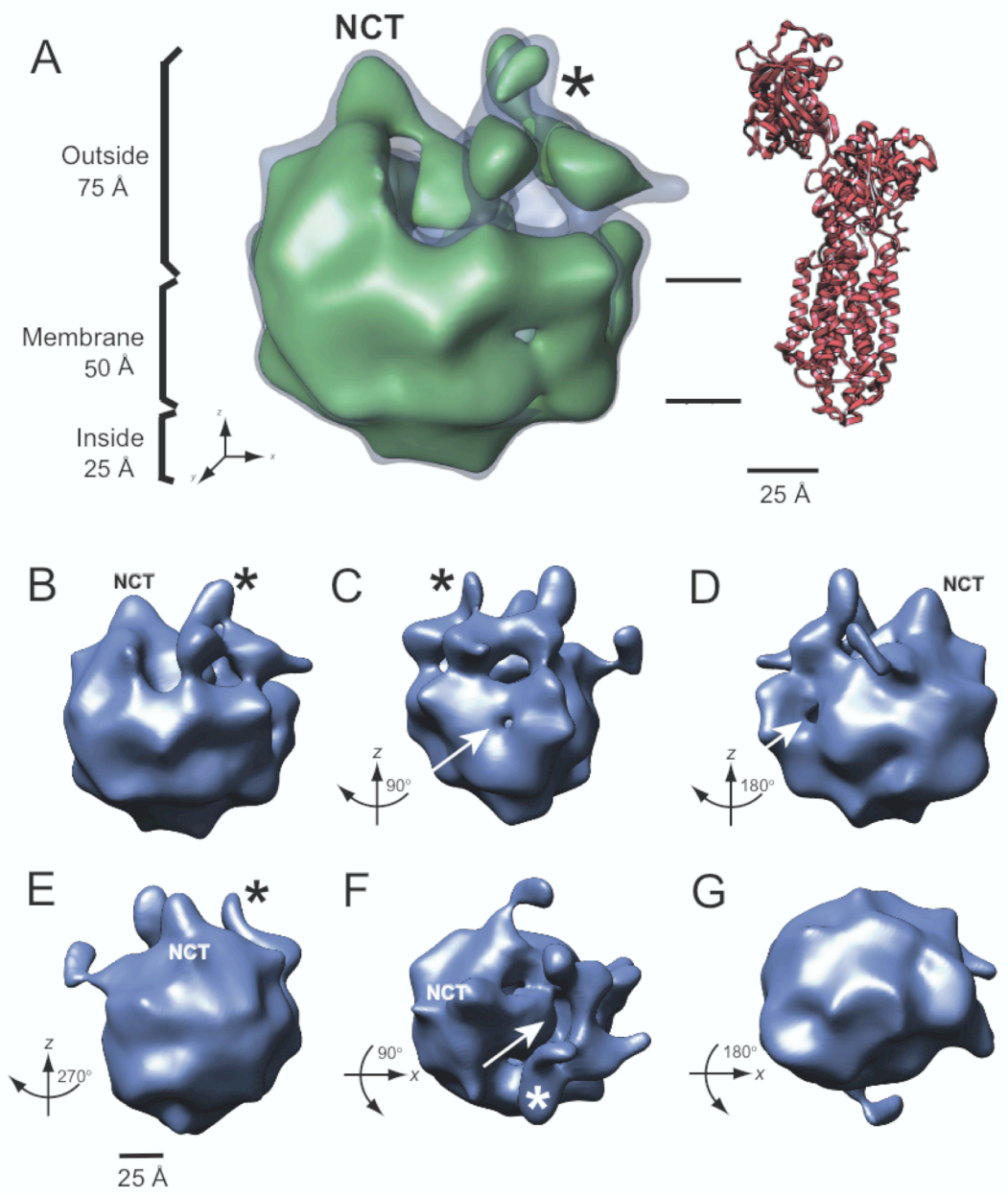

A
NCT
*
Outside
75 Å
Membrane
50 Å
Inside
25 Å
25 Å
B
NCT
*
C
*
z
90°
D
NCT
z
180°
E
*
NCT
z
270°
25 Å
F
NCT
90°
x
*
G
180°
x

FIGURE 15

Figure 16. Horizontal slices through the 3D reconstruction of γ-secretase. (A) A side view (as in Fig. 15B) of the γ-secretase/NCT54 complex contoured at 3σ is given as reference. Discontinuous lines indicate the horizontal slices perpendicular to the membrane plane and separated by 25Å. **(B-E)** Views of the horizontal slices through the density map with the following display thresholds; blue at 3σ, orange at 2σ and red at 1σ. For clarity, the sections are shown with the map tilted 25° towards the reader relative to the side view in. The sections start at the extracellular region and finish at the cytosolic region. The scale bar is shown.

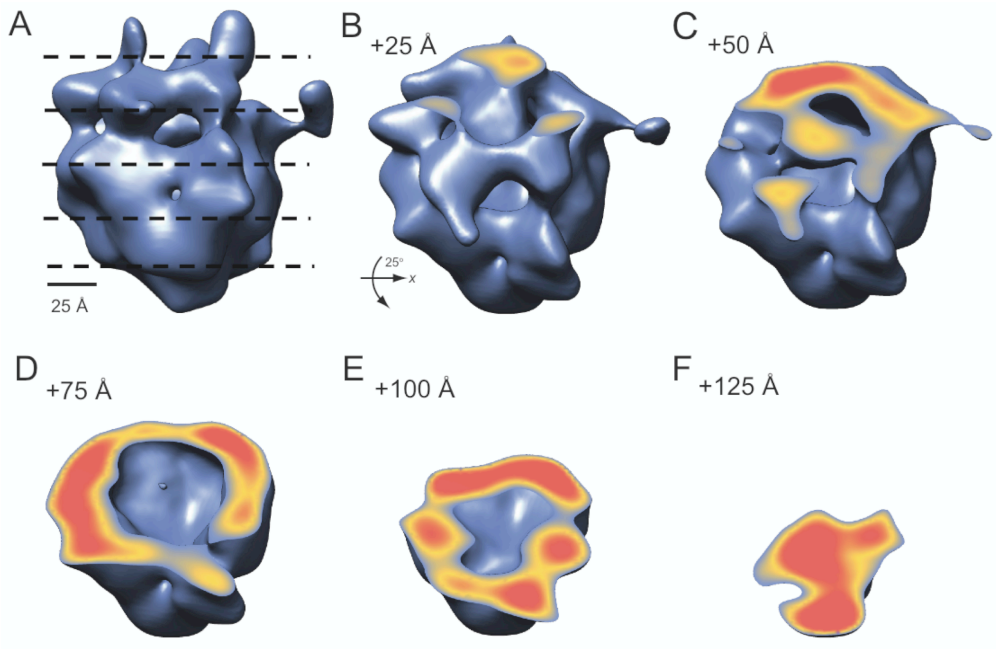


FIGURE 16

Figure 17. Gold labeling of the γ-secretase active site. (A) Gallery of images of γ-secretase particles labeled with gold-streptavidin complexed to the inhibitor. **(B)** Positions of the gold-streptavidin labels, indicated by red dots, overlaid with a view of the complex sliced (as in Fig. 16) by a horizontal plane perpendicular to the membrane and located at a height corresponding to the extracellular membrane surface. The map is contoured as in figure 16. The gold-labels are tethered by a 39.4Å spacer arm to the inhibitor warhead and distributed in an arc around a point that indicates the inhibitor location. **(C)** Front view (as in Fig. 15B) of the γ-secretase contoured at 3σ in which, a red circle shows the area populated by the gold-streptavidin label on the surface of the complex. **(D)** View of the complex as in **B** showing the putative location of PS and of the γ-secretase active site (AS) exposed to the internal cavity, as derived from the gold labeling experiments. Scale bars are shown.

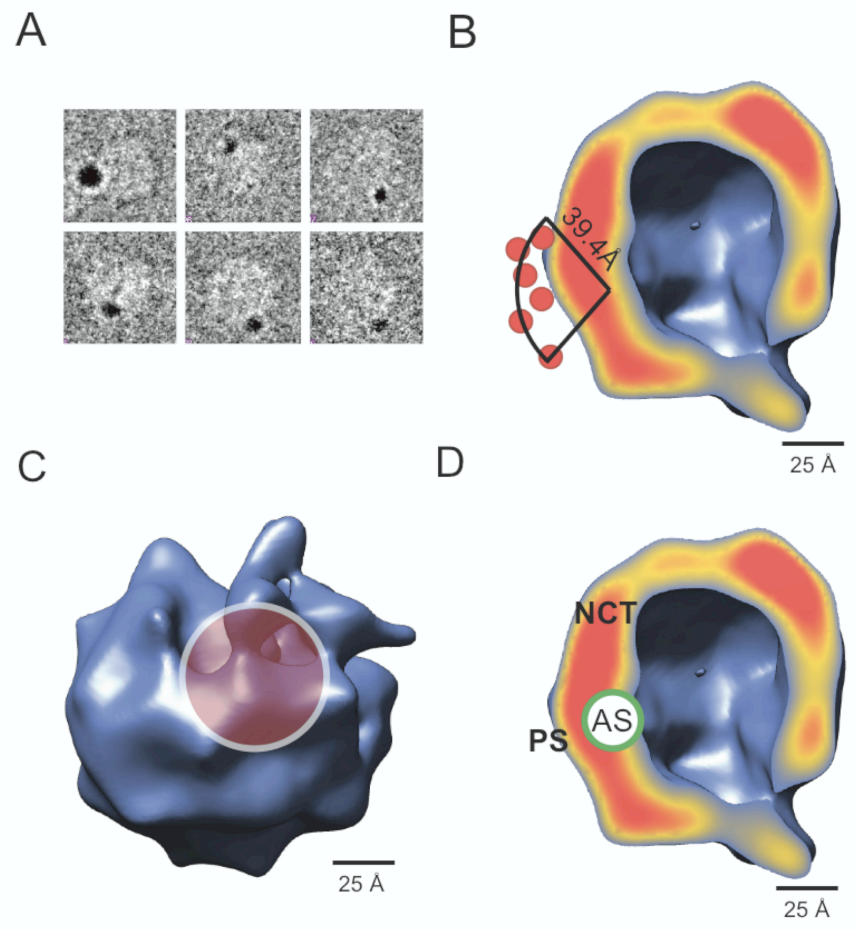


FIGURE 17

Figure 18. Model for transmembrane substrate hydrolysis by γ-secretase. A schematic for the hydrolysis of hydrophobic transmembrane substrates by catalytic aspartyl moieties at the interface of a water-accessible internal cavity is depicted. **1.** A cartoon of γ-secretase representing a cut-open view from the same side view as in fig. 15B, in which the water-accessible internal cavity, the putative location of the catalytic aspartates, and the ectodomain of NCT (marked by an asterisk) are shown. A schematic of the substrate APP-CTF (a green rectangle depicts its TMD and black-solid lines represent its N- and C-termini) is also shown. **2**. The substrate interacts with the membrane-exposed surface of γ-secretase and with the ectodomain of NCT. **3.** The substrate is transferred to the active site where, following partial unwinding of the transmembrane helix, water-accessible aspartyl moieties carry out peptide bond hydrolysis. **4.** The resulting N-terminal (Aβ) and C-terminal products (CD) are released into the lumen/extracellular space and cytosol, respectively.

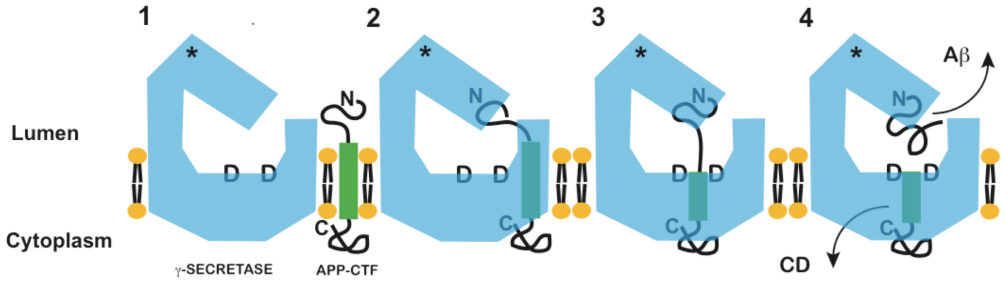


FIGURE 18

Figure 19. 3D structures of γ-secretase complex. (A) Negatively stained EM structure at 15 Å. **(A.1)** The first row displays a side view after rotating the map 180° and the second row shows tilted views by rotating around a horizontal axis. Two openings at the top and bottom are labeled H1 and H2, respectively. The top density is assigned as the ectodomain of NCT. **(A.2)** To the left, the potential transmembrane segment is colored green and outlined by two parallel dashed lines. To the right, a cut-open view of the γ-secretase complex revealing a central aqueous chamber. Two weak-density regions are labeled with asterisks. *Lazarov, V.K., et al., Electron microscopic structure of purified, active gamma-secretase reveals an aqueous intramembrane chamber and two pores. Proc. Natl. Acad. Sci. U S A, 2006. 103(18): p. 6889-94*. **(B)** Cryo-EM structure at 12 Å. Four domains are colored based on separations of electron densities observed and the arrow indicates the apparent cavity on the cytoplasmic side. As observed by the transmembrane region outlined by two parallel dashed lines, a larger extracellular domain is assigned compared to the negatively stained EM map. *Osenkowski, P., et al., Cryoelectron microscopy structure of purified gamma-secretase at 12 A resolution. J. Mol. Biol., 2009. 385(2): p. 642-52.*

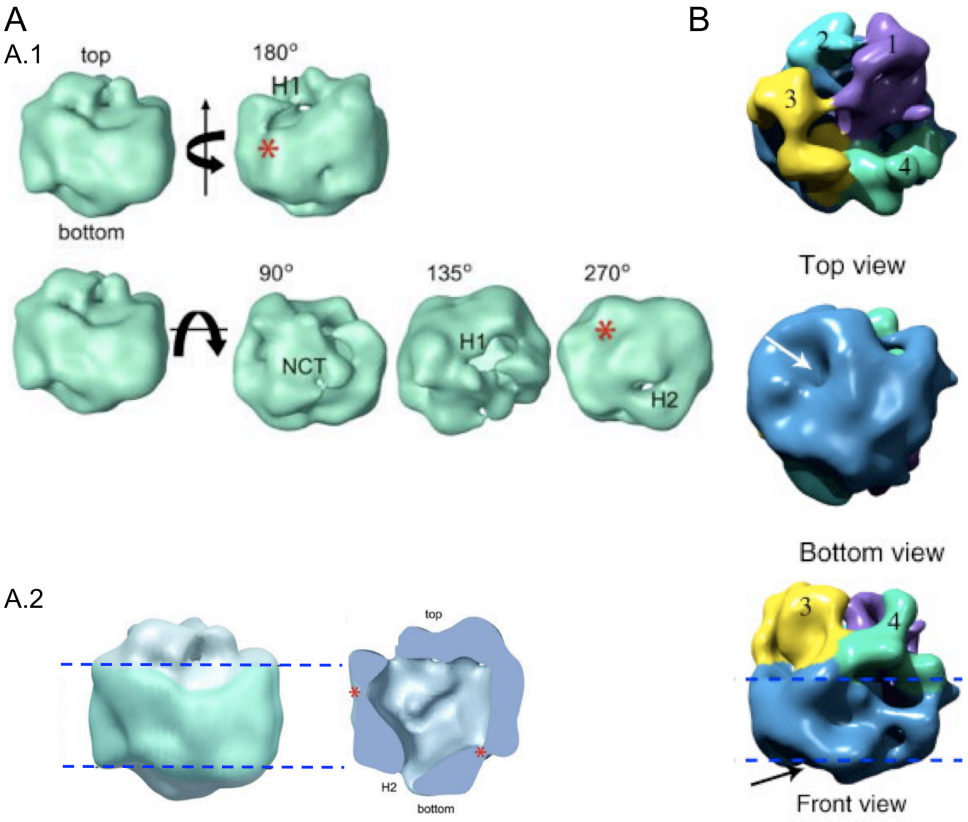

A
A.1
top
180°
H1
bottom
90°
135°
270°
NCT
H1
H2
A.2
top
H2
bottom
B
1
2
3
4
Top view
Bottom view
3
4
Front view

FIGURE 19

Figure 20. Expression and purification of γ-secretase from *P. pastoris*. (A) Samples of cell pellets from induction media BMMY were taken every 24 hours (24, 48 and 72 hrs). In the control lane (Co), cell sample was taken before resuspension in induction media. Anti-PS CTF immunoblot analysis confirmed the heterologous expression of mature γ-secretase in *P. pastoris* as evidenced by the generation of PS CTF (black arrowhead). **(B)** γ-Secretase complex bound to Ni-NTA affinity resin was eluted in buffer containing 8, 12, 16, 20 or 24 mM CHAPSO. Samples from each elution were analyzed by anti-myc immunoblot from which a more predominant band corresponding to APH-1 was detected when the elution buffer contained CHAPSO at a final concentration of 24 mM. **(C)** Samples from the elution (Elu) after Ni-NTA chromatography and the main fraction following SEC (Peak 1) were taken and analyzed by immunoblot against each individual subunit. **(D)** Elution profile of CHAPSO-purified γ-secretase over a TSK-G4000SWxl TOSOH SEC column.

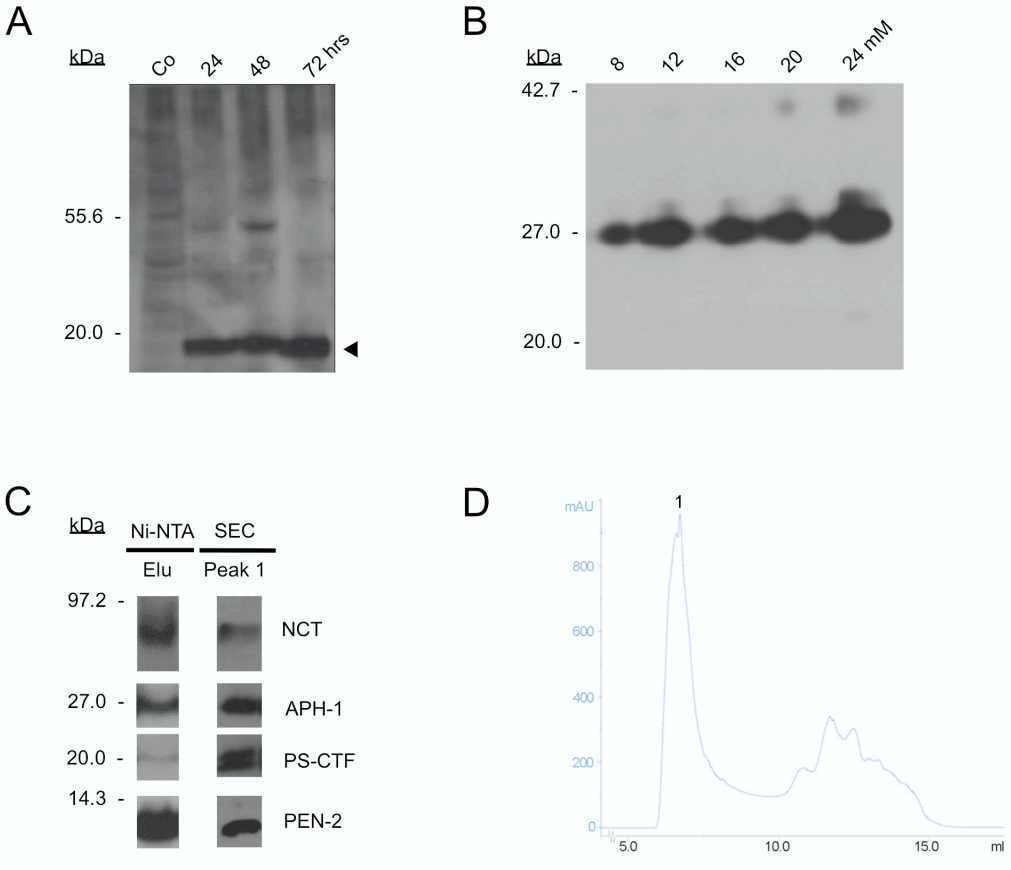

A
kDa
Co
24
48
72 hrs
55.6 -
20.0 -
B
kDa
8
12
16
20
24 mM
42.7 -
27.0 -
20.0 -
C
kDa
Ni-NTA
SEC
Elu
Peak 1
97.2 -
27.0 -
20.0 -
14.3 -
NCT
APH-1
PS-CTF
PEN-2
D
mAU
1
800
600
400
200
0
5.0
10.0
15.0
ml

FIGURE 20

Figure 21. *In vitro* activity assay of γ-secretase-containing fractions. The activity present in the CHAPSO solubilizate, the Ni-NTA eluate and the SEC eluate was measured based on the production of Aβ40 (blue bars). Full inhibition by the γ-secretase transition state analog inhibitor L-685, 458 was achieved demonstrating that the activity displayed is γ-secretase specific (red bars).

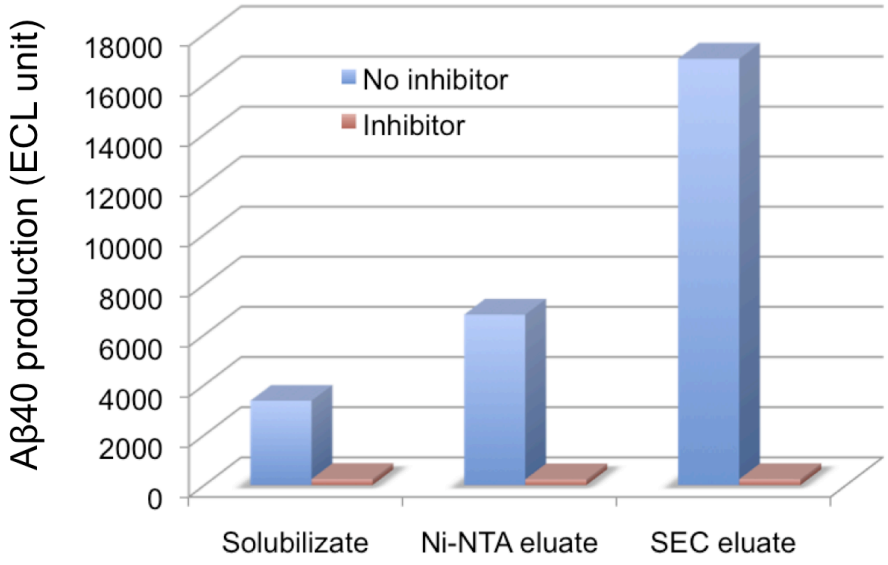

Aβ40 production (ECL unit)
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
No inhibitor
Inhibitor
Solubilizate
Ni-NTA eluate
SEC eluate

FIGURE 21

Figure 22. Small-scale trial of expression of PS. Samples of cell pellets corresponding to *E. coli* transformed with human PS gene cloned into pET15b (lane 1) and pET29b (lane 2) after protein induction were mixed with sample loading buffer and loaded onto a 12% SDS-PAGE. After anti-His immunoblot analysis, a band speculatively corresponding to full length, C-terminally histidine tagged PS cross-reacted at ~50 kDa (arrowhead).

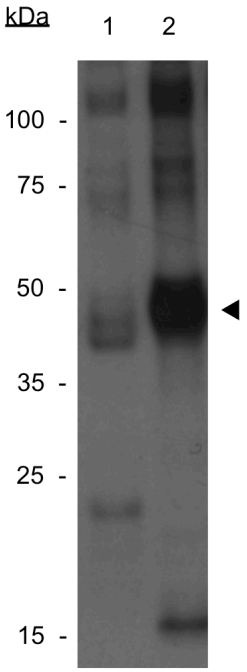


FIGURE 22

Figure 23. Purification of PEN-2 from Rosetta(DE3)pLysS. **(A)** A total of ~2 ug of each of the protein present before induction (lane 1), protein produced after induction with a IPTG (lane 2), in the broken cell suspension (lane 3), in the TX-100 solubilized membranes (lane 4), in the non-solubilized pellet (lane 5), eluted from Ni-NTA resin (lane 6), pooled fractions from ion exchange column (lane 7), and the pooled fractions from SEC (lane 8) were all loaded onto a 12% Tris-Glycine gel for SDS-PAGE and stained with coomassie blue. Black arrowhead indicates an enriched band corresponding to PEN-2 at ~15kDa. **(B)** Anti-His immunoblot analysis confirmed the presence of PEN-2 in the preparation (white arrowhead).

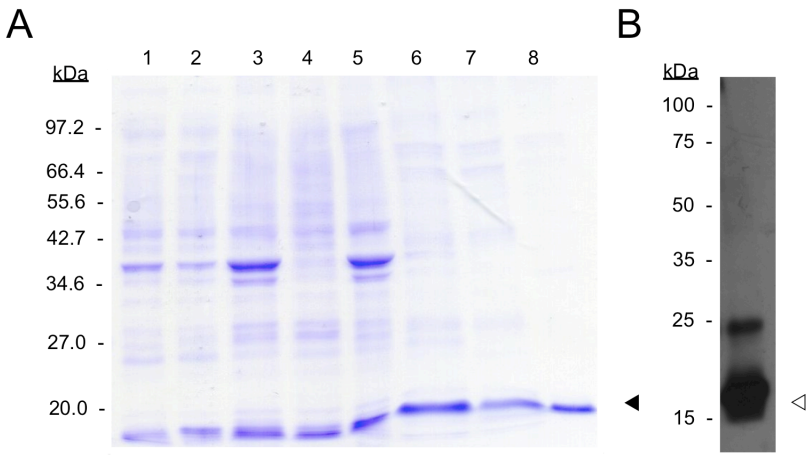

A
1 2 3 4 5 6 7 8
kDa
97.2
66.4
55.6
42.7
34.6
27.0
20.0
B
kDa
100
75
50
35
25
15

FIGURE 23

Figure 24. Partial purification of APH-1 from Rosetta(DE3)pLysS. **(A)** A sample of cell pellet corresponding to *E. coli* strain transformed with human APH-1 gene cloned into pET29b was taken after protein induction. Anti-His immunoblot analysis confirmed the expression of full-length APH-1 as shown by a band that cross-reacted at the corresponding molecular weight (black arrowhead). A band likely to correspond to a proteolytic C-terminal fragment from APH-1 also cross-reacted at <25kDa (white arrowhead). **(B)** From a small-scale expression trial of APH-1 induced at 20°C, samples from the detergent solubilizates were taken and analyzed by anti-C-terminus APH-1 immunoblot. The generation of a proteolytic fragment derived from full-length APH-1 (black arrowhead) was abolished upon induction at low temperatures as shown by the lack of cross-reactivity at low molecular weights. In addition, membrane solubilization using DDM (lane 2) instead of TX-100 (lane 1) increased the yield of chimeric APH-1 as shown by a more intense band that corresponded to full-length APH-1 (black arrowhead).

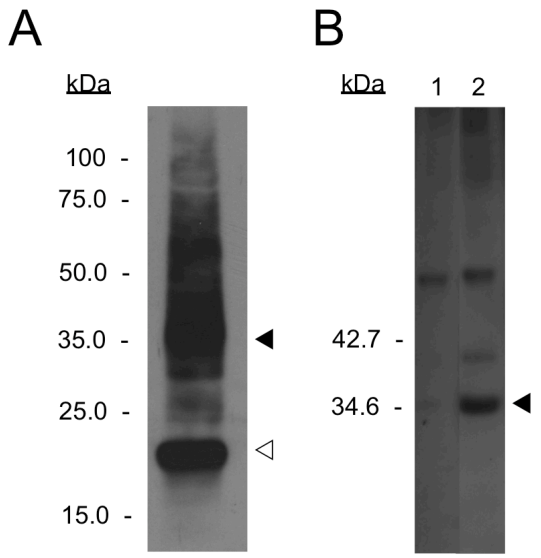


FIGURE 24

Figure 25. Identification of PSΔexon9 by mass spectrometry. (A) PSΔexon9 eluted from the M2 affinity resin corresponded to the predominant band detected in the coomassie-stained gel (black arrowhead). **(B)** Two peptides (boxed) located outside the predicted TMDs regions (underlined) of the PSΔexon9 sequence were identified by mass spectrometry from the excised band in the gel shown in (A). The trypsin cleavage sites are in red.


A
B
kDa
97.2
66.4
55.6
42.7
34.6
27.0
20.0
MTELPAPLSYFQNAQMSEDNHLSNTVRSQNDNRERQEH
NDRRSLGHPEPLSNGRPQGNSRQVVEQDEEEDEELTLK
YGAKHVIMLFVPVTLCMVVVVATIKSVSFYTRKDGQLIYT
PFTEDTETVGQRALHSILNAAIMISVIVVMTILLVVLYKYRC
YKVIHAWLIISSLLLLFFFSFIYLGEVFKTYNVAVDYITVAL
LIWNFGVVGMISIHWKGPLRLQQAYLIMISALMALVFIKYL
PEWTAWLIXAVISVYDLVAVLCPKGPLRMLVETAQERNET
LFPALIYSCTERESQDTVAENDDGGFSEEWEAQRDSHLG
PHRSTPESRAAVQELSSSILAGEDPEERGVKLGLGDFIFY
SVLVGKASATASGDWNTTIACFVAILIGLCLTLLLLAIFKKA
LPALPISITFGLVFYFATDYLVQPFMDQLAFHQFYIGTLVP
RGSMAISDPNSSSVDKLDYKDDDDKLEHHHHHH

FIGURE 25

Figure 26. Expression screening and small-scale production of putative di-aspartyl proteases. PS- (panel A) and TFPP-like genes (panel B) were amplified from 12 and 9 different archaeal genomes, respectively and protein expression screened. Proteins bearing either an N- or a C-terminal 10xHis tag were purified from a 100 ml bacterial culture by metal affinity chromatography. Purification was performed in either DDM or FC-12. Arrows show expression of tagged MHJF1 (A, lane 6), MCMJR1 (A, lane 8), MCV (B, lane 3) and MMS2 (B, lane 4), all four chosen for large-scale experiments. Small-scale production of sumo tagged MHJF1 and MCMJR1 yielded comparable results to those shown in panel A. Archaeal genomes from which PS-like genes were amplified were *Haloarcula marismortui* (lane 1), *Methanosarcina mazei* (lane 2), *Archaeoglobus fulgidus* (lane 3), *Ferroplasma acidarmanus* (lane 4), *Picrophilus torridus* (lane 5), *Methanospirillum hungatei* (MHJF1; lane 6), *Thermoplasma volcanium* (lane 7), *Methanoculleus marisnigri* (MCMJR1; lane 8), *Thermoplasma acidophilum* (lane 9), *Methanosarcina barkeri* (lane 10), *Methanococcoides burtonii* (lane 11) and *Methanosarcina acetivorans* (lane 12). TFPP-like genes were amplified from *Thermofilum pendens* Hrk 5 (lane 1), *Archaeoglobus fulgidus* DSM 4304 (lane 2), *Methanococcus voltae* (lane 3), *Methanococcus maripaludis* S2 (lane 4), *Sulfolobus solfataricus* P2 (lane 5), *Methanospirillum hungatei* JF-1 (lane 6), *Methanoculleus marisnigri* JR1 (lane 7) and finally two genes were amplified from *Methanosarcina acetivorans* C2A genome (lanes 8 and 9).

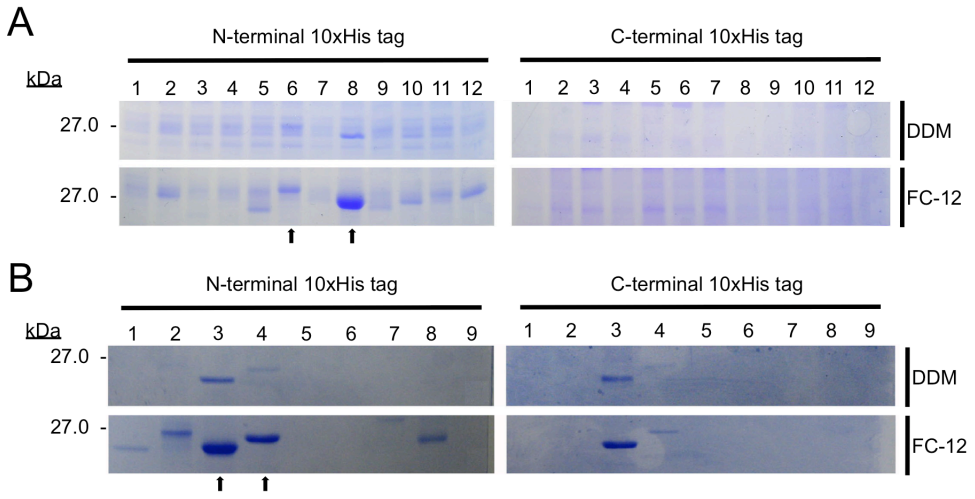

A
N-terminal 10xHis tag
C-terminal 10xHis tag
kDa
1 2 3 4 5 6 7 8 9 10 11 12
1 2 3 4 5 6 7 8 9 10 11 12
27.0
27.0
DDM
FC-12
B
N-terminal 10xHis tag
C-terminal 10xHis tag
kDa
1 2 3 4 5 6 7 8 9
1 2 3 4 5 6 7 8 9
27.0
27.0
DDM
FC-12

FIGURE 26

Figure 27. Large-scale production of putative di-aspartyl proteases. 6xHisSUMOMHJF, 6xHisSUMOMCMJR1, 10xHisMCV and 10xHisMMS2 were purified from DDM-solubilized membranes using Ni-NTA resin. After treating 6xHisSUMOMHJF and 6xHisSUMOMCMJR1 with SUMO protease and re-passing the preparation through Ni-NTA, a total of 1-2 ug of MHJF (lane 1) and MCMJR1 (lane 2) were loaded onto a 12% Tris-glycine SDS-PAGE and stained with coomassie blue. Eluted 10xHisMCV and 10xHisMMS2 from the Ni-NTA resin were dialyzed against PBS buffer and an equivalent to 1-2 ug of protein were loaded in lanes 3 and 4, respectively. Proteins purified in FC-12 yielded comparable results.

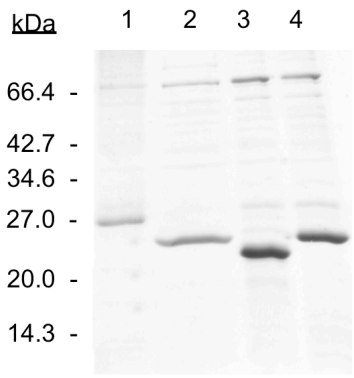


FIGURE 27

Figure 28. *In vitro* activity exhibited by wild-type MCMJR1. An aliquot equivalent to 4 ng of chimeric substrate was taken from the negative control mixture (lane 1) and from the reaction mixtures of MBPSpitz, MBPKeren and MBPGurken incubated with DDM-purified MCMJR1 (lane 2), MHJF1 (lane 3), 10xHisMCV (lane 4) and 10xHisMMS2 (lane 5), loaded onto a 10% SDS-PAGE and analyzed by anti-MBP immunoblot. The generation of a N-terminal product (white arrowhead) when incubated the chimeric substrate with MCMJR1 was distinctively detected from the full-length chimeric substrate (black arrowhead). Incubation with DDM-purified MJFH1, 10xHisMCV, 10xHisMMS2 failed to generate a N-terminal product under the conditions tested.

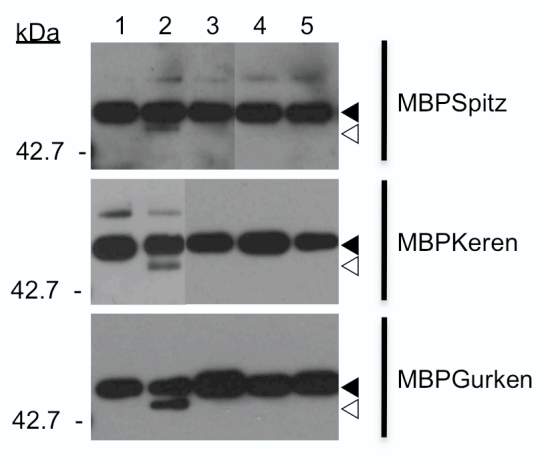
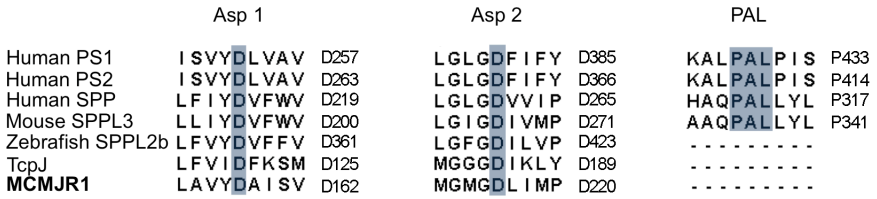


FIGURE 28

Figure 29. Sequence alignment and topology comparison between MCMJR1 and PS. (A) Sequence alignment as shown on figure 8 including now MCMJR1 (bold). **(B)** Cylinders depict the predicted TMD regions of Presenilin (PS) and MCMJR1. Stars show the predicted positions of the catalytic aspartate residues and for each, the surrounding amino acid residues are also indicated. A continuous slab depicts the membrane.

A


Asp 1
Asp 2
PAL
Human PS1 ISVYDLVAV D257 LGLGDFIFY D385 KALPALPIS P433
Human PS2 ISVYDLVAV D263 LGLGDFIFY D366 KALPALPIS P414
Human SPP LFIYDVFWV D219 LGLGDVVIP D265 HAQPALLYL P317
Mouse SPPL3 LLIYDVFWV D200 LGIGDIVMP D271 AAQPALLYL P341
Zebrafish SPPL2b LFVYDVFFV D361 LGFGDILVP D423 - - - - - - - - -
TcpJ LFVIDFKSM D125 MGGGDIKLY D189 - - - - - - - - -
MCMJR1 LAVYDAISV D162 MGMGDLIMP D220 - - - - - - - - -

B

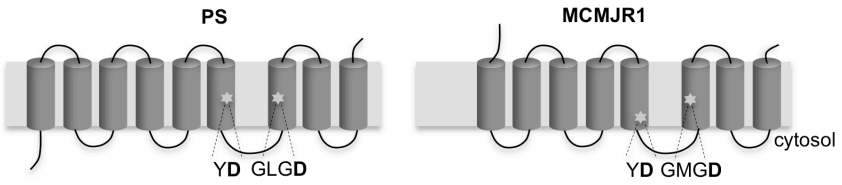

PS
MCMJR1
cytosol
YD GLGD
YD GMGD

FIGURE 29

Figure 30. Membrane topology model of MCMJR1. A prediction of the membrane-spanning regions and their orientations was obtained from the amino acid sequence of MCMJR1 using the TMpred analysis tool on the ExPASy server. Horizontal lines delimit the lipid bilayer. All aspartic amino acid residues (D5, D40, D128, D162, D195, D220 and D236) in MCMJR1 sequence are depicted as dark circles. Amino acid residues flanking the candidate catalytic aspartic acid residues are also shown.

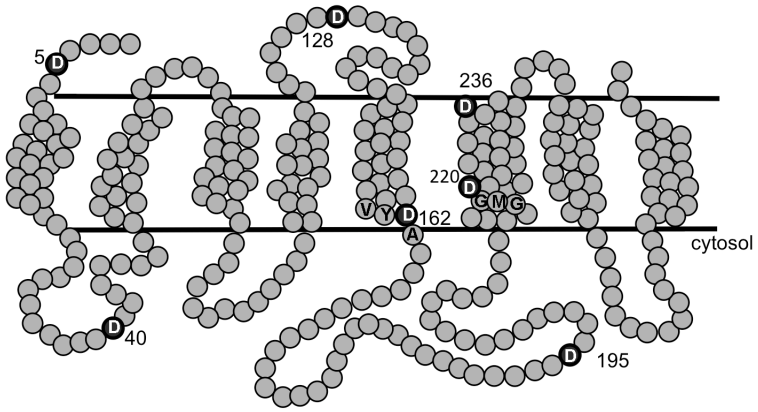


FIGURE 30

Figure 31. Identification of catalytic amino acid residues in MCMJR1. (A) Aliquots (~1.5 ug of MBP substrate per well) from the reaction mixtures of MBPGurken with wild-type (wt) MCMJR1 and aspartate-to-alanine MCMJR1 mutant variants were analyzed by PAGE. In the MCMJR1-negative control lane (MBPGurken), an aliquot from the incubation mixture of MBPGurken in PBS buffer was loaded. Additionally, an aliquot from the incubation mixture of MCMJR1 in PBS buffer was also loaded (lane MCMJR1). Bands at ~75 and ~40 kDa are minor contaminants present in the MCMJR1 preparation. N-terminal products (white arrowhead) after cleavage by wild-type MCMJR1 and D5A, D40A, D128A, D195A and D236A MCMJR1 mutant variants were detected. Full-length MBPGurken (black arrowhead) remained uncleaved after incubation with D162A and D220A MCMJR1 mutant variants (stars). Analysis of similar reaction mixtures with MBPKeren yielded comparable results. **(B)** Reaction mixtures of MBPGurken with wt MCMJR1 and MCMJR1 mutant variants were incubated with 10 ul of pre-equilibrated Ni-NTA resin. An aliquot equivalent to 4 ng of chimeric substrate was taken from the flow through and analyzed by anti-MBP immunoblot. The generation of a N-terminal product by wt MCMJR1, D5A, D40A, D128A, D195A and D236A MCMJR1 mutant variants was confirmed by the band that cross-reacted at ~46 kDa (white arrowhead). Residual undigested chimeric substrate is still detectable after the incubation with Ni-NTA resin of the reaction mixtures with D236A, D220A and D162A mutant variants, being more evident in the last two (black arrowhead).

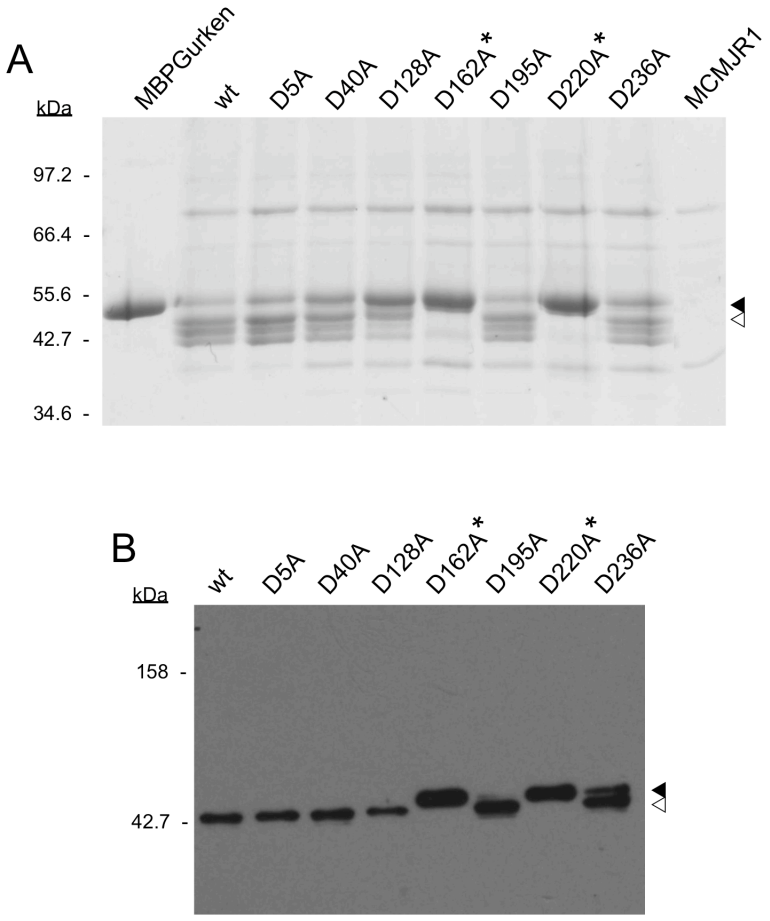

A
MBPGurken
wt
D5A
D40A
D128A
D162A*
D195A
D220A*
D236A
MCMJR1
kDa
97.2
66.4
55.6
42.7
34.6
B
wt
D5A
D40A
D128A
D162A*
D195A
D220A*
D236A
kDa
158
42.7

FIGURE 31

Figure 32. Site-directed mutagenesis of the catalytic aspartates in MCMJR1 Aliquots (~4 ng of MBP substrate per well) from the reaction mixtures of MBPGurken with wild-type (wt) MCMJR1, aspartate-to-asparagine (D162N and D220N) and aspartate-to-glutamate (D162E and D220E) MCMJR1 mutant variants were analyzed by anti-MBP immunoblot. In the MCMJR1-negative control lane (MBPGurken), an aliquot from the incubation mixture of MBPGurken in PBS buffer was loaded. N-terminal product (white arrowhead) was detected only after incubation with wt MCMJR1. Full-length MBPGurken (black arrowhead) remained uncleaved after incubation with D162N, D162E, D220N and D220E MCMJR1 mutant variants. Analysis of similar reaction mixtures with MBPKeren yielded comparable results.

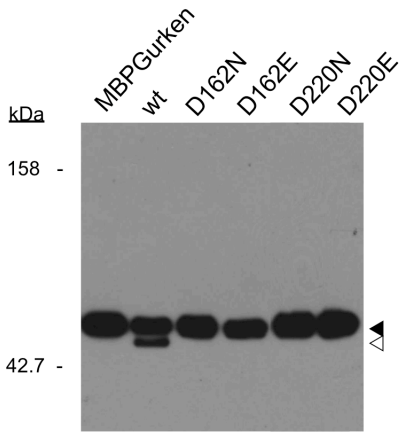


FIGURE 32

Figure 33. Steps of isolation of the N-terminal product for MS analysis. (A) Schematic of the processing of MBP chimeric substrate by MCMJR1 and generation of samples for MS analysis. Wider cylinder represents the TMD region. **(B)** Aliquots equivalent to ~1.5 ug of MBPGurken after incubation in PBS buffer (lane 1), after incubation with MCMJR1 (lane 2), after removal of uncleaved full-length MBPGurken from the reaction mixture using Ni NTA resin (lane 3) and after treatment of the purified N-terminal product with thrombin (lane 4) were analyzed by PAGE. Full-length MBPGurken (black arrowhead) and N-terminal product (white arrowhead) generated after cleavage of MBPGurken by wild-type MCMJR1 are indicated. The band at ~42kDa in lane 4 corresponds to the MBP moiety released from the N-terminal product after cleavage by thrombin.

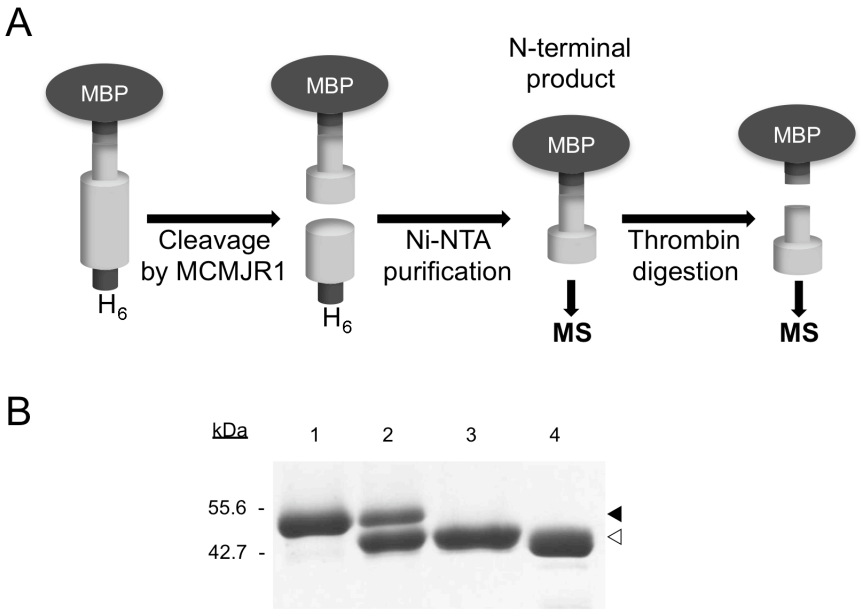

A
MBP
H6
Cleavage
by MCMJR1
MBP
H6
Ni-NTA
purification
N-terminal
product
MBP
MS
Thrombin
digestion
MBP
MS
B
kDa
1
2
3
4
55.6 -
42.7 -

FIGURE 33

Figure 34. Intramembrane cleavage of chimeric MBP substrates. (A, B) MALDI-TOF MS analyses of control (panel A) and TT samples (panel B) isolated from the reaction mixture of MCMJR1 with MBPGurken. Observed masses (Da) are shown for the most representative peaks. Singly, doubly, triple and quadruple charged species are indicated for the control sample. The peptides between the thrombin cleavage site and the MCMJR1 cleavage site are simplified in brackets for the TT sample. **(C)** Amino acid sequence of MBP chimeric substrates. Amino acid residues within the predicted soluble region of wild type Gurken and Keren are colored gray and the TMD region is underlined. The thrombin cleavage site following the MBP tag is also shown. Arrows indicate the cleavage positions along with the calculated (calc) and observed (obs) masses in daltons taken from the MS analysis carried out for the TT sample.

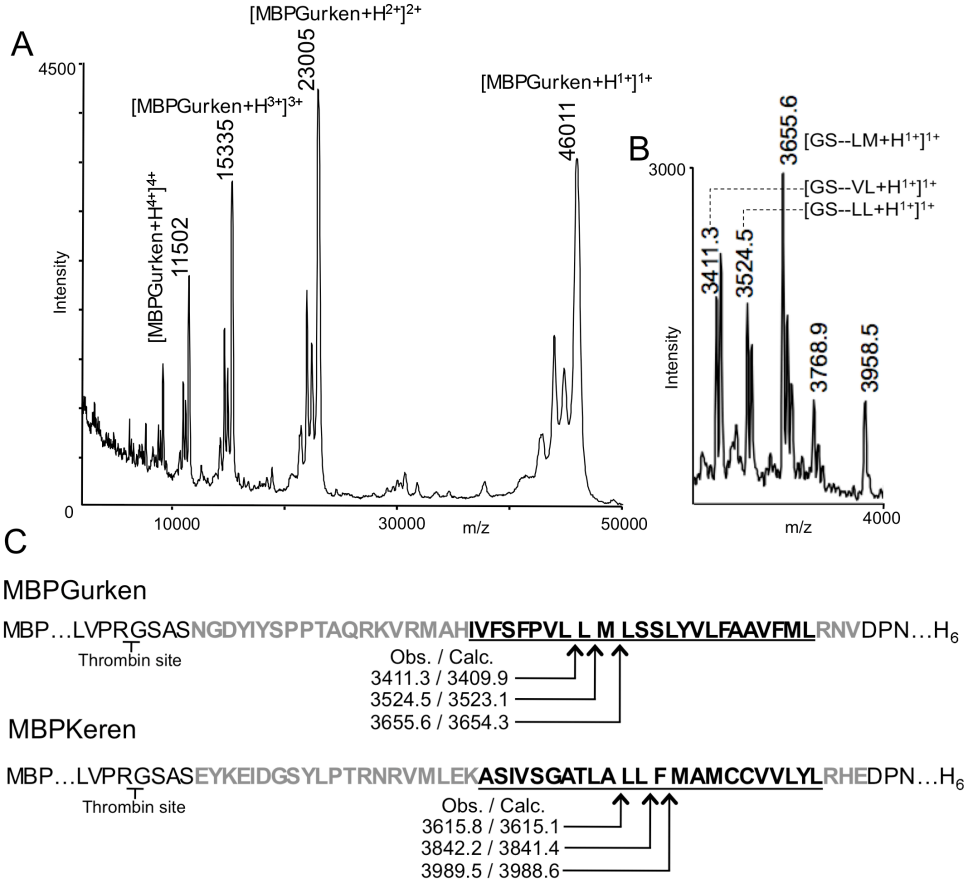

A
[MBPGurken+H2+]2+
23005
[MBPGurken+H3+]3+
15335
[MBPGurken+H1+]1+
46011
[MBPGurken+H4+]4+
11502
4500
Intensity
0
10000
30000
m/z
50000
B
3655.6
[GS--LM+H1+]1+
[GS--VL+H1+]1+
[GS--LL+H1+]1+
3411.3
3524.5
3768.9
3958.5
3000
Intensity
m/z
4000
C
MBPGurken
MBP…LVPRGSASNGDYIYSPPTAQRKVRMAHIVFSFPVL L M LSSLYVLFAAVFMLRNVDPN…H6
Thrombin site
Obs. / Calc.
3411.3 / 3409.9
3524.5 / 3523.1
3655.6 / 3654.3
MBPKeren
MBP…LVPRGSASEYKEIDGSYLPTRNRVMLEKASIVSGATLA LL F MAMCCVVLYLRHEDPN…H6
Thrombin site
Obs. / Calc.
3615.8 / 3615.1
3842.2 / 3841.4
3989.5 / 3988.6

FIGURE 34

Figure 35. MCMJR1 cleaves γ-secretase chimeric substrate C100Flag. (A) Amino acid sequence of C100Flag chimeric substrate. The amino acid residues within the predicted soluble region of βCTF are colored gray and the TMD region is underlined. Ellipsis (…) simplifies a number of amino acid residues. The flag tag (FLAG) is indicated as linked to the C-terminus. Calculated molecular weights corresponding to fragments generated after cleavage of C100Flag approximately close to the middle of the TMD are also shown. **(B)** Aliquots equivalent to 1 ug of C100Flag alone (lane 1) and from the reaction mixture of C100Flag with DDM-purified MCMJR1 (lane 2) were loaded onto a 15% Tris-Glycine gel SDS-PAGE. In addition, MCMJR1 incubated under the same conditions as the reaction mixture was also loaded (Co). A C-terminal product of approximately 8 kDa (white arrowhead) was generated after incubation of C100Flag with MCMJR1 and distinctively detected from the full-length chimeric substrate (black arrowhead) **(C)**. Anti-Flag immunoblot analysis confirmed the latter. A band corresponding to C100Flag dimmer also cross-reacted at <27 kDa.

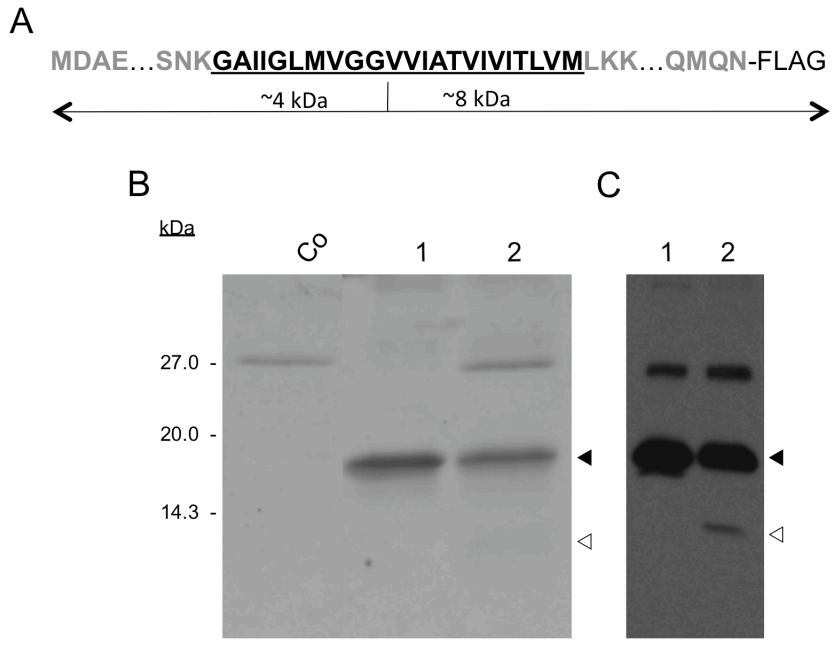

A
MDAE…SNKGAIIGLMVGGVVIATVIVITLVMLKK…QMQN-FLAG
~4 kDa
~8 kDa
B
C
kDa
Co
1
2
1
2
27.0 -
20.0 -
14.3 -

FIGURE 35

Figure 36. MCMJR1 cleaves chimeric MBPAPP within the TMD region. (A) Aliquots equivalent to 1 ug of MBPAPP from the incubation corresponding to MBPAPP alone (lane 1) and from the reaction mixture of MBPAPP with DDM-purified MCMJR1 (lane 2) were loaded onto a 10% Tris-Glycine SDS-PAGE. The N-terminal product (white arrowhead) generated after processing by MCMJR1 was distinctively detected from the full-length chimeric substrate (black arrowhead) and its identity confirmed by anti-MBP immunoblot (lane 3). **(B)** MALDI-TOF MS analysis of control sample from the reaction mixture of MCMJR1 with MBPAPP. Observed masses (Da) are shown for the most representative peaks. Singly, doubly, triple and quadruple charged species are indicated. **(C)** Amino acid sequence of MBPAPP substrate. Amino acid residues corresponding to the TMD region of APP are underlined. The thrombin cleavage site following the MBP tag is also shown. Arrows indicate the cleavage positions along with the calculated (calc) and observed (obs) masses in daltons taken from the MS analysis carried out for the control sample.

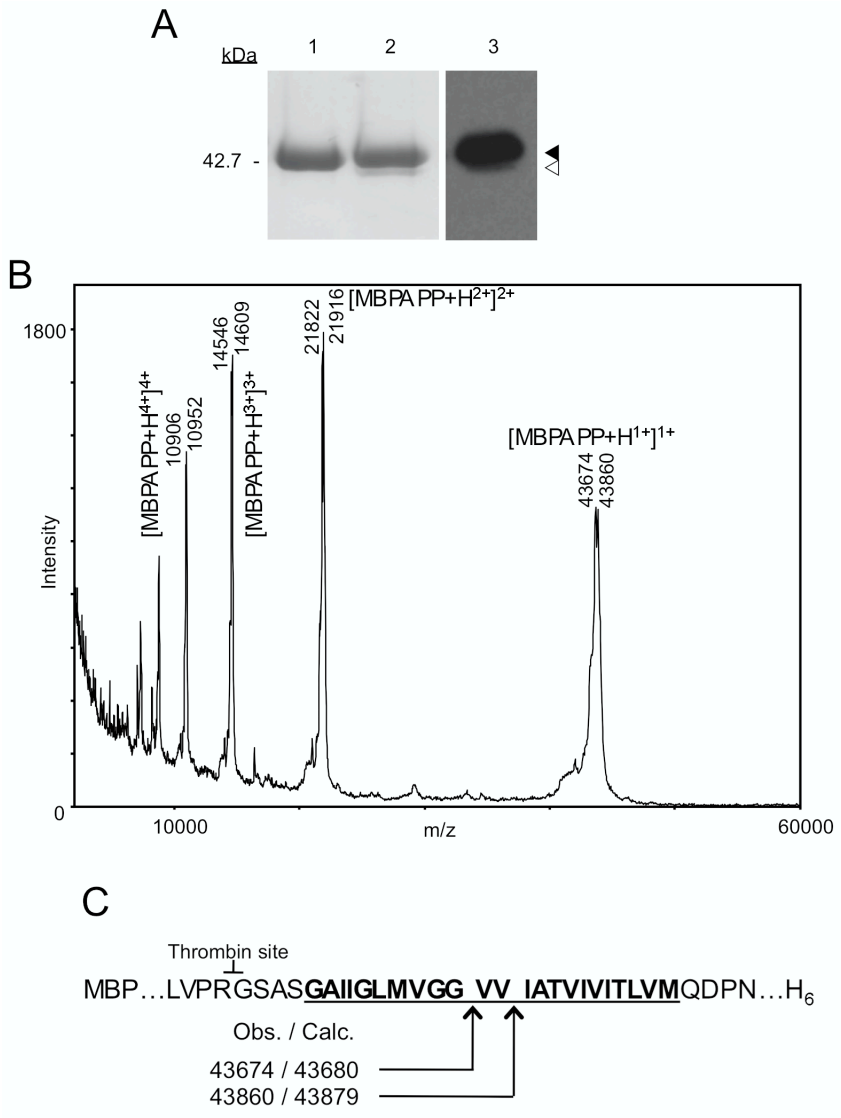

A
kDa
1
2
3
42.7 -
B
1800
0
Intensity
10000
m/z
60000
[MBPAPP+H4+]4+
10906
10952
14546
14609
[MBPAPP+H3+]3+
21822
21916
[MBPAPP+H2+]2+
[MBPAPP+H1+]1+
43674
43860
C
Thrombin site
MBP…LVPRGSASGAIIGLMVGG VV IATVIVITLVMQDPN…H6
Obs. / Calc.
43674 / 43680
43860 / 43879

FIGURE 36

Figure 37. Chemical structures of inhibitors. The structures of common serine (PMSF) and cysteine (E-64) protease inhibitors and that of the aminopeptidase bestatin are shown. Inhibitors of aspartyl proteases based on the isosteres statine (pepstatin), hydroxyethylene (L-685,458) and hydroxyethylamine (31C), the last two known as transition-state analogs that specifically inhibit γ-secretase activity, are also given. Added to this list are two other highly potent γ-secretase inhibitors, non-transition-state analogs, Compound E and DAPT. Isostere moieties are boxed.

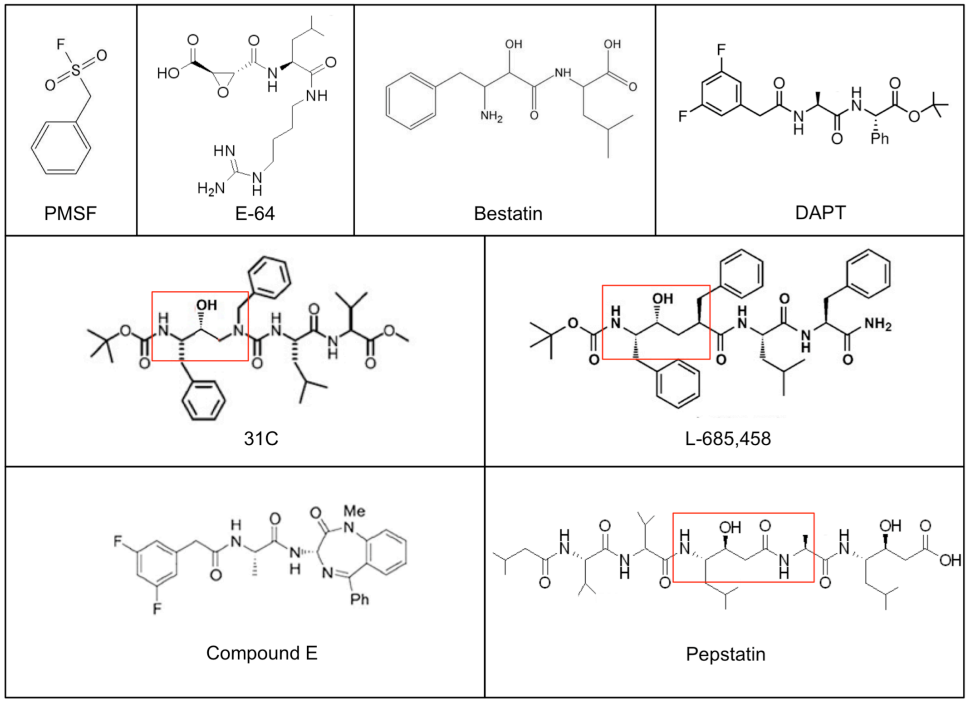


FIGURE 37

Figure 38. 31C abrogates the generation of the N-terminal product. Aliquots from the reaction mixtures of MBPGurken with DDM-purified MCMJR1 pre-incubated over night at 37 °C with L-685,458 (20 uM), DAPT (235 uM), 31C (100 uM), compound E (235 uM) and pepstatin A (235 uM) were analyzed by anti-MBP immunoblot. In the lane corresponding to MCMJR1-negative control (Co-), an aliquot from the incubation after 8 (panel A) and 24 hours (panel B) at 37 °C of MBPGurken in NaHepes buffer was loaded. The N-terminal product (white arrowhead) generated after cleavage by MCMJR1 (Co+) and MCMJR1 pre-incubated with DMSO (Co_{DMSO}) can be distinctively detected from the full-length MBPGurken (black arrowhead). Generation of the N-terminal product was also observed in the reaction mixtures where MCMJR1 was pre-incubated with L-685,458, DAPT, Compound E and pepstatin A. MCMJR1 pre-incubated with 31C failed to generate the N-terminal product after 8 hours (panel A) although a very faint band was detected after 24 hours (panel B) of the beginning of the reaction.

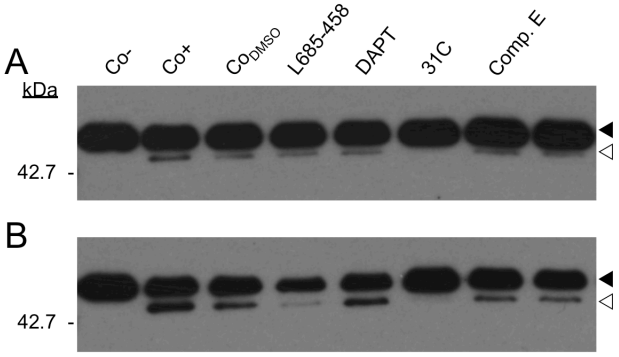


FIGURE 38

Figure 39. 31C inhibits the activity of MCMJR1 in a concentration-dependant manner. Aliquots were taken from the reaction mixtures of MBPGurken with DDM-purified MCMJR1 incubated over night at 37 °C with 31C (panel A) at final concentrations of 10, 50, 75 and 100 uM and with L-685,458 (panel B) at final concentrations of 10, 25, 50, 75, 100 and 200 uM and analyzed by anti-MBP immunoblot. In the lane corresponding to MCMJR1-negative control (Co-), an aliquot from the incubation after 24 hours (panels A and B) at 37 °C of MBPGurken in NaHepes buffer was loaded. The N-terminal product (white arrowhead) generated after cleavage by MCMJR1 pre-incubated with 31C (10uM; panel A) and with L-685,458 (panel B) can be distinctively detected from the full-length MBPGurken (black arrowhead). At higher concentrations of inhibitors, the capacity of MCMJR1 to generate the N-terminal product was affected by 31C as shown by the weak band that cross-reacted at the corresponding molecular weight.

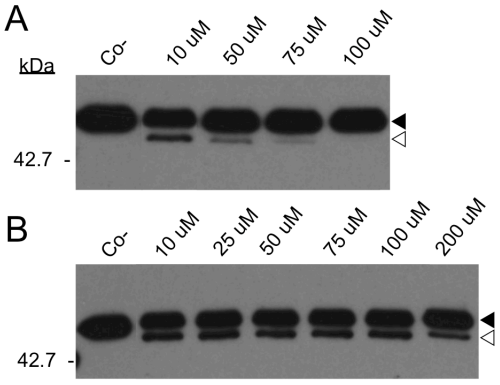

A
kDa
Co-
10 uM
50 uM
75 uM
100 uM
42.7
B
Co-
10 uM
25 uM
50 uM
75 uM
100 uM
200 uM
42.7

FIGURE 39

Figure 40. Slowed intramembrane proteolysis of G129A MCMJR1 mutant variant. Aliquots were taken from the reaction mixtures of MBPGurken with DDM-purified MCMJR1 wild-type (wt) and G219A MCMJR1 mutant variant (G219A) after 30 min, 2, 3, 4, 6, 8 and 24 hrs and analyzed by anti-MBP immunoblot. In the lane corresponding to MCMJR1-negative control (Co-), an aliquot from the incubation after 24 hours at 37 °C of MBPGurken in NaHepes buffer was loaded. The N-terminal product (white arrowhead) generated after digestion of MBPGurken by wild-type MCMJR1 can be distinctively detected from the full-length MBPGurken (black arrowhead) after 2 hours. Instead, the generation of the N-terminal product by MCMJR1 G219A occurred after 8 hours.

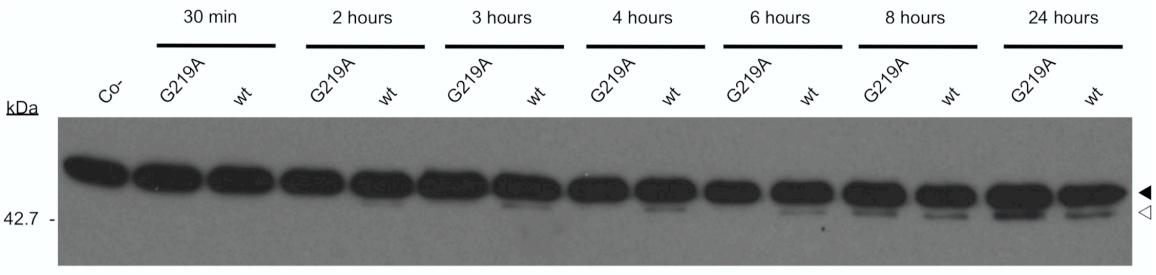

30 min
2 hours
3 hours
4 hours
6 hours
8 hours
24 hours
Co-
G219A
wt
kDa
42.7 -

FIGURE 40

Figure 41. Improved substrate specificity by MCMJR1 towards MBPSpitz mutant variant (A) Amino acid sequence of MBPKeren and MBPSpitz. Amino acid residues within the predicted soluble region of wild type Keren and Spitz are colored gray and those within the TMD region are in bold. Dotted lines indicate MCMJR1 cleavage positions in MBPKeren. In MBPSpitz, smaller fonts indicate the residues that were substituted for the amino acid situated right above in caps. Numbering shown according to the full-length sequence of wild-type Spitz. **(B)** The capacity of MCMJR1 to generate an N-terminal product (white arrowhead) from MBPKeren and MBPSpitz (Co+) was compared to that when incubated with M149T/C150L/V153L (3x149) and L156A/F157M/V158C (3x156) MBPSpitz triple mutant variants. Also loaded were MPBKeren and MBPSpitz alone, incubated under the same conditions as the reaction mixture (Co-). **(C)** The cleavage capacity of MCMJR1 towards M149T/C150L (2x149/150), M149T/V153L (2x149/153) and C150L/V153L (2x1150/153) MBPSpitz double mutant variants was compared to that towards wild-type MBPSpitz (wt) and M149T/C150L/V153L (3x149) MBPSpitz triple mutant.

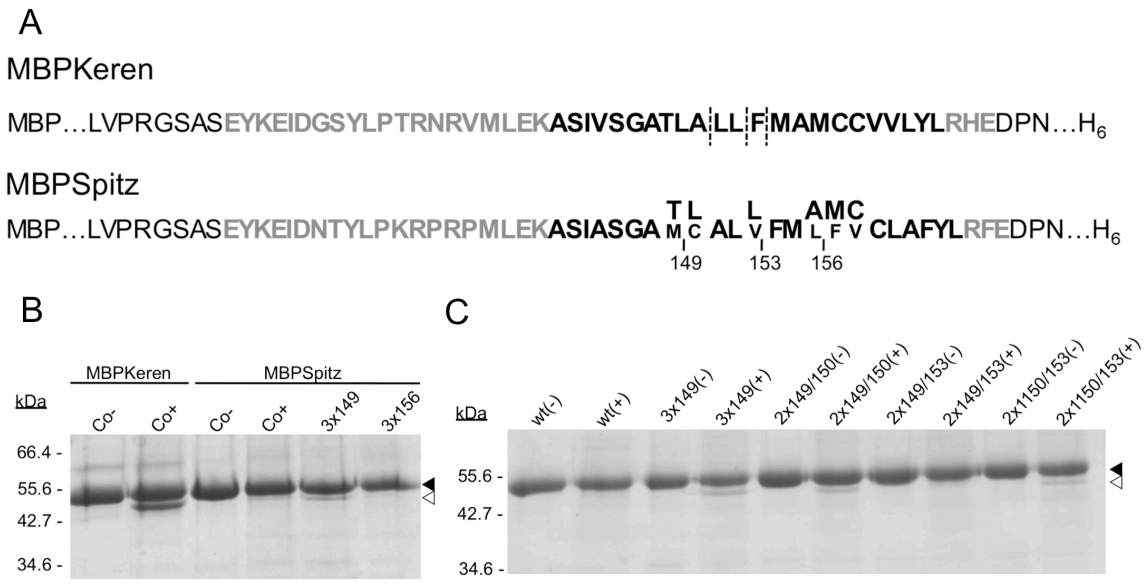

A
MBPKeren
MBP…LVPRGSASEYKEIDGSYLPTRNRVMLEKASIVSGATLALLFMAMCCVVLYLRHEDPN…H6
MBPSpitz
MBP…LVPRGSASEYKEIDNTYLPKRPRPMLEKASIASGAM/TIC/LALV/LFM/AL/MC/FV CLAFYLRFEDPN…H6
149
153
156
B
MBPKeren
MBPSpitz
kDa
Co-
Co+
Co-
Co+
3x149
3x156
66.4 -
55.6 -
42.7 -
34.6 -
C
kDa
wt(-)
wt(+)
3x149(-)
3x149(+)
2x149/150(-)
2x149/150(+)
2x149/153(-)
2x149/153(+)
2x1150/153(-)
2x1150/153(+)
55.6 -
42.7 -
34.6 -

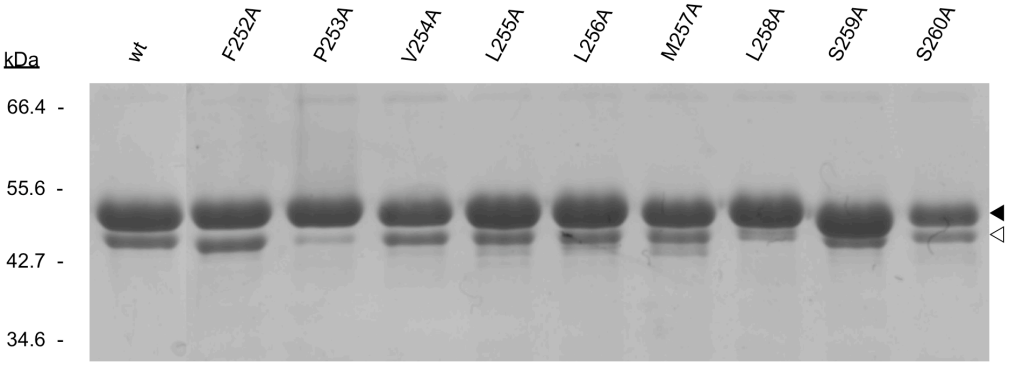
FIGURE 41

Figure 42. Insights into peptide bond selectivity of MCMJR1 towards MBPGurken. **(A)** Amino acid sequence of MBPGurken. Amino acid residues within the predicted soluble region of wild-type Gurken are colored gray and the TMD is underlined. Amino acid residues flanking the cleavage site and that were exchanged to alanine residues are boxed. Dotted lines indicate MCMJR1 cleavage positions. **(B)** Aliquots equivalent to ~1.5 ug of MBP substrate were taken from the reaction mixtures of DDM-purified MCMJR1 with wild-type (wt) MBPGurken and X-to-A MBPGurken mutant variants (where X represents amino acid residues 252 to 260 in the full-length sequence of wild-type Gurken) and analyzed by PAGE. From among all N-terminal products (white arrowhead), the intensity corresponding to that generated after processing of P253A MBPGurken mutant variant by MCMJR1 was distinctively lower. Undigested full-length MBP chimeric substrate is also shown (black arrowhead). **(C)** Amino acid sequence of the TMD region of MBPGurken. A smaller font was assigned for M257 to depict its substitution for an alanine residue (in bold). MS analysis revealed the abrogation of the cleavage at the position crossed when the wild type residue was mutated to alanine.

A

MBP…LVPRGSASNGDYIYSPPTAQRKVRMAHIVFSFPVLLMLSSLYVLFAAVFMLRNVDPN…H_6

B



C

Figure 43. Cleavage sites identified for MCMJR1 and PS/γ-secretase. Depicted are the transmembrane regions of chimeric MBP substrates and APP. Arrows indicate the cleavage sites identified for MCMJR1 and PS/γ-secretase. For the latter, γ sites that generate Aβ38, Aβ40, Aβ42, Aβ46 and Aβ49 (numbered accordingly) are also shown.

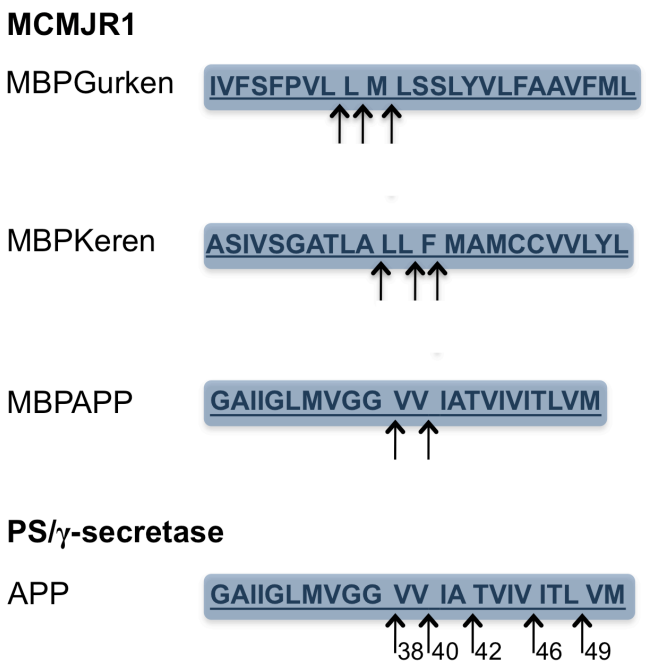

MCMJR1
MBPGurken
IVFSFPVL L M LSSLYVLFAAVFML
MBPKeren
ASIVSGATLA LL F MAMCCVVLYL
MBPAPP
GAIIGLMVGG VV IATVIVITLVM
PS/γ-secretase
APP
GAIIGLMVGG VV IA TVIV ITL VM
38 40 42 46 49

FIGURE 43

Figure 44. Hypothetical scheme of the active site of MCMJR1. Active site pockets (arcs) are shown binding to the inhibitor 31C. Hydrogen bond interactions between the carboxylic acid moieties of the two catalytic aspartate residues (D162 and D220), the critical hydroxyl group and the isostere amine are also indicated (dotted lines).

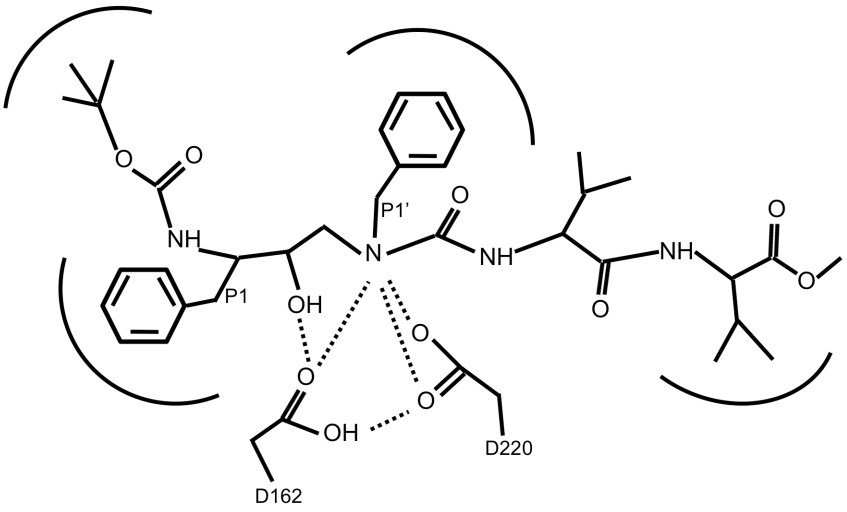


FIGURE 44

Figure 45. Effect of loop truncations on MCMJR1 activity. A coomassie stained 12% SDS-PAGE analysis (left) of purified MCMJR1 loop deletion variants Δ169-203 (the amino acids 169-203 have been deleted), Δ169-217, Δ176-196 and Δ176-203. The wild-type enzyme (wt) was included for comparison. The right panel shows a coomassie stained 10% SDS-PAGE analysis of the incubations of the loop deletion variants with MBPGurken. Protein bands corresponding to the undigested (black arrowhead) and digested (white arrowhead) substrates are indicated on the right and the molecular weight marker positions are shown on the left side.

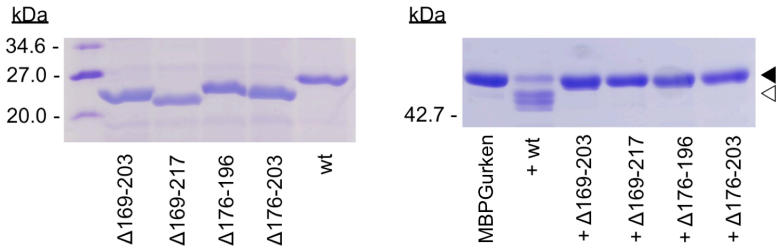

kDa
34.6 -
27.0 -
20.0 -
Δ169-203
Δ169-217
Δ176-196
Δ176-203
wt
kDa
42.7 -
MBPGurken
+ wt
+ Δ169-203
+ Δ169-217
+ Δ176-196
+ Δ176-203
◀
△

FIGURE 45

Figure 46. Multiple sequence alignment of human presenilin, MCMJR1 and putative archaeal homologs of MCMJR1. Partial alignment of the sequences is represented using the ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2/) based on the predicted TM topology for MCMJR1. The positions of three of the predicted 8 TMD of MCMJR1 are shown (TMD5, TMD 6 and TMD8) including the large hydrophilic loop (LOOP). Aminoacid positions indicated with a star (*) denote their potential for mutagenesis studies due to their conservancy. Catalytic aspartic acid residues in presenilin (PS) and MCMJR1 within the YD and GXGD motifs are also shown (+). The UniProtKB accession number of the sequences corresponding to the putative archaeal homologs of MCMJR1 are shown preceding the alignment.

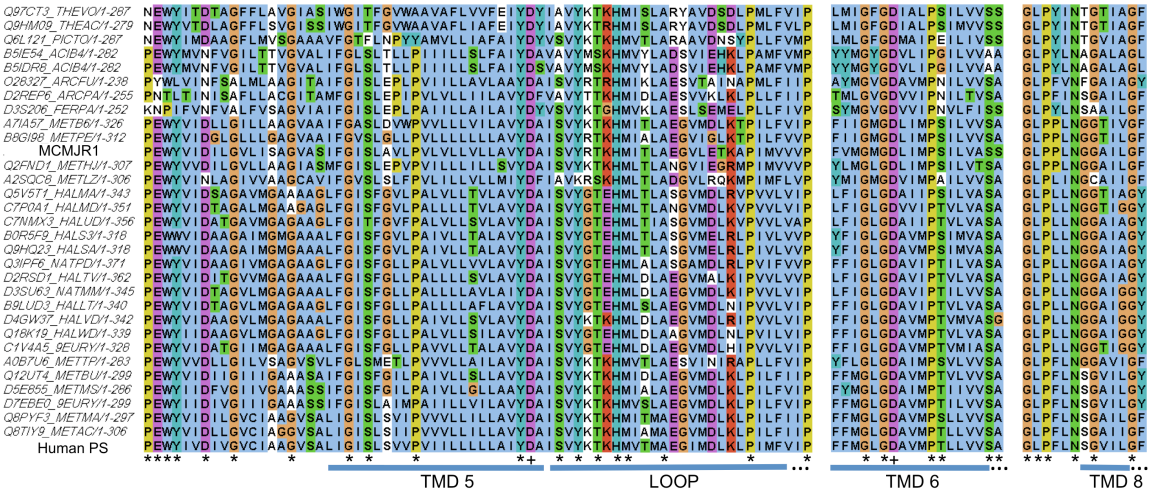


FIGURE 46

References

1. Neurath, H. and G.H. Dixon, *Structure and activation of trypsinogen and chymotrypsinogen.* Fed Proc, 1957. **16**(3): p. 791-801.
2. Neurath, H. and K.A. Walsh, *Role of proteolytic enzymes in biological regulation (a review).* Proc Natl Acad Sci U S A, 1976. **73**(11): p. 3825-32.
3. Rawlings, N.D., et al., *MEROPS: the peptidase database.* Nucleic Acids Res, 2008. **36**(Database issue): p. D320-5.
4. Saito, S., et al., *Granzyme B induces apoptosis of chondrocytes with natural killer cell-like cytotoxicity in rheumatoid arthritis.* J Rheumatol, 2008. **35**(10): p. 1932-43.
5. Page, M.J. and E. Di Cera, *Serine peptidases: classification, structure and function.* Cell Mol Life Sci, 2008. **65**(7-8): p. 1220-36.
6. Martel-Pelletier, J., D.J. Welsch, and J.P. Pelletier, *Metalloproteases and inhibitors in arthritic diseases.* Best Pract Res Clin Rheumatol, 2001. **15**(5): p. 805-29.
7. Lemaitre, V. and J. D'Armiento, *Matrix metalloproteinases in development and disease.* Birth Defects Res C Embryo Today, 2006. **78**(1): p. 1-10.
8. Heron, M., et al., *Deaths: final data for 2006.* Natl Vital Stat Rep, 2009. **57**(14): p. 1-134.
9. Berdowska, I., *Cysteine proteases as disease markers.* Clin Chim Acta, 2004. **342**(1-2): p. 41-69.
10. Netzel-Arnett, S., et al., *Membrane anchored serine proteases: a rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer.* Cancer Metastasis Rev, 2003. **22**(2-3): p. 237-58.
11. Diamandis, E.P., et al., *The new human kallikrein gene family: implications in carcinogenesis.* Trends Endocrinol Metab, 2000. **11**(2): p. 54-60.
12. Talesa, V.N., *Acetylcholinesterase in Alzheimer's disease.* Mech Ageing Dev, 2001. **122**(16): p. 1961-9.
13. Selkoe, D.J., *Translating cell biology into therapeutic advances in Alzheimer's disease.* Nature, 1999. **399**(6738 Suppl): p. A23-31.
14. Sinha, S. and I. Lieberburg, *Cellular mechanisms of beta-amyloid production and secretion.* Proc Natl Acad Sci U S A, 1999. **96**(20): p. 11049-53.
15. Cooper, J.B., *Aspartic proteinases in disease: a structural perspective.* Curr Drug Targets, 2002. **3**(2): p. 155-73.
16. Lehninger, A.L., D.L. Nelson, and M.M. Cox, *Lehninger principles of biochemistry*. 4th ed. 2005, New York: W.H. Freeman. 1 v. (various pagings).
17. Berg, J.M., J.L. Tymoczko, and L. Stryer, *Biochemistry*. 5th ed. 2002, New York: W.H. Freeman. 1 v. (various pagings).
18. Nagle, J.F. and S. Tristram-Nagle, *Structure of lipid bilayers.* Biochim Biophys Acta, 2000. **1469**(3): p. 159-95.
19. Simons, K. and W.L. Vaz, *Model systems, lipid rafts, and cell membranes.* Annu Rev Biophys Biomol Struct, 2004. **33**: p. 269-95.

20. Rajendran, L. and K. Simons, *Lipid rafts and membrane dynamics.* J Cell Sci, 2005. **118**(Pt 6): p. 1099-102.
21. Pike, L.J., *Growth factor receptors, lipid rafts and caveolae: an evolving story.* Biochim Biophys Acta, 2005. **1746**(3): p. 260-73.
22. Carruthers, A. and D.L. Melchior, *Effects of lipid environment on membrane transport: the human erythrocyte sugar transport protein/lipid bilayer system.* Annu Rev Physiol, 1988. **50**: p. 257-71.
23. Szabo, G., *Dual mechanism for the action of cholesterol on membrane permeability.* Nature, 1974. **252**(5478): p. 47-9.
24. Hinz, H.J. and J.M. Sturtevant, *Calorimetric investigation of the influence of cholesterol on the transition properties of bilayers formed from synthetic L- -lecithins in aqueous suspension.* J Biol Chem, 1972. **247**(11): p. 3697-700.
25. Creighton, T.E., *Proteins : structures and molecular properties*. 2nd ed. 1993, New York: W.H. Freeman. xiii, 507 p.
26. Cordes, F.S., J.N. Bright, and M.S. Sansom, *Proline-induced distortions of transmembrane helices.* J Mol Biol, 2002. **323**(5): p. 951-60.
27. Popot, J.L. and D.M. Engelman, *Helical membrane protein folding, stability, and evolution.* Annu Rev Biochem, 2000. **69**: p. 881-922.
28. Nyholm, T.K., S. Ozdirekcan, and J.A. Killian, *How protein transmembrane segments sense the lipid environment.* Biochemistry, 2007. **46**(6): p. 1457-65.
29. Hulko, M., et al., *The HAMP domain structure implies helix rotation in transmembrane signaling.* Cell, 2006. **126**(5): p. 929-40.
30. Fuh, G., et al., *Rational design of potent antagonists to the human growth hormone receptor.* Science, 1992. **256**(5064): p. 1677-80.
31. Matthews, E.E., M. Zoonens, and D.M. Engelman, *Dynamic helix interactions in transmembrane signaling.* Cell, 2006. **127**(3): p. 447-50.
32. Alberts, B., *Molecular biology of the cell*. 4th ed. 2002, New York: Garland Science. xxxiv, [1548] p.
33. Landman, N. and T.W. Kim, *Got RIP? Presenilin-dependent intramembrane proteolysis in growth factor receptor signaling.* Cytokine Growth Factor Rev, 2004. **15**(5): p. 337-51.
34. Brown, M.S., et al., *Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans.* Cell, 2000. **100**(4): p. 391-8.
35. McNaught, A.D., A. Wilkinson, and International Union of Pure and Applied Chemistry., *Compendium of chemical terminology : IUPAC recommendations*. 2nd ed. 1997, Oxford [England] ; Malden, MA, USA: Blackwell Science. vii, 450 p.
36. Cardone, M.H., et al., *Regulation of cell death protease caspase-9 by phosphorylation.* Science, 1998. **282**(5392): p. 1318-21.
37. Beynon, R.J. and J.S. Bond, *Proteolytic enzymes : a practical approach*. 2nd ed. The practical approach series. 2001, Oxford ; New York: Oxford University Press. xviii, 340 p.
38. Hedstrom, L., *Serine protease mechanism and specificity.* Chem Rev, 2002. **102**(12): p. 4501-24.

39. Groll, M., et al., *Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome.* Structure, 2006. **14**(3): p. 451-6.
40. Seemuller, E., et al., *Proteasome from Thermoplasma acidophilum: a threonine protease.* Science, 1995. **268**(5210): p. 579-82.
41. Nguyen, J.T., et al., *Design of potent aspartic protease inhibitors to treat various diseases.* Arch Pharm (Weinheim), 2008. **341**(9): p. 523-35.
42. Brik, A. and C.H. Wong, *HIV-1 protease: mechanism and drug discovery.* Org Biomol Chem, 2003. **1**(1): p. 5-14.
43. Hofmann, T., A.L. Fink, and B.M. Dunn, *Cryoenzymology of penicillopepsin; with an appendix: mechanism of action of aspartyl proteinases.* Biochemistry, 1984. **23**(22): p. 5247-56.
44. Hernick, M. and C.A. Fierke, *Zinc hydrolases: the mechanisms of zinc-dependent deacetylases.* Arch Biochem Biophys, 2005. **433**(1): p. 71-84.
45. Fujinaga, M., et al., *The molecular structure and catalytic mechanism of a novel carboxyl peptidase from Scytalidium lignicolum.* Proc Natl Acad Sci U S A, 2004. **101**(10): p. 3364-9.
46. Schechter, I. and A. Berger, *On the size of the active site in proteases. I. Papain.* Biochem Biophys Res Commun, 1967. **27**(2): p. 157-62.
47. Perona, J.J. and C.S. Craik, *Structural basis of substrate specificity in the serine proteases.* Protein Sci, 1995. **4**(3): p. 337-60.
48. Gillmor, S.A., C.S. Craik, and R.J. Fletterick, *Structural determinants of specificity in the cysteine protease cruzain.* Protein Sci, 1997. **6**(8): p. 1603-11.
49. Berger, A. and I. Schechter, *Mapping the active site of papain with the aid of peptide substrates and inhibitors.* Philos Trans R Soc Lond B Biol Sci, 1970. **257**(813): p. 249-64.
50. Pop, C. and G.S. Salvesen, *Human caspases: activation, specificity, and regulation.* J Biol Chem, 2009. **284**(33): p. 21777-81.
51. Tompa, P., et al., *On the sequential determinants of calpain cleavage.* J Biol Chem, 2004. **279**(20): p. 20775-85.
52. Tozser, J., et al., *Studies on the role of the S4 substrate binding site of HIV proteinases.* FEBS Lett, 1991. **279**(2): p. 356-60.
53. Griffiths, J.T., et al., *Different requirements for productive interaction between the active site of HIV-1 proteinase and substrates containing -hydrophobic*hydrophobic- or -aromatic*pro- cleavage sites.* Biochemistry, 1992. **31**(22): p. 5193-200.
54. Phylip, L.H., et al., *Hydrolysis of synthetic chromogenic substrates by HIV-1 and HIV-2 proteinases.* Biochem Biophys Res Commun, 1990. **171**(1): p. 439-44.
55. Kageyama, H., et al., *Differences in the P1' substrate specificities of pepsin A and chymosin.* J Biochem. **147**(2): p. 167-74.
56. Tang, J., *Specificity of Pepsin and Its Dependence on a Possible 'Hydrophobicbinding Site'.* Nature, 1963. **199**: p. 1094-5.
57. Dunn, B.M., *Structure and mechanism of the pepsin-like family of aspartic peptidases.* Chem Rev, 2002. **102**(12): p. 4431-58.
58. Gelb, M.H. and R.H. Abeles, *Mechanism of inactivation of chymotrypsin by 3-benzyl-6-chloro-2-pyrone.* Biochemistry, 1984. **23**(26): p. 6596-604.

59. Terada, S., et al., *Purification and amino acid sequence of brevilysin L6, a non-hemorrhagic metalloprotease from Agkistrodon halys brevicaudus venom.* J Biochem, 1999. **125**(1): p. 64-9.
60. Kopan, R. and M.X. Ilagan, *Gamma-secretase: proteasome of the membrane?* Nat Rev Mol Cell Biol, 2004. **5**(6): p. 499-504.
61. Weihofen, A., et al., *Release of signal peptide fragments into the cytosol requires cleavage in the transmembrane region by a protease activity that is specifically blocked by a novel cysteine protease inhibitor.* J Biol Chem, 2000. **275**(40): p. 30951-6.
62. Haass, C. and H. Steiner, *Alzheimer disease gamma-secretase: a complex story of GxGD-type presenilin proteases.* Trends Cell Biol, 2002. **12**(12): p. 556-62.
63. Steiner, H., et al., *Glycine 384 is required for presenilin-1 function and is conserved in bacterial polytopic aspartyl proteases.* Nat Cell Biol, 2000. **2**(11): p. 848-51.
64. LaPointe, C.F. and R.K. Taylor, *The type 4 prepilin peptidases comprise a novel family of aspartic acid proteases.* J Biol Chem, 2000. **275**(2): p. 1502-10.
65. Strom, M.S., D.N. Nunn, and S. Lory, *A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family.* Proc Natl Acad Sci U S A, 1993. **90**(6): p. 2404-8.
66. Ng, S.Y., et al., *Archaeal signal peptidases.* Microbiology, 2007. **153**(Pt 2): p. 305-14.
67. Maupin-Furlow, J.A., et al., *Archaeal proteasomes and other regulatory proteases.* Curr Opin Microbiol, 2005. **8**(6): p. 720-8.
68. Szabo, Z., S.V. Albers, and A.J. Driessen, *Active-site residues in the type IV prepilin peptidase homologue PibD from the archaeon Sulfolobus solfataricus.* J Bacteriol, 2006. **188**(4): p. 1437-43.
69. Dupuy, B., et al., *Neisseria gonorrhoeae prepilin export studied in Escherichia coli.* J Bacteriol, 1991. **173**(23): p. 7589-98.
70. Sato, C., et al., *The C-terminal PAL motif and transmembrane domain 9 of presenilin 1 are involved in the formation of the catalytic pore of the gamma-secretase.* J Neurosci, 2008. **28**(24): p. 6264-71.
71. Wang, J., et al., *Conserved "PAL" sequence in presenilins is essential for gamma-secretase activity, but not required for formation or stabilization of gamma-secretase complexes.* Neurobiol Dis, 2004. **15**(3): p. 654-66.
72. Wang, J., et al., *C-terminal PAL motif of presenilin and presenilin homologues required for normal active site conformation.* J Neurochem, 2006. **96**(1): p. 218-27.
73. Iben, L.G., et al., *Signal peptide peptidase and gamma-secretase share equivalent inhibitor binding pharmacology.* J Biol Chem, 2007. **282**(51): p. 36829-36.
74. Friedmann, E., et al., *Consensus analysis of signal peptide peptidase and homologous human aspartic proteases reveals opposite topology of catalytic domains compared with presenilins.* J Biol Chem, 2004. **279**(49): p. 50790-8.
75. Martoglio, B. and T.E. Golde, *Intramembrane-cleaving aspartic proteases and disease: presenilins, signal peptide peptidase and their homologs.* Hum Mol Genet, 2003. **12 Spec No 2**: p. R201-6.

76. Xia, W. and M.S. Wolfe, *Intramembrane proteolysis by presenilin and presenilin-like proteases.* J Cell Sci, 2003. **116**(Pt 14): p. 2839-44.
77. Haass, C., et al., *Mutations associated with a locus for familial Alzheimer's disease result in alternative processing of amyloid beta-protein precursor.* J Biol Chem, 1994. **269**(26): p. 17741-8.
78. Citron, M., et al., *Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production.* Nature, 1992. **360**(6405): p. 672-4.
79. Kimberly, W.T., et al., *Gamma-secretase is a membrane protein complex comprised of presenilin, nicastrin, Aph-1, and Pen-2.* Proc Natl Acad Sci U S A, 2003. **100**(11): p. 6382-7.
80. Lee, S.F., et al., *A conserved GXXXG motif in APH-1 is critical for assembly and activity of the gamma-secretase complex.* J Biol Chem, 2004. **279**(6): p. 4144-52.
81. Niimura, M., et al., *Aph-1 contributes to the stabilization and trafficking of the gamma-secretase complex through mechanisms involving intermolecular and intramolecular interactions.* J Biol Chem, 2005. **280**(13): p. 12967-75.
82. Zhang, Y.W., et al., *Nicastrin is critical for stability and trafficking but not association of other presenilin/gamma-secretase components.* J Biol Chem, 2005. **280**(17): p. 17020-6.
83. Shah, S., et al., *Nicastrin functions as a gamma-secretase-substrate receptor.* Cell, 2005. **122**(3): p. 435-47.
84. Kim, S.H. and S.S. Sisodia, *A sequence within the first transmembrane domain of PEN-2 is critical for PEN-2-mediated endoproteolysis of presenilin 1.* J Biol Chem, 2005. **280**(3): p. 1992-2001.
85. Prokop, S., C. Haass, and H. Steiner, *Length and overall sequence of the PEN-2 C-terminal domain determines its function in the stabilization of presenilin fragments.* J Neurochem, 2005. **94**(1): p. 57-62.
86. Wolfe, M.S., et al., *Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity.* Nature, 1999. **398**(6727): p. 513-7.
87. Selkoe, D. and R. Kopan, *Notch and Presenilin: regulated intramembrane proteolysis links development and degeneration.* Annu Rev Neurosci, 2003. **26**: p. 565-97.
88. Greenwald, I., *LIN-12/Notch signaling: lessons from worms and flies.* Genes Dev, 1998. **12**(12): p. 1751-62.
89. Sardi, S.P., et al., *Presenilin-dependent ErbB4 nuclear signaling regulates the timing of astrogenesis in the developing brain.* Cell, 2006. **127**(1): p. 185-97.
90. De Strooper, B. and W. Annaert, *Proteolytic processing and cell biological functions of the amyloid precursor protein.* J Cell Sci, 2000. **113 (Pt 11)**: p. 1857-70.
91. Haass, C., et al., *The Swedish mutation causes early-onset Alzheimer's disease by beta-secretase cleavage within the secretory pathway.* Nat Med, 1995. **1**(12): p. 1291-6.
92. Qi-Takahara, Y., et al., *Longer forms of amyloid beta protein: implications for the mechanism of intramembrane cleavage by gamma-secretase.* J Neurosci, 2005. **25**(2): p. 436-45.

93. Chandu, D., S.S. Huppert, and R. Kopan, *Analysis of transmembrane domain mutants is consistent with sequential cleavage of Notch by gamma-secretase.* J Neurochem, 2006. **96**(1): p. 228-35.
94. Lammich, S., et al., *Presenilin-dependent intramembrane proteolysis of CD44 leads to the liberation of its intracellular domain and the secretion of an Abeta-like peptide.* J Biol Chem, 2002. **277**(47): p. 44754-9.
95. Schroeter, E.H., et al., *A presenilin dimer at the core of the gamma-secretase enzyme: insights from parallel analysis of Notch 1 and APP proteolysis.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 13075-80.
96. Tolia, A., L. Chavez-Gutierrez, and B. De Strooper, *Contribution of presenilin transmembrane domains 6 and 7 to a water-containing cavity in the gamma-secretase complex.* J Biol Chem, 2006. **281**(37): p. 27633-42.
97. Vidal, G.A., et al., *Presenilin-dependent gamma-secretase processing regulates multiple ERBB4/HER4 activities.* J Biol Chem, 2005. **280**(20): p. 19777-83.
98. Marambaud, P., et al., *A presenilin-1/gamma-secretase cleavage releases the E-cadherin intracellular domain and regulates disassembly of adherens junctions.* Embo J, 2002. **21**(8): p. 1948-56.
99. Beel, A.J. and C.R. Sanders, *Substrate specificity of gamma-secretase and other intramembrane proteases.* Cell Mol Life Sci, 2008. **65**(9): p. 1311-34.
100. Lichtenthaler, S.F., et al., *Mutations in the transmembrane domain of APP altering gamma-secretase specificity.* Biochemistry, 1997. **36**(49): p. 15396-403.
101. Lichtenthaler, S.F., et al., *Mechanism of the cleavage specificity of Alzheimer's disease gamma-secretase identified by phenylalanine-scanning mutagenesis of the transmembrane domain of the amyloid precursor protein.* Proc Natl Acad Sci U S A, 1999. **96**(6): p. 3053-8.
102. Kienlen-Campard, P., et al., *Amyloidogenic processing but not amyloid precursor protein (APP) intracellular C-terminal domain production requires a precisely oriented APP dimer assembled by transmembrane GXXXG motifs.* J Biol Chem, 2008. **283**(12): p. 7733-44.
103. Munter, L.M., et al., *GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of Abeta42.* Embo J, 2007. **26**(6): p. 1702-12.
104. Hass, M.R., et al., *Presenilin: RIP and beyond.* Semin Cell Dev Biol, 2009. **20**(2): p. 201-10.
105. Ponting, C.P., et al., *Identification of a novel family of presenilin homologues.* Hum Mol Genet, 2002. **11**(9): p. 1037-44.
106. Weihofen, A., et al., *Identification of signal peptide peptidase, a presenilin-type aspartic protease.* Science, 2002. **296**(5576): p. 2215-8.
107. Lemberg, M.K., et al., *Intramembrane proteolysis of signal peptides: an essential step in the generation of HLA-E epitopes.* J Immunol, 2001. **167**(11): p. 6441-6.
108. McLauchlan, J., et al., *Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets.* Embo J, 2002. **21**(15): p. 3980-8.
109. Targett-Adams, P., et al., *Signal peptide peptidase cleavage of GB virus B core protein is required for productive infection in vivo.* J Biol Chem, 2006. **281**(39): p. 29221-7.

110. Robakis, T., et al., *An internal signal sequence directs intramembrane proteolysis of a cellular immunoglobulin domain protein.* J Biol Chem, 2008. **283**(52): p. 36369-76.
111. Chang, H.T., et al., *Signal peptide of eosinophil cationic protein upregulates transforming growth factor-alpha expression in human cells.* J Cell Biochem, 2007. **100**(5): p. 1266-75.
112. El Hage, F., et al., *Preprocalcitonin signal peptide generates a cytotoxic T lymphocyte-defined tumor epitope processed by a proteasome-independent pathway.* Proc Natl Acad Sci U S A, 2008. **105**(29): p. 10119-24.
113. Okamoto, K., et al., *Intramembrane proteolysis and endoplasmic reticulum retention of hepatitis C virus core protein.* J Virol, 2004. **78**(12): p. 6370-80.
114. Weihofen, A. and B. Martoglio, *Intramembrane-cleaving proteases: controlled liberation of proteins and bioactive peptides.* Trends Cell Biol, 2003. **13**(2): p. 71-8.
115. Lemberg, M.K. and B. Martoglio, *Requirements for signal peptide peptidase-catalyzed intramembrane proteolysis.* Mol Cell, 2002. **10**(4): p. 735-44.
116. Golde, T.E., M.S. Wolfe, and D.C. Greenbaum, *Signal peptide peptidases: a family of intramembrane-cleaving proteases that cleave type 2 transmembrane proteins.* Semin Cell Dev Biol, 2009. **20**(2): p. 225-30.
117. Nyborg, A.C., et al., *Intramembrane proteolytic cleavage by human signal peptide peptidase like 3 and malaria signal peptide peptidase.* Faseb J, 2006. **20**(10): p. 1671-9.
118. Tsachaki, M., J. Ghiso, and S. Efthimiopoulos, *BRI2 as a central protein involved in neurodegeneration.* Biotechnol J, 2008. **3**(12): p. 1548-54.
119. Fluhrer, R., et al., *A gamma-secretase-like intramembrane cleavage of TNFalpha by the GxGD aspartyl protease SPPL2b.* Nat Cell Biol, 2006. **8**(8): p. 894-6.
120. Feng, L., et al., *Structure of a site-2 protease family intramembrane metalloprotease.* Science, 2007. **318**(5856): p. 1608-12.
121. Wang, Y., Y. Zhang, and Y. Ha, *Crystal structure of a rhomboid family intramembrane protease.* Nature, 2006. **444**(7116): p. 179-80.
122. Lazarov, V.K., et al., *Electron microscopic structure of purified, active gamma-secretase reveals an aqueous intramembrane chamber and two pores.* Proc Natl Acad Sci U S A, 2006. **103**(18): p. 6889-94.
123. Osenkowski, P., et al., *Cryoelectron microscopy structure of purified gamma-secretase at 12 A resolution.* J Mol Biol, 2009. **385**(2): p. 642-52.
124. Zhu, J., et al., *Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: procedure and application to the 70S Escherichia coli ribosome.* J Struct Biol, 1997. **118**(3): p. 197-219.
125. Frank, J., *Single-particle imaging of macromolecules by cryo-electron microscopy.* Annu Rev Biophys Biomol Struct, 2002. **31**: p. 303-19.
126. Caffrey, M., *Membrane protein crystallization.* J Struct Biol, 2003. **142**(1): p. 108-32.
127. Puig, O., et al., *The tandem affinity purification (TAP) method: a general procedure of protein complex purification.* Methods, 2001. **24**(3): p. 218-29.

128. Bottcher, B., S.A. Wynne, and R.A. Crowther, *Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy.* Nature, 1997. **386**(6620): p. 88-91.
129. Malhotra, A., et al., *Escherichia coli 70 S ribosome at 15 A resolution by cryo-electron microscopy: localization of fMet-tRNAfMet and fitting of L1 protein.* J Mol Biol, 1998. **280**(1): p. 103-16.
130. Cheng, Y. and T. Walz, *The advent of near-atomic resolution in single-particle electron microscopy.* Annu Rev Biochem, 2009. **78**: p. 723-42.
131. Valle, M., et al., *Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy.* Nat Struct Biol, 2003. **10**(11): p. 899-906.
132. Ohi, M., et al., *Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy.* Biol Proced Online, 2004. **6**: p. 23-34.
133. Frank, J., *Cryo-electron microscopy as an investigative tool: the ribosome as an example.* Bioessays, 2001. **23**(8): p. 725-32.
134. Wang, L. and F.J. Sigworth, *Cryo-EM and single particles.* Physiology (Bethesda), 2006. **21**: p. 13-8.
135. Van Heel, M., *Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction.* Ultramicroscopy, 1987. **21**(2): p. 111-23.
136. Penczek, P.A., R.A. Grassucci, and J. Frank, *The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles.* Ultramicroscopy, 1994. **53**(3): p. 251-70.
137. Penczek, P.A., J. Zhu, and J. Frank, *A common-lines based method for determining orientations for N > 3 particle projections simultaneously.* Ultramicroscopy, 1996. **63**(3-4): p. 205-18.
138. Radermacher, M., *Three-dimensional reconstruction of single particles from random and nonrandom tilt series.* J Electron Microsc Tech, 1988. **9**(4): p. 359-94.
139. Reeves, P.J., R.L. Thurmond, and H.G. Khorana, *Structure and function in rhodopsin: high level expression of a synthetic bovine opsin gene and its mutants in stable mammalian cell lines.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11487-92.
140. Chen, C. and H. Okayama, *High-efficiency transformation of mammalian cells by plasmid DNA.* Mol Cell Biol, 1987. **7**(8): p. 2745-52.
141. Li, Y.M., et al., *Presenilin 1 is linked with gamma-secretase activity in the detergent solubilized state.* Proc Natl Acad Sci U S A, 2000. **97**(11): p. 6138-43.
142. Kim, S.H., et al., *Regulated hyperaccumulation of presenilin-1 and the "gamma-secretase" complex. Evidence for differential intramembranous processing of transmembrane subatrates.* J Biol Chem, 2003. **278**(36): p. 33992-4002.
143. Li, Y.M., et al., *Photoactivated gamma-secretase inhibitors directed to the active site covalently label presenilin 1.* Nature, 2000. **405**(6787): p. 689-94.
144. Ludtke, S.J., P.R. Baldwin, and W. Chiu, *EMAN: semiautomated software for high-resolution single-particle reconstructions.* J Struct Biol, 1999. **128**(1): p. 82-97.
145. Frank, J., et al., *SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields.* J Struct Biol, 1996. **116**(1): p. 190-9.

146. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.
147. Price, D.L. and S.S. Sisodia, *Mutant genes in familial Alzheimer's disease and transgenic models.* Annu Rev Neurosci, 1998. **21**: p. 479-505.
148. Fischer, H., I. Polikarpov, and A.F. Craievich, *Average protein density is a molecular-weight-dependent function.* Protein Sci, 2004. **13**(10): p. 2825-8.
149. Tarassishin, L., et al., *Processing of Notch and amyloid precursor protein by gamma-secretase is spatially distinct.* Proc Natl Acad Sci U S A, 2004. **101**(49): p. 17050-5.
150. Boisset, N., et al., *Three-dimensional reconstruction of human alpha 2-macroglobulin and refinement of the localization of thiol ester bonds with monomaleimido nanogold.* Ann N Y Acad Sci, 1994. **737**: p. 229-44.
151. Wang, M.C., et al., *The three-dimensional structure of the cardiac L-type voltage-gated calcium channel: comparison with the skeletal muscle form reveals a common architectural motif.* J Biol Chem, 2004. **279**(8): p. 7159-68.
152. Sato, T., et al., *Active gamma-secretase complexes contain only one of each component.* J Biol Chem, 2007. **282**(47): p. 33985-93.
153. Stewart, P.L., et al., *Review: resolution issues in single-particle reconstruction.* J Struct Biol, 1999. **128**(1): p. 58-64.
154. Sato, C., et al., *Structure of the catalytic pore of gamma-secretase probed by the accessibility of substituted cysteines.* J Neurosci, 2006. **26**(46): p. 12081-8.
155. Erez, E., D. Fass, and E. Bibi, *How intramembrane proteases bury hydrolytic reactions in the membrane.* Nature, 2009. **459**(7245): p. 371-8.
156. Ubarretxena-Belandia, I. and D.M. Engelman, *Helical membrane proteins: diversity of functions in the context of simple architecture.* Curr Opin Struct Biol, 2001. **11**(3): p. 370-6.
157. Das, C., et al., *Designed helical peptides inhibit an intramembrane protease.* J Am Chem Soc, 2003. **125**(39): p. 11794-5.
158. Kornilova, A.Y., et al., *The initial substrate-binding site of gamma-secretase is located on presenilin near the active site.* Proc Natl Acad Sci U S A, 2005. **102**(9): p. 3230-5.
159. Fraering, P.C., et al., *Purification and characterization of the human gamma-secretase complex.* Biochemistry, 2004. **43**(30): p. 9774-89.
160. Crystal, A.S., et al., *Membrane topology of gamma-secretase component PEN-2.* J Biol Chem, 2003. **278**(22): p. 20117-23.
161. Fortna, R.R., et al., *Membrane topology and nicastrin-enhanced endoproteolysis of APH-1, a component of the gamma-secretase complex.* J Biol Chem, 2004. **279**(5): p. 3685-93.
162. Henricson, A., L. Kall, and E.L. Sonnhammer, *A novel transmembrane topology of presenilin based on reconciling experimental and computational evidence.* Febs J, 2005. **272**(11): p. 2727-33.
163. Cacquevel, M., et al., *Rapid purification of active gamma-secretase, an intramembrane protease implicated in Alzheimer's disease.* J Neurochem, 2008. **104**(1): p. 210-20.

164. Cereghino, J.L. and J.M. Cregg, *Heterologous protein expression in the methylotrophic yeast Pichia pastoris.* FEMS Microbiol Rev, 2000. **24**(1): p. 45-66.
165. Grinna, L.S. and J.F. Tschopp, *Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, Pichia pastoris.* Yeast, 1989. **5**(2): p. 107-15.
166. Higgins, D.R., *Overview of protein expression in Pichia pastoris.* Curr Protoc Protein Sci, 2001. **Chapter 5**: p. Unit5 7.
167. Cregg, J.M., et al., *Functional characterization of the two alcohol oxidase genes from the yeast Pichia pastoris.* Mol Cell Biol, 1989. **9**(3): p. 1316-23.
168. Jungo, C., I. Marison, and U. von Stockar, *Regulation of alcohol oxidase of a recombinant Pichia pastoris Mut+ strain in transient continuous cultures.* J Biotechnol, 2007. **130**(3): p. 236-46.
169. Cregg, J.M., *Introduction: distinctions between Pichia pastoris and other expression systems.* Methods Mol Biol, 2007. **389**: p. 1-10.
170. Cregg, J.M., et al., *Pichia pastoris as a host system for transformations.* Mol Cell Biol, 1985. **5**(12): p. 3376-85.
171. Lin Cereghino, G.P., et al., *New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of Pichia pastoris.* Gene, 2001. **263**(1-2): p. 159-69.
172. Vuorela, A., et al., *Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast pichia pastoris: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase.* Embo J, 1997. **16**(22): p. 6702-12.
173. Steiner, H., et al., *The biological and pathological function of the presenilin-1 Deltaexon 9 mutation is independent of its defect to undergo proteolytic processing.* J Biol Chem, 1999. **274**(12): p. 7615-8.
174. Thinakaran, G., et al., *Evidence that levels of presenilins (PS1 and PS2) are coordinately regulated by competition for limiting cellular factors.* J Biol Chem, 1997. **272**(45): p. 28415-22.
175. Edbauer, D., et al., *Reconstitution of gamma-secretase activity.* Nat Cell Biol, 2003. **5**(5): p. 486-8.
176. Levitan, D., et al., *Assessment of normal and mutant human presenilin function in Caenorhabditis elegans.* Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14940-4.
177. Borchelt, D.R., et al., *Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1-42/1-40 ratio in vitro and in vivo.* Neuron, 1996. **17**(5): p. 1005-13.
178. Page, R.M., et al., *Generation of Abeta38 and Abeta42 is independently and differentially affected by familial Alzheimer disease-associated presenilin mutations and gamma-secretase modulation.* J Biol Chem, 2008. **283**(2): p. 677-83.
179. Parks, T.D., et al., *Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase.* Anal Biochem, 1994. **216**(2): p. 413-7.
180. Sato, T., et al., *Signal peptide peptidase: biochemical properties and modulation by nonsteroidal antiinflammatory drugs.* Biochemistry, 2006. **45**(28): p. 8649-56.

181. Bardy, S.L. and K.F. Jarrell, *FlaK of the archaeon Methanococcus maripaludis possesses preflagellin peptidase activity.* FEMS Microbiol Lett, 2002. **208**(1): p. 53-9.
182. Albers, S.V., Z. Szabo, and A.J. Driessen, *Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity.* J Bacteriol, 2003. **185**(13): p. 3918-25.
183. Bardy, S.L. and K.F. Jarrell, *Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon Methanococcus voltae.* Mol Microbiol, 2003. **50**(4): p. 1339-47.
184. Cunningham, B.C. and J.A. Wells, *High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis.* Science, 1989. **244**(4908): p. 1081-5.
185. le Coutre, J., et al., *Proteomics on full-length membrane proteins using mass spectrometry.* Biochemistry, 2000. **39**(15): p. 4237-42.
186. Zhang, N. and L. Li, *Effects of common surfactants on protein digestion and matrix-assisted laser desorption/ionization mass spectrometric analysis of the digested peptides using two-layer sample preparation.* Rapid Commun Mass Spectrom, 2004. **18**(8): p. 889-96.
187. Yeung, Y.G., et al., *Removal of detergents from protein digests for mass spectrometry analysis.* Anal Biochem, 2008. **382**(2): p. 135-7.
188. Mirza, S.P., et al., *Improved method for the analysis of membrane proteins by mass spectrometry.* Physiol Genomics, 2007. **30**(1): p. 89-94.
189. Norris, J.L., N.A. Porter, and R.M. Caprioli, *Mass spectrometry of intracellular and membrane proteins using cleavable detergents.* Anal Chem, 2003. **75**(23): p. 6642-7.
190. Cruts, M., et al., *Molecular genetic analysis of familial early-onset Alzheimer's disease linked to chromosome 14q24.3.* Hum Mol Genet, 1995. **4**(12): p. 2363-71.
191. Fluhrer, R., et al., *Intramembrane proteolysis of GXGD-type aspartyl proteases is slowed by a familial Alzheimer disease-like mutation.* J Biol Chem, 2008. **283**(44): p. 30121-8.
192. Koonin, E.V., et al., *The rhomboids: a nearly ubiquitous family of intramembrane serine proteases that probably evolved by multiple ancient horizontal gene transfers.* Genome Biol, 2003. **4**(3): p. R19.
193. Lewis, A.P. and P.J. Thomas, *A novel clan of zinc metallopeptidases with possible intramembrane cleavage properties.* Protein Sci, 1999. **8**(2): p. 439-42.
194. Tandon, A. and P. Fraser, *The presenilins.* Genome Biol, 2002. **3**(11): p. reviews3014.
195. Levitan, D. and I. Greenwald, *Facilitation of lin-12-mediated signalling by sel-12, a Caenorhabditis elegans S182 Alzheimer's disease gene.* Nature, 1995. **377**(6547): p. 351-4.
196. Li, X. and I. Greenwald, *HOP-1, a Caenorhabditis elegans presenilin, appears to be functionally redundant with SEL-12 presenilin and to facilitate LIN-12 and GLP-1 signaling.* Proc Natl Acad Sci U S A, 1997. **94**(22): p. 12204-9.
197. Lefevre, F., M.H. Remy, and J.M. Masson, *Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function.* Nucleic Acids Res, 1997. **25**(2): p. 447-8.

198. Murthy, K.H., et al., *The crystal structures at 2.2-A resolution of hydroxyethylene-based inhibitors bound to human immunodeficiency virus type 1 protease show that the inhibitors are present in two distinct orientations.* J Biol Chem, 1992. **267**(32): p. 22770-8.
199. Dohnalek, J., et al., *Hydroxyethylamine isostere of an HIV-1 protease inhibitor prefers its amine to the hydroxy group in binding to catalytic aspartates. A synchrotron study of HIV-1 protease in complex with a peptidomimetic inhibitor.* J Med Chem, 2002. **45**(7): p. 1432-8.
200. De Jonghe, C., et al., *Evidence that Abeta42 plasma levels in presenilin-1 mutation carriers do not allow for prediction of their clinical phenotype.* Neurobiol Dis, 1999. **6**(4): p. 280-7.
201. Deng, Y., et al., *Deletion of presenilin 1 hydrophilic loop sequence leads to impaired gamma-secretase activity and exacerbated amyloid pathology.* J Neurosci, 2006. **26**(14): p. 3845-54.
202. Warne, T., et al., *Development and crystallization of a minimal thermostabilised G protein-coupled receptor.* Protein Expr Purif, 2009. **65**(2): p. 204-13.
203. Selkoe, D.J. and M.S. Wolfe, *In search of gamma-secretase: presenilin at the cutting edge.* Proc Natl Acad Sci U S A, 2000. **97**(11): p. 5690-2.
204. Li, S.C. and C.M. Deber, *A measure of helical propensity for amino acids in membrane environments.* Nat Struct Biol, 1994. **1**(6): p. 368-73.
205. Williams, K.A. and C.M. Deber, *Proline residues in transmembrane helices: structural or dynamic role?* Biochemistry, 1991. **30**(37): p. 8919-23.
206. Hurley, J.H., D.A. Mason, and B.W. Matthews, *Flexible-geometry conformational energy maps for the amino acid residue preceding a proline.* Biopolymers, 1992. **32**(11): p. 1443-6.
207. Smith, J.A. and L.G. Pease, *Reverse turns in peptides and proteins.* CRC Crit Rev Biochem, 1980. **8**(4): p. 315-99.
208. Deber, C.M., M. Glibowicka, and G.A. Woolley, *Conformations of proline residues in membrane environments.* Biopolymers, 1990. **29**(1): p. 149-57.
209. Tieleman, D.P., et al., *Proline-induced hinges in transmembrane helices: possible roles in ion channel gating.* Proteins, 2001. **44**(2): p. 63-72.
210. Barlow, D.J. and J.M. Thornton, *Helix geometry in proteins.* J Mol Biol, 1988. **201**(3): p. 601-19.
211. Sansom, M.S. and H. Weinstein, *Hinges, swivels and switches: the role of prolines in signalling via transmembrane alpha-helices.* Trends Pharmacol Sci, 2000. **21**(11): p. 445-51.
212. Orzaez, M., et al., *Influence of proline residues in transmembrane helix packing.* J Mol Biol, 2004. **335**(2): p. 631-40.
213. Gouaux, E. and R. Mackinnon, *Principles of selective ion transport in channels and pumps.* Science, 2005. **310**(5753): p. 1461-5.
214. Cruts, M. and C. Van Broeckhoven, *Presenilin mutations in Alzheimer's disease.* Hum Mutat, 1998. **11**(3): p. 183-90.
215. Sato, T., et al., *Distinct pharmacological effects of inhibitors of signal peptide peptidase and gamma-secretase.* J Biol Chem, 2008. **283**(48): p. 33287-95.
216. Barten, D.M., et al., *Dynamics of {beta}-amyloid reductions in brain, cerebrospinal fluid, and plasma of {beta}-amyloid precursor protein transgenic*

mice treated with a {gamma}-secretase inhibitor. J Pharmacol Exp Ther, 2005. **312**(2): p. 635-43.

217. Yan, X.X., et al., *Binding sites of gamma-secretase inhibitors in rodent brain: distribution, postnatal development, and effect of deafferentation.* J Neurosci, 2004. **24**(12): p. 2942-52.
218. Prasad, C.V., et al., *Discovery of (S)-2-((S)-2-(3,5-difluorophenyl)-2-hydroxyacetamido)-N-((S,Z)-3-methyl-4- oxo-4,5-dihydro-3H-benzo[d][1,2]diazepin-5-yl)propanamide (BMS-433796): a gamma-secretase inhibitor with Abeta lowering activity in a transgenic mouse model of Alzheimer's disease.* Bioorg Med Chem Lett, 2007. **17**(14): p. 4006-11.
219. Weggen, S., et al., *A subset of NSAIDs lower amyloidogenic Abeta42 independently of cyclooxygenase activity.* Nature, 2001. **414**(6860): p. 212-6.
220. Weggen, S., et al., *Abeta42-lowering nonsteroidal anti-inflammatory drugs preserve intramembrane cleavage of the amyloid precursor protein (APP) and ErbB-4 receptor and signaling through the APP intracellular domain.* J Biol Chem, 2003. **278**(33): p. 30748-54.
221. Weggen, S., et al., *Evidence that nonsteroidal anti-inflammatory drugs decrease amyloid beta 42 production by direct modulation of gamma-secretase activity.* J Biol Chem, 2003. **278**(34): p. 31831-7.
222. Kinch, L.N., K. Ginalski, and N.V. Grishin, *Site-2 protease regulated intramembrane proteolysis: sequence homologs suggest an ancient signaling cascade.* Protein Sci, 2006. **15**(1): p. 84-93.
223. Urban, S., *Rhomboid proteins: conserved membrane proteases with divergent biological functions.* Genes Dev, 2006. **20**(22): p. 3054-68.
224. Makinoshima, H. and M.S. Glickman, *Site-2 proteases in prokaryotes: regulated intramembrane proteolysis expands to microbial pathogenesis.* Microbes Infect, 2006. **8**(7): p. 1882-8.
225. Sakai, J., et al., *Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment.* Cell, 1996. **85**(7): p. 1037-46.
226. Rawson, R.B., et al., *Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs.* Mol Cell, 1997. **1**(1): p. 47-57.
227. Selkoe, D.J., *Cell biology of the amyloid beta-protein precursor and the mechanism of Alzheimer's disease.* Annu Rev Cell Biol, 1994. **10**: p. 373-403.
228. Dobrosotskaya, I.Y., et al., *Regulation of SREBP processing and membrane lipid production by phospholipids in Drosophila.* Science, 2002. **296**(5569): p. 879-83.
229. Ye, J., et al., *ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs.* Mol Cell, 2000. **6**(6): p. 1355-64.
230. Rudner, D.Z., P. Fawcett, and R. Losick, *A family of membrane-embedded metalloproteases involved in regulated proteolysis of membrane-associated transcription factors.* Proc Natl Acad Sci U S A, 1999. **96**(26): p. 14765-70.
231. Alba, B.M. and C.A. Gross, *Regulation of the Escherichia coli sigma-dependent envelope stress response.* Mol Microbiol, 2004. **52**(3): p. 613-9.

232. Makinoshima, H. and M.S. Glickman, *Regulation of Mycobacterium tuberculosis cell envelope composition and virulence by intramembrane proteolysis.* Nature, 2005. **436**(7049): p. 406-9.
233. Urban, S., J.R. Lee, and M. Freeman, *Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases.* Cell, 2001. **107**(2): p. 173-82.
234. Gallio, M., et al., *A conserved mechanism for extracellular signaling in eukaryotes and prokaryotes.* Proc Natl Acad Sci U S A, 2002. **99**(19): p. 12208-13.
235. Tolia, A. and B. De Strooper, *Structure and function of gamma-secretase.* Semin Cell Dev Biol, 2009. **20**(2): p. 211-8.
236. Weihofen, A., et al., *Targeting presenilin-type aspartic protease signal peptide peptidase with gamma-secretase inhibitors.* J Biol Chem, 2003. **278**(19): p. 16528-33.
237. Nyborg, A.C., et al., *Signal peptide peptidase forms a homodimer that is labeled by an active site-directed gamma-secretase inhibitor.* J Biol Chem, 2004. **279**(15): p. 15153-60.
238. Narayanan, S., T. Sato, and M.S. Wolfe, *A C-terminal region of signal peptide peptidase defines a functional domain for intramembrane aspartic protease catalysis.* J Biol Chem, 2007. **282**(28): p. 20172-9.
239. Wu, Z., et al., *Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry.* Nat Struct Mol Biol, 2006. **13**(12): p. 1084-91.
240. Sisodia, S.S. and P.H. St George-Hyslop, *gamma-Secretase, Notch, Abeta and Alzheimer's disease: where do the presenilins fit in?* Nat Rev Neurosci, 2002. **3**(4): p. 281-90.