



TECHNISCHE UNIVERSITÄT MÜNCHEN

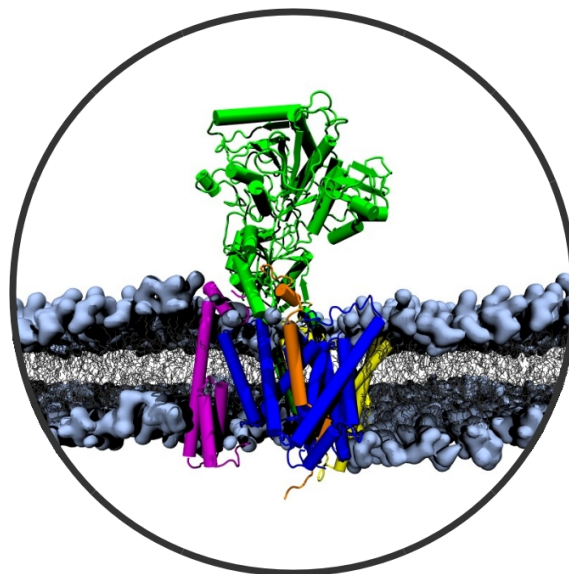
TUM SCHOOL OF NATURAL SCIENCES

Doctoral Thesis

**Interaction and Modulation of  $\gamma$ -Secretase and  
Amyloid Precursor Protein Studied by  
Molecular Dynamics Simulations**

**Shu-Yu Chen**

Supervisor: Prof. Dr. Martin Zacharias





TECHNISCHE UNIVERSITÄT MÜNCHEN

TUM SCHOOL OF NATURAL SCIENCES

**Interaction and Modulation of  $\gamma$ -Secretase and  
Amyloid Precursor Protein Studied by  
Molecular Dynamics Simulations**

Shu-Yu Chen

Vollständiger Abdruck der von der TUM School of Natural Sciences der Technischen  
Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften genehmigten Dissertation

Vorsitz: Prof. Dr. Friedrich C. Simmel  
Prüfer\*innen der Dissertation: 1. Prof. Dr. Martin Zacharias  
2. Prof. Dr. Karen Alim

Die Dissertation wurde am 04.04.2023 bei der Technischen Universität München eingereicht  
und durch die TUM School of Natural Sciences am 23.06.2023 angenommen.



I confirm that this doctoral thesis is my own work and I have documented all sources and material used.

Munich, 31.March.2023

Shu-Yu Chen

# Abstract

The intramembrane proteolysis of the substrate C99 by  $\gamma$ -secretase generates amyloid-beta ( $A\beta$ ) peptides, which are believed to be one of the main causes of Alzheimer's disease (AD). Inhibition and modulation of  $\gamma$ -secretase are strategies for AD prevention and treatment. Although several experiments have reported the cleavage behavior upon single-point mutations, changes in the membrane environment, or responses toward inhibitors and modulators, little is known about the molecular details of the cleavage process. Recent advances in cryo-EM imaging technology have resolved the 3-dimensional structures of  $\gamma$ -secretase in complex with its substrates and ligands, providing scientists in the field with a new aspect to understand the proteolytic process.

The thesis uses molecular dynamics simulations to investigate the dynamics of  $\gamma$ -secretase in substrate-free, substrate-bound, and ligand-bound forms. Stability and energetic analysis suggested three main interactions between  $\gamma$ -secretase and its substrate. First, the hydrophobic patch on the transmembrane domain 6a helix of the presenilin subunit plays an essential role in the stability of the enzyme-substrate complex and the stability of the active catalytic site. However, the helical structure can be disturbed when presenilin is placed in a detergent environment. Second, the enzyme-substrate hybrid  $\beta$ -sheet C-terminal to the substrate scissile bond is crucial in controlling the water molecules to the cleavage center and is essential for the formation of a catalytically active geometry. Third, an internal docking site at the C99- $\gamma$ -secretase interface inside the membrane environment was identified as the strongest interacting region. The balance between the interaction at this region and the membrane distortion induced by a lysine located at the N-terminal juxtamembrane region of C99 can largely control the length of the released  $A\beta$  peptides. Based on this model, a putative binding mode of a  $\gamma$ -secretase modulator is proposed. The purposed enzyme-substrate-modulator configuration not only successfully explains the conserved chemical motifs across a wide range of other modulators but also offers an explanation of how it promotes the secretion of shorter  $A\beta$  peptides. As a continuation of the study on the ternary complex, 19 protein-protein-stabilizer complexes were studied *in silico* and the dual-binding mechanism was suggested.

Notably, the molecular models proposed for the  $\gamma$ -secretase system such as the influence of lipid environment, the hybrid  $\beta$ -sheet, and membrane hydrophobic mismatch on proteolysis, are likely transferrable to other membrane-embedded systems and a wide range of other proteases. The results reported in the thesis not only facilitate the understanding of the mechanism of AD-causing mutations on  $\gamma$ -secretase and C99 but also open new aspects of interpreting similar biological processes.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>1. Introduction</b>	<b>1</b>
<b>2. <math>\gamma</math>-secretase and A<math>\beta</math> peptides</b>	<b>3</b>
2.1. Biological function and subunits of $\gamma$ -secretase . . . . .	3
2.2. 3D-Structures of $\gamma$ -secretase resolved by X-ray crystallography and cryo-EM .	3
2.3. A $\beta$ peptides secretion through interactions between C99 and $\gamma$ -secretase . . . .	7
2.3.1. Substrate Recognition . . . . .	7
2.3.2. Substrate Positioning . . . . .	9
2.3.3. Processive Cleavage . . . . .	9
2.3.4. Effects of mutations on APP and PS1 . . . . .	10
2.4. Inhibition and Modulation of $\gamma$ -Secretase as Therapeutic Strategies against Alzheimer's Disease . . . . .	12
2.4.1. $\gamma$ -Secretase Inhibitor . . . . .	12
2.4.2. $\gamma$ -Secretase Modulator . . . . .	13
<b>3. Theory and Methods</b>	<b>14</b>
3.1. Molecular Dynamics simulation . . . . .	14
3.1.1. From quantum mechanism to classical mechanics . . . . .	14
3.1.2. Force field . . . . .	15
3.1.3. Simulation box and periodic boundary condition . . . . .	16
3.1.4. Integrating the equations of motion . . . . .	18
3.2. Free Energy Calculation . . . . .	20
3.2.1. Biasing method . . . . .	23
3.3. Simulation protocol for $\gamma$ -secretase . . . . .	26
3.3.1. Protonation state of the catalytic center . . . . .	26
3.3.2. Simulation setup . . . . .	29
<b>4. Dynamics of wild type <math>\gamma</math>-secretase</b>	<b>30</b>
4.1. Introduction . . . . .	30
4.2. Method and Materials . . . . .	32
4.3. Result . . . . .	33
4.3.1. Bending motion of the NCT ectodomain and substrate conformational change upon E-S complexation . . . . .	33
4.3.2. E-S interactions stabilize the $\gamma$ -secretase catalytic unit PS1 . . . . .	35
4.3.3. Open and close geometries at the catalytic center . . . . .	39

4.4. Discussion . . . . .	41
<b>5. Influence of Membrane Environment on C83-PSH Interaction</b>	<b>42</b>
5.1. Introduction . . . . .	42
5.2. Methods and Materials . . . . .	43
5.3. Result . . . . .	46
5.3.1. Structural modeling shows key features and active geometry of substrate-bound $\gamma$ -secretase in PSH . . . . .	46
5.3.2. DDM micelle destabilizes the catalytic site geometry . . . . .	46
5.3.3. Increase in flexibility and unwinding of PSH TMD6a helix induced by detergent-enzyme interaction . . . . .	48
5.3.4. Lysine mutations in TMD6a lead to helix unwinding and reduced activity	51
5.4. Discussion and outlook . . . . .	54
<b>6. Hybrid <math>\beta</math>-Sheet between APP and <math>\gamma</math>-Secretase</b>	<b>58</b>
6.1. Introduction . . . . .	58
6.2. Methods and Materials . . . . .	61
6.3. Result . . . . .	64
6.3.1. Stability of the active geometry is sensitive to substrate length . . . . .	64
6.3.2. Sampling the $\beta$ 3-strand association/dissociation by Hamiltonian replica exchange simulations . . . . .	68
6.3.3. APP M51P and L52P mutations decrease $\gamma$ -secretase cleavage efficiency	71
6.3.4. Simulations explain why P1' $\rightarrow$ F enhances while P2' $\rightarrow$ F mutation weakens the formation of the catalytically active geometry . . . . .	74
6.4. Discussion and outlook . . . . .	77
<b>7. An internal docking site stabilizes substrate binding to <math>\gamma</math>-secretase</b>	<b>83</b>
7.1. Introduction . . . . .	83
7.2. Methods and Materials . . . . .	86
7.3. Results . . . . .	86
7.3.1. Weakening the interaction at PS1 internal docking site leads to re-adjustment of substrate helical interface . . . . .	86
7.3.2. Modeling the $\epsilon$ 48 binding pose of C99 . . . . .	88
7.3.3. Local hydrophobic mismatch induced by K28 in A $\beta$ n- $\gamma$ -secretase binding complexes. . . . .	91
7.3.4. APP mutations K28A and G33I stabilize substrate binding and catalytic geometry of A $\beta$ 40 . . . . .	93
7.4. Discussion and outlook . . . . .	96
<b>8. A Putative Binding Mode of <math>\gamma</math>-Secretase Modulator</b>	<b>100</b>
8.1. Introduction . . . . .	100
8.2. Methods . . . . .	102
8.3. Results . . . . .	104

8.4. Discussion . . . . .	111
<b>9. Analysis and Application of the Dual-Binding Mechanism</b>	<b>113</b>
9.1. Introduction . . . . .	113
9.2. Methods and Materials . . . . .	115
9.3. Results and Discussion . . . . .	116
<b>Conclusion and Outlook</b>	<b>128</b>
<b>Acknowledgments</b>	<b>130</b>
<b>Bibliography</b>	<b>132</b>
<b>A. Supplementary Information for Chapter 4</b>	<b>162</b>
A.1. Reconstruction of the Notch1-bound form . . . . .	162
A.2. Supplementary Figures . . . . .	163
<b>B. Supplementary Information for Chapter 6</b>	<b>167</b>
B.1. Supplementary Figures . . . . .	167
<b>C. Supplementary Information for Chapter 7</b>	<b>177</b>
C.1. Supplementary Figures . . . . .	177
<b>D. Supplementary Information for Chapter 8</b>	<b>194</b>
D.1. Supplementary Figures . . . . .	194
<b>E. Supplementary Information for Chapter 9</b>	<b>209</b>
E.1. Supplementary Figures . . . . .	209
<b>List of Publication</b>	<b>209</b>

# 1. Introduction

Amyloid precursor protein (APP) is a type I transmembrane protein, serving as a cell surface receptor and a regulator of synapse formation, neural plasticity, antimicrobial activity, and iron export[1, 2]. Besides its regulatory functions, APP is best known for the pathological role it plays in the development of Alzheimer's disease (AD)[1, 3, 4]. APP is first cleaved by either  $\alpha$ -secretase or  $\beta$ -secretase (BACE1) at the extracellular side, producing a soluble N-terminal product and a membrane-embedded C-terminal product[5, 6, 7, 8]. The C-terminal product upon BACE1 cleavage with 99 amino acids, called C99, is later recognized and trimmed by  $\gamma$ -secretase within the lipid bilayer, generating the APP intracellular domain (AICD) and amyloid- $\beta$  ( $A\beta$ ) peptides of different lengths[9]. Species of  $A\beta$  peptides are conventionally named according to the number of amino acids they consist of. For example,  $A\beta$ 42 contains 42 amino acids and  $A\beta$ 40 contains 40 amino acids. Major species of  $A\beta$  products are  $A\beta$ 43,  $A\beta$ 42,  $A\beta$ 40 and  $A\beta$ 38[9]. According to the amyloid hypothesis, longer species of  $A\beta$  such as  $A\beta$ 42[3, 10] and  $A\beta$ 43[11, 12], due to their hydrophobic nature, are prone to aggregate and form amyloid plaques, which are known to be neurotoxic and are associated with neural degeneration and Alzheimer's disease[13, 14].

While aging is known to be a factor in the accumulation of amyloid plaques, presumably because of the diminished  $A\beta$  clearance[15] and alternation in BACE1[16] and  $\gamma$ -secretase[17] activity, a genetic mutation, also called Familial AD (FAD) mutation, can as well cause early-onset of AD, at the age of 30. FAD mutations are mostly located in the genes of APP and the catalytic subunits of  $\gamma$ -secretase presenilin (PS), of which many lead to the decrease in  $A\beta$  processivity, i.e. the ability to secrete shorter  $A\beta$  products[18, 19]. To reduce the production of  $A\beta$  peptides, PS-targeting  $\gamma$ -secretase inhibitors (GSIs) are developed as a strategy for AD treatment[20, 21]. However, due to the off-targeting effect, GSI inhibits also the cleavages of over 130 other substrates[22], including the Notch receptor, which regulates cell-fate decision[23, 24], and causes severe side effects. Thus,  $\gamma$ -secretase modulators (GSM), which aims to increase the  $A\beta$  processivity without altering the enzymatic activity, have become a promising therapeutic strategy against AD[25, 26].

In order to come up with a more potent therapeutic strategy to reduce the amount of longer forms of  $A\beta$  peptides, it is of fundamental importance to understand the molecular details of how they are generated by  $\gamma$ -secretase. Although extensive wet-lab experiments on  $\gamma$ -secretase investigating their response to mutations and GSM have been reported, the molecular mechanism of these results remains elusive. By performing molecular dynamics (MD) simulations in an atomic resolution, we aim to answer the following four questions:



- 
1. What are the main interactions between  $\gamma$ -secretase and its substrates?
  2. Why does  $\gamma$ -secretase cleave C99 in a step of 3 residues?
  3. What determines the length of the released A $\beta$  peptides?
  4. How do GSMs promote A $\beta$  processivity?

This thesis is split into 10 Chapters and organized as follows:

- An introduction to  $\gamma$ -secretase and A $\beta$  peptides as well as experimental results relevant to the present thesis are given in Chapter 2.
- In Chapter 3, principles and computational methods used in this thesis will be introduced.
- In Chapter 4, the dynamics of the substrate-free (apo-form) and substrate-bound (holo-form)  $\gamma$ -secretase and its interaction with C99 are investigated as the foundation for the following chapters.
- The dynamics of the holo-form presenilin homolog (PSH) within the micelles and bilayer environment are studied in collaboration with an experimental group (Chapter 5).
- In Chapter 6, the functional role of the  $\beta$ -sheet formed between  $\gamma$ -secretase and its substrate C-terminal to the catalytic active site is studied using *in silico* mutation and Hamiltonian replica-exchange coupled umbrella sampling (HREUS) method.
- To identify the structural discrepancies between different stages along the processive cleavage, homology modeling is extensively employed to construct all binding poses along the  $\epsilon 49$ - $\zeta 46$ - $\gamma 43$ - $\gamma 40$ - $\gamma 37$  production line and presented in Chapter 7.
- In Chapter 8, a putative binding of a GSM and how it modulates  $\gamma$ -secretase processivity is reported.
- As a side project during the doctoral research period, a dual-binding mechanism in the protein-protein stabilization process studied with a mathematical model applied to 19 case studies is presented in Chapter 9.
- The thesis is concluded in Chapter 10 by answering the questions raised in this chapter and future perspectives of the study of  $\gamma$ -secretase and molecular design.

## 2. $\gamma$ -secretase and A $\beta$ peptides

Generation of A $\beta$  peptides through  $\gamma$ -secretase intramembrane proteolysis is a complicated process and is believed to be involved in AD pathology. In this chapter, an overview of the function and structure of  $\gamma$ -secretase as well as its interaction with its substrate and small molecules is introduced.

### 2.1. Biological function and subunits of $\gamma$ -secretase

$\gamma$ -Secretase is an aspartyl intramembrane protease that cleaves over 130 type-I single-pass integral membrane proteins[22]. Among several substrates, APP and Notch receptors are the most studied proteins because of their association with the pathology of AD[27, 4, 28] and cancer[23, 24], respectively.  $\gamma$ -Secretase is composed of four subunits —Nicastrin (NCT), presenilin enhancer protein 2 (PEN-2), anterior-pharynx-defective protein 1 (APH-1), and the catalytic subunit presenilin (PS1 or PS2)[29, 30, 31]. Upon the complexation of the four proteins, an autocatalytic cleavage[29], splitting PS into an N-terminal fragment (NTF), containing transmembrane domain (TMD) 1-6, and a C-terminal fragment (CTF), containing TMD7-9, is required for the complex maturation. The autoproteolysis, as well as substrate cleavages, are executed via hydrolysis mediated by the two aspartates located on TMD6 and TMD7, respectively[29]. However, PS cannot conduct its function on its own and needs NCT, PEN-2, and APH-1 to stabilize the topology[32, 33]. In addition, NCT is also suggested to serve as a filter, avoiding transmembrane peptides with large N-terminal domains from accessing PS[34]. Two conserved motifs are observed across the PS family. The GXGD motif on the N-terminus of TMD7 is suggested to be directly involved in catalytic function and substrate identification [35, 36, 37] whereas the PALP motif on the N-terminus of TMD9 is suggested to play a critical role in complex stabilization and proteolytic activity[38, 39, 40].

### 2.2. 3D-Structures of $\gamma$ -secretase resolved by X-ray crystallography and cryo-EM

To understand the biological functions and facilitate the development of  $\gamma$ -secretase-targeting drugs, it is of significant importance to reveal the tertiary structures of NCT, PEN-2, APH-1, and PS and how they assemble together within the lipid bilayer in both substrate-free (apo) and substrate-bound (holo) form. However, due to the crystallization complexity of membrane protein and the large molecular weight, the atomic structure of  $\gamma$ -secretase was for a long time undetermined despite its important role in AD pathology[41, 42]. In 2013, a series of presenilin homolog (PSH) structures were crystallized and resolved with an X-ray diffraction

## 2.2. 3D-STRUCTURES OF $\gamma$ -SECRETASE RESOLVED BY X-RAY CRYSTALLOGRAPHY AND CRYO-EM

technique (Figure 2.1A-C), offering the first structural information about PS[43]. As predicted previously for human PS, PSH contains 9 TMDs and the catalytic aspartates D162 and D220 are located on TMD6 and TMD7, respectively[29, 44, 45]. Two years after, the 3D structure of a GSI-bound form PSH was also resolved with no significant conformational difference to the ligand-free PSH (Figure 2.1D)[46]. The GSI L-682,679 located between TMD6 and TMD7 indicated an inhibitory mechanism by the separation of the two catalytic aspartic acids. Note, tetramerization of PSH was observed in all available structures with TMD3 inserted into the TMD2-TMD6 gap and the TMD2 and TMD3 are oriented toward the tetrameric interior (Figure 2.1E). This conformation, however, blocks all possible substrate entries if PSH adopts the same substrate binding mode as PS1-containing  $\gamma$ -secretase (Figure 2.2G, H), and is likely not the native state where PSH conducts its biological function.

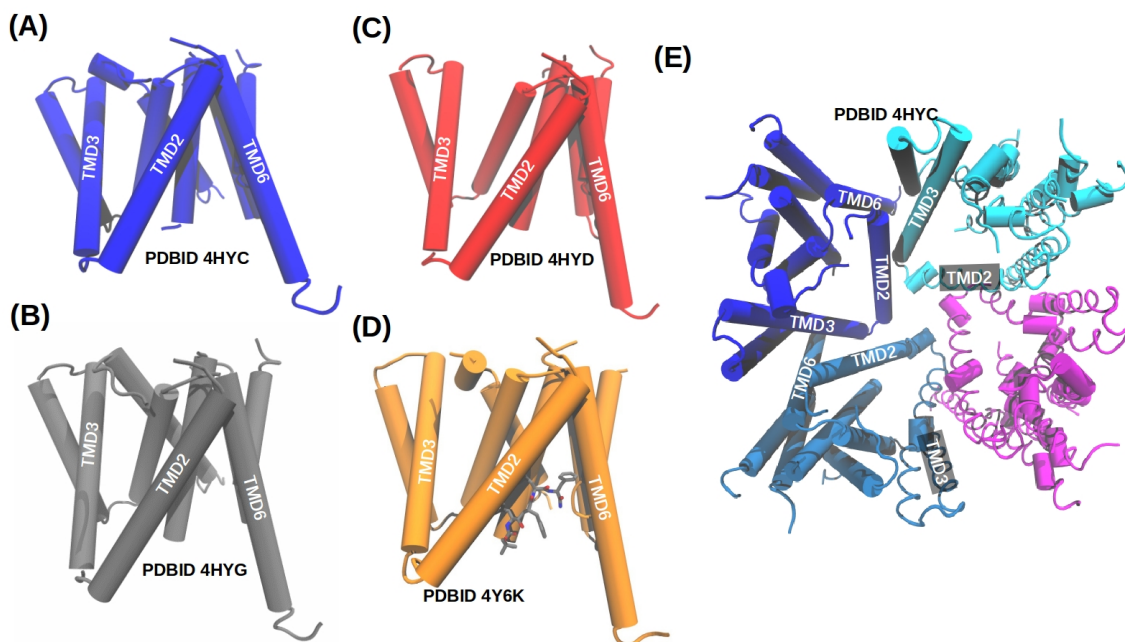


Figure 2.1.: Structures of PSH resolved with X-ray diffraction. (A) Structure of PSH in P2 space group. (PDBID 4HYC) (B) Structure of PSH in C2221 space group. (PDBID 4HYD) (C) Structure of PSH in C222 space group. (PDBID 4HYG) (D) Structure of PSH bound with L-682,679. (PDBID 4Y6K) (E) Tetramer conformation of PSH. (PDBID 4HYC)

With the rapid advances in cryogenic electron microscopy (cryo-EM), structures of large membrane proteins are gradually resolvable in recent years[47]. However, the first few attempts to uncover the 3D structure of the  $\gamma$ -secretase complex only yield low resolution and no converged conformations were reached. The first cryo-EM-based  $\gamma$ -secretase structure was reported in 2014, showing a horseshoe-shaped transmembrane domain and a large NCT ectodomain[48, 49, 50, 51, 52]. However, the resolution of the image was too low to

## 2.2. 3D-STRUCTURES OF $\gamma$ -SECRETASE RESOLVED BY X-RAY CRYSTALLOGRAPHY AND CRYO-EM

be deposited in the Protein Data Bank. In 2015, Sun et. al. and Bai et. al. resolved six 3D structures of the human apo-form PS1-containing  $\gamma$ -secretase with an atomic resolution (Figure 2.2A-F)[53, 54, 55]. These structures were able to distinguish the TMDs from NCT, PEN-2, APH-1, and the NTF and CTF of PS1 for the first time. However, the C-terminal part of TMD6, the C-terminal part of TMD2, and the N-terminal part of TMD3 were poorly resolved without the presence of ligand, indicating their high mobility in the apo-form and stabilization effect upon ligand binding.

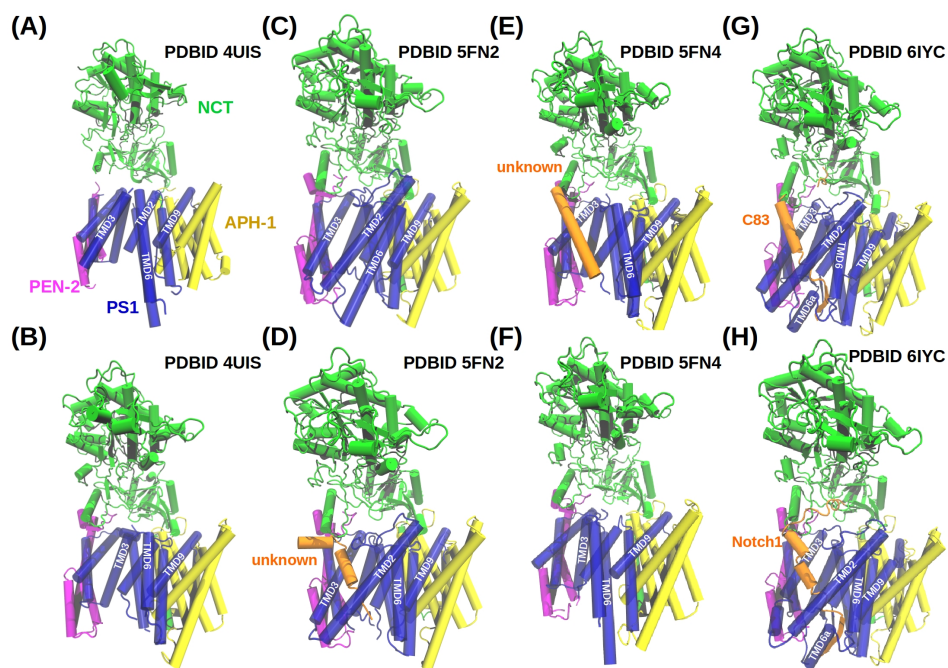


Figure 2.2.: Structures of PS1-containing  $\gamma$ -secretase in apo-form and holo-form. (A) Structure of the apo-enzyme (PDBID 4UIS). The T4 lysozyme fused to the N-terminal of PS1 is not shown. (B) Structure of the apo-enzyme (PDBID 5A63). (C) Structure of the DAPT-bound enzyme with (PDBID 5FN2). The coordinates of DAPT are not determined. (D) Structure of a helix-bound enzyme (PDBID 5FN3). The sequence of the bound helix is unknown. (E) Structure of a helix-bound enzyme (PDBID 5FN4). The sequence of the bound helix is unknown. (F) Structure of an apo-enzyme (PDBID 5FN5). (G) Structure of the C83-bound enzyme (PDBID 6IYC). (H) Structure of the Notch1-bound enzyme (PDBID 6IDF). Subunits of  $\gamma$ -secretase are colored green (NCT), yellow (APH-1), magenta (PEN2), and blue (PS1). The bound peptides are colored in orange and their biological names are noted next to them.

Four years later, excellent work done by Zhou. et. al.[39] and Yang et. al.[40] revealed the binding mode of C83, the C-terminal product of APP upon  $\alpha$ -secretase cleavage, and Notch1 to  $\gamma$ -secretase, providing for the first time the structural information about the in-

teraction between  $\gamma$ -secretase and its substrate (Figure 2.2 G, H). In order to maintain the enzyme-substrate(E-S) complex without cleavage, the proteolytic activity was impaired by PS1 mutation D385A. In addition, cysteine was introduced in loop1 of PS1-NTF (Q112C), Notch1 (P1728C), and C83 (V24C) to stabilize the E-S complex by disulfide bonds. Importantly, they revealed a similar docking pose shared by both C83 and Notch1 with a helix-loop-strand conformation at the TMD, which largely differed from the previously resolved NMR structures of APP (PDBID 2LLM)[56] and Notch1 (PDBID 5KZO)[57], and the substrate cleavage site situated at the loop-strand transition region. Furthermore, a small helix shortly following TMD6, called TMD6a (L268-E277), is formed and stays in contact with the substrate P1' and P3' residue. The substrate  $\beta$ -strand, called  $\beta$ 3, is stabilized by the formation of a hybrid  $\beta$ -sheet with Y288-S290 from PS1-NTF ( $\beta$ 1), R377-L381 from PS1-CTF ( $\beta$ 2) and PS1 L432. These components together form an E-S hybrid  $\beta$ -sheet and are shown to be indispensable for the enzymatic activity of  $\gamma$ -secretase[39, 40].

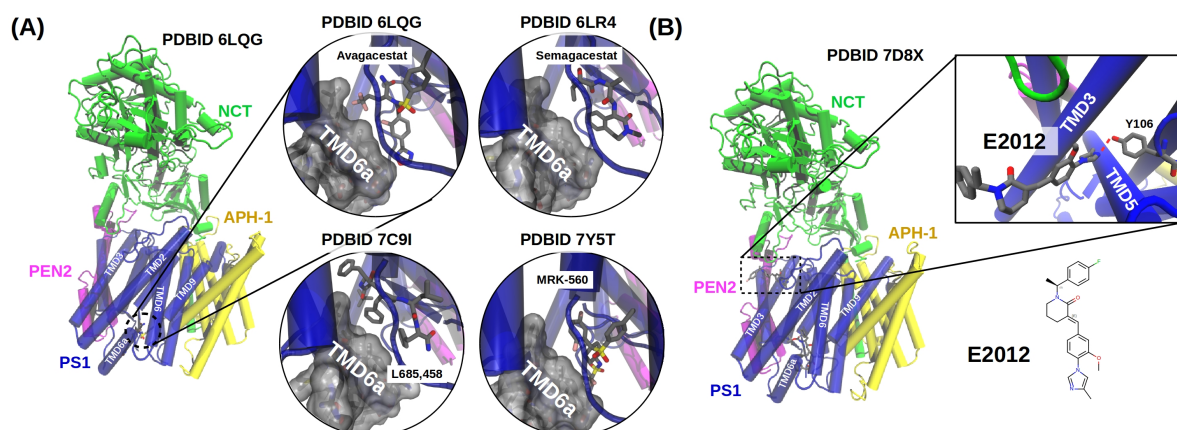


Figure 2.3.: Structures of the (A) GSI-bound and (B) GSM-bound  $\gamma$ -secretase complexes. Color-coding of the  $\gamma$ -secretase subunits is the same as in Figure 2.2. The white surface area shown in (A) represents the van der Waals surface of TMD6a of PS1.

Recently, the cryo-EM structures of  $\gamma$ -secretase bound with GSI and GSM are also resolved[58, 59]. All GSIs consistently bind at the region where the substrate  $\beta$ 3 is located, implicating a catalytically important function this region conducts. Additionally, they all contact with the hydrophobic patch formed by PS1 TMD6a (Figure 2.3A). The GSM- $\gamma$ -secretase cryo-EM structure reveals a GSM binding pocket located 25Å away from the catalytic center[40]. The binding pocket is formed by loop1 and TMD3 of PS1 and I243 and N243 of NCT (Figure 2.3B). In particular, a hydrogen bond was found between the GSM and PS1 Y106, which was reported to play an important role in the binding of a wide range of imidazole-based GSIs. However, the coordinate of the GSM revealed an overlay with the coordinate of APP and Notch solved in 2019[39, 40] (Figure 2.2F, G), leaving the modulation mechanism of GSM obscure.

### 2.3. $A\beta$ PEPTIDES SECRETION THROUGH INTERACTIONS BETWEEN C99 AND $\gamma$ -SECRETASE

The first 3D structure of PS2-containing  $\gamma$ -secretase was solved in 2022 by Guo et. al (Figure 2.4)[59]. Unlike in the PS1-containing isoform, (Figure 2.3A), the electron density of MRK-560 could not be observed in the PS2-containing  $\gamma$ -secretase in the co-incubation experiment, showing its inhibitory selectivity. Similar to the structures of PSH and the PS1-containing  $\gamma$ -secretase, TMD2, and TMD6 are too mobile to be determined without the presence of ligands.

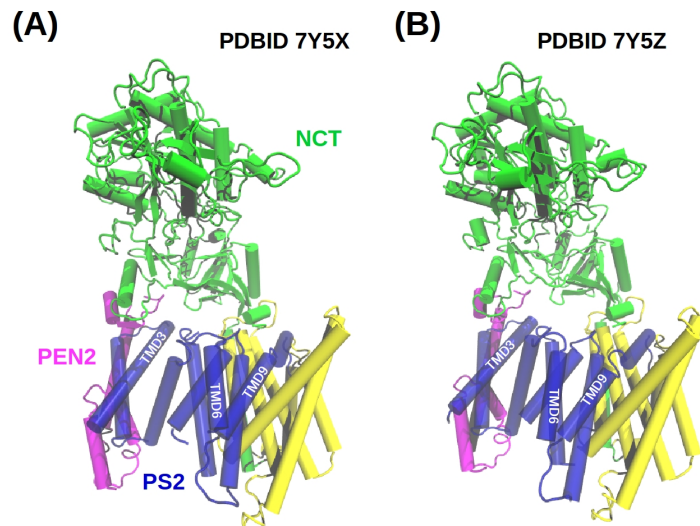


Figure 2.4.: Structures of the PS2-containing  $\gamma$ -secretase complexes. Color-coding of the  $\gamma$ -secretase subunits is the same as in Figure 2.2.

### 2.3. $A\beta$ peptides secretion through interactions between C99 and $\gamma$ -secretase

After the extracellular domain of APP is shedded by  $\beta$ -secretase[5, 6], the C-terminal product C99 remains still within the membrane bilayer and is processed by  $\gamma$ -secretase to secrete  $A\beta$  peptides in different length through a series of interactions (Figure 2.5)[9]. Although the detailed C99- $\gamma$ -secretase interaction pathway is not yet completely understood, it can be briefly discussed in three consecutive stages: substrate recognition, substrate positioning, and processive cleavage.

#### 2.3.1. Substrate Recognition

Because the intramembrane protease  $\gamma$ -secretase and its substrates are integrated into the membrane environment, the enzyme-substrate (E-S) recognition is restricted to a two-dimensional space with a little degree of freedom of helix tilting or bending. The NCT

### 2.3. $A\beta$ PEPTIDES SECRETION THROUGH INTERACTIONS BETWEEN C99 AND $\gamma$ -SECRETASE

subunit of  $\gamma$ -secretase at the first stage is suggested to serve as a filter that sterically hinders substrates with a large ectodomain from interacting with  $\gamma$ -secretase and only the substrates which are not excluded by NCT can approach  $\gamma$ -secretase and form possible contacts[34]. A comprehensive study using the photo-affinity labeling technique suggests that before C99 reaches the catalytic center of PS, it binds to several exosites of  $\gamma$ -secretase subunits with an order of NCT/PEN2  $\rightarrow$  PS1-NTF  $\rightarrow$  PS1-NTF/PS1-CTF[60]. A computational work using coarse grained simulation also suggested PEN2 as a non-selective binder that binds both substrates and non-substrates of  $\gamma$ -secretase but only substrates can be transferred and interact with PS-NTF[61].

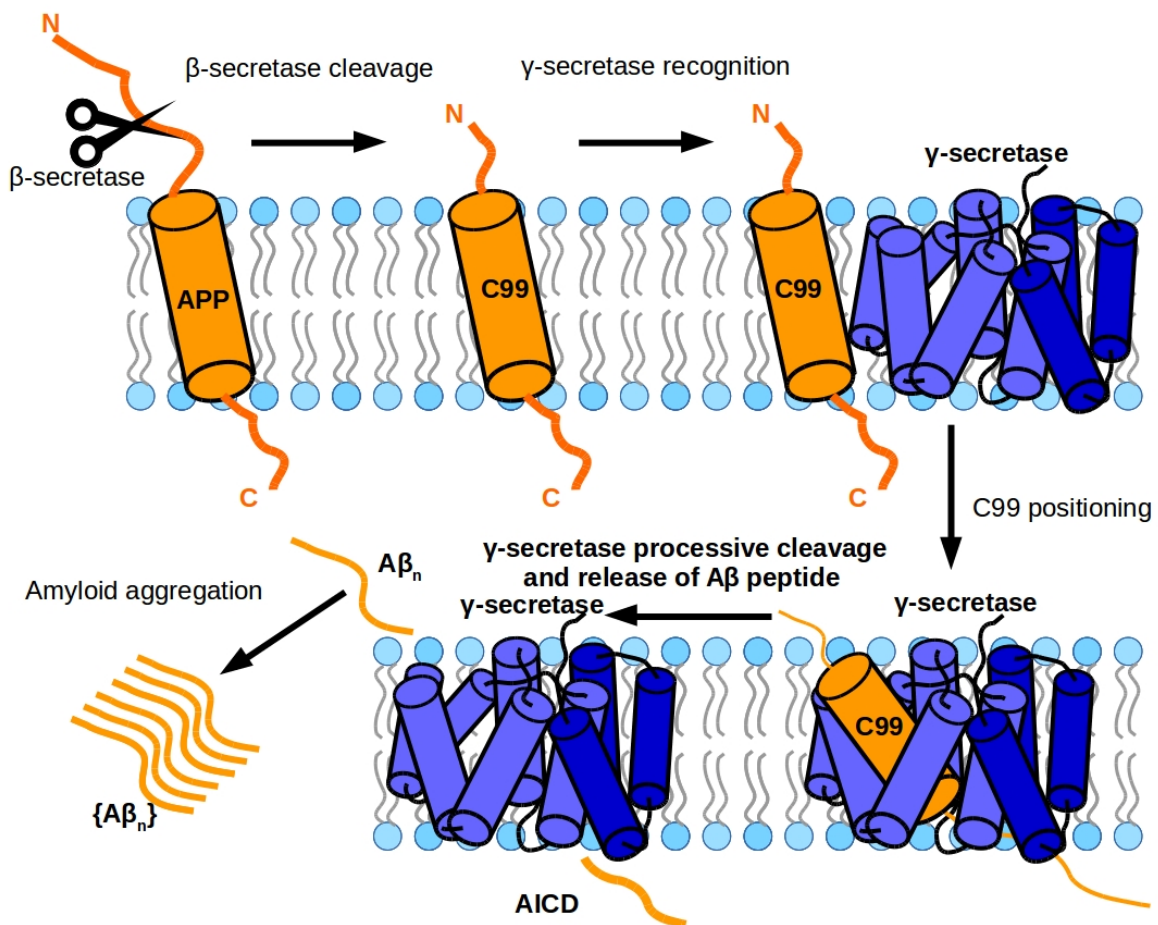


Figure 2.5.: Schematics of the process of APP processing by  $\beta$ -secretase and  $\gamma$ -secretase inside the membrane. Eventually, the  $A\beta$  products can form the neurotoxic amyloid aggregates in the solution.

### 2.3.2. Substrate Positioning

Although the cryo-EM structures of substrate-bound  $\gamma$ -secretase have already provided great insight into the E-S interaction[39, 40], the binding pose they revealed did not answer how substrates are recruited from the  $\gamma$ -secretase exosites to the PS interior. Initially, TMD2-TMD3, TMD2-TMD6, and TMD6-TMD9 were proposed to serve as the substrate entry through which the cleavage site can move to the catalytic center[62]. The possibility of the TMD6-TMD9 entry became implausible after the substrate-bound  $\gamma$ -secretase was resolved in 2019. However, the other two entries also face structural challenges. If the resolved C83-bound and Notch-bound cryo-EM structures, with several artificially introduced mutations, reflect the actual ready-for-cleavage binding mode, substrates need to bring the whole intracellular domain (more than 50 amino acids in C83 and Notch) over the short loop between TMD2 and TMD3 in order to enter from the TMD2-TMD3 gate or bend the whole extracellular domain (more than 20 amino acids in C83 and Notch) over PS1 loop1 to enter from the TMD2-TMD6 gate to end up in the binding position(Figure 2.2 G, H). Both models remain possible due to the high mobility of loop1 and TMD2 in the apo-form PS1 and PS2 and it is yet unclear whether the N-terminal and C-terminal domain of TMD2 can unwind and extend the loop1 or the TMD2-TMD3 loop to facilitate the putative substrate entry sites. Experimental works have also suggested a two-step substrate positioning[63, 64]. That is, a bending motion also plays a critical role in the substrate entering in addition to the translational helix movement after the initial binding. Despite the unclear PS recruitment process, the 3D structures of holo-form PS1-containing  $\gamma$ -secretase offer an excellent starting point for computational studies and interpretations of experimental findings[39, 40]. Both C83 and Notch share a highly similar contact pattern with PS1, including TMD6a hydrophobic patch (substrate P1', P3' using the general nomenclature formulated by Schechter and Berger[65]), hybrid  $\beta$ -sheet (substrate P1'-P4'), and PS1 internal docking site (substrate P5, P6). The present thesis reports extensive computational work focusing on these interacting regions and how APP mutations and lipid environment influence these interactions in Chapter 4, Chapter 5, Chapter 6, and Chapter 7.

### 2.3.3. Processive Cleavage

After the substrate is positioned to the read-to-cleave binding pose, presumably the binding pose presented in the holo-form  $\gamma$ -secretase cryo-EM structure (Figure 2.2G), it is cleaved processively by  $\gamma$ -secretase[66, 67] and the ability of  $\gamma$ -secretase to successively trim the A $\beta$  peptide is typically termed the processivity. Since this is the most critical step in producing the AD-related A $\beta$  peptides, many studies have put their focus on how mutations on APP and  $\gamma$ -secretase influence cleavage in terms of processivity and total activity. An overview of processive cleavage is depicted in Figure 2.6. The first cut of  $\gamma$ -secretase on C99 usually occurs at the amide bond between residue 49 and 50 ( $\epsilon$ 49 cleavage), generating A $\beta$  49 and AICD 50-99, or at the amide bond between residue 48 and 49 ( $\epsilon$ 48 cleavage), generating A $\beta$  48 and AICD 49-99[68, 69]. While the soluble C-terminal product AICD is released to the intracellular side upon the first cleavage, A $\beta$ 49 or A $\beta$ 48 is further cleaved into shorter products. Based on the initial  $\epsilon$  cleavage, the following cleavages follow two production



lines, namely the  $\epsilon 49$ - $\zeta 46$ - $\gamma 43$ - $\gamma 40$ - $\gamma 37$  one and the  $\epsilon 48$ - $\zeta 45$ - $\gamma 42$ - $\gamma 38$  one[70, 67]. Therefore, the position of the initial cleavage has a substantial influence on the species of A $\beta$  products. Unfortunately,  $\gamma$ -secretase does not always trim C99 until the end of each production line and the intermediates can also be secreted to the solution. In the *in vitro* experiments, AICD, and A $\beta$  species are commonly identified by mass spectroscopy where molecules with different mass-to-charge ratios ( $m/z$ ) are identified. Usually,  $\gamma$ -secretase activity is measured by measuring the amount of AICD or total A $\beta$  products. Note that the activity measured here depends on not only the efficiency of processive cleavage but also all preceding steps, including substrate recognition and positioning. On the other hand,  $\gamma$ -secretase processivity is commonly presented by the ratio between the shorter A $\beta$  species and longer A $\beta$  species, for example: A $\beta$ 40/A $\beta$ 43, A $\beta$ 38/A $\beta$ 42, or (A $\beta$ 37+A $\beta$ 38+A $\beta$ 40)/(A $\beta$ 42+A $\beta$ 43). Since the A $\beta$ 40 and A $\beta$ 42 are the most abundant species, the processivity is also represented by the A $\beta$ 40/A $\beta$ 42 ratio[71]. Notably, evidence has suggested that the total activity of  $\gamma$ -secretase does not always correlate with its processivity.[72]

#### 2.3.4. Effects of mutations on APP and PS1

Typically, the major products of  $\gamma$ -secretase cleavage toward C99 are composed of majorly A $\beta$ 40 (~80%), A $\beta$ 42 (~10%), and A $\beta$ 38 (~10%) and minorly A $\beta$ 37 (~1%) and A $\beta$ 43 (~1%)[9]. However, the composition can be altered by a single-point mutation on PS1 or APP. In particular, the mutations that lead to the early onset of AD, called FAD mutations, are shown to predominantly increase the longer, pathogenic peptides A $\beta$ 42 and A $\beta$ 43, and decrease the total  $\gamma$ -secretase activity[73, 74, 75, 71]. An overview of the FAD position on APP, PS1, and PS2 is well-organized on [alzforum.org/mutations](http://alzforum.org/mutations). Thermoactivity experiment suggests that while the E-S stability becomes progressively less stable along the processive cleavage the pathogenic PS1 and APP mutations further destabilize the E-S complex and lead to an early dissociation of the longer form A $\beta$  peptide[76]. Other data indicate that several FAD mutations on PS1 can directly destabilize the active site geometry[77, 78]. In comparison to the mutations on PS, the mechanism of APP is less clear because the same residue might play different roles at the different phases along the  $\gamma$ -secretase recognition, initial cleavage, and processive cleavage processes. For instance, the increase in the A $\beta$ 42 product can be induced by directly shifting the initial cleavage from  $\epsilon 49$  to  $\epsilon 48$  or inhibiting the  $\beta 42 \rightarrow$  A $\beta$ 38 cleavage. The hybrid effect can be best exemplified by APP mutation V44F and I45F[66]. V44F prefers the  $\epsilon 48$  cleavage but results in a low A $\beta$ 42/A $\beta$ 40 ratio presumably by inhibiting the A $\beta$ 45  $\rightarrow$  A $\beta$ 42 cleavage whereas I45F prefers the  $\epsilon 49$  cleavage but results in a high A $\beta$ 42/A $\beta$ 40 ratio presumably by inhibiting the A $\beta$ 46  $\rightarrow$  A $\beta$ 43 cleavage. Therefore, understanding where each residue locates and interacts with  $\gamma$ -secrease at each processing step is fundamentally crucial to interpreting the experimentally measured AICD and A $\beta$  profiles. To date, however, only the size-limiting S1'-S3' pocket is a well-accepted structural model that can be used to explain the cleavage behavior of a wide range of Phenylalanine-containing APP mutants[66].

2.3.  $A\beta$  PEPTIDES SECRETION THROUGH INTERACTIONS BETWEEN C99 AND  $\gamma$ -SECRETASE

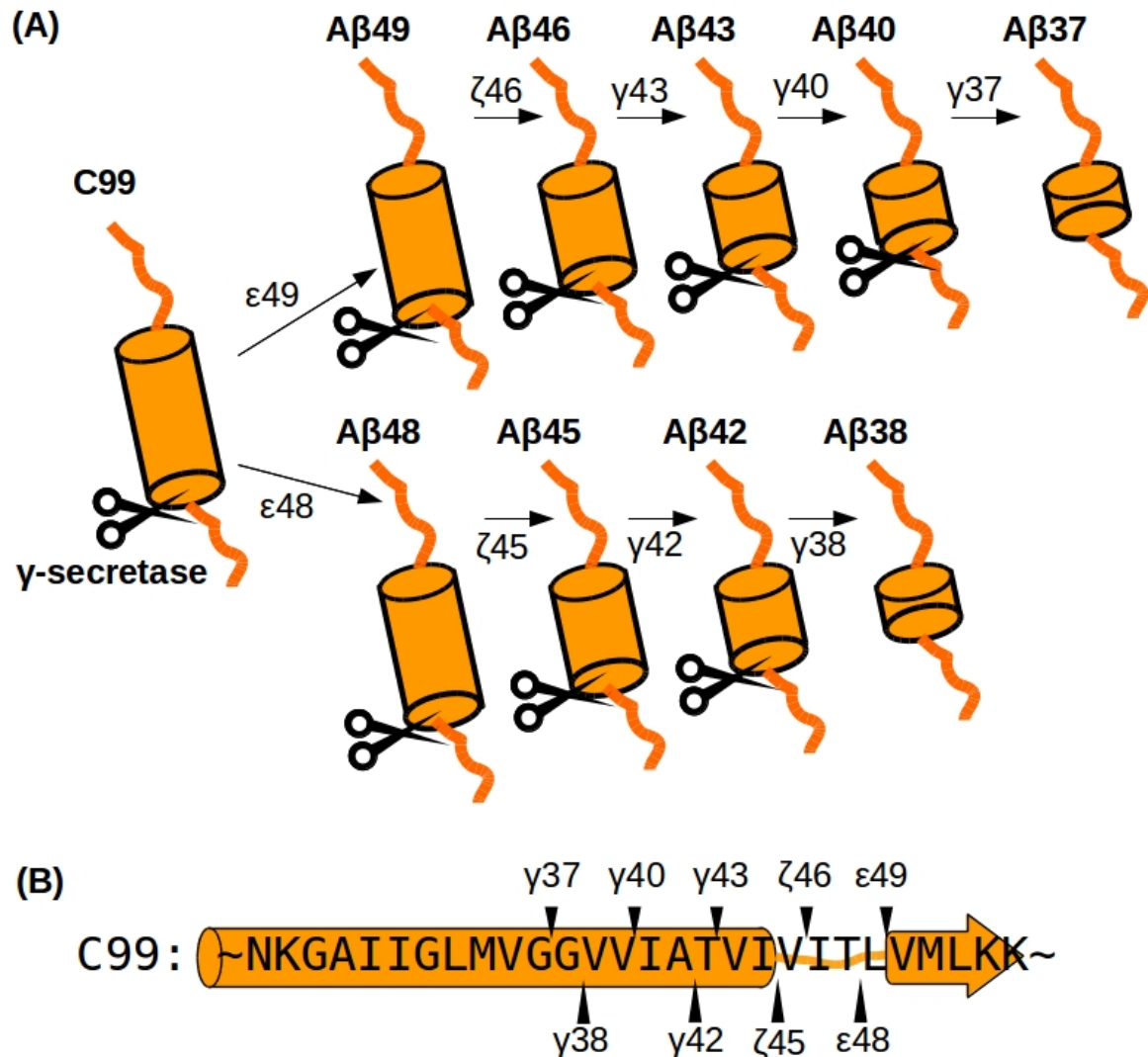


Figure 2.6.: Schematic of the processive cleavage of  $\gamma$ -secretase against C99. (A) Two production lines with the  $\epsilon_{49}$ - $\zeta_{46}$ - $\gamma_{43}$ - $\gamma_{40}$ - $\gamma_{37}$  one in the upper row and the  $\epsilon_{48}$ - $\zeta_{45}$ - $\gamma_{42}$ - $\gamma_{38}$  one in the lower row. (B) The sequence of C99 from N27 to K54. The orange cartoon represents the secondary structure of helix, loop, and  $\beta$ -strand shown in the enzyme-bound form (Figure 2.2G). Two production lines are drawn upper and lower to the sequence. The cleavage position at each cleavage is depicted with a black triangle sign.

## 2.4. Inhibition and Modulation of $\gamma$ -Secretase as Therapeutic Strategies against Alzheimer's Disease

According to the amyloid hypothesis, aggregation of the longer forms of A $\beta$  peptides is related to the AD-causing senile plaques[4, 28]. Reducing the generation of longer A $\beta$  species is therefore believed to be an effective way to cure AD or alleviate AD progress and small molecules  $\gamma$ -secretase inhibitor (GSI) and  $\gamma$ -secretase modulator (GSM) are designed to achieve such a goal. Several comprehensive reviews have covered critical progress and detailed comparison of existing GSI and GSMs. As an introduction[20, 26, 25, 21], only a selected set of small molecules are discussed in this section.

### 2.4.1. $\gamma$ -Secretase Inhibitor

One of the most intuitive ways to reduce the generation of catalytic products is to inhibit the correspondent protease[79]. At the time of writing this thesis, 3D structures of four GSIs have been determined, including GSI transition state analog (TSA) L458, non-selective GSI Semagacestat, Notch-sparing GSI Avagacestat[40], and PS2-Sparing GSI MRK-560[59]. Based on the structures, all GSIs form hydrophobic contact with the TMD6a and hydrogen bonds with L432 and/or  $\beta$ 2 of PS (Figure 2.3A). Although another non-selective GSI DAPT was also attempted for imaging (PDBID 5FN2), the electron density was too low to determine its actual coordinate[55]. TSA inhibitors are the inhibitors that directly mimic the substrate cleavage site with an uncleavable motif[80, 81]. All  $\gamma$ -secretase TSAs inhibitors contain a transition-state-mimicking hydroxyl isostere followed by a tripeptide fragment binding at the size-selecting PS S1'-S2'-S3' subpockets[66]. It has been shown that the inhibition efficiency depends highly on the sizes of the amino acids that fit these subpockets. While the binding pose of TSA overlays perfectly with both the backbone and sidechains of substrate P1-P3', non-TSA GSIs do not occupy the catalytic center[58]. However, severe side effects are observed because of the off-targeting effects of non-selective GSIs[82, 83, 84, 79]. In particular, inhibition of cleavage of the Notch receptor through non-selective GSIs is known to cause gastrointestinal, immune, and cutaneous adverse effects[82]. Notch-sparing TSIs such as avagacestat thus offer new potentials to reduce A $\beta$  secretion with low toxicity[85]. Nonetheless, the structural factors that differentiate Notch-sparing TSIs from non-selective TSIs still remain elusive, establishing barriers to their further development. While several small molecule GSIs have been submitted to clinical trials, none of them has been approved for AD treatment. In particular, clinical trials on semagacestat and avagacestat were terminated at phase III and phase II, respectively, due to their increased risk of skin cancer and cognitive worsening[83].

Besides small molecule inhibitors that target the catalytic center, helical peptide inhibitors (HPI) were also developed to block the substrate from binding to  $\gamma$ -secretase[86, 87]. The initial design of helical inhibitors makes use of aminoisobutyric acid (Aib), a  $3_{10}$ -helix-prone amino acid[88, 89], and was thought to bind at the initial surface binding site. In the later competition experiment with TSA-GSI, it became more convincing that HPIs bind to PS1 interior docking site. Furthermore, the inhibitor concatenating HPI with TSA GSI reaches a sub-nanomolar inhibition concentration ( $K_i = 0.42$  nM) and is competitive with both TSA GSI

and HPI, suggesting a similar binding mode to the APP-bound cryo-EM structure[86].

### 2.4.2. $\gamma$ -Secretase Modulator

With the failure of GSI, other opportunities to reduce the amyloid aggregates are in demand. Since the longer forms of A $\beta$  peptides are more closely related to the formation of neurotoxic amyloid plaques[3, 10, 11, 12], enhancing  $\gamma$ -secretase processivity is also considered a therapeutic strategy against AD. The first generation of GSMs is discovered in early 2001[90]. A subset of nonsteroidal anti-inflammatory drugs (NSAIDs) such as ibuprofen, indomethacin, and sulindac sulfide impede the secretion of A $\beta$ 42 peptide while increasing the A $\beta$ 38 form without interfering with the Notch1 cleavage. However, these compounds suffer from weak A $\beta$ 42 inhibitory potencies ( $IC_{50} > 10\mu M$ ) and poor brain penetration performance[91, 92]. Second-generation GSMs are developed to increase the A $\beta$ 42 inhibitory potency and brain availability. According to the structural similarity, second-generation GSMs can be categorized into three classes: NSAID-derived carboxylic acid GSMs, imidazole-based GSMs, and natural product-derived GSMs.

NSAID-derived carboxylic acid GSMs, exemplified by GSM-1, exclusively decrease the A $\beta$ 42 peptides without much interfering with the A $\beta$ 40 products with a significantly improved cell-based A $\beta$ 42 inhibitory potency ( $IC_{50} < 500$  nM)[93, 94]. Interestingly, modifying the carboxylic acid to an ester or amide shows an inverse action on the  $\gamma$ -secretase processivity. That is, the concentration of A $\beta$ 42 is increased whereas the concentration of A $\beta$ 38 is decreased. Although several studies have reported how mutations on APP[95, 96], PS1[97, 96], and PS2[98] respond to the modulation effect of GSM-1, the binding mode as well as the mechanism of modulation remain unclear. In contrast, imidazole-based GSMs decrease the concentration of both A $\beta$ 42 and A $\beta$ 40 ( $IC_{50} < 100$ nM)[99]. Imidazole-based GSM is exemplified by E2012 (Figure 2.3B) which entered the first clinical trial in 2006 and showed a dose-dependent reduction in A $\beta$ 42 and A $\beta$ 40 level[100]. The binding pose of E2012 on PS1-containing  $\gamma$  secretase was revealed by cryo-EM in 2021, where the imidazole group forms a stable hydrogen bond with Y106 of PS1 (Figure 2.3B), paving a promising road for further structure-based drug design[58]. However, the cryo-EM structures of E2012 and C83 show an overlay in their coordinates, suggesting either E2012 or the substrate needs to undergo conformational adaption along the processive cleavage[40, 39]. Another unresolved question is the mechanism of the modulation effect. Surprisingly, the modulation effect of E2012, but not GSM-1 is inverted when K28 of APP is substituted into a negatively charged glutamic acid[101]. Similarly, although no residues on loop1 of PSH were identified with a similar biophysical property as Y106 of PS1, E2012 showed a strong modulation effect in PSH whereas GSM-1 did not [46]. Indeed, distinct allosteric binding sites of GSM-1 and E2012 are suggested by photoaffinity probing experiments[102]. Taken together, the NSAID-derived and imidazole-based GSMs might have distinct binding sites and modulate the processivity of  $\gamma$ -secretase with different mechanisms. Natural product-derived GSMs exhibit a distinct modulation profile where the levels of A $\beta$ 42 and A $\beta$ 38 are decreased and the levels of A $\beta$ 39 and A $\beta$ 37 are increased[103]. Nonetheless, this class of GSM is lack consistent chemical moiety and is not as well studied as GSMs from the other two classes.

## 3. Theory and Methods

In the present dissertation, different computational approaches were implemented to study the biomolecular systems of interest, including molecular dynamics (MD) simulations, molecular docking homology modeling, and pocket detection techniques. However, since these computational methods were developed to model real-world behavior of biomolecular systems, it is of fundamental importance to understand what assumptions are made when exploiting these methods. In this chapter, the fundamental theories of MD simulations and several routines and algorithms used in the following chapters are introduced.

### 3.1. Molecular Dynamics simulation

#### 3.1.1. From quantum mechanism to classical mechanics

In reality, interactions between atoms can be most accurately described by quantum mechanics (QM). In the scheme of QM, the probability to find a particle at position  $\vec{r}$  at time  $t$  can be described by the absolute square of its wave function  $|\Psi(\vec{r},t)|^2$ . The time-dependent wave function obeys the Schrödinger equation

$$\hat{H}\Psi(\vec{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\vec{r},t) \quad (3.1)$$

, where  $\hat{H}$  is the Hamiltonian of the system,  $i$  is the imaginary unit, and  $\hbar$  is the reduced Plank constant. However, calculating all atomic interactions from first principle requires computational resources far beyond the current hardware limit. Since the nuclei are much heavier than the electrons, in practice, the wave functions of atomic nuclei and electrons can be treated separately (the Born-Oppenheimer approximation[104]) so that the total wave function  $\Psi_{total}$  can be expressed as

$$\Psi_{total} = \psi_{electron}\psi_{nuclear} \quad (3.2)$$

. By approximating that the movement of electrons is much faster than the nuclei, the time-dependent Schrödinger equation can thus be solved by treating the nuclei as fixed and the electronic state is only determined by the configuration of the nuclei  $\psi_{nuclear}(\vec{r},t)$  so that the electron part of the Schrödinger equation can be re-written as

$$\hat{H}\psi_{electron}(\vec{r}) = [\hat{T} + \hat{V} + \hat{U}]\psi_{electron} \quad (3.3)$$

where  $\hat{H}$  is the Hamiltonian, composed of  $\hat{T}$  the kinetic energy,  $\hat{V}$  the external field due to the nuclei, and  $\hat{U}$  the electron-electron repulsive interaction. While this equation can

be solved numerically by quantum chemistry methods such as Hartree-Fock[105, 106, 107], post-Hartree-Fock[108], or density functional theory (DFT[109]) with a selection of basis sets, pure QM simulation is extremely slow (in a timescale of femtoseconds per day for less than 10 atoms) and is usually used for single-point calculation and geometry optimization. Although QM calculation combined with molecular mechanics (QM/MM[110]) enables simulation with a larger system with higher efficiency, one usually reaches a few picoseconds per day even with GPU-accelerated simulation engines.

Since the majority of molecular interactions of interest, except for catalysis, do not require the formation or breaking of chemical bonds, it is usually enough to treat the electrons around an atomic nucleus as a point charge fixed at the center of the atom. With this approximation, the dynamics of a biomolecular system can be simulated using classical mechanics described by the Newtonian equation of motion

$$\vec{F}_i = m_i \frac{d^2 \vec{r}_i}{dt^2} \quad (3.4)$$

where  $\vec{F}_i$  is an external force action on a particle with mass  $m_i$  with a Cartesian coordinate  $\vec{r}_i$ . Molecular dynamics (MD) is a computer simulation method that numerically solves the Newtonian equation of motion (equation 3.4) with interatomic forces described by forcefields (see Section 3.1.2). By constructing a simulation box, the biomolecules of interest are solvated in its biological environments such as water solution or lipid environment (see Section 3.1.3). By iteratively solving the Newtonian equation of motion(see Section 3.1.4), the dynamics of the particles within the simulation box can be captured.

### 3.1.2. Force field

In MD simulations, classical approximations are applied and the interactions between atoms can be described by pair-wise interactions and the energy terms can be expressed as a functional form of the sum of individual potential energy terms, including the bonded terms and non-bonded terms

$$U_{\text{total}} = U_{\text{bonded}} + U_{\text{nonbonded}} \quad (3.5)$$

. Although the functional forms of the energy terms are slightly different in different MD engines, they usually adopt a similar form and their parameters are usually inter-transferable. In the Assisted Model Building with Energy Refinement (AMBER[111]) MD engine [111], the bonded functional terms are formulated by

$$\begin{aligned} U_{\text{bonded}} &= U_{\text{bond}} + U_{\text{angle}} + U_{\text{dihedral}} \\ &= \sum_{r \in \text{bonds}} K_r (r - r_{eq})^2 + \sum_{\theta \in \text{angles}} K_\theta (\theta - \theta_{eq})^2 + \sum_{\phi \in \text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \end{aligned} \quad (3.6)$$

. The first two terms describe the stretching energy of a covalent bond between two atoms at length  $r$  and the bending energy between two covalent bonds between three atoms at angle  $\theta$  using harmonic potentials with force constant  $K_r$  and  $K_\theta$ , equilibrium length  $r_{eq}$  and

equilibrium angle  $\theta_{eq}$ , respectively. The third term describes the potential of the dihedral angle  $\phi$  between two planes formed by four neighboring atoms. The potential is approximated as a truncated Fourier series with energy barrier height  $V_n$ , phase factor  $\gamma$ , and multiplicity  $n$ . The non-bonded functional terms are formulated by

$$\begin{aligned}
 U_{\text{non-bonded}} &= U_{\text{electrostatic}} + U_{\text{van der Waals}} \\
 &= \sum_{i<j} \frac{q_i q_j}{\epsilon R_{ij}} + \sum_{i<j} 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{R_{ij}} \right)^6 \right] \quad (3.7)
 \end{aligned}$$

. The electrostatic interaction between two particles is described by Coulomb's law with particle charges  $q_i$ ,  $q_j$ , particle distance  $R_{ij}$  within a continuum with dielectric constant  $\epsilon$ . The Lennard-Jones potential is used as an approximate model to describe the van der Waals contributions [112] between two particles with a potential well depth  $\epsilon_{ij}$  and van der Waals radius  $\sigma_{ij}$ . In practice,  $\epsilon_i$  and  $\sigma_i$  are assigned to each atom type, and  $\epsilon_{ij}$  and  $\sigma_{ij}$  can be calculated as the geometrical mean  $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$  (Berthelot rule), and arithmetic mean  $\sigma_{ij} = (\sigma_i + \sigma_j)/2$  (Lorentz rule), respectively.

According to the chemical and physical properties of atoms, different atom types are assigned to each atom with a set of fixed parameters. For example, the oxygen attached to a carbonyl group is usually assigned as a different atom type than the oxygen attached to a hydroxyl group. A forcefield is a library that assigns each atom with an atom type and a set of fixed parameters. Depending on the molecular system of interest, different forcefields were developed to simulate systems such as water, membrane, ion, DNA, protein, DNA-protein complex, and intrinsically disordered proteins. In addition to the classical and atomistic forcefields described by the formalism similar to equation 3.6 and equation 3.7, polarizable forcefields[113] and coarse-grained models[114, 115] were developed to capture the polarization effects and large-scale (both in time and space) dynamics, respectively.

### 3.1.3. Simulation box and periodic boundary condition

To initiate an MD simulation, one usually attempts to prepare an environment that can reflect real-world circumstances within reasonable computational costs. First, a biomolecule can be prepared by taking the experimental structure from online databases, ab-initio calculation, or structural prediction with homology [116] or machine learning[117, 118]. The biomolecule is then solvated either implicitly by the reaction field or explicitly by solvent molecules such as water, salt, or lipid molecules. In the case of explicit solvent, periodic boundary condition (PBC) is often used to simulate a large number of solvent molecules without restricting them in a confined space (Figure 3.1A). Depending on the shape of the simulation box, usually in a cubic or truncated octahedron shape, PBC is applied by imaging the simulation box (called a unit cell in the content of PBC) in each dimension. Under PBC, when a particle approaches the boundary of the original box, it moves to the neighboring box and is re-imaged into the original box from its opposite boundary side (Figure 3.1B).

In theory, such box replication would generate an infinite number of atoms through infinite

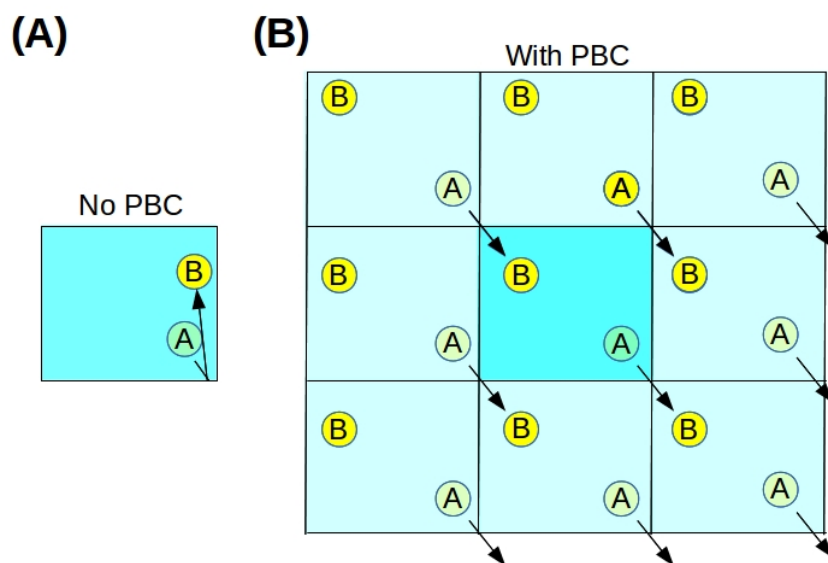


Figure 3.1.: Schematics of particle movement, from point A to point B (black arrow), with and without boundary condition (PBC) when approaching the edge of the simulation box. (A) Particle from position A reflects from the box boundary when PBC is not applied. (B) The original unit cell (colored in light blue) is replicated in each dimension with PBC (transparent blue) and the particle moving from A to B is re-imaged back to the simulation box without colliding with the boundary.



replication along each dimension, and the computational cost would increase quadratically for the non-bonded interactions. To cope with the issue, the Coulomb interaction in equation 3.7 can be split into four terms (The Ewald formula[119]), namely the real space contribution  $U^{(r)}$ , reciprocal space contribution  $U^{(k)}$ , self-energy  $U^{(self)}$ , and  $U^{(dipolar)}$  the polar correction

$$U_{\text{electrostatic}} = U^{(r)} + U^{(k)} + U^{(self)} + U^{(dipolar)} \quad (3.8)$$

$$U^{(r)} = \frac{1}{2} \sum_i^N \sum_j^N \sum_{n \in \mathbb{Z}^3} q_i q_j \frac{\text{erfc}(\alpha |\vec{r}_{ij} + \vec{n}L|)}{|\alpha |\vec{r}_{ij} + \vec{n}L||} \quad (3.9)$$

$$U^{(k)} = \frac{1}{2L^3} \sum_{\vec{k} \in \mathbb{K}^3, k \neq 0} \frac{4\pi}{k^2} e^{-k^2/4\alpha^2} |\hat{\rho}(\vec{k})|^2 \quad (3.10)$$

$$U^{(self)} = -\frac{\alpha}{\sqrt{\pi}} \sum_i q_i \quad (3.11)$$

$$U^{(dipolar)} = \frac{2\pi}{(1 + 2\epsilon_\infty)L^3} \left( \sum_i q_i r_i \right)^2 \quad (3.12)$$

where  $\alpha$  is the Ewald splitting parameter and  $L$  is the periodic box size. Usually, the  $U^{(dipolar)}$  term can be ignored due to the large effective dielectric constant at the large distance  $\epsilon_\infty = \infty$ . Note that the real space contribution  $U^{(r)}$  with an error function in equation 3.9 and the reciprocal space contribution  $U^{(k)}$  with a Gaussian function in equation 3.10 decrease exponentially with the interatomic distance  $|\vec{r}_{ij}|$  (short-range contribution) and the wave number  $|\vec{k}|$  (long-range contribution). In practice, a cut-off distance (in real space) is chosen to truncate the short-range electrostatic interaction and the van der Waals interaction (equation 3.7), which also decays with  $(\frac{1}{r})^6$  and a cut-off wave number (in the reciprocal space) is chosen to truncate the long-range interaction. By using the fast Fourier transformation (FFT) method on pre-set grid points and a multidimensional piecewise-interpolation approach, particle-mesh Ewald ) method reduces the complexity of the non-bonded interactions from  $\mathcal{O}(N^2)$  to  $\mathcal{O}(N \ln N)$ (PME[120]).

### 3.1.4. Integrating the equations of motion

With a biomolecule embedded within a simulation box and the atomic interactions described by a chosen force field, one can calculate the dynamics of the system by solving the Newtonian equations of motion (equation 3.4). Typically, this is numerically solved with finite difference methods such as the Verlet algorithm[121] or the leap-frog algorithm[122].

In the scheme of the Verlet algorithm, the position of particle  $i$  with mass  $m_i$  has a velocity  $\vec{v}_i(t)$  and experiences force  $\vec{F}_i(t)$ . Its position at time  $t + \Delta t$  can be Taylor expanded around time  $t$  so that

$$\vec{x}(t + \Delta t) = x(t) + \vec{v}_i(t)\Delta t + \vec{F}_i \frac{\Delta t^2}{2m_i} + \mathcal{O}(\Delta t^3) + \mathcal{O}(\Delta t^4) \quad (3.13)$$

and similarly its position at time  $t - \Delta t$  can be expressed as

$$\vec{x}(t - \Delta t) = x(t) - \vec{v}_i(t)\Delta t + \vec{F}_i \frac{\Delta t^2}{2m_i} - \mathcal{O}(\Delta t^3) + \mathcal{O}(\Delta t^4) \quad (3.14)$$

. By summing up equation 3.13 and equation 3.14, we get

$$\vec{x}(t + \Delta t) = 2\vec{x}(t) - \vec{x}(t - \Delta t) + \vec{F}(t) \frac{\Delta t^2}{2m_i} + \mathcal{O}(\Delta t^4) \quad (3.15)$$

$$(3.16)$$

where the force can be calculated by taking the derivative of the potential energy function  $U$ , as described in equation 3.5,

$$\mathbf{F}(\mathbf{x}) = -\nabla U(\mathbf{x}) \quad (3.17)$$

. In MD simulations,  $\Delta t$  is the time step used for trajectory integration. While a small time step requires a massive amount of steps to reach the same amount of simulation time, a large time step can lead to system instability and losses the resolution to capture the finest motion of the system of interest. In other words, the time step is limited by the motion of the quickest-moving atoms, namely the hydrogens. Typically, a time step of 0.5fs or 1fs is used for MD simulations. However, applying the SHAKE algorithm and the Hydrogen mass repartitioning method can be used to increase the time step. The SHAKE algorithm [123] constraints the bond lengths to the hydrogens and enables a time step of 2fs. Further use of the hydrogen mass repartitioning method partitions the mass of the heavy atom to the connecting hydrogens and enables a time step of 4fs[124].

To model the physiological condition, barostats, and thermostats are usually coupled to the systems to control the pressure and the temperature, respectively, to generate a canonical ensemble (NPT). A common choice of barostat is the Berendsen barostat[125] by first calculating the pressure of the system with a Volume  $V$  by

$$P = \frac{1}{V} \frac{1}{3} \sum_{i,j>i}^N (\vec{F}_{ij}|\vec{r}_i - \vec{r}_j| + \frac{|\vec{p}_i|^2}{m_i}) \quad (3.18)$$

where  $m_i$  and  $\vec{p}_i$  are the mass and momentum of a particle, respectively. With a pressure relaxation constant  $\tau_P$  and the isothermal compressibility  $\beta$ , the targeted pressure  $P_0$  can be reached by rescaling the box lengths  $d$  ( $d=V^{1/3}$  for a cubic box) and the center of mass coordinates of each molecule by a scaling factor  $\mu$

$$\mu = d_{new}/d_{old} = [1 - \frac{\beta\Delta t}{\tau_P}(P_0 - P)]^{1/3} \quad (3.19)$$

. A targeted temperature  $T_0$  of the system can as well be controlled by the Berendsen thermostat[125] in a similar approach by rescaling the velocity

$$\vec{v}_{new} = \sqrt{1 + \frac{\Delta t}{\tau_T}(\frac{T_0}{T} - 1)} \cdot \vec{v}_{old} \quad (3.20)$$

where  $\tau_T$  is the temperature relaxation constant.

Although controlling the pressure and temperature using the Berendsen scheme is relatively simple to implement, such integration will generate deterministic trajectories and is not convenient for the calculation of statistical observables. Instead, a Langevin thermostat is a more robust way to control the temperature through a stochastic Langevin equation of motion[126]

$$m \frac{d^2 \vec{r}_i(t)}{dt^2} = \vec{F} - \gamma \frac{d\vec{r}_i(t)}{dt} m_i + \vec{R}(t). \quad (3.21)$$

The equation of motion, compared to the Newtonian one (equation 3.4), includes two additional terms: the frictional forces mediated by the collision frequency  $\gamma$  and the Brownian random force  $\vec{R}$ . The Brownian random force is a delta-correlated stationary Gaussian process with zero-mean  $\langle \delta R(t) \rangle = 0$  and  $\langle \delta R_i(t) \delta R_j(t') \rangle = 2\gamma k_B T m \delta(t - t')$  where  $k_B$  is the Boltzmann constant.

## 3.2. Free Energy Calculation

MD simulation provides a way to understand the dynamics of a given biomolecular system with an atomistic resolution. With the trajectories generated under physiological conditions, one can derive several dynamical and thermodynamical properties using statistical mechanics. An important thermodynamical property of a biomolecular system is the free energy required to bring a system from one state to the second state. For example, the free energy required to bring a molecule from a vacuum to a water environment (solvation-free energy), the free energy required to bring a ligand from an unbound state to a receptor-bound state (binding free energy), and the free energy required to bring a protein from a close conformation to an open conformation (conformational free energy). Under the isothermal-isobaric condition (NPT), the free energy of a certain state with pressure  $P$ , volume  $V$ , and temperature  $T$ , is usually described in the form of Gibbs free energy

$$G = H - TS = U + PV - TS \quad (3.22)$$

where  $G$  is the Gibbs free energy,  $H$  is the enthalpy,  $U$  is the internal energy that can be directly computed from the force field, and  $S$  is the entropy of the system. The Gibbs free energy, abbreviated as free energy in the following context, as the difference between two states is usually calculated by end-point methods, alchemical transformation, or biasing MD simulations. Since only the binding energy was calculated in the present work using end-point method and the biasing MD simulation, more detailed information about alchemical transformation is not addressed here.

### End-point methods: MMPB/SA and MMGB/SA

In the binding energy calculation, a ligand-receptor pair is considered. The end-point method appropriately re-evaluates the trajectories generated at the bound state and at the unbound

state using the molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) method and the molecular mechanics generalized Born surface area (MM/GBSA) method[127, 128]. Essentially, the binding free energy in the solvated environment is desired. With this, the binding free energy difference can be written as

$$\Delta G_{\text{bind,solv}} = G_{\text{bound,solv}} - G_{\text{unbound,solv}} = \Delta H_{\text{bind,solv}} - T\Delta S_{\text{bind,solv}} \quad (3.23)$$

Where  $H_{\text{bind,solv}}$  and  $S_{\text{bind,solv}}$  are the changes in enthalpy and entropy of the system between the solvated unbound and bound state, respectively. However, because of the large number of degrees of freedom of the solvent molecules, the sampling of the solvent-solvent interactions results in fluctuations (noise) that are often orders of magnitude larger than the desired ligand binding free energy. To reduce these fluctuations, the mean solvent response based on a continuum solvent model is considered for each conformation of the trajectory. According to the thermodynamic cycle shown in Figure 3.2, the change in binding free energy is split into a change of ligand-receptor interaction energy in the vacuum state  $\Delta G_{\text{bind, vacuum}}$ , a change in solvation free energy ( $\Delta G_{\text{solv, receptor}}$  and  $\Delta G_{\text{solv, ligand}}$ ) and a change in conformational entropy  $\Delta S_{\text{bind, vacuum}}$  of the receptor and the ligand. Note, the change in solvation-free energy already includes entropic contributions of the solvent. While the binding enthalpy change in the vacuum phase simply relates to the average change in force field energy, calculating the change in conformational entropy poses a problem. The simulation trajectories are usually too short to estimate conformational entropy changes directly. Instead, approximate methods such as normal mode analysis (NMA[129]) and interaction entropy (IE[130]) methods can be employed.

The solvation-free energy change is typically further split into a polar (electrostatic) and non-polar (hydrophobic) contribution.

$$\Delta G_{\text{solv}} = G_{\text{solv, npol}} + G_{\text{solv, pol}} \quad (3.24)$$

The non-polar solvation energy  $\Delta G_{\text{solv, npol}}$  calculates the energy cost of creating a cavity with the size of the solute and can be approximated using a linear function of the solvent-accessible surface area (SASA)

$$\Delta G_{\text{npol}} = \gamma \cdot \text{SASA} \quad (3.25)$$

. In the MM/PBSA approach the polar solvation free energy change  $\Delta G_{\text{solv, pol}}$  is calculated by numerically solving the Poisson-Boltzmann equation (PB) equation[106], which can be written in a linearized form

$$\nabla \cdot [\epsilon(\vec{r}) \nabla \phi(\vec{r})] = -4\pi\rho(\vec{r}) - 4\pi\lambda(\vec{r}) \sum_i z_i c_i \exp(-z_i \phi(\vec{r})/k_B T) \quad (3.26)$$

where  $\epsilon(\vec{r})$  is the dielectric constant,  $\phi(\vec{r})$  is the electrostatic potential,  $\rho(\vec{r})$  is the solute charge,  $\lambda(\vec{r})$  is the Stern layer masking function.  $z_i$  is the charge of an ion type with concentration  $c_i$ . Essentially, the PB equation calculates how ions are distributed in the solution due to the charge distribution of the solute and how the corresponding distribution interacts with the

$$\Delta G_{bind,water} = \Delta G_{bind,vacuum} + \Delta G_{solv,complex} - (\Delta G_{solv,receptor} + \Delta G_{solv,ligand})$$

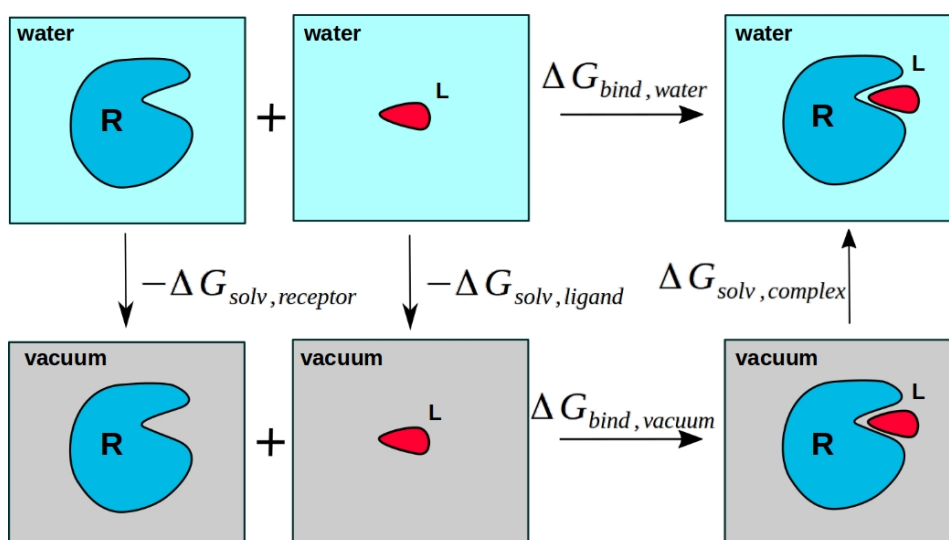


Figure 3.2.: Thermodynamic cycle in the end-point free-energy calculation methods. The receptor (R) is colored in blue and the ligand (L) is colored red. Note that in the end-point approaches the free energy required for the receptor and ligand conformational changes are not included.

solute through electrostatic interaction. The change in solvation-free energy can be obtained by substrating the polar solvation energy in the water state ( $\epsilon=80$ ) with the polar solvation energy in the vacuum state ( $\epsilon=1$ ), as depicted in Figure 3.2. In the computationally less demanding MM/GBSA, the PB equation is approximated by a pair-wise generalized Born (GB) formula[131, 132]

$$\Delta G_{\text{pol,GB}} = \frac{1}{2} \sum_{ij} \frac{q_i q_j}{f_{\text{GB}}} \left( \frac{1}{\epsilon_{\text{in}}} - \frac{1}{\epsilon_{\text{out}}} \right) \quad (3.27)$$

$$f_{\text{GB}} = \sqrt{(r_{ij}^2 + R_i R_j e^{-\alpha_{ij}})} \quad (3.28)$$

$$\alpha_{ij} = r_{ij} / 4R_i R_j \quad (3.29)$$

where  $r_{ij}$  is the distance between atom pair  $i$  and  $j$  with Born Radius  $R_i$  and charge  $q_i$  of atom  $i$ .

In practice, MMPB/SA and MMGB/SA can be also used to evaluate the binding free energy through a single bound-state trajectory, and the unbound state can be generated simply by assuming no conformational changes occur between the bound state and the unbound state.

### 3.2.1. Biasing method

#### Umbrella sampling

Although the end-point methods provide a rapid way to evaluate the binding free energy without additional sampling efforts, the assumption that the receptor and ligand do not undergo conformational changes does not hold true in most cases. For example, both the induced-fit and the conformational selection binding mechanisms require the receptor to undergo several conformational changes. Alternatively, the binding free energy can also be derived using the biasing potential to sample the possible states along a certain reaction coordinate (RC, mathematically denoted as  $\xi$ ). An RC is a variable or a set of variables that describe the binding process of the concerning system. In the case of the ligand-receptor pair, a straight-forward option is to select the center of mass (COM) distance between the ligand and the receptor binding pocket, as depicted in Figure 3.3A.

Formally, the free energy change associated with the spatial dissociation of the ligand along the RC is given by the following formula

$$G(\xi) = G(\xi = \infty) + \int_{\infty}^{\xi} d\xi' \langle -dG/d\xi' \rangle \quad (3.30)$$

where  $\xi$  is the value of the RC and  $\langle -dG/d\xi \rangle$  is the mean force along the RC. Averaging is done over all remaining degrees of freedom. The free-energy profile  $G(\xi)$  obtained by taking the integral of the mean force is called the potential of mean force (PMF) along  $\xi$ , from which the binding free energy  $\Delta G_{\text{bind}}$  can in principle be derived as  $\Delta G_{\text{bind}} = G(\xi=\text{bound}) - G(\xi = \text{unbound} = \infty)$ . The PMF profile can be estimated using equilibrium and non-equilibrium sampling methods. The most common equilibrium sampling method is umbrella sampling [133] (US) which typically brings the ligand from its associated to the dissociated

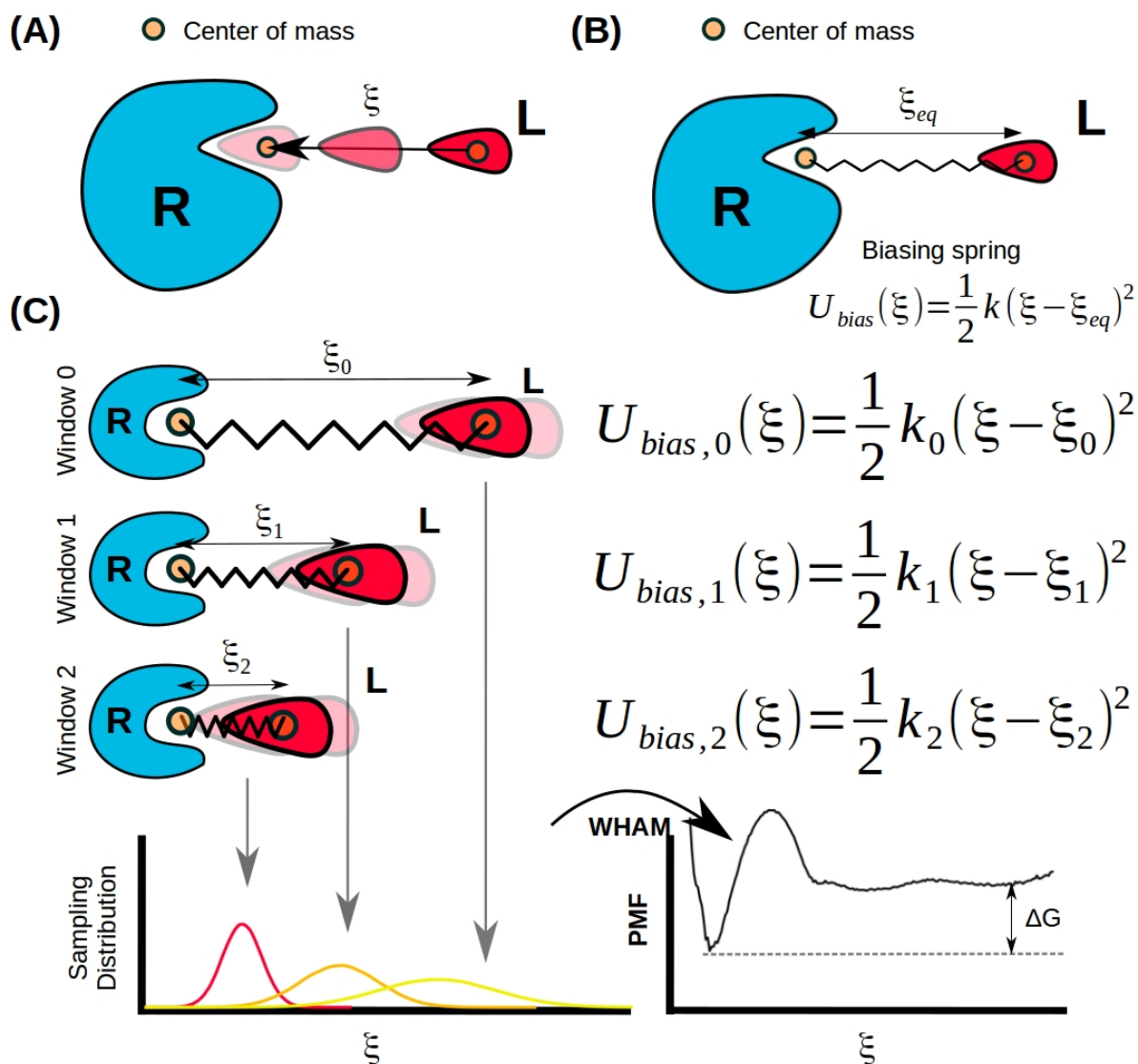


Figure 3.3.: Calculation of potential of mean force (PMF) along a pre-defined reaction coordinate (RC, denoted as  $\zeta$ ). (A) Schematic of a simple ligand binding process where the RC can be determined by the distance between the center of masses of the binding pocket and the ligand. (B) Schematic of introducing a biasing potential by a biasing spring with equilibrium spring length  $\zeta_{eq}$  and spring constant  $k$ . (C) Schematics of umbrella sampling with three umbrella windows (window 0 to window 2) with spring constant  $k_i$  and equilibrium spring length  $\zeta_i$  in the  $i^{th}$  window. The sampled distributions are collected and converted into the PMF profile using WHAM. The binding free energy is denoted as the free energy difference  $\Delta G$  between the bound form ( $\zeta = \zeta_{bound}$ ) and the unbound form ( $\zeta = \infty$ ).

state (or vice versa) by adding a harmonic biasing potential to the force field along the RC. In MD simulation, one can restrain the RC at a certain value, in this case, distance, by attaching a harmonic spring with a spring constant  $k$  and equilibrium distance  $\xi_{eq}$  (Figure 3.3B). By generating a set of springs with different  $k$ s and  $\xi_{eq}$ s in different replicas, so-called windows, along the RC, one can restrict the motion of the biomolecular system within a certain phase space in each window (Figure 3.3C). In the scheme of umbrella sampling (US), the conformation of the system is first sampled with  $\xi_{eq}$  near the  $\xi$  of the starting structure, as depicted as window 0 in Figure 3.3C, and configurations in this window are sampled until the system is considered equilibrated. Next, the system is sampled in a neighboring window to guide the RC toward the targeted value. US records the work that is needed to keep the RC in a certain window by counting the sampled  $\xi$  along the biasing process and iteratively calculates PMF with self-consistent equations such as the weighted histogram algorithm method (WHAM[134]). WHAM discretizes a continuous free profile into a set of discrete bins  $G_j = \text{PMF}_j = \text{PMF}(\xi_i < \xi \leq \xi_{i+1})$  and counts the sampled RC values within the bin range during the simulation of the  $i^{\text{th}}$  umbrella  $n_i(\xi_j) = \sum_{\xi_i < \xi \leq \xi_{i+1}} 1$ . With this, the probability distribution in the unbiased simulation,  $P_j$ , can be estimated using the reweighting method

$$P_j = \frac{\sum_{i=1}^{N_{\text{windows}}} n_i(\xi_j)}{\sum_{i=1}^{N_{\text{windows}}} N_i \exp([G_j - U_{\text{bias},i}(\xi_j)]/k_B T)} \quad (3.31)$$

, and  $G_j$  can be computed as

$$G(\xi) = G(\xi = \infty) + \int_{\infty}^{\xi} d\xi' \langle dG/d\xi \rangle \quad (3.32)$$

$$G_i = -k_B T \ln \left\{ \sum_{j=1}^{N_{\text{bins}}} P_j \exp[U_{\text{bias},i}(\xi_j)] / k_B T \right\} \quad (3.33)$$

. As shown in equation 3.31 and equation 3.33, WHAM equations contain two unknown quantities  $P_j$  and  $G_j$  which can be solved iteratively until  $G_j$  of all bins do not differ by more than a pre-defined tolerance parameter (usually around the scale of  $10^{-5}$  kcal/mol) between consecutive iterations and the error of  $G_j$  can be estimated by the bootstrapping method [135, 136]. A demonstration of free energy estimation using WHAM from a set of  $\xi$  sampled in the US is shown in Figure 3.3C. In a more general case where one wants to compare the free energy between two states A and B, the energy difference is usually denoted as  $\Delta\Delta G_{A \rightarrow B} = \Delta G_A - \Delta G_B$ .

### Hamiltonian Replica Exchange coupled Umbrella Sampling Method

In the conventional US scheme, one can calculate the energy difference between A and B by applying biasing potentials to bring the system from state A to state B or vice versa. However, the calculated PMF profiles by biasing the system from A to B or from B to A can be very different, known as the convergence problem. Usually, the convergence problem can be solved



by performing long US simulations (convergence and Sampling in Determining Free Energy Landscapes for Membrane Protein Association). To circumvent the issue, the Hamiltonian replica exchange coupled US (HREUS) method has been developed to bias the system in both directions at the same time[137]. First, the conventional US method is applied to bias the system into several windows with targeted  $\zeta$ s to prepare the starting conformation for HREUS. Second, the configuration at each window is simulated for a certain amount of time, called exchange frequency  $\tau_{ex}$ , before a replica exchange trial. For every  $\tau_{ex}$ , the neighboring windows attempt to exchange their configurations with an acceptance probability following the metropolis's criterion

$$P_{acc,ij} = \min\left(1, \frac{e^{-\beta[U_i(\vec{r}_j)+U_j(\vec{r}_i)]}}{e^{-\beta[U_i(\vec{r}_i)+U_j(\vec{r}_j)]}}\right) \quad (3.34)$$

where  $\beta$  is  $1/k_B T$  and  $U_i(\vec{r}_i)$  is the energy of the  $i^{th}$  replica with the system in the configuration  $\vec{r}_i$ . Essentially, the nominator and denominator in equation 3.34 represent the total energy before and after the exchange attempt, respectively. This equation favors the attempt if the attempt results in lower total energy but allows a probability to still allow the exchange with a certain probability. The HREUS method has been shown to converge the PMF profile faster than the conventional US sampling method[138].

### 3.3. Simulation protocol for $\gamma$ -secretase

Besides Chapter 9, the main part of the present dissertation is about the MD simulations of  $\gamma$ -secretase. Therefore, a short section is presented here to describe the simulation protocol for  $\gamma$ -secretase.

#### 3.3.1. Protonation state of the catalytic center

$\gamma$ -Secretase is an aspartyl protease that cleaves its substrates with two aspartic acids (Asp) from the PS subunit. A schematic of aspartyl protease exemplified by  $\beta$ -secretase is depicted in Figure 3.4. An active geometry is formed by a protonated Asp and an unprotonated Asp from the protease [139, 140, 141, 142], the substrate scissile bond, and a catalytic water molecule. Upon the formation of the active geometry, the substrate is cleaved into the N-terminal and C-terminal products via a series of proton transfer processes.

However, the protonation state of Asp needs to be determined during the simulation box preparation in a regular MD. Therefore, usually, simulations for both possible protonation states need to be conducted to understand the influence of the choice of protonation state. Although no consensus on the choice of protonation state for  $\gamma$ -secretase has been reached, it is possible to predict which Asp is more likely to be in the protonated state based on its microenvironment using the pKa calculation technique[144]. The predicted pKa values for the catalytic Asp using the available 3D structures of PS homolog, PS1, and PS2 are listed in Table 3.1. Besides the cryo-EM structure 5FN5, the Asp located on TMD7 is predicted to be more likely in the protonated state than the Asp on TMD6. Indeed, recent work using pH

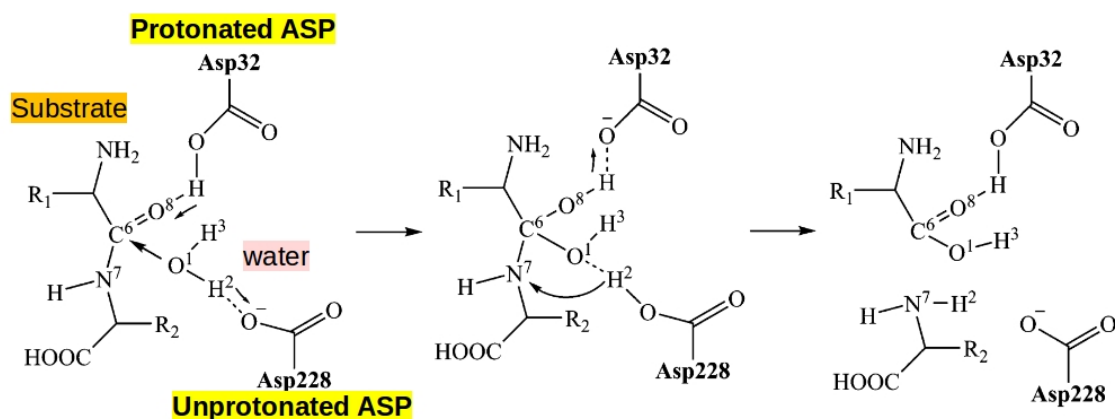


Figure 3.4.: A putative pathway of the hydrolysis reaction in aspartyl protease, exemplified by  $\beta$ -secretase. The diagram is reproduced from reference [143].)

replica exchange molecular dynamics method has also suggested that D385 of PS1 is more likely to be protonated than D257[145]. However, both possible protonation states are used in Chapter 4, Chapter 7, and Chapter 6 to study the influence of the choice of protonation state on the dynamics of the systems.

#### *In silico* mutations

An advantage of computer simulation is that the mutations on the targeted biomolecules can be introduced relatively easily by direct atom deletion or growing in the structure file. Protein mutations in the study are performed by deleting all side chain atoms besides the first carbon, the  $\beta$  carbon in the file, from the old amino acid, and the missing atom from the new residue are grown automatically using tleap from AmberTools20[111]. However, this *in silico* mutation method assumes that the introduced mutation does not alter the folding of the protein. Therefore, caution needs to be taken when introducing mutations, especially when a set of mutations is introduced simultaneously to a protein.

Instead of investigating the effects of mutations on  $\gamma$ -secretase, this dissertation focuses on the mutant points on APP for three major reasons. First, full-length  $\gamma$ -secretase consists of more than 1500 amino acids and 20 TMDs and how a single point mutation affects the intramolecular topology, e.g. protein misfolding or trafficking, and intermolecular topology, e.g.  $\gamma$ -secretase complexation and maturation, requires a simulation timescale that is unfeasible with the current technique. In contrast, APP contains only 1 TMD with 24 amino acids (G29-L52), and the influence on its conformational space by a single mutation can be more efficiently sampled. Secondly, according to the holo-form cryo-EM structure, individual residues on substrate TMD fit into different substrate binding pockets composed of multiple residues of  $\gamma$ -secretase. In other words, introducing a single mutation on  $\gamma$  secretase is not sufficient to alter the overall pocket microenvironment, whereas a single mutation on the substrate TMD can significantly alter the E-S interaction. Lastly and most importantly experimental results reporting mutations' effect on the processivity are conducted with APP

Table 3.1.:  $pK_a$  predictions by PORPKA3.1 for published PSH and  $\gamma$ -secretase structures. The catalytic aspartate residues with higher  $pK_a$  are more likely to be in the protonated state and are highlighted in red. The number indicated in the parenthesis show the catalytic residues on TMD6 and TMD7.

PDBID	Enzyme	Ligand	Prediced pKa (D162, D220)
4HYC	PSH	None	6.14, <b>6.35</b>
4HYD	PSH	None	5.64, <b>6.57</b>
4HYG	PSH	None	5.04, <b>6.62</b>
4Y6K	PSH	None	5.63, <b>7.52</b>
PDBID	Enzyme	Ligand	Prediced pKa (D257, D385)
4UIS	PS1	None	3.18, <b>6.16</b>
5A63	PS1	None	4.42, <b>6.16</b>
5FN5	PS1	None	<b>4.98</b> , 3.63
5FN4	PS1	Unknown helix	4.70, <b>4.71</b>
5FN3	PS1	Unknown helix	4.90, <b>7.13</b>
5FN2	PS1	DAPT	5.13, <b>9.93</b>
6IYC	PS1	C83	6.39, X
6IDF	PS1	Notch1	6.21, X
6LR4	PS1	Semagacestat	6.12, <b>7.94</b>
6LQG	PS1	Avagacestat	6.08, <b>7.22</b>
7V9I	PS1	L-685,458	7.11, <b>8.90</b>
7D8X	PS1	L-685,458 and E2012	7.01, <b>8.69</b>
7Y5T	PS1	MRK-560	6.37, <b>7.70</b>
PDBID	Enzyme	Ligand	Prediced pKa (D263, D366)
7Y5X	PS2	None	4.88, <b>6.07</b>
7Y5Z	PS2	None	4.10, <b>4.76</b>

mutations. Thus, it is more convenient to computationally investigate those mutations which have been found to cause significant influence on the activity or the processivity.

### 3.3.2. Simulation setup

After the simulation box of  $\gamma$ -secretase with or without the ligands is constructed for each system, MD simulations are conducted following energy minimization, equilibration, and a production run. The procedures of the energy minimization of equilibrium are listed in Table 3.2).

First, the energy of the system is minimized by using the steepest decent in the first half and conjugate gradient in the second half of the minimization steps with positional restraints applied on the proteins and the membrane molecules (Min-1 and Min-2 in Table 3.2 using the MPI version of the pmemd program in AMBER[111]. Next, the system is equilibrated in 6 steps with a time step of 2fs with SHAKE algorithm for 400 ps at 303.15K controlled by Langevin dynamics using the cuda accelerated pmemd of AMBER with positional restraints listed in 3.2 (Eq1-Eq6). The first three equilibration steps are conducted within an isovolumic-isothermal ensemble (NVT) while the later three steps are conducted within an isobaric-isothermal ensemble (NPT) with pressure controlled at 1 bar using Berendsen barostat with a relaxation time of 0.5ps. In addition, the semi-isotropic Berendsen algorithm is applied on the xy plane to maintain zero surface tension of the membrane layer. The equilibrated system is then submitted for production run with controlled pressure at 1 bar and zero surface tension with Berendsen thermostat and 303.15K with Langevin dynamics. A time-step of 4fs was used with the SHAKE algorithm and the hydrogen mass repartitioning method. Simulation time and force fields differ from system to system and are specified in each corresponding methods section. All production runs are carried out by the cuda accelerated pmemd of AMBER with a non-bonded cutoff of 9 Å.

Table 3.2.: Simulation setup for  $\gamma$ -secretase systems.  $k_{protein}$  and  $k_{membrane}$  are the positional restraint strength applied on the protein and membrane in a unit of  $\text{kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-2}$ , respectively.

Name	Steps	$k_{protein}$	$k_{membrane}$
Energy Miminization			
Min-1	20,000	10	2.5
Min-2	50,000	10	2.5
NVT Equilibration			
Eq-1	25,000	10	2.5
Eq-2	25,000	5	2.5
Eq-3	25,000	2.5	1
NPT Equilibration			
Eq-4	25,000	2.5	1
Eq-5	50,000	0.5	0.1
Eq-6	50,000	0.1	0

## 4. Dynamics of wild type $\gamma$ -secretase

MD simulations allow us to investigate the dynamics of a set of molecules in an atomic resolution. Starting with either a predicted structure or an experimentally determined structure, one can investigate the function of the molecules of interest by collecting the conformational space from conventional MD simulation or implementing advanced sampling methods combined with statistical mechanics theory to discover rare events or free energy landscapes along a certain reaction coordinate. Furthermore, by comparing the dynamics of the wild-type protein and its mutants, one can try to interpret the corresponding experimental findings and even propose a molecular mechanism of a certain phenotype. In this section, the dynamics and interactions between  $\gamma$ -secretase and its substrates are studied with MD simulations. Note, the results from this chapter are mostly taken from our previously published work[146].

### 4.1. Introduction

$\gamma$ -Secretase and APP are relatively large intramembrane proteins and contain over 1500 amino acids and 770 amino acids, respectively. Surprisingly, a single mutation (FAD mutation) on either protein can cause an early onset of AD[19, 18]. Many experimental efforts in the past 20 years have focused on conformational changes and the change in enzyme activity as well as substrate cleavability resulting from these mutations[147, 77, 148, 149, 71]. With the advances in cryo-EM techniques, 3D structures of  $\gamma$ -secretase in the substrate-free (apo) form[53, 55, 54] and the substrate-bound (holo) form[39, 40, 58, 59] resolved in the past 8 years have allowed scientists to interpret the experimental findings from a single snapshot. In particular, many FAD mutations are found to have close contact with their substrates. Although only the binding poses of C83 and Notch1 are resolved to the date of the dissertation, the highly overlapped electron density shared by these two substrates suggested a universal binding mode of  $\gamma$ -secretase substrates. However, since all crystal structures are resolved at cryogenics temperature (lower than 120K) and solvated in detergent micelles, some also contain artificial mutations, they might not reflect the real wild-type structure at room temperature within the lipid bilayer[39, 40]. By assuming that the effects of temperature, lipid environment, and mutation can be restored after short *in silico* equilibration, computational methods have been applied to study the dynamics of  $\gamma$ -secretase by MD simulations.

Before any 3D structure of  $\gamma$ -secretase was available, Somavarapu and Kepp attempted to build a homology model of PS1 by comparative modeling with the PS homolog structure as the template[150]. They suggested that the loop between TMD6 and TMD7 (exon9 loop) serves as a "plug" which controls the access of the substrate entering PS1. However, in-

consistent with their computational prediction, several structures of  $\gamma$ -secretase resolved by cryo-EM in the following years failed to identify the coordinates of the exon9 loop, indicating their high flexibility[53, 55, 54]. But their effort provided structural insight into how PS1 is embedded inside the membrane bilayer. Based on the apo-state structure published in 2015, many computational works have consistently discovered the bending motion of NCT using an anisotropic network model[151], coarse-grained model[152], and atomic simulations[153, 152]. This bending motion is interpreted as the mechanism of how  $\gamma$ -secretase excludes intramembrane peptides with a large ectodomain observed in experiments. The atomistic simulation also captured, for the first time, two geometries at the catalytic center, namely the open state and the close state in different protonation states[152]. Combining MD simulations of the apo-form  $\gamma$ -secretase and molecular docking, a fit-stay-trim (FIST) model is proposed[154]. That is,  $\gamma$ -secretase substrate can only fit into the semi-open state of PS1 and undergo the proteolysis process. Extensive coarse-grained simulations[61] studying the E-S interaction revealing how substrates bind to the surface of  $\gamma$ -secretase agrees with the photoaffinity labeling experiment[60] that the substrates have a higher propensity to bind at PS1-NTF than PS1-CTF. If one compares the structures of apo- and holo-form  $\gamma$ -secretase, a question arises: how does substrate enter the enzyme and reach its final read-to-cut position? Hitztenberger and Zacharias performed biasing MD simulations to pull the C99 substrate into the PS1 active site through three putative gates: TMD2-TMD3 gate, TMD2-TMD6 gate, and TMD6-TMD9 gate, before the structures of the holo-enzyme have been published and suggested that the TMD2-TMD3 gate is energetically the most favorable one[62]. This question remains nonetheless unsolved even after the structure of the holo-enzyme is released because all three putative entries implicate that the substrate has to overcome either the extracellular loop (loop 1) or the intracellular loop (loop 2 or exon 9).

Shortly after the release of the C83-bound and Notch1-bound  $\gamma$ -secretase structures[39, 40], Hitztenberger and Zacharias proposed a binding mode of TSA GSI L-685,458[155] which agreed perfectly with the binding pose later resolved by cryo-EM[58], and pointed out important residues forming the GSI binding pockets S1', S2', and S3'. However, another computational works exploiting Gaussian accelerated MD (GaMD) technique [156] suggested that the S2' pocket resolved by cryo-EM and predicted by Hitztenberger et. al.[155] exists only when bound by APP mutant I45F and T48P. The same computational group conducted another work using peptide GaMD (Pep-GaMD)[157], suggesting a tilting motion of the APP substrate without translational movement and hybrid  $\beta$ -sheet is enough for the processive cleavage from  $\epsilon$ 49 to  $\zeta$ 46 cleavage. Nonetheless, it is speculative whether the conformational space sampled with the the large boosting potential ( $\Delta V_{avg} > 100$  kcal/mol) can reflect real-world scenarios. Similar to this, MD simulations conducted with D257 and D385 both unprotonated revealed that the hybrid  $\beta$ -sheet resolved in the substrate-bound structures is likely just an artifact of the cryo temperature and is unstable at room temperature[158]. In contrast, our work using homology modeling[146] and free energy calculation using biasing-exchange MD simulations (under review) suggested that the hybrid  $\beta$ -sheet should also appear C-terminal to the catalytic center in both  $\epsilon$  and  $\zeta$  cleavages.

MD simulations were also applied to elucidate the molecular mechanism of GSM. Prior to the imidazole-based GSM-bound  $\gamma$ -secretase being solved in early 2021[58], several GSM-binding sites are suggested[159, 151]. Interestingly, efforts combining molecular docking and machine learning were claimed to correlate significantly with experimental activity even by targeting an incorrect binding site[160]. How GSM binds together with the substrate, however, remains elusive even when the GSM-bound structure is resolved[58] because it overlaps with the structure of the substrate-bound  $\gamma$ -secretase structure[39], implicating either substrate/GSM needs to deviate from its bind-alone pose or GSM does not bind at the first cleavage. Combining *in silico* and *in vitro* evidence, Petit et. al. suggested a slightly different binding GSM binding mode while preserving the Y106 hydrogen bonds[161].

In this chapter, the dynamics of the wild type  $\gamma$ -secretase in the apo- and -holo forms are studied and compared using MD simulations. Upon substrate binding, several regions are stabilized, including the active site geometry and the stability of TMD2, TMD3, TMD6a, and the hybrid- $\beta$  sheet regions. Energetic analysis with the MM/PBSA method reveals the interaction profile between  $\gamma$ -secretase and its substrate, serving as a foundation for further computational and experimental studies reported in the following chapters.

## 4.2. Method and Materials

Two structures of apo-form  $\gamma$ -secretase derived from two different cryo-EM structures are presented in the study. The first apo-structure is taken from RCSB PDBID 5FN2[55] with mutation Y256T restored with AmberTools18[162]. The second apo-structure is taken from RCSB PDBID 6IYC [39] by removing the APP substrate. The APP-bound structure is taken from RCSB PDBID 6IYC and the Notch1-bound structure is taken from RCSB PDBID 6IDF[40]. To model the correct S3 cleavage pose, the structure is reconstructed as described in Appendix A. Only the D385-protonated state (D385<sup>H</sup>) of PS1 is simulated in the apo-system and both D385<sup>H</sup> and D257<sup>H</sup> PS1 are simulated in the holo-systems. For simplicity reasons, only the result performed under the D385<sup>H</sup> protonation state in the main text and the results obtained under the D257<sup>H</sup> protonation state is shown in Appendix A. Starting structures of free-form APP and Notch1 are taken from RCSB PDBID 2LLM[56] and PDBID 5KZ0[57], respectively. Structures are solvated in POPC membrane (503 POPC molecules for  $\gamma$ -Secretase complexes and 200 POPC molecules for free-form substrates) with 0.15M potassium chloride through CHARMMGUI server. Atomic interactions are described by tip3p for water[163], Joung & Cheatham parameters for monovalent ions, ff14SB for proteins, and lipid17 [164] for POPC membrane. Energy minimization, equilibration, and NPT production run are performed with the procedure described in Chapter 3 section 3.3. Each apo-system was simulated for 600ns 3 times and each holo-system was simulated for 600ns 2 times. One 1-microsecond long simulation was performed for each free-form substrate. All protein stays stable during the simulation, as indicated by low RMSD values shown in Figure A.2

Residue-wise RMSFs are calculated using the averaged structure across each simulation. The secondary structure was analyzed using the DSSP[165] method with CPPTRAJ. Secondary structures of 'Alpha', '3-10', and 'Pi' are classified as 'Helix', and only 'Anti' is classified as

' $\beta$ -strand' in the main text. Water accessibilities of each substrate residue are calculated by counting the number of waters within 5Å of the side chain of the corresponding residue. The mean binding energy of the substrate is computed using the molecular mechanics energies combined with the Poisson-Boltzmann and surface area continuum solvation (MM/PBSA) method. Only the last 500 frames of each trajectory, in a total of 100ns, are processed for energy evaluation. Dielectric constants are set to  $\epsilon_{lipid}=2$  representing the membrane region between  $z=-18.5\text{\AA}$  and  $z=+18.5\text{\AA}$ ,  $\epsilon_{protein}=1$  within the protein interior, and  $\epsilon_{water}=80$  otherwise to represent the aqueous environment. Noticing that K28 creates an uneven membrane thickness, only the substrate TMD is used for the calculation to avoid energy divergence. Note that this approach approximates the whole membrane region with the implicit regions and therefore interactions involving water molecules such as the scissile bond at the catalytic center cannot be precisely calculated. Principle component analysis (PCA) is conducted by aligning the trajectory to the PS1 component and calculating the covariance matrix of the whole  $\gamma$ -secretase. Eigenvectors and correspondent eigenvectors are calculated and visualized using the normal mode wizard plugin in VMD.

## 4.3. Result

### 4.3.1. Bending motion of the NCT ectodomain and substrate conformational change upon E-S complexation

To begin with, the dynamics of the apo-form and holo-form of wt  $\gamma$ -secretase complexes were studied within the lipid bilayer at room temperature. We first analyzed the collective motion of the whole complex by performing PCA on the simulation trajectories. Consistent with the computational and experimental findings conducted previously by other groups, we observed a concertedly bending motion of the NCT ectodomain component in both apo- and holo-structures in the top 2 principle component (PC, Figure 4.1A). This bending motion changes the size of the cavity formed between the NCT ectodomain and the intracellular domain of  $\gamma$ -secretase and the distance between the NCT ectodomain to the membrane surface through an exchange between the compact and the open state. While the former influence was related to the substrate filtering function of NCT, the second influence might be important for substrate recognition. In our simulations, the extracellular domain of APP (G13-K28), but not Notch1, remains very close to the membrane surface without  $\gamma$ -secretase (Figure 4.1B), and the recognition of APP by NCT, thus, requires the NCT ectodomain in the compact state. However, a comparison between the free-form substrate (Figure 4.1B) and enzyme-bound substrate (Figure 4.1A) shows that the substrate tilting angle with respect to the membrane normal is much smaller when substrates are bound to the enzyme. The up-bending movement upon substrate is likely facilitated by the compact-to-open state transition of the NCT ectodomain.

As already indicated by the experimentally determined crystal structures[39, 40], the substrate TMD undergoes a significant change in secondary structure when binding to  $\gamma$ -secretase. For simplicity reasons, the general nomenclature of the substrate cleavage site



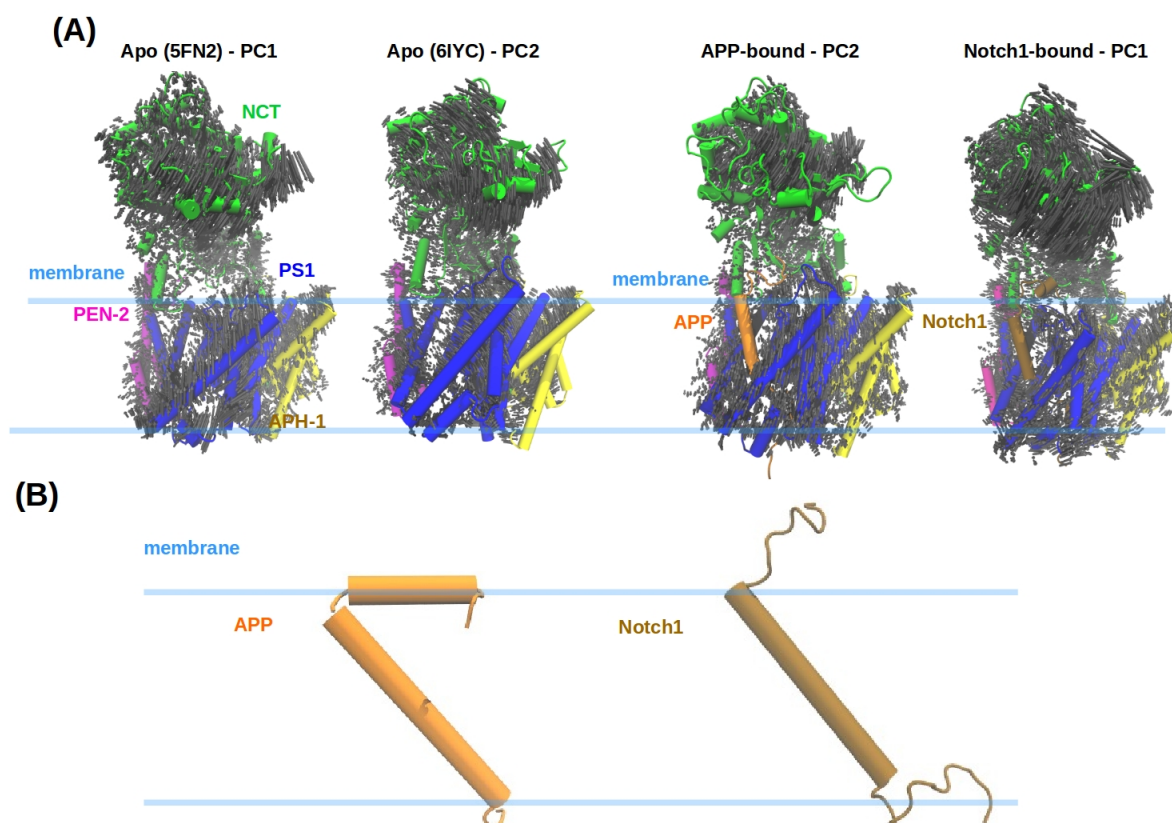


Figure 4.1.: Global motion of  $\gamma$  secretase in the apo- and holo-form and the change in the secondary structure of substrates upon enzyme binding. (A) Up-down bending motion observed apo- and holo-state  $\gamma$  secretase simulations. Collective motions are calculated by PCA and the direction and magnitude of PCs are shown with the gray arrows. (B) TMD of APP and Notch1 tilted in POPC membrane during the simulations.  $\gamma$ -Secretase components NCT, PEN2, APH-1, and PS1 are colored in green, magenta, yellow, and dark blue, respectively. APP is colored orange and Notch1 is colored brown. The upper and lower leaves of the POPC membrane are indicated by the light blue lines, respectively.

positions is denoted as formulated by Schechter and Berger[65] (Table 4.1). In the simulation, the TMD of both Notch1 (Q1733-L1756) and APP (G29-L52) formed in a complete helix in the free state but adapted to a helix-loop-strand configuration when binding at the interior of  $\gamma$ -secretase (Figure 4.2A and B). A stable  $\beta$ -strand, called the  $\beta$ 3-strand from the literature, was formed between substrate P1'-P4'. Furthermore, a  $3_{10}$ -helix was occasionally formed during the bound-state simulation in both substrates between the  $\alpha$ -helical region and the loop region (P7-P5). Overall, large conformational changes were observed between the free state and the enzyme-bound state, especially at the C-terminal half of the substrate TMD. Indeed, in the following sections and chapters, our data suggested that conformational changes are necessary for  $\gamma$ -secretase proteolysis.

Table 4.1.: Nomenclature of the substrate position, from P10 to P4', with respect to the cleavage site in proteolysis. The cleavage of  $\gamma$ -secretase is executed at the amine bone between the P1 and P1' residue and is highlighted with double margins.

	P10	P9	P8	P7	P6	P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'
APP	V40	I41	A42	T43	V44	I45	V46	I47	T48	L49	V50	M51	L52	K53
Notch1	F1744	V1745	L1746	L1747	F1748	F1749	V1750	G1751	C1752	G1753	V1754	L1755	L1756	S1757

#### 4.3.2. E-S interactions stabilize the $\gamma$ -secretase catalytic unit PS1

Next, we aimed to see which part of  $\gamma$ -secretase is mostly responsible for substrate binding by comparing the residue-wise RMSF of each  $\gamma$ -secretase component with and without the presence of substrate binding. While NCT, PEN2, and APH-1 do not show significant stabilization effects upon substrate binding (Figure 4.3A-C), the catalytic subunit PS1 shows a dramatic reduction in flexibility in several distinct regions (Figure 4.4A). The reduction in RMSF can be best interpreted by the E-S interactions that stabilize different parts of PS1. To gain more structural and energetic insight into the E-S interaction, residue-wise binding energy decomposition was performed to quantify the substrate residues interacting with the  $\gamma$ -secretase (Figure 4.4B). Overall, the N-terminal half of both substrates do not form strong interaction with the enzyme, which can be also visualized by the little E-S contact from the binding pose (Figure 4.4C) and the E-S interaction is dominated by the C-terminal half of the substrate TMD.

According to the geometry of each stabilized PS1 region and their contacts with the substrate, we highlight three main E-S interactions: the PS1 internal docking site, the TMD6a hydrophobic patch, and the hybrid  $\beta$ -sheet (Figure 4.4C). The 3 residues at the C-terminal end of substrate TMD (P1'-P3') contribute the largest backbone and side chain binding energy. The backbone contribution of P1'-P3' comes from the hydrogen-bond network at the hybrid  $\beta$ -sheet region (Figure 4.4C) and is responsible for the PS1 RMSF reduction at the TMD7 N-terminus and the loop connecting TMD8 and TMD9 (Figure 4.4A). It has been shown that mutations on PS1 that deliberately disrupt the formation of the hybrid  $\beta$ -sheet impair the  $\gamma$ -secretase activity. Atomic geometries at the hybrid  $\beta$ -sheet regions in the apo-state, APP-bound state, and Notch1-bound state are shown in Figure A.5. In Chapter 6 we will

### 4.3. RESULT

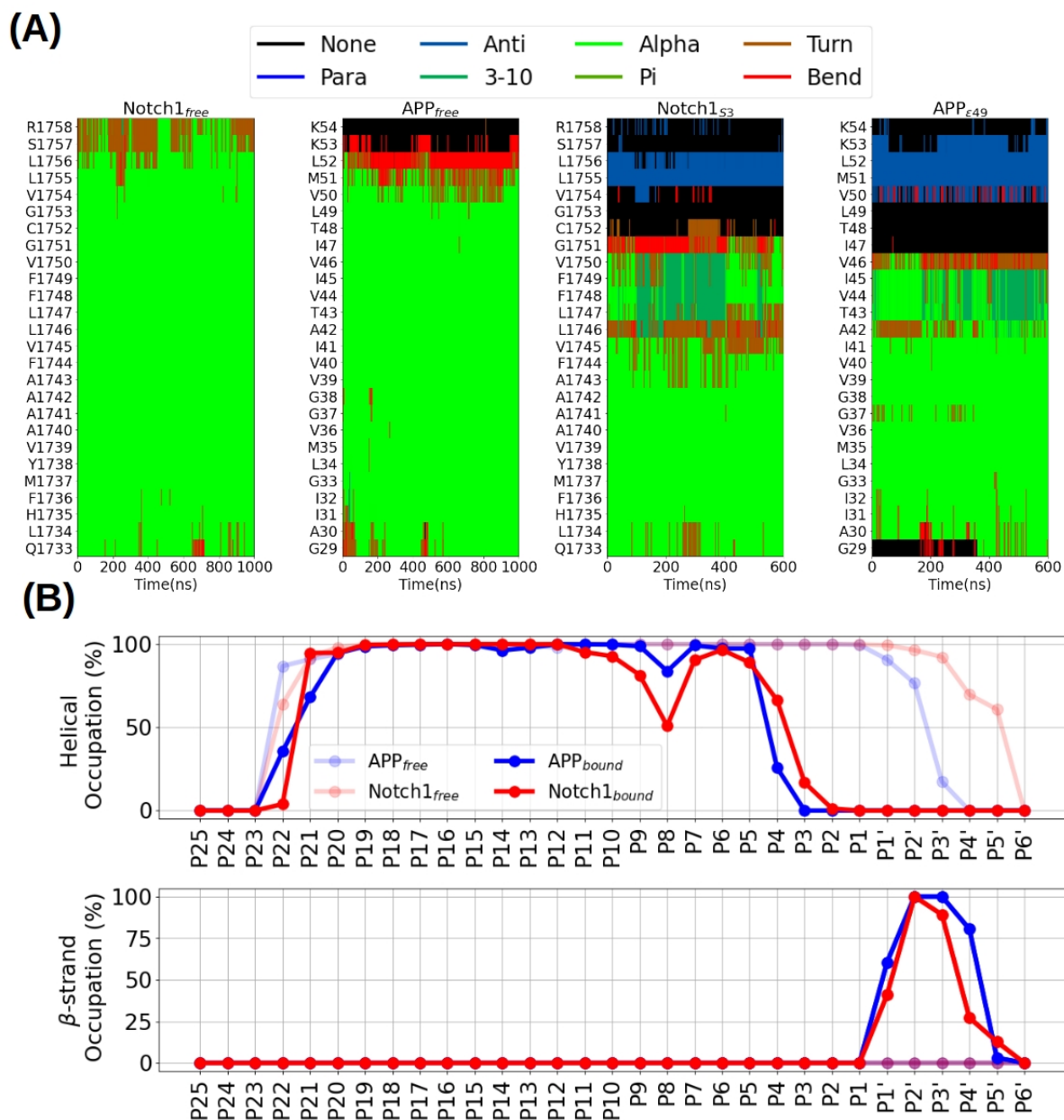


Figure 4.2.: Change in the secondary structure of substrates upon enzyme binding. (A) Time series of the substrate secondary structure in the simulations of (from left to right) free state Notch1, free state APP, enzyme-bound state Notch1, and enzyme-bound state APP. (B) Average occupation of helical (up, including  $\alpha$ -helix,  $3_{10}$ -helix, and  $\pi$ -helix) and  $\beta$ -strand (down, including parallel and anti-parallel strand) secondary structures of the APP (blue) and Notch (red) TMD over time and simulation replicas. Occupation of secondary structures of the substrate TMDs binding to PS1-D257<sup>H</sup> is shown in Figure A.3A

### 4.3. RESULT

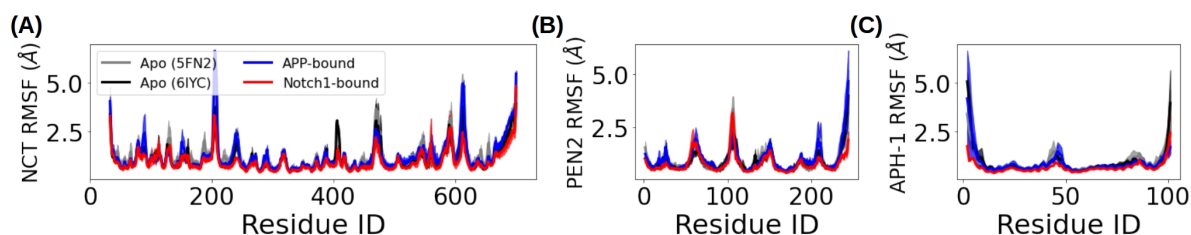


Figure 4.3.: Residue-wise flexibility of the  $\gamma$ -secretase non-catalytic subunits calculated by RMSF. (A) RMSF of NCT. (B) RMSF of PEN2. (C) RMSF of APH-1. RMSF of the 5FN2-derived apo-form, 6IYC-derived apo-form, APP-bound form, and Notch1-bound form is colored in gray, black, blue, and red, respectively. Standard deviations of each RMSF across two simulation trajectories are shown with transparent shades. Corresponding RMSFs of the holo-form in the PS1-D257<sup>H</sup> protonation state are shown in Figure A.3B-D

also discuss how the length and sequence C-terminal to the substrate cleavage site affects the  $\beta$ -sheet formation and the stability of the active site geometry. On the other hand, the side chain contributions of P1'-P3' are mediated by hydrophobic interaction with PS1 TMD6a and S2' pocket (Figure (Figure 4.4C) and is responsible for the stabilization effect observed at PS1 TMD6a (Figure 4.4A). The hydrophobic contact with TMD6a is also widely observed in the structure of all GSI-bound  $\gamma$ -secretase complexes (Figure 2.3A). With collaboration with the experimental group, we have presented that interruption of TMD6a helix formation of PSH by introducing mutations or change in lipid environment can lead to a reduction in  $\gamma$ -secretase activity and weaken the binding of GSI (Chapter 5). The second strongest interacting region locates at the C-terminal end of the TMD helical domain (side chains of P5 and P6, Figure 4.4B), where  $3_{10}$  helix was observed (Figure 4.2C). P5 and P6 residues of APP and Notch1 are compactly surrounded by several hydrophobic residues located on TMD2, TMD3, and TMD5 (Figure 4.4C) and this explains the reduction in the flexibility of these TMDs (Figure 4.4A). Importantly, plenty of residues of PS1 and APP involved in this interaction are reported as FAD mutations (Table 4.2 and Table 4.3). We note that large and hydrophobic amino acids at positions P5 and P6 are found not only in APP and Notch1 but also in a wide range of, although not all, known  $\gamma$ -secretase substrates. We suggest that this interaction is decisive for the substrate to reside inside the enzyme, especially when trimmed to the short A $\beta$  form, and will be extensively studied with homology modeling and *in silico* mutagenesis methods. Atomic geometries at PS1 internal docking site regions in the apo-state, APP-bound state, and Notch1-bound state are shown in Figure A.7. Notably, the  $3_{10}$  helix at this position nicely aligns A42 to the P5' docking pocket and I41 to the P6' docking pocket (Figure A.6), indicating a high possibility that APP can undergo  $i \rightarrow i+3$  processive cleavage through a translational movement at the helical part.

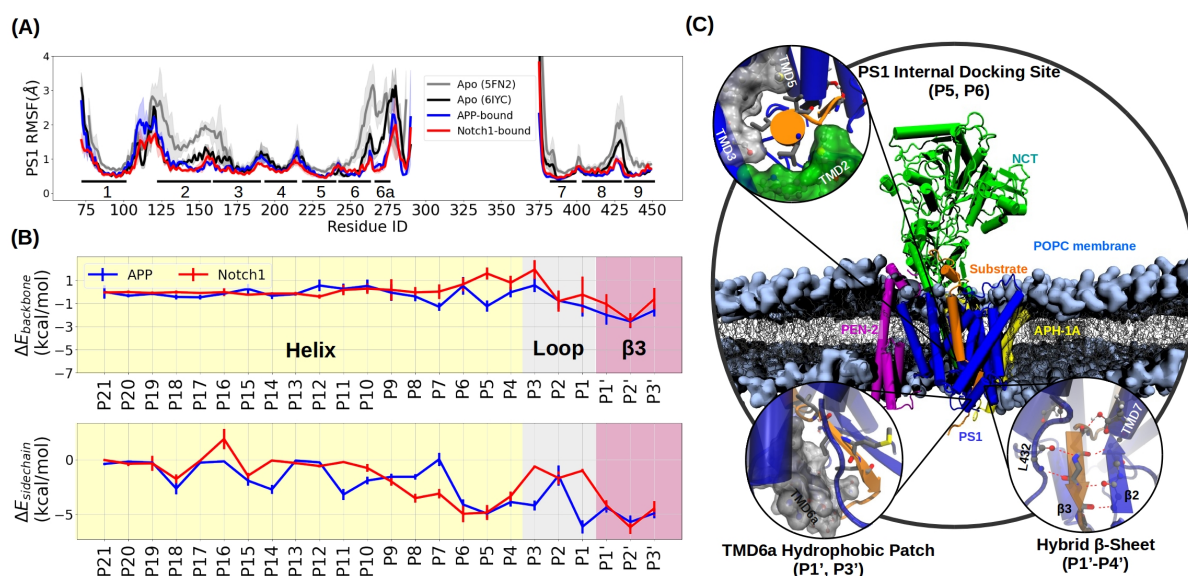


Figure 4.4.: Interactions between  $\gamma$ -secretase and its substrates. (A) Residue-wise RMSF of the catalytic subunit PS1 in the apo- and holo-state. RMSF curves are represented the same way as in Figure 4.3. Numbers under each line indicate the TMD numbers each residue locates on. (B) Residue-wise interaction energy of APP (blue) and Notch1 (red) was analyzed by MM/PBSA. Energy contribution from the backbone and side chain atoms are plotted in the upper and lower panels, respectively. Helical, loop, and  $\beta$ -3 domains are determined from the secondary structure analysis from Figure 4.2B and colored in yellow, gray, and red, respectively. (C) Schematics of the three E-S interaction sites highlighted in the dissertation, including the PS1 internal docking site (upper left), the TMD6a hydrophobic patch (lower left), and the hybrid  $\beta$ -sheet (lower right). The substrate residues involved in the corresponding interactions written in the parenthesis. PS1 residues surrounding P5 and P6 are listed in Table 4.2 and their Van der Waals surface are represented by the white and green surfaces, respectively, in the upper left highlighting circle. Van der Waals surface of PS1 residues R268-R278 is represented by the white surface in the lower left highlighting circle. PS1 RMSF of the holo-state and the substrate binding energy contribution in the PS1-D257<sup>H</sup> protonation state are shown in Figure A.3E and F, respectively.

Table 4.2.: PS1 residues in the proximity of substrate P5 and P6 and their related FAD mutations reported on alzforum.org.

	Residues surrounding substrate P5	Residues surrounding substrate P6
PS1 residues	S169, S170, L173, I213, L226, I229, M233, L286, I287, L383, I387	M146, L150, I162, W165
Related FAD mutations	S169P, DS169, S170F, L173W, I213L, I229F, M233T, M233L, L286V	M146L, L150P, W165G

Table 4.3.: FAD mutations around P5 and P6 of APP reported on alzforum.org.

Position	P4	P5	P6	P7
APP FAD mutation	V46F, V46G, V46I, V46L	I45F, F45M, I45T, I45V	V44I, V44M	T43A, T43I

### 4.3.3. Open and close geometries at the catalytic center

Upon the association of the substrate with the enzyme interior, the substrate scissile bond has to be exposed to the catalytic center. The catalytic center of  $\gamma$ -secretase is formed by two aspartic acids located on TMD6 and TMD7 of the PS subunit (D257 and D385 on PS1 and D263 and D366 on PS2). As an aspartyl protease, the proteolysis of  $\gamma$ -secretase is executed via hydrolysis reaction through a chain of the proton transfer process. As depicted in Figure 3.4, the hydrolysis process requires the recruitment of water molecules to the proximity of the catalytic site and a catalytic hydrogen bond formed between the protonated aspartic acid and the carbonyl group of the scissile bond. Indeed, in the simulations of the substrate-bound systems, around one water molecule can access the loop region of the substrate while most residues at the helical domain are isolated from the water environment (Figure 4.5A). Furthermore, the catalytic hydrogen bond was found constantly in the APP-bound system and occasionally in the Notch1-bound system (Figure 4.5B). Interestingly, when no substrate is present in the  $\gamma$ -secretase active site, the catalytic center exhibited two geometrical states: a close state and an open state, distinguished by the  $C_{\gamma}$ - $C_{\gamma}$  distance between the catalytic aspartates (Figure 4.5C, D). The close state, with a  $C_{\gamma}$ - $C_{\gamma}$  distance around 4.5 Å, is stabilized by the direct hydrogen bond between D257 and D385 whereas the open state, with a  $C_{\gamma}$ - $C_{\gamma}$  distance around 6.5 Å, is maintained by a water molecule which forms hydrogen bonds with both aspartic acids. Interestingly, a  $C_{\gamma}$ - $C_{\gamma}$  distance of 6.5 Å was also frequently sampled with the presence of the substrate (Figure 4.5C and D), indicating that the open state geometry at the catalytic center is well-capable of accommodating the substrate scissile bond. In our published work[78], we have shown that the open state is energetically the more favorable state in both apo- and holo-form  $\gamma$ -secretase and that FAD mutations that skewed the balance toward the close state or deviate the  $C_{\gamma}$ - $C_{\gamma}$  distance are closely associated with the loss in enzymatic function. Because the detailed calculation of the free energy landscape along the close-open state transition and its response to FAD mutations were already reported in the author's Master dissertation, they are not repeated in the present chapter.

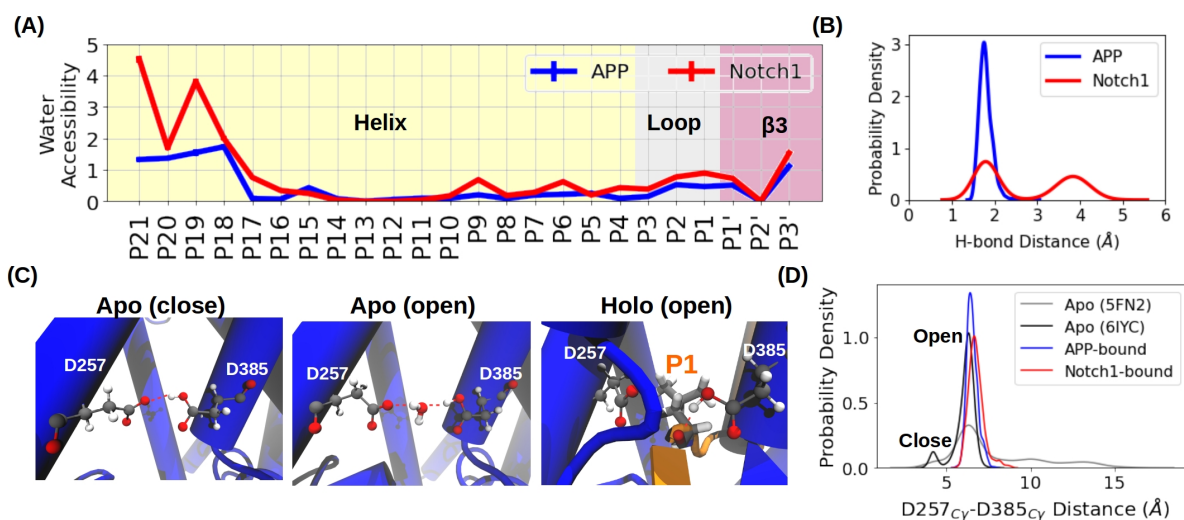


Figure 4.5.: Water accessibility and geometry at the catalytic center. (A) Residue-wise water accessibility of the TMDs of APP (blue) and Notch1 (red). (B) Probability density of the distance of the catalytic hydrogen bond distance ( the distance between the protonated aspartate and the carbonyl group of the scissile bond) in APP-bound (blue) and Notch1-bound (red) simulations. (C) Schematics of the close and open states in the apo-enzyme and the open state in the holo-enzyme. PS1 is colored in blue and the substrate is colored in orange. (D) Probability density of the  $C_{\gamma}$ - $C_{\gamma}$  distance between D257 and D385 sampled the simulations. Distribution curves are represented the same way as in Figure 4.3. Water accessibility of substrate TMD, probability densities of the catalytic hydrogen bond and  $C_{\gamma}$ - $C_{\gamma}$  distance between D257 and D385 in the PS1-D257<sup>H</sup> protonation state are shown in Figure A.4A, B, and C, respectively.

## 4.4. Discussion

$\gamma$ -Secretase is an intramembrane protease that cleaves its substrates with a hydrolysis reaction within the hydrophobic membrane environment. Among many substrates, APP and Notch1 are clinically the most studied substrates because of their roles in Alzheimer's disease[28, 4] and cancer[23, 24], respectively. Because of their pathological importance, numerous efforts have been invested in understanding the biological pathway of  $\gamma$ -secretase proteolysis[9]. However, due to the high molecular weight and high fragility of membrane proteins, the 3D structure of  $\gamma$ -secretase is difficult to crystallize for X-ray diffraction. With the advances of cryo-EM technology, 3D structures of the apo- and holo-form  $\gamma$ -secretase were finally resolved in recent years, providing valuable structural information about the topology of  $\gamma$ -secretase and how its biological substrates and small molecules bind to it.

In this chapter, we reported the results obtained from free simulations and highlighted several important findings. Regardless of the presence of the substrate, the up-down bending motion of the NCT ectodomain was observed. By comparing the topology of the substrates in the enzyme-free simulations, we speculate that the bending motion plays an essential role in the initial substrate recognition and adjusting the tilting angle of the substrate. In order to bind into the interior of  $\gamma$ -secretase and form a catalytically active geometry, the substrates have to undergo a structural transition from the complete helix to a helix-loop-strand conformation. In particular, a  $3_{10}$ -helix was observed at the C-terminal end of the helical domain, which rationalizes the use of Aib in HIP design. By looking at the stabilized regions of PS1 and the residue-wise interaction profiles of APP and Notch1, we identified three main interaction sites: the TMD6a hydrophobic patch, the hybrid  $\beta$ -sheet, and PS1 internal docking site. These three interactions constitute the foundation of the present dissertation and will be discussed in detail in Chapter 5, Chapter 6, and Chapter 7, respectively. Furthermore, we also studied the geometry of the catalytic center and found two possible states, namely the close and the open states, distinguished by the distance between the catalytic aspartates. In comparison to the close state, the open state is thermodynamically the more favorable state in the apo-form and is compatible with accommodating the substrate scissile bond.



## 5. Influence of Membrane Environment on C83-PSH Interaction

Intramembrane proteases convey their functional roles by cleaving their substrates within the lipid bilayer. Although the constitution of the lipid environments is known to modulate the enzyme activity, the underlying mechanism of this modulation remains unclear. Conventionally, experiments on the cleavages of these insoluble proteases are conducted in a detergent-based environment instead of the biological amphiphilic phospholipid molecules. In collaboration with our experimental partner, we studied the interactions between the E-S complex and the micelles/bilayer environment using homology modeling and MD simulations. The computational result suggested that the decrease in APP processivity and activity of PSH can be linked to the instability of the TMD6a helix in the micelles environment. The hypothesis is verified by *in vitro* and *in silico* experiments. Our data provide an in-depth molecular insight into how a detergent-based environment can modulate the dynamics of the membrane-embedded proteins. The work is published as reference [166] with a massive reduction of experimental part.

### 5.1. Introduction

Membrane-embedded proteins, also called integral membrane proteins (IMPs), are proteins residing in the biological membranes and are involved in numerous crucial functions such as signaling, transporting, and cell adhesion [167, 168]. It is estimated that around 30% of human proteins are IMPs[167]. Compared to the other types of proteins, their biological functions and structures are poorly understood, despite the abundance of human genes and the critical roles they play in our lives. Since IMPs are generally not soluble in aqueous solution, studying membrane proteins requires the protein to be extracted from their native membrane environment. Because of the difficulties in carrying the purified proteins into the reconstituted membrane environment, experiments on IMPs are usually conducted in detergent-based or mixed lipid-detergent systems[169, 170], although it has been shown that substituting the membrane environment with detergent can affect the functions and structures of membrane proteins[171, 172, 76]. To correctly translate the *in vitro* experimental result obtained using the detergent-based environment to the natural scenario, it is important to understand the molecular mechanism of how the change in environmental factors affects enzyme activity.

Several studies have indicated the the constitution of the lipid environment has a critical impact on the function of  $\gamma$ -secretase[173, 174, 175, 176, 177]. Although  $\gamma$ -secretase cleaves its substrate with PS, its catalytic function is only active through the complicated maturation

process with the help of the other three components: NCT, PEN-2, and APH-1[9]. That is, conformational changes in the other subunit of  $\gamma$ -secretase induced by the change in lipid constitution can be the cause of the modulation effect. As shown in Figure 2.1 and Figure 2.2, and Figure 2.4, PSH folds into a tertiary structure as PS without the presence of other components of  $\gamma$ -secretase. Besides the structural similarity, PSH also cleaves C99 in a sequential manner starting at the initial  $\epsilon$  (either  $\epsilon$ 48 or  $\epsilon$ 49) cleavage sites and releasing similar  $A\beta$  species by  $\gamma$ -secretase, making it an attractive model for the intrinsic protease activity of PS[46]. In collaboration with the experimental partner, we constructed the C83-bound PSH structure *in silico* using homology modeling. Experimental data showed that PSH possesses stronger processivity and activity in the membrane bilayer than in the detergent micelles. By solvating the constructed E-S complex with these two environments, we studied the dynamical properties of the complex using MD simulations. In the most promising homology model, we found an increase in the flexibility of TMD6a when the membrane bilayer is substituted by detergent micelles. To verify whether the flexibility of TMD6a is responsible for the enzyme activity, we introduced lysine on TMD6a *in silico* to disrupt both the secondary structure and the biophysical property of the hydrophobic patch. In good agreement with the experiment data, our result suggests that the interactions between TMD6a and C83 play an important role in substrate stabilization and enzyme activity.

## 5.2. Methods and Materials

The available PSH structures (PDBID 4HYC, 4HYD, 4HYG[43], and 4Y6K[46]) are crystallized with several mutations and the removal of the loops between TMD1 and TMD2 and between TMD6 and TMD7. In addition, as discussed the Chapter 2 and depicted in Figure 2.1, the tetrameric structure observed in these structures is likely not the natural form where PSH conducts its biological function. Under the assumption that PSH binds C83 in a similar way as the PS1-containing  $\gamma$ -secretase, we took the structural information of C83-bound  $\gamma$ -secretase (PDBID 6IYC[39]) and apo-form PSH (PDBID 4HYG[43]) as templates to generate three C83-bound PSH models using the SWISS online server[178] and the comparative modeling program MODELLER[116] based on sequence alignment shown in Figure 5.1. Model 1 was built by taking only PDB 6IYC as the template as generated by the SWISS online server (Table 5.1). Model 2 was generated by using residues ranging from L7 to D162 and D220 to L292 from chain B of PDB 4HYG and model 1 as templates with the MODELLER multi-template method (Table 5.1). Model 3 was built by taking all residues resolved in chain B of PDB 4HYG and model 1 with the MODELLER multi-template method (Table 5.1). Similar to the PS1 template structure, the final holo PSH structure is composed of two fragments with an N-terminal fragment from L7-R193 and a C-terminal fragment from E210 to A293.

Each C83-bound PSH model was embedded in the bilayer and micelles environments with tip3p water and 0.15M potassium chloride through the CHARMM-GUI server[179]. The bilayer environment constitutes 302 POPC molecules (152 in the upper leaf and 150 in the lower leaf) and the micelle environment constitutes 150 n-dodecyl  $\beta$ -D-maltoside (DDM) detergent molecules. To further understand how the size of micelles alters the formation of

micelles and influences the DDM-PSH interaction, a larger micelle consisting of 50% more DDM molecules (225 DDM in total) is prepared for the best-performing model (model 2). Lysine mutations M172K, I173K, L175K, and A176K were constructed via AmberTools18 based on model 2 with RMSD (WT vs mutant) < 0.1 Å and embedded in a membrane bilayer system with 302 POPC molecules using the CHARMM-GUI online server. Because the forcefield parameters of DDM are not determined in the lipid17 forcefield, the atomic interaction of POPC, DMM, water, ions, and proteins are all described by the charmm36m forcefield[180]. Energy minimization, equilibration, and NPT production run are performed with the procedure described in Chapter 3 section 3.3. Three simulations with 600ns each were performed for each system, in total 33 NPT trajectories were generated for further analysis. The non-bonded cutoff was set to 12 Å with a force-based switching distance of 10 Å. D220 was selected to be protonated while D162 was unprotonated according to the pKa prediction on the existing PSH and PS1 structures by PROPKA3.1 (Table 3.1).

Residue-wise RMSFs are calculated with respect to the averaged structure in each simulation and residue-wise water accessibilities are calculated by counting the average number of water molecules within 5 Å of the corresponding residue. The catalytic hydrogen bond is defined by the distance between the carbonyl group of C83 L49 and the catalytic proton of D220. The lipid tail order parameter SCH was computed in model 2 in DDM and POPC environments to show the orientation and the ordering of the concerning CH vector with respect to the protein principle axis, which was aligned to the lipid normal in the POPC environment. Secondary structure analysis is conducted using the DSSP method[165].

Table 5.1.: Templates used for the model building of PSH in complex with C83. Residues of PSH used for model building are indicated. \* PSH residues L7-R193 and E210-A293 were modeled based on the template.

	Template		
	6IYC (PS1, C83)	4HYG (Chain B)	Model 1
Model 1	all*	-	-
Model 3	-	L7-D162, D220-L292	all
Model 3	-	L7-A176, D210-L292	all

Table 5.2.: RMSD values between the available crystal structures (PDB 4HYG and 4Y6K) and the three different PSH models.

RMSD (Å)	4HYG (Chain B)	4Y6K (Chain B)	Model 1	Model 2	Model 3
4HYG (Chain B)	0				
4Y6K (Chain B)	0.349	0			
Model 1	3.379	3.349	0		
Model 2	1.573	1.602	1.355	0	
Model 3	1.416	1.513	1.484	0.163	0

## 5.2. METHODS AND MATERIALS



Figure 5.1.: Alignment of PS1 and PSH used for homology modeling based on the TMD annotations of the available cryo-EM or crystal structures, respectively (PDB 6IYC for PS1 and PDB 4HYG for PSH). Active site motifs including the catalytic aspartate residues (red) are highlighted in bold. The  $\beta$ 2-strands of  $\gamma$ -secretase and PSH are indicated as orange arrows and residues that were mutated in this study are highlighted in violet.

## 5.3. Result

### 5.3.1. Structural modeling shows key features and active geometry of substrate-bound $\gamma$ -secretase in PSH

Because the biological role of proteins is highly related to how they fold in 3D space, structural information of a protein is very important to understand how it conducts its function in solution. Since no holo-state PSH structure has been determined experimentally until the date of the dissertation, we generated three putative C83-bound PSH structures under the assumption that C83 binds to PSH in an analogous fashion as in the PS1 homolog (Figure 5.2A). Because the first model was constructed by only taking the information of the C83-bound PS1[39] structure, it is structurally largely deviated from the X-ray determined PSH apo-structure[43] (Table5.1 and Table5.2). The other two models were constructed on the basis of the first model and the available apo-PSH structure (Table5.1 and Table5.2). With the additional information from the C83-bound  $\gamma$ -secretase structure, the generated models reproduced key structural features responsible for substrate binding of  $\gamma$ -secretase, including the bending motion of the C-terminal of TMD6 (Figure 5.1B-D) and the hybrid  $\beta$ -sheet C-terminal to the active site (Figure 5.2E). In the  $\gamma$ -secretase structure, TMD6a is stabilized by a hydrogen bond between Y159 on TMD3 and R278 on TMD6a (Figure 5.1C). This structural feature is reconstructed by the salt bridge between R70 and E181 of PSH in our models (Figure 5.1D). The  $\beta$ 3-strand of C83 is stabilized by the  $\beta$ 2-strand (A213-G217) and residue Q272 of PSH, forming a hybrid  $\beta$ -sheet.

The structure generated through homology modeling is sensitive to the sequence alignment strategy and the choice of template. Therefore, it is necessary to verify the generated structures using experimental tools. It has been shown that the activity of  $\gamma$ -secretase toward C83 is impaired when the  $\beta$ 2-strand (R377-L381) is deleted. To verify the reliability of the modeled structures, the residues located on the putative  $\beta$ 2-strand are mutated or deleted, as done for PS1(Figure 5.1E). With the strong decrease in the enzyme activity in comparison to the WT enzyme, our data suggest that these residues are indeed critical for the substrate cleavage of PSH and likely form into a  $\beta$ 2-strand as predicted in our homology models(Figure 5.1F).

### 5.3.2. DDM micelle destabilizes the catalytic site geometry

To study the environmental influence of PSH conformation and E-S interaction, the constructed C83-bound PSH models were embedded in the POPC and DDM micelle environment (Figure 5.3A-C) and three 0.6 $\mu$ s trajectories were generated with random initial velocity distribution. Among the homology models, the first model which was created purely with the  $\gamma$ -secretase template exhibited the highest RMSD whereas most trajectories in models 2 and 3 showed an RMSD of around 3 Å, indicating the instability of model 1. Despite a lower overall RMSD, TMD6a in model 3 failed to form an interaction with the substrate in the POPC environment (Figure 5.3G) and showed a relatively high flexibility at TMD4 and residues C-terminal to TMD6a (Figure 5.3H). Taken together, model 2 showing high stability

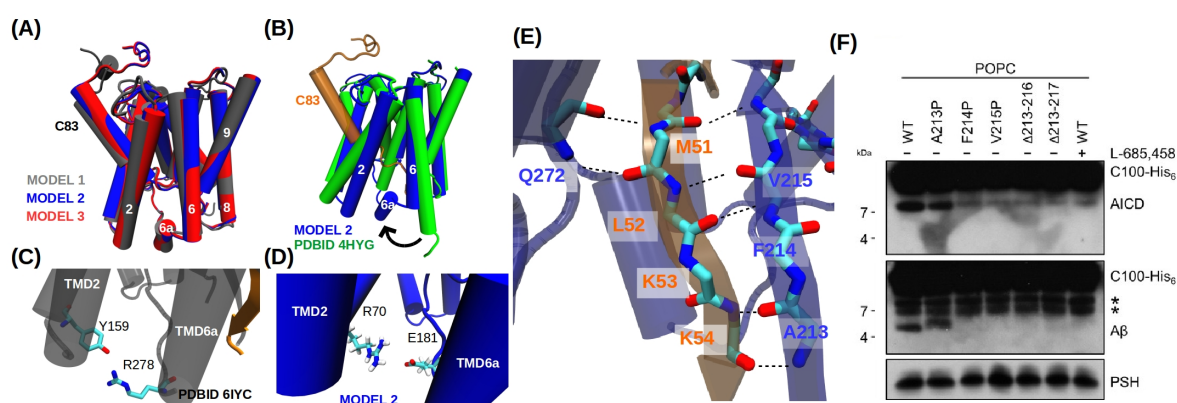


Figure 5.2.: Homology modeling reconstructed key features verified by experiments. (A) Comparison between the three constructed models in this study. (B) Comparison between model 2 (blue and orange) and the experimentally determined apo-PSH structure (light green, PDBID 4HYG). (C) The hydrogen bond between R278 and Y159 resolved in the C83-bound  $\gamma$ -secretase structure (PDBID 6IYC). (D) A salt bridge formed between R70 and E181 in model 2. (E) Reconstructed hybrid  $\beta$ -sheet formed by  $\beta$ 2-strand (A213, F214, V215) and Q272 or PSH (blue) and M51-K54 of C83 (orange). (F) Analysis of WT and mutant PSH activity in POPC environment. GSI L-685,458 is used as the control group. Experiments are done by the collaboration partner instead of the author of the dissertation. Results obtained in the DDM environments and the methods and materials used in the experiments can be found in our published work.

and form expected C83-TMD6a contact was considered as the most realistic structure and was used for the following analysis. In order to avoid bias from the choice of micelle size, simulations of model 2 embedded in a micelle constituting 25% more DDM molecules (DDM 225, Figure 5.3C). Indeed, dynamics of the C83-bound PSH structure from model 2 showed a similar stability (Figure 5.3E), substrate water accessibility (Figure 5.4A, catalytic geometry (Figure 5.4B, C), and PSH RMSF profile (Figure 5.5A). In the following context, only model 2 is used for analysis and is referred to directly without explicit statements.

We next had a closer look at the catalytic center of the E-S complex in different environments. Consistent with simulations in the holo-state  $\gamma$ -secretase (Figure 4.5A), the substrate remained dry in the TMD domain (G29-K52) but the residues around the scissile bond are well solvated with water molecules (Figure 5.4A) with no significant difference between POPC and DDM environments. In general, the formation of catalytic hydrogen bonds in PSH is much less frequently formed than in  $\gamma$ -secretase (Figure 4.5B and Figure 5.4B). However, this hydrogen bond is even less stable when the C83-PSH complex is embedded in DDM micelles. In contrast, the active site geometry, defined by the  $C_{\gamma}$ - $C_{\gamma}$  distance between D162 and D220 (Figure 5.4C, D) showed high stability in an open-state ( $C_{\gamma}$ - $C_{\gamma}$  distance around 6.5 Å. An active site geometry with an even larger distance at around 8.2 Å, mediated by two water molecules between D162 and D220 (Figure 5.4E), was also occasionally sampled (Figure 5.4C). This indicates that although the ability of PSH to form a catalytic hydrogen bond might not be as good as in  $\gamma$ -secretase, the catalytic geometry is nevertheless stable in the POPC environment. In contrast, the DDM micelle might destabilize both the catalytic hydrogen bond and the active site geometry.

To see how the lipid environment influences PSH cleavage, the cell-free assay was conducted by our collaboration partner. Interestingly, as seen by the decreased A $\beta$ 38 level and the increase in  $\beta$ 42, the processivity of PSH is weakened in the DDM micelles environment (Figure 5.4F). Furthermore, secretion of A $\beta$  species longer than A $\beta$ 42 was observed in the DDM micelles, indicating an early release of A $\beta$  substrate during the processive cleavage. To validate whether the decrease in processivity comes from the instability at the active site geometry, as predicted by MD simulations, the L685,458-based biotinylated affinity ligand Merck C was used to probe the stability of the catalytic center. Indeed, quantification of specifically Merck C-bound PSH was markedly less pronounced in the DDM micelles than in the POPC bilayer.

### 5.3.3. Increase in flexibility and unwinding of PSH TMD6a helix induced by detergent-enzyme interaction

With the decrease in processivity and stability of the active site geometry, we next aim to understand the molecular mechanism of the modulation effect from the lipid environment. By comparing the RMSF profiles of PSH in different environments, an increase in TMD6a flexibility is observed in both medium size (with 150 DDM) or larger (with 225) DDM micelles (Figure 5.2A, B). In our C83-bound PSH models, the TMD6a helix creates a hydrophobic

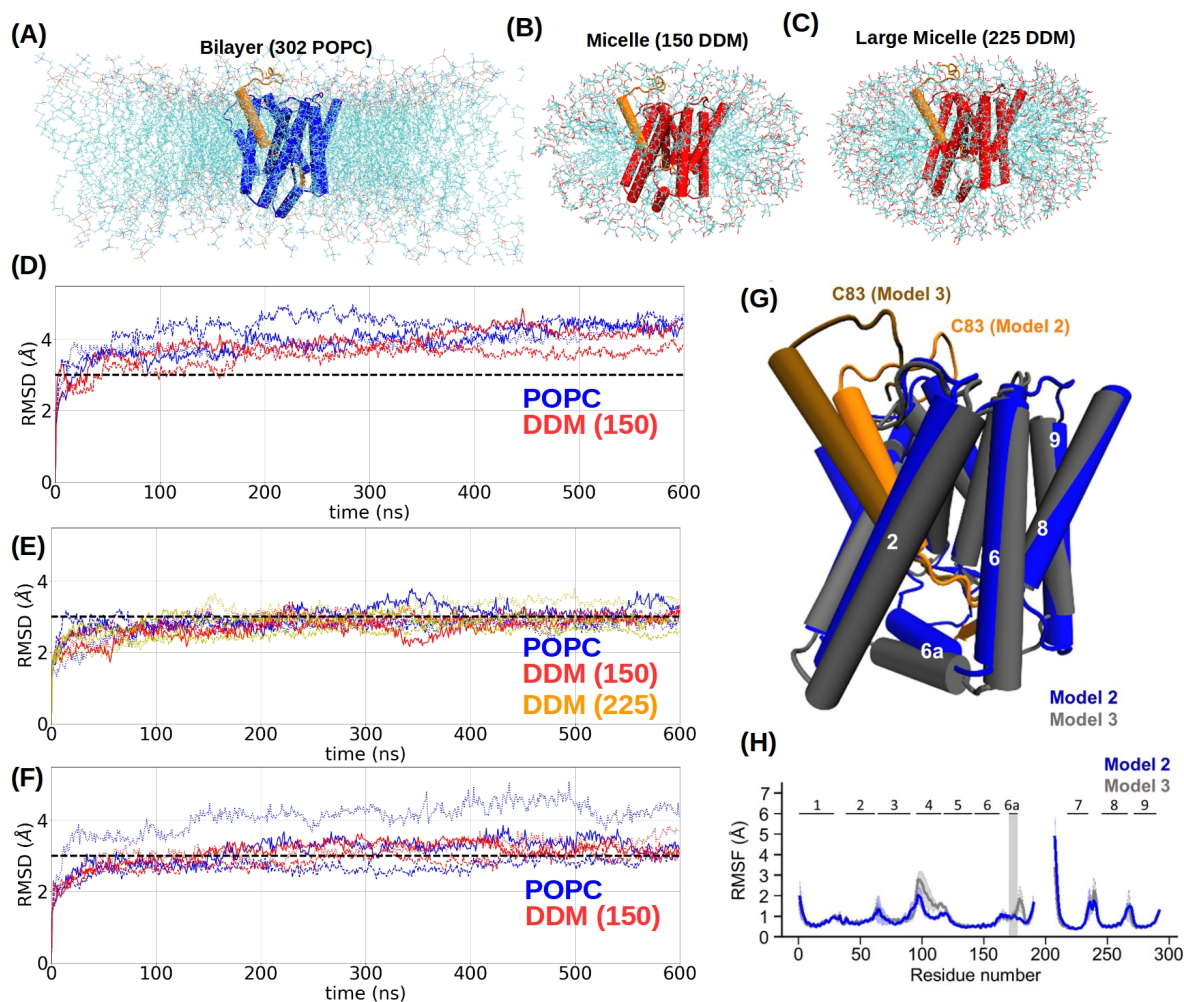


Figure 5.3.: Comparison between different homology models using MD simulations. (A)-(C) Schematic view of the model 2 C83-PSH complex embedded in (A) POPC bilayer, (B) DDM micelle, and (C) larger DDM micelle. (D)-(F) Structural stability of PSH calculated by RMSD in POPC (blue), medium-size DDM (red), and large DDM (yellow) environments using homology (D) model 1, (E) model 2, and (F) model 3. (G) Structure superposition between the C83-bound PSH from model 2 (PSH in blue and C83 in orange) and model 3 (PSH in gray and substrate in dark brown). (H) Residue-wise RMSF of PSH in model 2 (blue) and model 3 (gray). The shaded areas represent the standard deviation computed across three simulations and numbers in the plot indicate the residue range of each TMD.



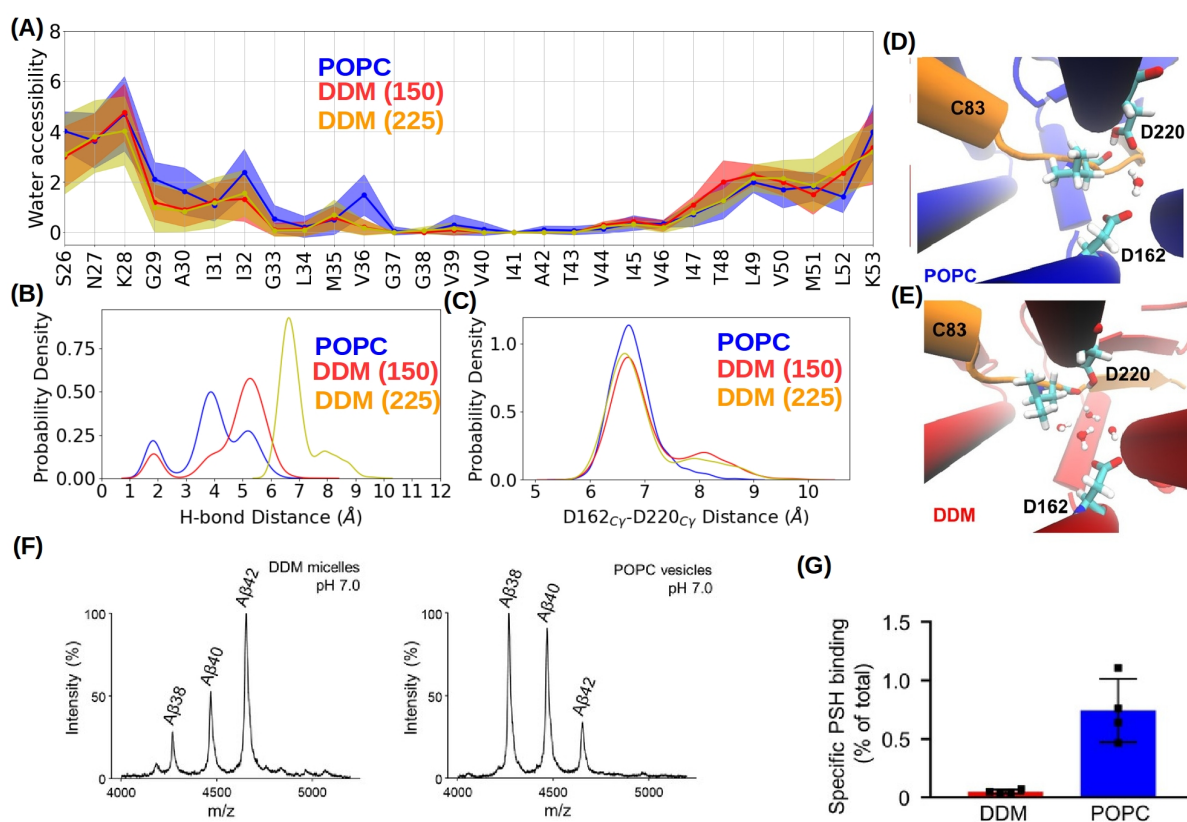


Figure 5.4.: Water accessibility of C83 in the PSH-bound form and the stability of the PSH active site geometry. (A) Reisdidue-wise water accessibility of C83. The modeled scissile bond locates between L49 and V50. The shaded areas represent the standard deviation computed across three simulations. (B) Probability density of the distance of the catalytic hydrogen bond sampled in different environments. (C) Probability density of the D162<sub>C $\gamma$</sub> -D220<sub>C $\gamma$</sub>  distance sampled in different environments. (D) Schematic of the active site geometry with D162<sub>C $\gamma$</sub> -D220<sub>C $\gamma$</sub>  distance of 6.8 Å sampled in the POPC environment. (E) Schematic of the active site geometry with D162<sub>C $\gamma$</sub> -D220<sub>C $\gamma$</sub>  distance of 8.2 Å sampled in the DDM environment. (F) A $\beta$  species generated by PSH in DDM micelles and POPC vesicles at pH 7.0. Representative mass spectra from four independent biological replicates are shown. The intensity of the highest peak was set to 100%. (G) Quantitation of PSH binding by Merck C. Specific binding was defined as the difference of PSH signals in the absence or presence of L-685,458 after additional subtraction of unspecific background binding signals. (F) and (G) are not generated by the dissertation of the author. Methods and materials used in the experiments can be found in our published work.

patch formed by residues M172, I173, L175, and A176, in contact with V50 and L52 of C83. Importantly, this interaction was observed in all substrate-bound and GSI-bound  $\gamma$ -secretase structures (Figure 2.3A and Figure 4.4C). To gain the intrinsic difference between the micelles and bilayer, we calculated the lipid order parameters, which indicates how well lipids are aligned to the principle axis of the protein, and how they interact with PSH. As expected from its planer geometry (Figure 5.3A), POPC molecules are well aligned with the membrane normal and TMDs of PSH, as indicated by the high order parameter (Figure 5.5C, D). In contrast, a much lower order parameter was observed in the DDM-based systems (Figure 5.5C, D), as also expected from its spherical shape (Figure 5.3B, C). In the POPC environment, TMD6 remained in the stable helix and the hydrophobic patch is in contact with V50 and L52 of C83 in all simulations (Figure 5.6 and Figure 5.5E). We note that the helical conformation is necessary for residues M172, I173, L175, and A176 to face in the same direction and form the hydrophobic patch. Nonetheless, TMD6a was unwound in the simulations of DDM-embedded systems (Figure 5.6). When looking into the structural detail, we observed that the DDM head groups can insert into the gap between TMD2 and TMD6, forming unspecific hydrogen bonds with the backbone atoms of TMD6a (Figure 5.5F, G). The insertion of the lipid molecules is facilitated by its intrinsically low order parameter that enables the DDM molecules to align perpendicular to PSH TMDs. The hydrogen bond interactions from the DDM molecules compete with the intramolecular  $i \rightarrow 3.6$  hydrogen bond required for the formation of  $\alpha$ -helix and therefore unwound the TMD6a helix. Collectively, our data suggest that replacing the POPC bilayer with DDM micelles destabilizes the stability of the TMD6a helix, an important structural element involved in substrate stabilization.

#### 5.3.4. Lysine mutations in TMD6a lead to helix unwinding and reduced activity

Our computational results suggest that the interactions between PSH TMD6a and C83 are disrupted in the DDM micelles and might be responsible for the experimentally observed reduced enzyme activity. Notably, the corresponding hydrophobic patch in the TMD6a of PS1 (L271, V272, T274, and A275) is also affected by FAD mutations and mutations such as L271V and T274R were reported to impair the  $\gamma$ -secretase cleavage, supporting the idea that TMD6a has an important function in substrate cleavage. In order to specifically study and examine the functional role of TMD6a, we introduced lysine mutations, positively charged residues, at residues M172, I173, L175, and A176 in the TMD6a hydrophobic patch (Figure 5.5E). These mutations are expected to destruct the apolar characteristic of the TMD6a helix and weaken the interaction with C83.

Because the TMD6a was found unstable already in the wt PSH in the DDM environment, we studied the dynamics of the mutated PSH only in the POPC bilayer. In all lysine mutants, no significant difference was found in the water accessibility to the C83 residues compared to the wt PSH (Figure 5.7A). However, the catalytic hydrogen bond and the active site geometry are destabilized in the mutated system (Figure 5.7B, C). In the cases of the L175K and A176K mutants, the catalytic hydrogen bond was rarely sampled during the simulations. Although a similar population of catalytic hydrogen bonds was formed in the M172K and I173 mutants,

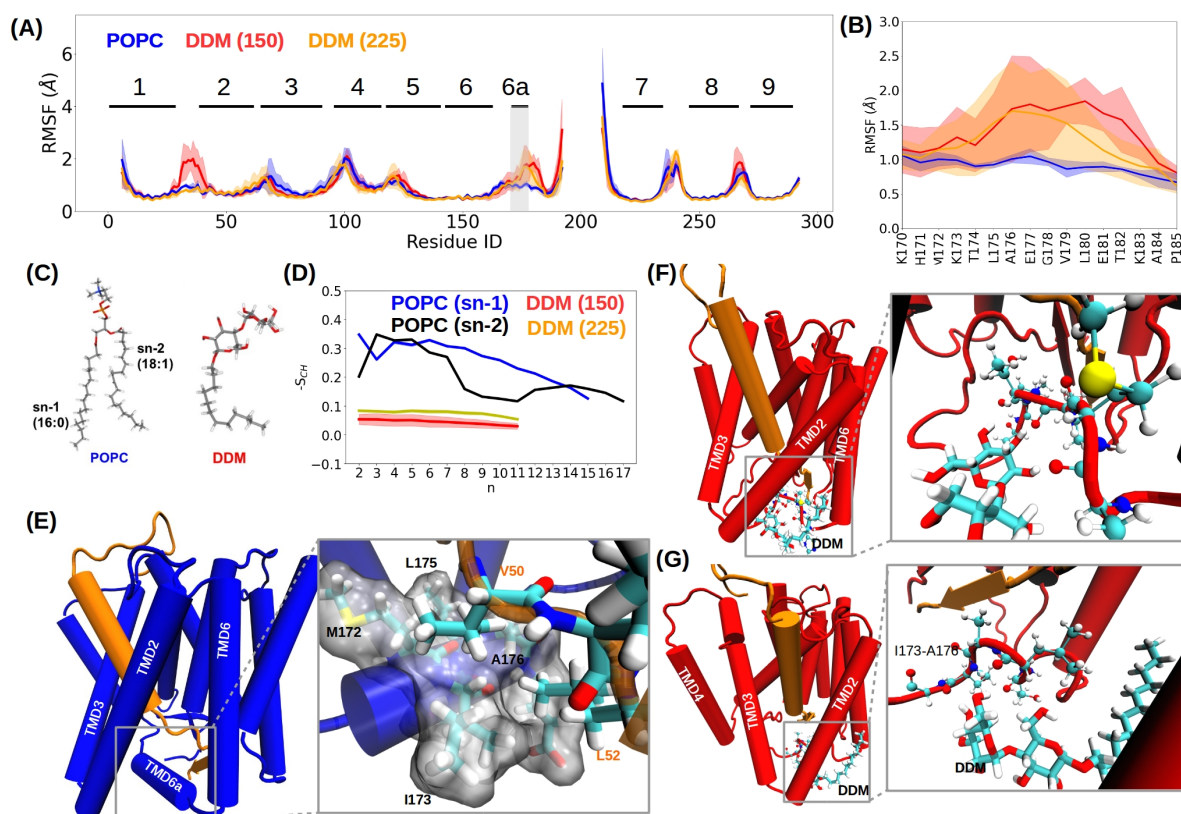


Figure 5.5.: Influence of lipid environments on the PSH stability. (A) Residue-wise RMSF of C83-bound PSH in the POPC (blue), medium-size DDM (red), and large DDM (yellow) environments. Numbers in the plot indicate the residue range of each TMD. (B) Zoom-in view of RMSF values of the TMD6a residues. Color codes are adopted from (A). (C) Structures of POPC and DDM molecules. (D) Lipid order parameter calculated in the MD simulations. The shaded areas in (A), (B), and (D) represent the standard deviation computed across three simulations. (E) Structural snapshot of C83-bound PSH embedded in the POPC bilayer at the end of the simulation. The zoom-in view shows the hydrophobic contact between PSH TMD6a and V50 and L52 of C83. (F)-(G) structural snapshots of C83-bound PSH embedded in the medium-size DDM micelles at 400ns in (F) the first and (G) the second simulation. The zoom-in views show the atomic interactions between DDM and the backbone atoms of TMD6a.

### 5.3. RESULT

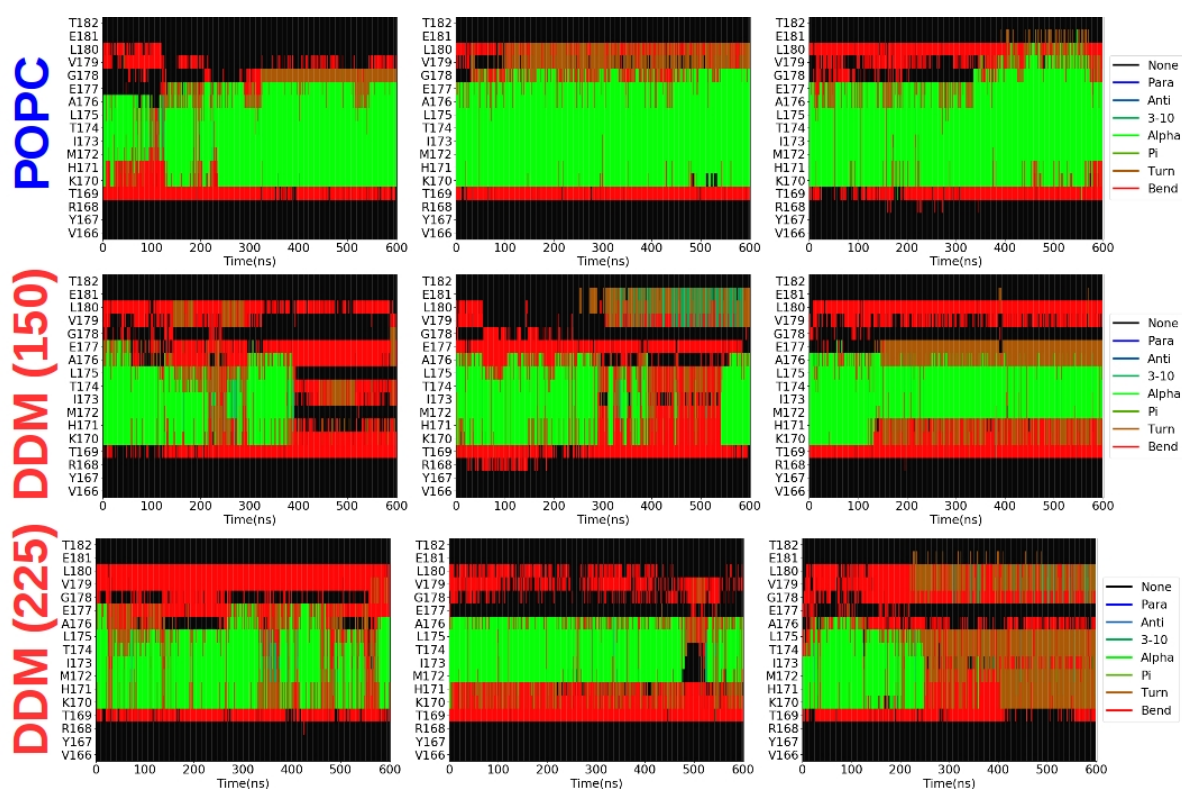


Figure 5.6.: Time-plot of the secondary structure of PSH TMD6a in (top) POPC, (middle) medium-size DDM, and (bottom) large DDM environments. Results obtained from three  $0.6 \mu\text{s}$ -long simulations are shown from left to right.

the active site geometry is destabilized. Interestingly, a D162<sub>C $\gamma$</sub> -D220<sub>C $\gamma$</sub>  distance of 8.2 Å sampled in the DDM-based environment (Figure 5.4C) was also found in the M172K and A176K mutants (Figure 5.7C). As shown by the PSH RMSF plots, TMD6a was destabilized in all mutants while only smaller effects in other regions were affected in all PSH mutants, in comparison to WT (Figure 5.7D). To different degrees, at TMD6a helix was found distorted in at least one simulation of all mutants (Figure 5.8). Overall, our computational works suggest that introducing lysine mutations at TMD6a hydrophobic patch is able to interfere with the secondary structure of TMD6 and destabilize the active site geometry. This theoretical work was also verified by our experimental partner by examining the enzyme activity of these PSH mutants *in vitro*. As predicted, the activity of PSH showed a strongly decreased, nearly abolished cleavage of C99 in all four mutants in both DDM micelles and POPC bilayer conditions (Figure 5.7E, F). This result supports that TMD6a is indeed important for substrate cleavage.

## 5.4. Discussion and outlook

Detergent has been long used in the experimental setup of IMPs because of their high availability and reproducibility. However, measuring the structures and functions of IMPs in an environment different from its natural one can be sometimes questionable. Here, we performed MD simulations and *in vitro* experiments to study the modulation effects of the lipid constitution. With a similar C83 and C99 cleavage profile, we demonstrated PSH as a good surrogate for  $\gamma$ -secretase. This allows us to directly study the proteolytic activity of  $\gamma$ -secretase catalytic presenilin subunit in the absence of its complex partner. In close collaboration with the experimental partner, we performed MD simulations to study the dynamics of the C83-PSH complex in different environments. By assuming a similar substrate binding mode between PSH and the PS1-containing  $\gamma$ -secretase, we constructed three homologous C83-bound PSH models with different strategies. The constructed models include the key features observed in the substrate-bound  $\gamma$ -secretase structures, namely the TMD6a helix and the hybrid  $\beta$ -sheet. As observed for substrate-bound PS1, our mutational analysis supports that the hybrid  $\beta$ -sheet also exists in the C83-PSH complex. By stability and conformational evaluation, we identified the model constructed with structural information from both holo-PS1 and apo-PSH as the optimal model, with which we performed further analysis and *in silico* experiments.

MD simulations revealed that in both POPC bilayer and DDM micelles, C83 obtained a hydration profile in PSH similar to what was observed in PS1. However, the catalytic hydrogen bond required for the hydrolysis reaction is less frequently formed. Furthermore, the active site geometry was destabilized in the DDM micelles, and a geometry with a D162<sub>C $\gamma$</sub> -D220<sub>C $\gamma$</sub>  distance of 6.8 Å was sampled. This geometry is, however, considered to be incompatible with catalysis because of the large distance between the unprotonated aspartate residue and the scissile bond. In the cell-free assay, a decrease in PSH processivity was observed in the DDM micelles. In addition, the destabilization effect on the active site geometry was also

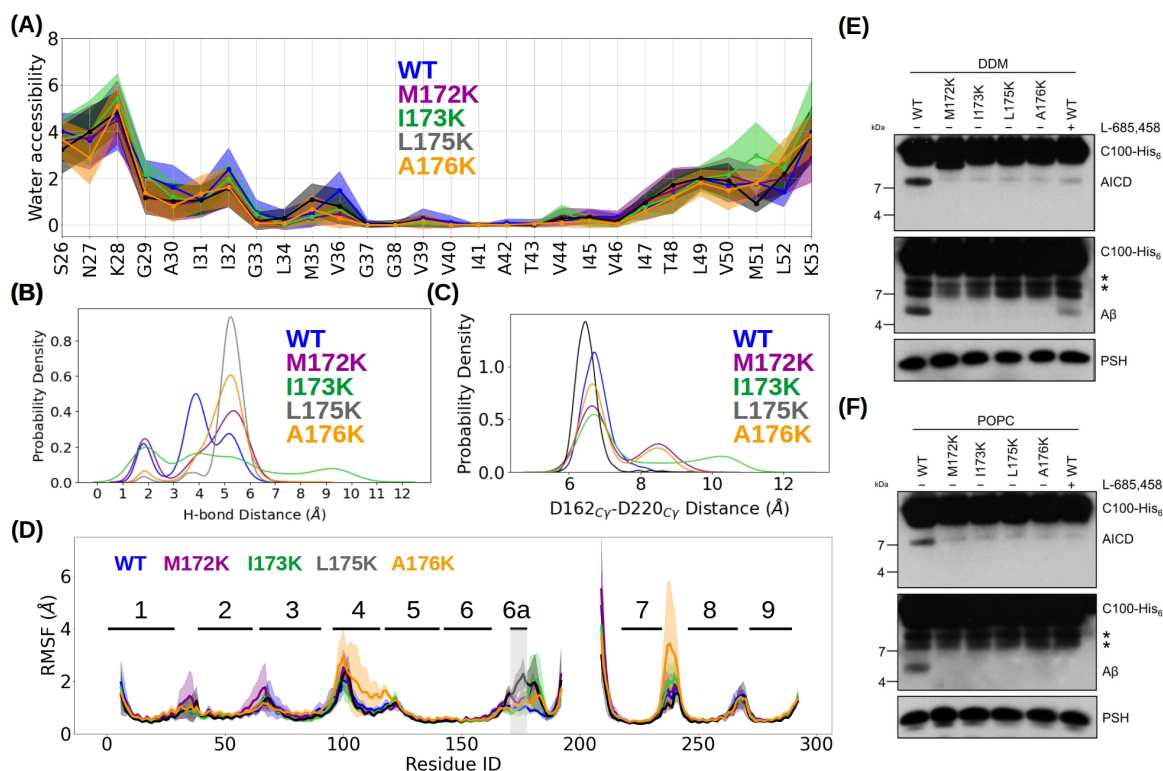


Figure 5.7.: MD simulations and *in vitro* experiments on the PSH mutants M172K, I173K, L175K, and A176K. (A) Residue-wise water accessibility of C83. The modeled scissile bond locates between L49 and V50. (B) Probability density of the distance of the catalytic hydrogen bond sampled in the POPC bilayer. (C) Probability density of the D162<sub>C $\gamma$</sub> -D220<sub>C $\gamma$</sub>  distance sampled in POPC bilayer. (D) Residue-wise RMSFs of WT and PSH mutant in POPC bilayer. The shaded areas in (A) and (D) represent the standard deviation computed across three simulations. (E),(F) Analysis of WT and lysine-mutant PSH activity in (E) DDM and (F) POPC environments. Experimental data in (E) and (F) are not generated by the dissertation of the author. Methods and materials used in the experiments can be found in our published work.

## 5.4. DISCUSSION AND OUTLOOK

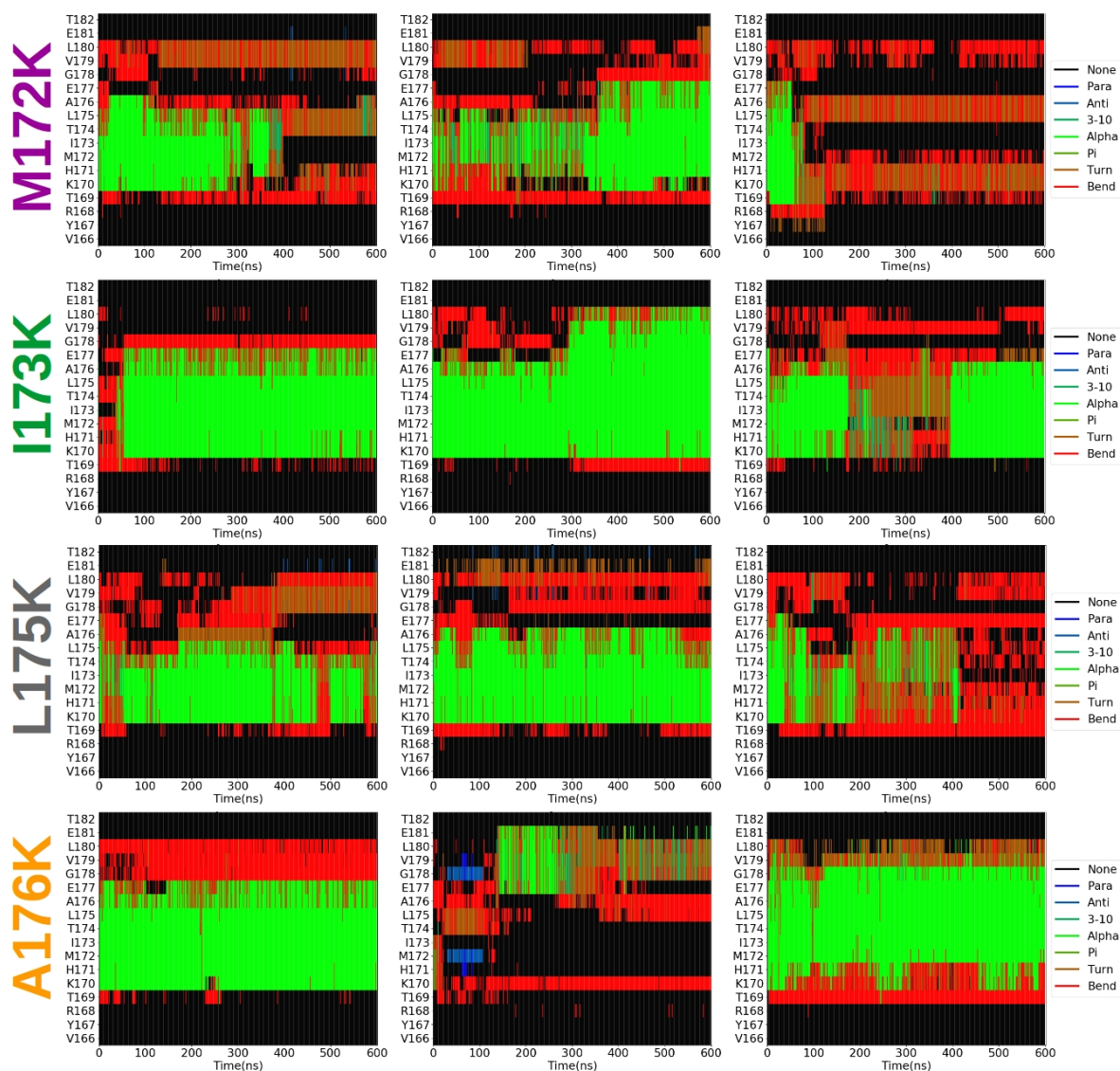


Figure 5.8.: Time-plot of the secondary structure of PSH TMD6a in mutant (top down) M172K, I173K, L175K, and A176K in POPC bilayer. Results obtained from three 0.6  $\mu$ s-long simulations are shown from left to right.

verified by the TSA-inhibitor-based immunoblot analysis.

Our simulation work suggested that the major discrepancy between the POPC and DDM environments lies in the lipid order parameter. While POPC is well-aligned with the principle axis of the enzyme TMDs, DDM molecules can turn a direction perpendicular to the PSH normal and insert into the gap between TMDs. The insertion of DDM molecules between the gap between TMD2 and TMD6 attacked the  $\alpha$ -helical structure of TMD6a through unspecific hydrogen bonding and destabilized TMD6a. As suggested by mutagenesis data and several substrate-bound and GSI-bound structures, the interaction between GSI/substrate and the TMD6a hydrophobic patch might play a crucial role in both GSI binding and substrate processing. Therefore, the observed distortion of TMD6a was suspected to be the culprit of the loss of PSH processivity and active site stability, and aimed to see if lysine mutation can also influence the function of PSH. Indeed, the instability and unwinding of TMD6a and the destabilized active site geometry were successfully reproduced in the simulations when the residues of the TMD6a hydrophobic patch, including M172, I173, L175, and A176, were mutated into lysines. In agreement with the computational study, the PSH activity was significantly impaired when lysine is introduced into the hydrophobic patch.

In fact, the active site geometry with two distant catalytic residues, despite never being observed in the  $\gamma$ -secretase simulation inside the POPC membrane, was also seen in the available structures of GxGD-type proteases such as the  $\gamma$ -secretase (10.6 Å[53]) or Flak (12 Å[181]), presumably due to the non-native environment during the sample preparation procedure. Upon substrate association, the two catalytic aspartates decrease their distance and approach closer to the initial cleavage sites[39, 40]. Our study showed a direct relationship between the active site geometry, TMD6a stability, and how they are affected by the change in environment. As a general implication for intramembrane proteolysis, our data suggest that a lipid bilayer-mediated stabilization of the active-site geometry might also be observable for other intramembrane proteases of different catalytic types.

Taken together, in a good correlation between experimental and simulation data, our results with PSH as a model intramembrane protease highlight an important role of the membrane lipid environment in providing a stabilized E-S conformation that is crucial for substrate processing in intramembrane proteolysis. Our data further underscore the key role of the conformational flexibility of presenilin/PSH TMD6a for substrate interactions and proteolytic cleavage of presenilin-type proteases. Most importantly, they provide evidence that the lipid bilayer promotes the formation of a conformationally stable active site geometry, which is of general importance for an efficient catalytic operation of intramembrane proteases.



## 6. Enzyme-Substrate Hybrid $\beta$ -Sheet Controls Geometry and Water Access to the $\gamma$ -Secretase Active Site

A characteristic[182, 183] of  $\gamma$ -secretase cleavages toward C99 is the 3-residue-wise processive cleavage. That is, C99 and its C-terminal products are most of the time, except for  $\beta$ 42, cleaved sequentially with an interval of three residues. With the aim to decipher the underlying mechanism of the tripeptide trimming, we performed restraint-free and biasing MD simulations to study how the length of the substrate is correlated with its cleavability. Our data suggested that the length-sensitive stability of the substrate  $\beta$ 3-strand might play an important role in controlling the water access and the stability of the  $\gamma$ -secretase catalytic center and is responsible for the tripeptide cleavage. The hypothesis was further supported by the impaired *in vitro*  $\gamma$ -secretase activity when proline is introduced to break down the  $\beta$ 3-strand stability. To our knowledge, our work provides the first molecular insight into the proteolytic role of the E-S hybrid  $\beta$ -sheet and the molecular mechanism of  $\gamma$ -secretase tripeptide cleavage.

### 6.1. Introduction

Proteolysis of proteins is an essential biological process that breaks proteins into smaller polypeptides or amino acids[184]. Typically, the proteases degrade their protein substrate through hydrolysis with the nucleophilic attack of a catalytic water molecule. Intramembrane proteases are a particular type of protease that targets the TMD of IMPs and according to their catalytic mechanism, they can be classified into four groups: metalloprotease, an aspartyl protease, a serine protease, and glutamyl protease[30]. As the catalytic subunit of  $\gamma$ -secretase, PS is an aspartyl protease that cleaves its substrate through two catalytic aspartic acids on TMD6 and TMD7 inside the membrane environment[185, 29] and releases the C-terminal product of AICD and the N-terminal product A $\beta$  peptides[186]. Interestingly, the cleavage of PS is executed in a processive fashion[9, 1]. That is, the substrate is cleaved several times before it is released from the enzyme. According to the tripeptide and tetrapeptide released from the TMD of the substrate,  $\gamma$ -secretase is suggested to cleave C99 in a step of 3 or 4 (from A $\beta$ 42 to A $\beta$ 38) residues[67]. Despite this 3- or 4-residue-wise cleavage observed for more than a decade ago, no convincing mechanism has been proposed to explain the phenomenon. Although a computational work using gaussian-accelerated MD simulations was published during the period of this work, stating their simulation works enabled the elucidation of the tripeptide trimming mechanism, no specific molecular detail was provided to explain

the three residue difference between the  $\epsilon 49$  and  $\zeta 46$  cleavage[157]. In addition, the model that tilting of the helical domain and unwinding of the C-terminal helix of the substrate without the movement of the helical domain is difficult to be generalized to the subsequent  $\gamma$ -cleavages. Therefore, a more mechanistic picture of  $\gamma$ -secretase tripeptide cleavage that can be generalized for the subsequent processive cleavages is still in demand.

Despite a great insight on how  $\gamma$ -secretase targets the substrate at the  $\epsilon 49$  site was provided through the lately resolved C83-bound structure[39], how C99 or C83 moves toward its next cleavage binding pose is left unknown. In light of the high structural similarity between the C83-bound[39] and the Notch1-bound[40] structures, it is reasonable to assume that other substrates of  $\gamma$ -secretase might also bind to the enzyme in a similar binding pose, including A $\beta$  peptides. Under this assumption, we proposed a general processing model for the processive cleavage process, as exemplified by the A $\beta 49 \rightarrow$  A $\beta 46$  process (Figure 6.1). According to the results obtained from Chapter 4, the helical domain is embedded in a dry environment while the loop and  $\beta 3$ -strand are solvated by water molecules (Figure 4.5A). Upon the  $\epsilon 49$  cleavage on C99, the C-terminal soluble product AICD is released to the intracellular side and the N-terminal product A $\beta 49$ , carrying the negatively charged carboxylic acid group at the C-terminus, remains in the membrane bilayer for the next cleavage. If the position of the catalytic residues D257 and D385 are restricted within a certain space, the helical of the substrate domain needs to unwind at the C-terminal part to expose the next scissile bond to the catalytic center by either translational, bending, or tilting movement. Inevitably, the other amine bond such as the ones between T48-L49 and between I47-T48 would also be exposed to the cleavage site along the movement, causing unspecific  $\zeta 48$  and  $\zeta 47$  cleavage, respectively. However, only the  $\zeta 46$  cleavage between V46 and I47 was observed following the  $\epsilon 49$  cleavage in experiments. This implies a length-sensitive criterion must be fulfilled to control the accurate 3-residue cleaving profile. Therefore, we investigated how the length of the substrate influences the proteolysis of  $\gamma$ -secretase without the knowledge of the exact binding mode of the shorter A $\beta$  peptides by truncating APP into shorter peptides.

In this chapter, we performed restraint-free and biasing MD simulations to investigate how the catalytic geometry responds to the substrate in different lengths. In total, nine different substrates (WT and 8 mutants, sequences shown in Table 6.1) derived from C99 were chosen in this study. In the restraint-free simulation, a dramatic decrease in the formation of a catalytically active geometry, termed "active geometry" in the following context, was found when only one or two residues are present C-terminal to the scissile bond. Using the statistics and structural analysis, we suspected that the stability of the active geometry is closely related to the formation of substrate  $\beta 3$ -strand and the water accessibility to the catalytic center. From the dissociated pathway of  $\beta 3$ -strand observed in the A $\beta 52$  and A $\beta 51$  simulations, we choose the reaction coordinate for biasing MD using replica-exchange Hamiltonian sampling coupled umbrella sampling (HREUS) method to sample the structural response toward the association and dissociation of the hybrid  $\beta$ -sheet near the catalytic site. The biasing MD showed a clear correlation between the formation of the active site geometry, catalytic site water accessibility, and the formation of the substrate  $\beta 3$ -strand. With the wild-type sequence,

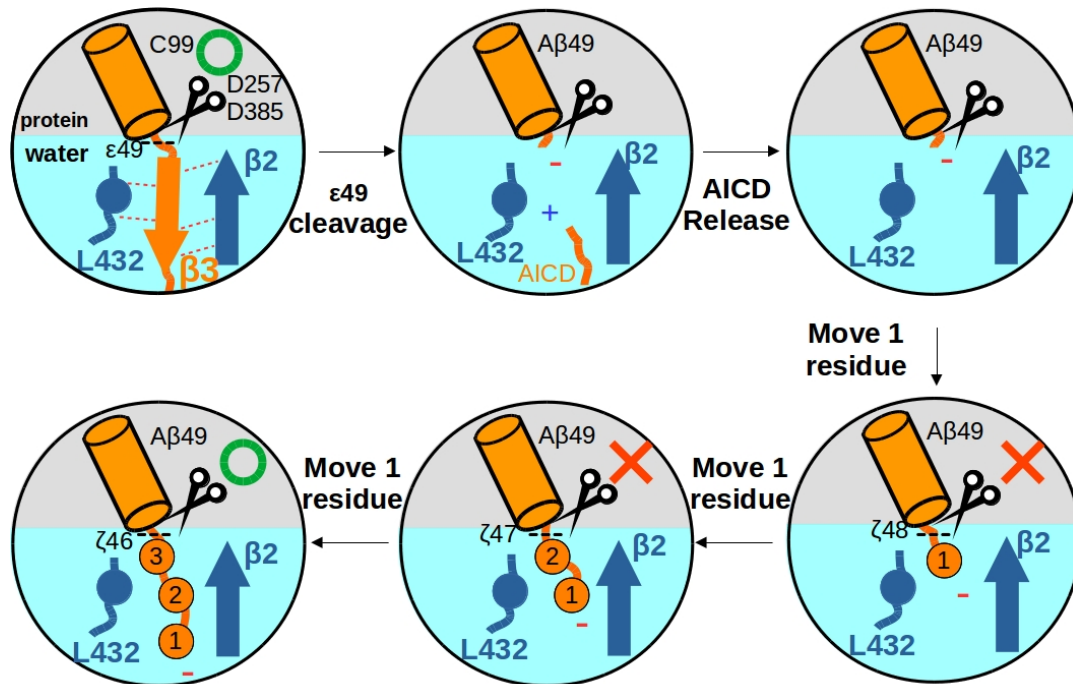


Figure 6.1.: Schematic view of a general C99 processive cleavage exemplified by the A $\beta$ 49  $\rightarrow$  A $\beta$ 46 process. The position of the protein-water interface is suggested by our previous result (Figure 4.5A). The substrate is colored in orange and  $\beta$ 2 and L432 of PS1 are colored blue. The catalytic aspartic acids are represented by the scissors symbol with the putative scissile bond shown in the dashed line. A green circle and red cross indicate whether the putative scissile bond can be cleaved. A red minus sign and a blue plus sign indicate the negatively charged A $\beta$  C-terminus and the positively charged AICD N-terminus, respectively.

the association binding free energy of the  $\beta 3$ -strand showed an anticorrelation with the length of the substrate. Importantly, a substrate with only two residues beyond the scissile bond is incapable of forming an active geometry even if biasing potential was implemented to force the association. A similar trend was found when proline (Pro) or phenylalanine (Phe) was introduced at the substrate P2' position. In good agreement with the *in vitro* experiment, we confirmed that distortion on the E-S hybrid  $\beta$ -sheet is indispensable for the cleavage of  $\gamma$ -secretase. Moreover, we found that the D385-protonated state is more capable of stabilizing the hybrid  $\beta$ -sheet while the D257-protonated state is more capable of recruiting water molecules towards the conserved GxGD motif near the catalytic center.

Table 6.1.: Sequence of the substrates present in this chapter. The \* start indicates the scissile bond ( $\epsilon$  cleavage site) modeled in the present study.

Substrate	Sequence
APP	...NKGAIIGLMVGGVVIATVIVITL*VMLKK...
APP <sub>M51P</sub>	...NKGAIIGLMVGGVVIATVIVITL*VPLKK...
APP <sub>L52P</sub>	...NKGAIIGLMVGGVVIATVIVITL*VMPLKK...
A $\beta$ 53	...NKGAIIGLMVGGVVIATVIVITL*VMLK-COO <sup>-</sup>
A $\beta$ 52	...NKGAIIGLMVGGVVIATVIVITL*VML-COO <sup>-</sup>
A $\beta$ 52 <sub>V50F</sub>	...NKGAIIGLMVGGVVIATVIVITL*VFL-COO <sup>-</sup>
A $\beta$ 52 <sub>M1F</sub>	...NKGAIIGLMVGGVVIATVIVITL*FML-COO <sup>-</sup>
A $\beta$ 51	...NKGAIIGLMVGGVVIATVIVITL*VM-COO <sup>-</sup>
A $\beta$ 50	...NKGAIIGLMVGGVVIATVIVITL*V-COO <sup>-</sup>

## 6.2. Methods and Materials

All substrate-bound  $\gamma$ -secretase structures in this chapter are constructed based on the C83-bound cryo-EM structure (PDBID 6IYC)[39]. First, the cryo-EM structure was pre-processed as described in Chapter 3 section 3.3 and embedded with 503 POPC membrane and 0.15M potassium chloride through the CHARMM-GUI [187, 179] server to generate the WT system. To get rid of the bias of the placement of lipid and water environments, all mutated substrates are constructed directly by substrate truncation and mutation on the WT systems. With both PS1-D385<sup>H</sup> and PS1-D257<sup>H</sup> being investigated, in total 18 systems are constructed for simulations and listed in Table 6.1. Energy minimization, equilibration, and production run were performed as described in Chapter 3 section 3.3. Two unrestrained 600 ns trajectories are generated for the APP, A $\beta$ 53, A $\beta$ 51, and A $\beta$ 50 systems in both protonation states and four unrestrained 600 ns trajectories are generated for the A $\beta$ 52 system because of its higher diversity in statistics. Atomic interactions were described by the amber14SB force field for the proteins, the tip3p for the water, and lipid17 for the POPC molecules.

To sample the  $\beta 3$  association pathway, HREUS was implemented. Furthermore, to sample a reasonable association/dissociation pathway, the chosen reaction coordinate (RC) should sample the similar configuration states as what was sampled in the restraint-free simulations.

With the knowledge of the dissociated  $\beta 3$  state sampled in the restraint-free simulations, the  $C\alpha$ - $C\alpha$  distance of PS1 L418 and substrate P2' are chosen as the reaction coordinate to capture the dissociation and re-association path. With the low RMSF observed in L418 in restraint-free APP-bound simulations (Supplementary Figure 5.5A), it is reasonable to introduce positional restraint of  $10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$  on the L418  $C\alpha$  atom to make sure that the measured RC relies totally on the movement of substrate P2'. In addition, a positional restraint of  $10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$  was applied on PS1 K380 to avoid the  $\beta 2$  distortion observed in the dissociated simulations. Following the energy minimization and equilibration protocol used in the restraint-free simulations, each system was first simulated for 10 ns with harmonic distance restraint forcing the formation of the catalytic hydrogen bond, namely  $d_1$ , at  $2.2 \text{ \AA}$ , and the  $C\alpha$ - $C\alpha$  distance between PS1 K380 and L432 was kept at  $12 \text{ \AA}$  to avoid  $\beta 3$  dissociation at a constant temperature of  $303.15 \text{ K}$  and the constant pressure of  $1 \text{ bar}$ . The structure is then sequentially submitted to regular US protocol ranging from  $RC = 8.0 \text{ \AA}$  to  $RC = 17.0 \text{ \AA}$  with an interval of  $0.6 \text{ \AA}$  to prepare the starting structure for the following HREUS sampling. The structure is sampled for 3 ns in each window, in total 16 windows are prepared with 48 ns NVT simulations. A force constant  $k = 6 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$  was applied on  $RC = 8.0, 8.6, 9.2, 9.8, 16.4,$  and  $17.0 \text{ \AA}$  and a force constant  $k = 8 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$  was applied on  $RC = 10.4, 11.0, 11.6, 14.6, 15.2, 15.8 \text{ \AA}$ . A stronger force constant  $k = 10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-2}$  was applied on  $RC = 12.2, 12.8, 13.4, 14.0 \text{ \AA}$  to sample the processes around the higher free energy barrier. The generated structures in different RC windows were together submitted to the HREUS protocol using `pmemd.cuda.mpi` of the Amber18 program [188]. Exchanges between the neighboring replica were attempted every 10ps, and in total, 60ns of NVT sampling was carried out. The first 20 ns of the HREUS trajectory was considered as the equilibrium process between the windows and only the later 40 ns were taken for the following analysis. The potential of mean force (PMF) profiles was calculated with the weighted histogram analysis method (WHAM[134]) [134] with error bars showing the standard deviation across the PMF curves calculated using data collected from 20ns-50ns, 20ns-55ns, and 20ns-60ns. Features such as water number,  $\beta 3$  fraction, active geometry fraction, and K380-L432 distance are binned into 9 intervals equally spread between  $RC = 8.0 \text{ \AA}$  and  $RC = 17.0 \text{ \AA}$  with error bars showing the standard errors. The sampled RC and convergence of PMF are shown in FigureB.1, FigureB.2, FigureB.3, and FigureB.4.  $\Delta\Delta G_{res}$  is taken from the PMF value at  $RC = 16 \text{ \AA}$  to describe the free energy difference between the associated and dissociated state if the global PMF minimal falls in the regime with  $RC < 11 \text{ \AA}$  and zero otherwise.

To better illustrate the quantities related to the geometry at the active site, the fraction of active site geometry and  $\beta 3$  formation, and the number of water molecules around the catalytic site are measured in both restraint-free and HREUS simulations. Geometry is only considered as a catalytically active one when a catalytic hydrogen bond is formed between the scissile carbonyl and the protonated aspartic acid (depicted as  $d_1 < 2.5 \text{ \AA}$  in Figure 6.2) and the distance between the unprotonated aspartic acid is not too far away from the scissile carbonyl so that the proton transfer is still achievable (depicted as  $d_2 < 5.6 \text{ \AA}$  in Figure 6.2). During the simulations of the APP-bound complex, P1'-P4', namely V50-K53, occasionally turned into an anti-parallel  $\beta$ -strand conformation (Figure 6.3). Compared to substrate P1'

and P4', P2' and P3' formed more stable hydrogen bonds with PS1 and stayed more steadily in the  $\beta$ -strand conformation in the APP-bound complex (Figure 6.3 and Figure B.5). The completeness of  $\beta 3$  is thus computed by taking the average  $\beta$ -strand occupation fraction of P2' and P3' calculated by the DSSP method, namely  $\beta 3 = [\beta(P2') + \beta(P3')]/2$ . Since A $\beta$ 51 does not have P3' residue, only the  $\beta$ -strand occupation fraction of P3' is taken into account, namely  $\beta 3 = \beta(M51)$ . A water molecule is considered being around the catalytic center if any atom of that water is within 5 Å of any atom from D257 or D385. The C $\alpha$ -C $\alpha$  distance between L432 and K380 is measured to indicate the width of the gap between L432 and  $\beta 2$ . The fraction of an active geometry and  $\beta 3$ -strand formation are calculated by averaging the quantities measured every 1 ns in each 600ns restraint-free simulation and every 50 ps in each 40ns HREUS simulation.

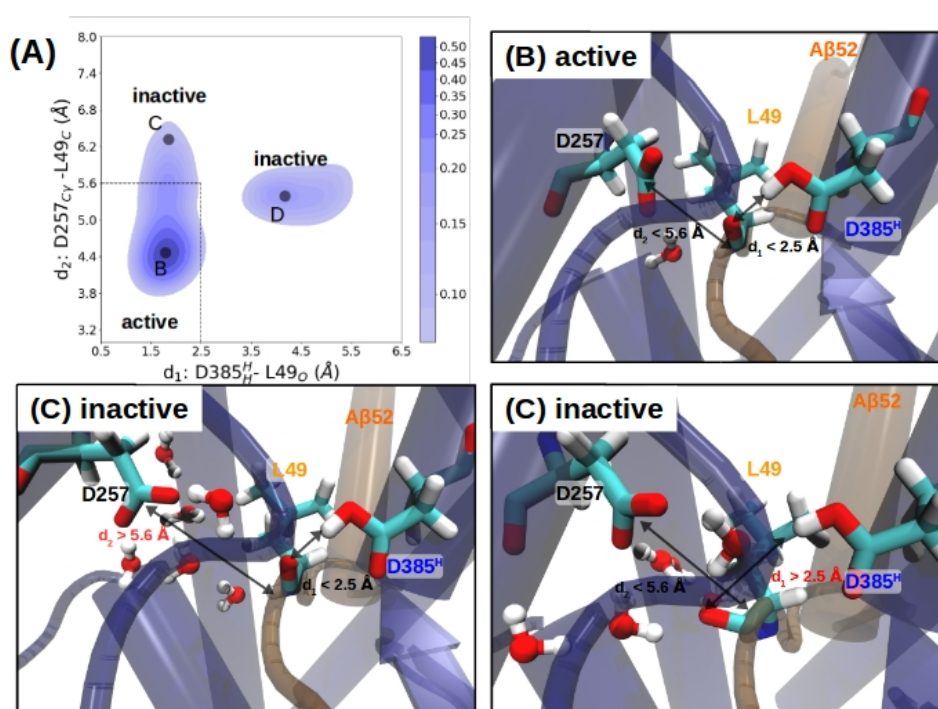


Figure 6.2.: (A) 2D distribution contour map with the distance of the catalytic hydrogen bond ( $d_2$ ) as the x-axis and vs. the distance between the substrate carbonyl and the C $\gamma$  of the deprotonated aspartate ( $d_1$ ) in a 600ns trajectory. Black dots indicate the examples of one active geometry and two inactive geometries. The 2 dashed lines represent the 2 distance-based criteria  $d_1 = 2.5$  Å and  $d_2 = 5.6$  Å and separate the active conformations (lower left) from the inactive states (rest). The color scale represents the distribution density. (B) Conformation of an active geometry fulfilling both distance-based criteria. (C) Conformation of an inactive geometry with a large  $d_2$  ( $> 5.6$  Å) filled with more than one water molecule. (D) Conformation of an inactive geometry without the catalytic hydrogen bond ( $d_1 < 2.5$  Å).

## 6.3. Result

### 6.3.1. Stability of the active geometry is sensitive to substrate length

Processive cleavage of  $\gamma$ -secretase toward C99 APP C-terminal fragment typically follows the A $\beta$ 49-A $\beta$ 46-A $\beta$ 43-A $\beta$ 40 or A $\beta$ 48-A $\beta$ 45-A $\beta$ 42-A $\beta$ 38 production lines, starting from the  $\epsilon$ 49 and  $\epsilon$ 48 cleavage sites [9]. Although the successive tripeptide cleavages have been characterized for more than a decade ago[67, 66], the molecular mechanism remains unclear. To investigate how the number of residues following the substrate scissile bond influences the active site dynamics and the substrate flexibility and cleavability, we generated several substrate-bound  $\gamma$ -secretase models *in silico* with the substrate being either APP or C-terminally shortened variants. The resulting shortened substrates are termed (in A $\beta$  numbering) A $\beta$ 53, A $\beta$ 52, A $\beta$ 51, or A $\beta$ 50, by truncating APP after K53, L52, M51, and V50, respectively (Table 6.1). Structural data from X-ray and neutron diffraction show that one of the Asp residues in aspartate proteases is protonated and the second one is unprotonated[139, 140, 141, 142]. However, no consensus has been reached on which Asp of PS1 is more likely to be protonated. Computational work with either D257 protonated[156, 157, 153], D385 protonated[78], or both unprotonated in PS1[189] has so far been conducted. Nevertheless, since this issue is not completely resolved, for the present study we always considered both possible PS1 protonation states, termed PS1-D257<sup>H</sup> and PS1-D385<sup>H</sup>, which may also give important insights into the influence of the selected protonation state. In total five models with wild-type (WT) substrate sequences were constructed and the dynamics were studied by multiple MD simulations in each case. Snapshots of the hybrid  $\beta$ -sheet cluster and the active site of each E-S complex after 600 ns simulation time are shown in Figure 6.4A-E.

In good agreement with our expectation, visual inspection at the catalytic center in each simulation already indicates that substrates with less than two residues C-terminal to the scissile bond, namely A $\beta$ 51 and A $\beta$ 50, cannot form a stable active geometry with the catalytic aspartic acids (Figure 6.4A-E). From the structural information, we observed that the position where the scissile bond was located in the APP, A $\beta$ 53, and A $\beta$ 52 was replaced by a water molecule. Note, the active site geometries observed in the A $\beta$ 51-bound and A $\beta$ 50-bound complexes correspond perfectly with the open state we have observed in Chapter 4.5C. That is, the catalytic hydrogen bond was out-competed by the water molecules coming from the C-terminal side. Indeed, excessive water molecules were found in the vicinity of the catalytic center when the substrate is truncated too short (shorter than A $\beta$ 52). From the structural information shown in Figure 6.4A-E, there are two factors leading to the high water accessibility at the catalytic site. First, in the longer substrates, the formation of hybrid  $\beta$ -sheet is efficiently blocking water molecules from access to the active site by the hydrogen bonds between substrate  $\beta$ 3 and L432 and K380 of PS1 (Figure 6.4A-C). Second, when the substrate is trimmed shorter, the negatively charged carboxylic acid is also brought closer to D257 and D385 of PS1, increasing the overall polarity of the catalytic center. In fact, the water molecules recruited from the carboxylic side of the substrate can also out-compete the hydrogen bonds at the hybrid  $\beta$ -sheet and cause the dissociation of the substrate  $\beta$ 3-strand.

### 6.3. RESULT

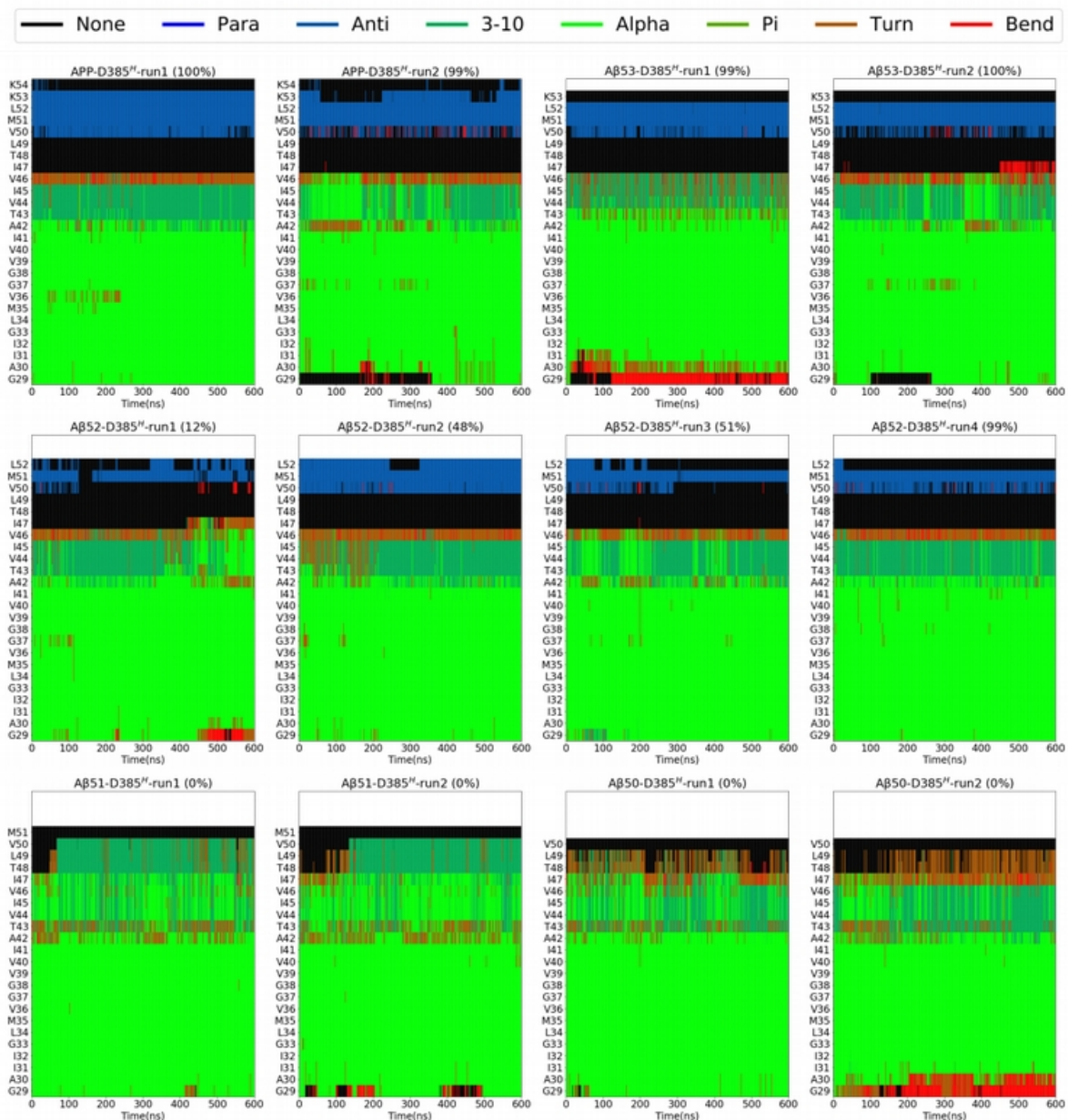


Figure 6.3.: Secondary structures over time calculated by DSSP method in APP and its truncated variants binding to D385<sup>H</sup>  $\gamma$ -secretase. Each 600ns-long simulation is shown with the sampled fraction of active geometry written in the parenthesis.



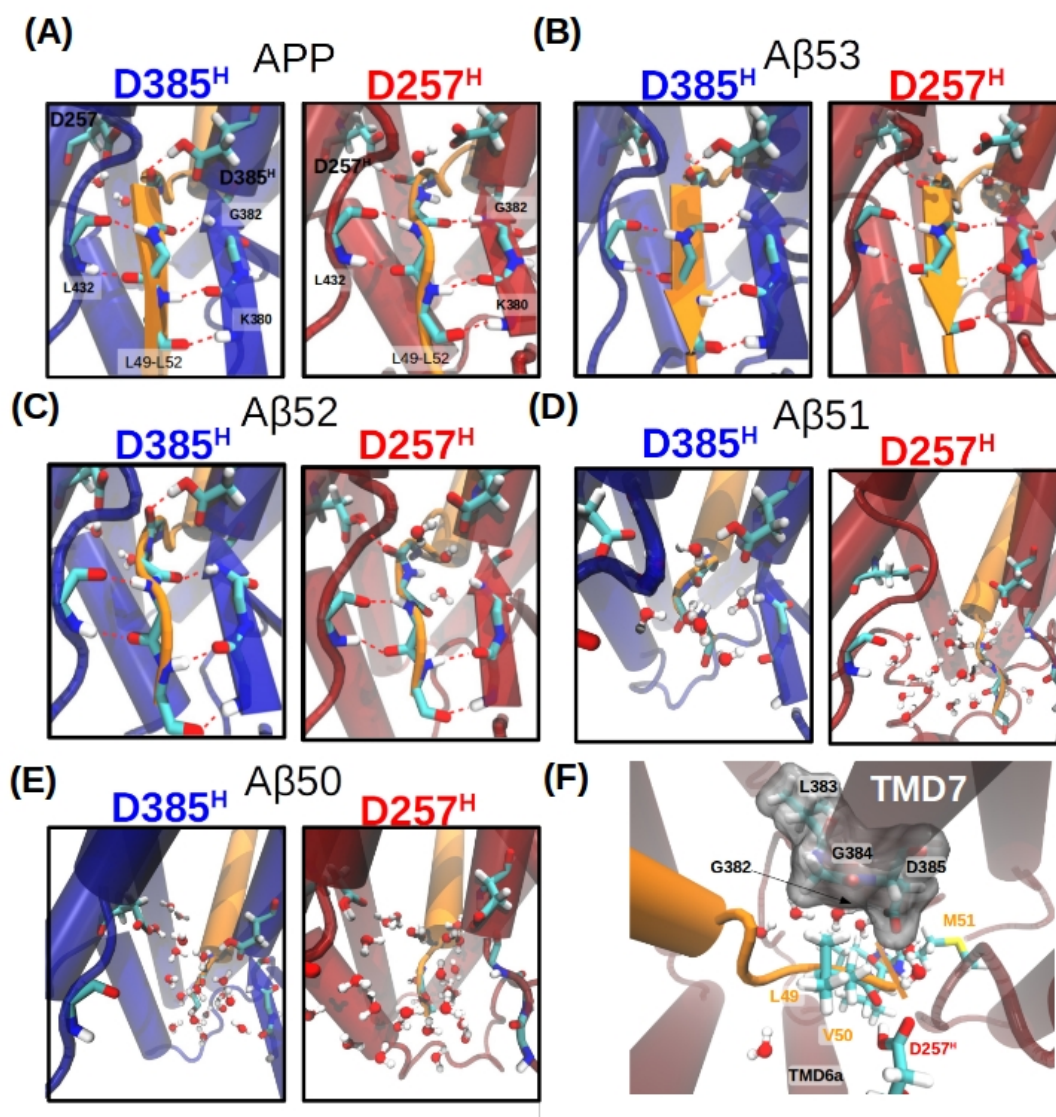


Figure 6.4.: Snapshots of hybrid  $\beta$ -strand of each E-S complex after 600ns simulation. (A)-(E) Detailed structure at the post-cleavage site region of  $\gamma$ -secretase in complex with (A) APP and (B)-(E) its truncated variants. PS1 D257, D385, and backbone atoms involved in the hydrogen network are shown in licorice representation with substrate-PS1 hydrogen bonds (red dashed lines) including L49-D257/D385, V50-G382, M51-L432, and L52-K380. Water within 5Å of L49-L52 of the substrate is shown in stick-ball representation. (F) Snapshot of the water-dwelling cavity formed by the GxGD motif (white van der Waals surface) in D257-protonated PS1. Waters within 5Å of D257 or D385 are shown in ball-stick representation. In each enzyme-substrate complex, D385<sup>H</sup>-PS1 is shown in blue in the left panel, D257<sup>H</sup>-PS1 is shown in red in the right panel, and substrates are shown in orange.

In addition to the effects related to the substrate length, we also observed that more water molecules were attracted to the unprotonated aspartic acid. However, in the PS1-D385<sup>H</sup> state, D257 is surrounded by hydrophobic residues such the TMD6a hydrophobic patch can only a few water molecules can approach the catalytic center. In contrast, the GxGD motif at the N-terminus of TMD7 compose of a small cavity for water to accommodate (Figure 6.4F).

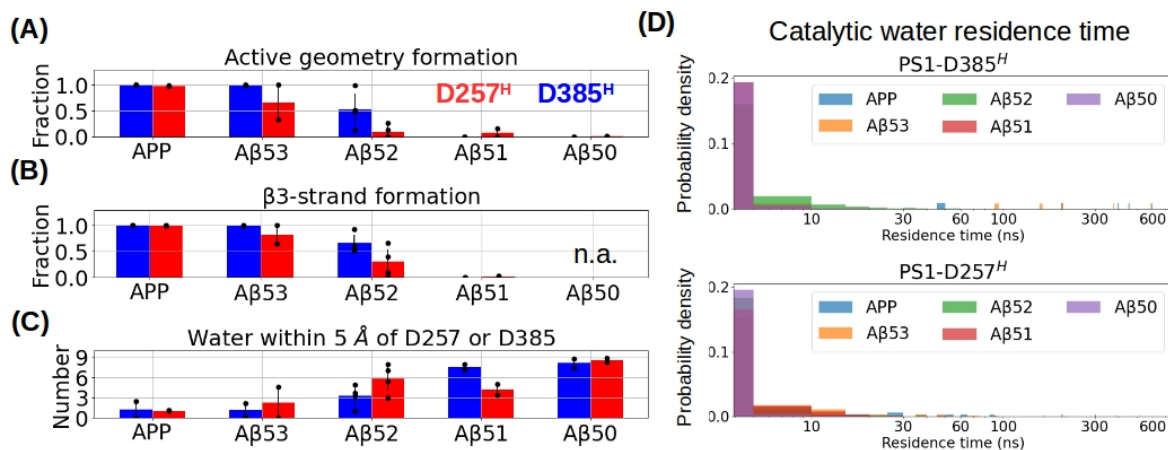


Figure 6.5.: Statistics obtained from the restraint-free simulations. (A) Fraction of sampled states that form an active site geometry compatible with cleavage. (B) Fraction of sampled states that form a hydrogen-bonded  $\beta$ 3-strand. Formation of  $\beta$ 3-strand is not applicable in the A $\beta$ 51 substrate. (See Methods and Materials) (C) Number of water molecules around the catalytic center in five different  $\gamma$ -secretase holo complexes (separate sets of simulations were performed for active site protonation states, color-coded red (D257<sup>H</sup>) or blue (D385<sup>H</sup>)). (D) Distribution of water residence time around the catalytic center when substrate binds to (top) PS1-D385<sup>H</sup> or (bottom) PS1-D257<sup>H</sup>. Histogram is binned with a width of 5 Å. The fraction of  $\beta$ 3-strand is not applicable for A $\beta$ 50 (see Methods and Materials).

To more quantitatively describe the substrate length influence on active geometry,  $\beta$ 3-strand stability, and catalytic site water accessibility, we performed statistics across multiple simulations in both PS1 protonation states. As shown in Figure 6.5A, we observed that the fraction of the active geometry formed in the simulations decreases for the C-terminally truncated substrate variants. Furthermore, simulations with PS1-D385<sup>H</sup> are statistically more prone to form an active geometry compared to those with D257<sup>H</sup>. In the case of a WT substrate, the active geometry is found in over 99% of the sampled conformations in both protonation states. While the fraction decreases to 53% for the bound A $\beta$ 52 case with PS1-D385<sup>H</sup> and drops to 10% in the case of PS1-D257<sup>H</sup>. For an even shorter substrate, the active geometry at the catalytic center can barely be found. This result indicates a strong coupling between the number of residues C-terminal of the cleavage site and the stability of the active site geometry ready for cleavage. Similar to the active geometry formation fraction, the fraction of the  $\beta$ 3-strand decreases as the substrate gets shorter and is no longer stable

in A $\beta$ 51 and A $\beta$ 50. (Figure 6.5B). Interestingly, in both formations of active geometry and substrate  $\beta$ 3-strand, we can clearly observe the initiation starting only when more than three residues are present C-terminal to the cleavage site. To see whether the instability of these two features is caused by the attack of the incoming water molecules, we also counted the number of water molecules found within 5Å of either D257 or D385. As shown in Figure 6.5C, on average around one water molecule can be reached the catalytic site in the APP-bound and A $\beta$ 53-bound complexes. In contrast, more than three water molecules are found around at the catalytic center in the A $\beta$ 52-, A $\beta$ 51- and A $\beta$ 50- bound complexes. Importantly, we observed that the water molecules around the catalytic site are trapped for a significantly longer period for APP and A $\beta$ 53 substrates, whereas shorter substrates like A $\beta$ 51 and A $\beta$ 50 usually result in a water residence times shorter than 10 ns (Figure 6.5D). Since the hydrolysis reaction of  $\gamma$ -secretase requires electron transfer between the enzyme, substrate, and the catalytic water molecule (Figure 3.4), a sufficiently long enough residence time is considered also necessary for the catalysis. Taken together, our structural data and simulation statistics not only reproduced the length-sensitive stability of the active geometry but also correlate with the formation of the substrate  $\beta$ 3-strand and the catalytic site water accessibility.

### 6.3.2. Sampling the $\beta$ 3-strand association/dissociation by Hamiltonian replica exchange simulations

Our multiple MD simulations with different substrates and protonation states indicate a strong coupling between the formation of an active geometry, the substrate  $\beta$ 3-strand formation, and the water distribution around the active site. However, how these features depend on each other is not totally clear due to the limited phase space sampled in the restraint-free simulations. For instance, the dissociated state and the associated state of the substrate  $\beta$ 3-strand are rarely sampled in the simulations of APP-bound and A $\beta$ 51-bound, respectively. To investigate how active site water distribution and formation of an active geometry respond to the association and dissociation  $\beta$ 3, umbrella sampling (US) coupled with Hamiltonian replica exchange between the US windows (HREUS) was applied to the APP-, A $\beta$ 53-, A $\beta$ 52- and A $\beta$ 51-bound complexes. It allows us to sample the intermediates along the  $\beta$ 3-strand association and dissociation processes and to extract associated free energy changes. Among several potential reaction coordinates (RC), the C $\alpha$ -C $\alpha$  distance between PS1 L418 on TMD8 and substrate M51 was found to correlate well with the  $\beta$ 3-strand dissociation event (Figure 6.6A) and provides a straightforward pathway of the  $\beta$ 3-strand dissociation without severely distorting the rest of the substrate (Figure 6.6B, C). In the restraint-free simulations, this selected RC distance is below 10 Å when the  $\beta$ 3-strand is formed and hydrogen bonds to the hybrid  $\beta$ -sheet with the PS1  $\beta$ -strands. An RC distance above 13 Å indicates  $\beta$ 3-strand dissociation from the PS1  $\beta$ -strands (Figure 6.6A). To cover the entire  $\beta$ 3-strand association and dissociation pathways, the chosen RC is sampled between 8 Å and 17 Å (see Methods and Materials). However, the enforced dissociation/association along the RC without additional restraints resulted in a distortion of the enzyme  $\beta$ 2-strand (in test simulations). To avoid such distortion, positional restraints were required on K380 of PS1, keeping PS1 in a ready-to-bind conformation, and on L418 in a stable geometry so that only the substrate strand is moving

while the RC is increased. It is important to note that such restraint is expected to artificially stabilize the  $\beta$ 3-strand associated state but the bias is present in all cases and hence should still allow us to obtain qualitative insight into the effect of shortening the C-terminus of the substrate. From the sampled configurations, the potential of mean force (PMF) profiles of  $\gamma$ -secretase complexes are calculated along the predefined RC using the weight histogram analysis method (WHAM). The free energy difference required to bring the  $\beta$ 3 away from the hybrid  $\beta$ -sheet region derived from the calculated PMF curve in the restrained system is denoted as  $\Delta\Delta G_{res}$ , as depicted in (Figure 6.7A). In addition, we also analyzed the fraction of sampled active geometry,  $\beta$ 3-strand formation rate, and the amount of water near the catalytic center denoted as water accessibility along the RC. For simplicity and clarity, we focus in the following on the binding processes with the PS1-D385<sup>H</sup> complex, and results for the PS1-D257<sup>H</sup> complex are illustrated in Chapter B (unless they deviate from the results for the PS1-D385<sup>H</sup> complexes).

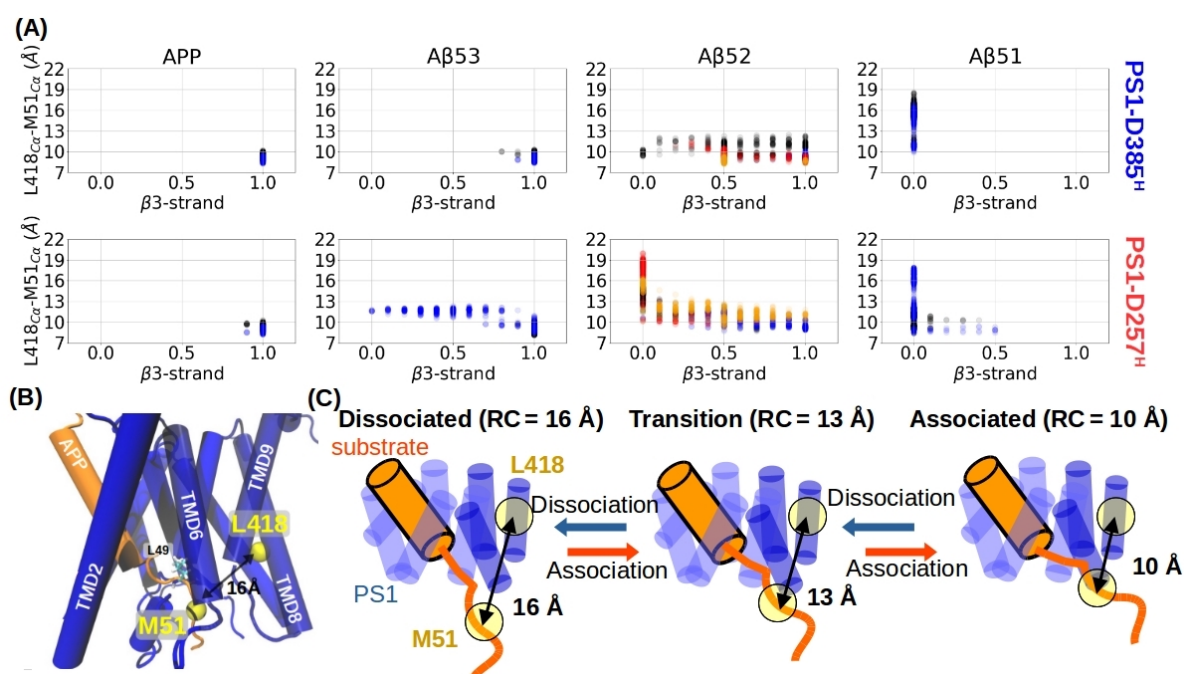


Figure 6.6.: Selection of reaction coordinate (RC) for HREUS setup inspired by the dissociation even sampled from the restraint-free MD simulations. (A) Correlation between the  $\beta$ 3-strand formation and the C $\alpha$ -C $\alpha$  distance between L418 of PS1 and substrate M51 in each MD run. Data points are averaged for every 5ns. (B) Structure of the APP-bound  $\gamma$ -secretase in a dissociated state. L418 of PS1 and substrate M51 are highlighted with yellow van der Waals beads. (C) Cartoon schematics of the HREUS sampling strategy along the reaction coordinate (RC) implemented in the study.

Comparison of the WT and truncated substrates shows a clear trend of decreasing free energy differences between  $\beta$ 3-associated (RC = 9-11 Å) and dissociated states (RC = 15-16 Å)

in the order  $\Delta\Delta G_{res,APP} > \Delta\Delta G_{res,A\beta53} > \Delta\Delta G_{res,A\beta52} > \Delta\Delta G_{res,A\beta51}$ , regardless of the chosen PS1 protonation state (Figure 6.7 and Figure B.6A). According to the change in the fraction of sampled active geometry and  $\beta3$ -strand formation rate, and the amount of water near the catalytic center, the conformations sampled along the RC can be roughly split into three regimes. These are the dissociated regime (regime I,  $RC > 16\text{\AA}$ ), transition regime (regime II,  $12\text{\AA} \leq RC \leq 16\text{\AA}$ ), and associated regime (regime III,  $RC < 12\text{\AA}$ ) (Figure 6.7B, representative snapshots of the E-S complex in each regime are shown in Figure 6.8). In the dissociated regime (regime I),  $\beta3$ -association and the active geometry are barely formed and around five water molecules can be found at the active site periphery in all complexes (Figure 6.7B). As the substrate is gradually approaching to form the hybrid  $\beta$ -sheet cluster (in regime II),  $\beta3$ -association and the active geometry begin to form, and neighboring water molecules are forced out of the active site region. In the associated regime (regime III), both the active geometry and  $\beta3$ -strand are frequently formed with only a few water molecules present around the catalytic site, except for the A $\beta51$  complex. A similar binding process is also observed in the case of the PS1-D257<sup>H</sup> complex, however, in this case, a few more water molecules are recruited into the catalytic center and the calculated binding affinity is weaker (Figure B.6A, B). Interestingly, we observed that the initiation of the PMF gradient and the  $\beta3$  association also correlate with the substrate length (Figure B.6A). While M51 and L52 start to turn into the  $\beta$ -strand conformation at  $RC > 15\text{\AA}$  in the APP substrate, this process started at  $RC 14.5\text{\AA}$  and  $RC 13.5\text{\AA}$  in A $\beta53$  and A $\beta52$ , respectively (Figure B.6B). The earlier stages of PMF gradient and  $\beta3$  association suggest that although the substrate residues C-terminal to L52 are not directly involved in  $\beta3$ -formation, they might indirectly facilitate the association process.

We next sought to understand how the formation of the  $\beta3$ -strand, active geometry, and water accessibility is influenced when M51 of the  $\beta3$ -strand is brought closer to the gap between  $\beta2$  and L432 of PS1. We found that the hydrogen bond between L432 of PS1 and V50 of APP and the hydrogen bond between K380 of PS1 and M51 of APP might play important roles in this process. In the dissociated state (regime I), these two hydrogen bonds are not formed and the gap between  $\beta2$  and L432 of PS1 is relatively flexible, indicated by the  $C\alpha$ - $C\alpha$  distance between L432 and K380 fluctuating within  $10\text{\AA}$  to  $14\text{\AA}$  (Figure 6.8A-D). This creates a gateway for water from the cytosolic side to access the catalytic center and perturbs the active site geometry. In this regime, a water-bridging hydrogen bonding network between D257 and D385 was found in all E-S complexes, which is a thermodynamically favorable state of the active site geometry that we have previously described for the apo-form  $\gamma$ -secretase (Figure 6.8 A-D and Figure 4.5B). In the transition state (regime II) M51 and L52 of APP can occasionally form a hydrogen bond with PS1 L432 and K380 on  $\beta2$ , respectively. The hydrogen bonds between the  $\beta3$ -strand and PS1 can efficiently block the water from accessing the catalytic center. When the  $\beta3$ -strand is brought to form a complete hybrid  $\beta$ -sheet cluster with PS1 (regime III), both hydrogen bonds are firmly formed and the  $C\alpha$ - $C\alpha$  distance between L432 and K380 of PS1 is confined to  $11\text{\AA}$  (Figure 6.8A-C), except for A $\beta51$  (Figure 6.8D). The firm hydrogen bonds on both sides of the  $\beta3$ -strand prevent additional water molecules from the intracellular side to access the catalytic center and in turn strongly

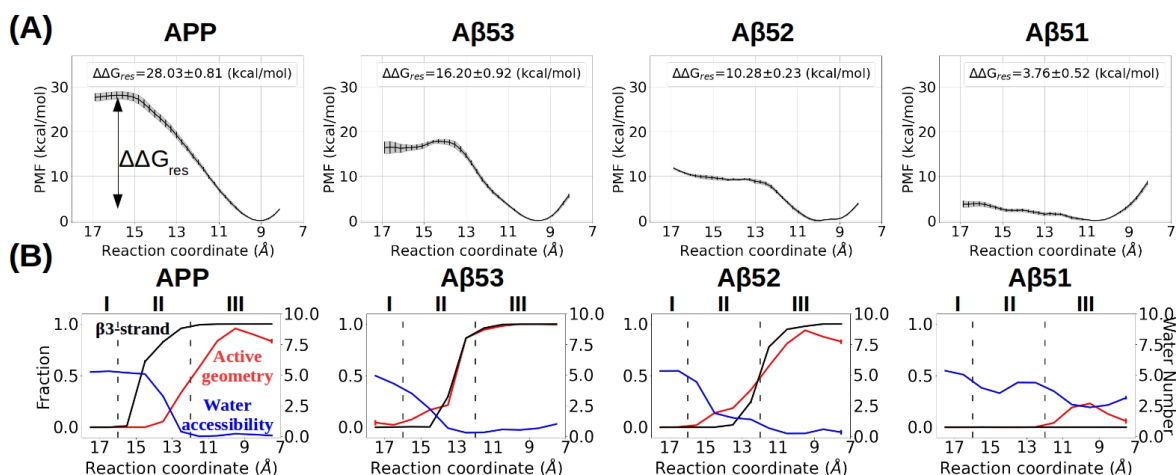


Figure 6.7.: Hamiltonian Replica exchange MD along the  $\beta$ 3-strand association pathway in APP and its truncated variants binding to  $\gamma$ -secretase-PS1-D385<sup>H</sup>. (A) Potential-of-mean-force (PMF) profiles were calculated with the HREUS method along the substrate  $\beta$ 3-strand association reaction of APP and its truncated substrates. (B) Fraction of sampled states that form an active site geometry compatible with cleavage (red),  $\beta$ 3-strand (black), and the number of water molecules around the catalytic center (blue), along the sampling pathway.

stabilize the active geometry (Figure 6.8A-C). The closing of the gap between  $\beta$ 2 and L432 of PS1 during the association process was also found in the PS1-D257<sup>H</sup> sampling (Figure B.6C). Note that a C $\alpha$ -C $\alpha$  distance of around 11 Å between L432 and K380 was sampled as well in all resolved  $\gamma$ -secretase structures when GSI or substrate forms hydrogen bonds with K380 and L432 of PS1 (Table 2).

Taken together, the free energy calculations on  $\gamma$ -secretase in complex with APP and the truncated variants allow us to understand the functional role of the hybrid  $\beta$ -sheet and how substrate length and PS1 protonation state influence the active site geometry. Our result shows that the formation of a sufficiently stable hybrid  $\beta$ -sheet is essential for the shaping of a stable active geometry and for limiting the number of water molecules at the catalytic center by forming hydrogen bonds with L432 and K380 of PS1. In particular, the Aβ51 substrate does not form a stable  $\beta$ 3-strand and coupled to this, also no active geometry at the catalytic site (Figure B.6B). Agreeing with the unrestrained MD simulations, HREUS simulations also indicate that PS1-D257<sup>H</sup> is more capable of water retention around the catalytic center while PS1-D385<sup>H</sup> binds the substrate  $\beta$ 3 more strongly.

### 6.3.3. APP M51P and L52P mutations decrease $\gamma$ -secretase cleavage efficiency

In the previous sections, we have demonstrated the essential role of a stable  $\beta$ 3-strand in forming a catalytically competent active site geometry of  $\gamma$ -secretase. To further verify this hypothesis, we next sought to see whether specifically disrupting the formation of a stable

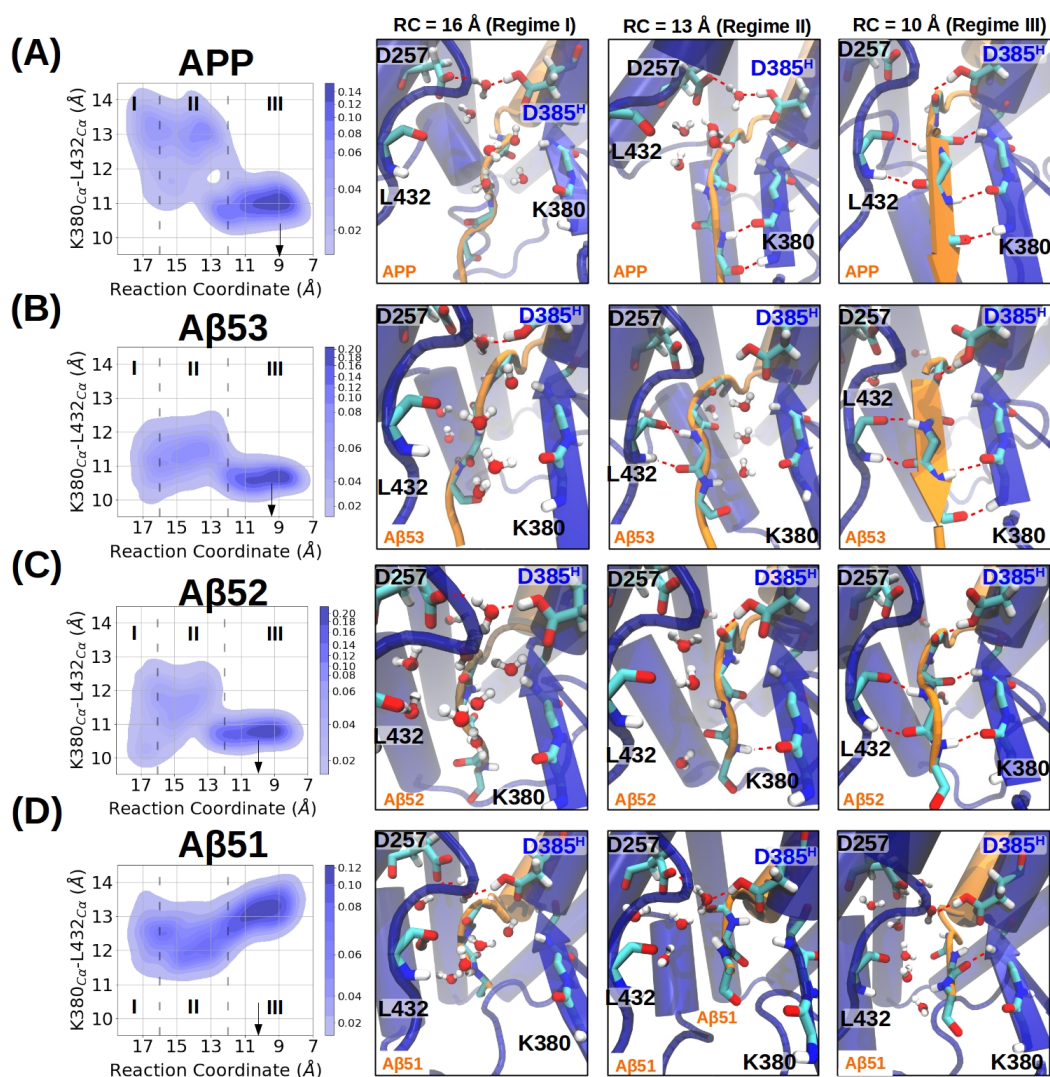


Figure 6.8.: Gap width between L432 and  $\beta 2$  during the  $\beta 3$ -strand association pathway in APP and its truncated variants binding to  $\gamma$ -secretase-PS1-D385H. Distribution of the gap width between L432 and  $\beta 2$  along the association pathway and representative snapshots in (left) regime I, (middle) regime II, and (right) regime III when (A) APP, (B) A $\beta$ 53, (C) A $\beta$ 52, or (D) A $\beta$ 51 binds to  $\gamma$ -secretase-PS1-D385H. The gap width between L432 and  $\beta 2$  is indicated by the C $\alpha$ -C $\alpha$  distance between L432 and K380. The black arrows point to the RC that corresponds to the PMF minimum. Three regimes are distinguished by the transparent dashed lines with regime I corresponding to the dissociated regime, regime II the transition regime, and regime III forming the associated regime. The atomic positions of L49 to L52 (from up to down) of the substrates are shown additionally in the licorice representation.

### 6.3. RESULT

Table 6.2.: C $\alpha$ -C $\alpha$  distances between PS1 K380 and L432 in the available cryo-EM structures. \* Residue K380 is not resolved in PDBID 5FN5 and PDBID 4UIS.

PDBID	Ligand	Hydrogen bond with PS1	K380 <sub>C<math>\alpha</math></sub> -L432 <sub>C<math>\alpha</math></sub> distance (Å)
4UIS	none	n.a.	n.a.*
5A63	none	n.a.	11.27
5FN2	DAPT (not resolved)	n.a.	8.42
5FN3	Unknown helix	No	5.09
5FN4	Unknown helix	No	11.13
5FN5	none	n.a.	n.a.*
6IYC	C83	L432, K380	11.24
6IDF	Notch1	L432, K380	11.07
6LQG	Avagacestat	No	12.11
6LR4	Semagacestat	L432, K380	10.68
7C91	L-685,458	L432, K380	11.20
7D8X	L-685,458 E2012	L432, K380	11.11
7Y5T	MRK-150	L432	12.10

hybrid  $\beta$ -sheet by substrate mutations may also influence  $\gamma$ -secretase cleavage. The amino acid proline can be used to disrupt hydrogen bonding in a  $\beta$ -sheet and we designed two APP substrate mutations M51P and L52P *in silico*. Similar to the free energy simulations on the bound wild-type (WT) substrate, the  $\beta$ 3-association/dissociation was studied using the HREUS simulations in the APP<sub>M51P</sub>- and APP<sub>L52P</sub>-  $\gamma$ -secretase complexes (see Methods). In comparison to the WT APP-bound complex, both substitutions lead to a decrease in  $\Delta\Delta G_{res}$  with the trend  $\Delta\Delta G_{res,APP} \gg \Delta\Delta G_{res,APP-L52P} > \Delta\Delta G_{res,APP-M51P}$  (Figure 6.9A). In the APP<sub>M51P</sub>-bound  $\gamma$ -secretase complex, both protonation states exhibit relatively flat PMF profiles (Figure 6.9B) and Figure B.7 compared to the WT APP substrate (Figure 6.7C and Figure B.6). In particular, the energy minimum no longer falls in the associated regime when APP<sub>M51P</sub> binds to PS1-D257<sup>H</sup> (Figure B.7A). In contrast, the L52P mutant shows a qualitatively similar (but reduced) PMF profile compared to WT (Figure 6.9B). Although the secondary structure of P51 and L52 in the APP<sub>M51P</sub> substrate can still stay in the  $\beta$ -strand conformations in the associated phase, the water molecules at the catalytic site are not effectively drained out as  $\beta$ 3-strand is brought to form the hybrid  $\beta$ -sheet segment (Figure 6.9C). Figure 6.9D, E show that APP<sub>M51P</sub> fails to close the gap between L432 and  $\beta$ 2 in the associated state (regime III). Despite hydrogen bond formation between L52 of APP<sub>M51P</sub> and K380 of PS1, the lack of a stable hydrogen bond between P51 of APP<sub>M51P</sub> and L432 of PS1 provides a water-accessible channel that consequently leads to a failure in forming a catalytically active geometry. In contrast, APP<sub>L52P</sub> is able to close the gap between L432 and  $\beta$ 2 by forming a stable hydrogen bond with PS1 L432 and a weak hydrogen bond with K380 (Figure 6.9E). This prevents the



water molecules from perturbing the active geometry and in the most thermodynamically favorable state ( $RC = 9 \text{ \AA}$ ), the active geometry was found in more than 90% of the sampled configurations (Figure 6.9C).

To prove our theoretical predictions, proline mutations were introduced at positions M51 and L52 of a C99-based recombinant C100-His6 substrate[190], and their cleavability by  $\gamma$ -secretase was assessed by incubation with a purified  $\gamma$ -secretase complex composed of PS1, NCT, PEN-2, and APH-1[191]. Indeed, compared to the WT substrate the proline mutants were less efficiently cleaved by  $\gamma$ -secretase, as judged from the decreased levels of AICD and A $\beta$  products (6.10A, B). Remarkably, as predicted the M51P mutation very strongly inhibited  $\gamma$ -secretase cleavage, while the L52P mutation was much better tolerated. In accordance with previous data[192], L52P was majorly cleaved after T48 ( $\epsilon$ 48) (Figure 6.10C) which can also be explained by the largely decreased  $\beta$ 3-strand association binding free energy in our simulations (Figure 6.9A). The P2' position is more sensitive to substitution compared to P3'. The cleavage site shift of the L52P mutant to  $\epsilon$ 48 places proline to the P4' position, resulting in a more favorable hybrid  $\beta$ -sheet formation. This explains why mutation L52P is efficiently cleaved at the  $\epsilon$ 48 site but M51P cleavage by  $\gamma$ -secretase is dramatically reduced. Importantly, a significant drop in cleavage efficiency was observed also in the shorter A $\beta$  peptides when Pro is placed at the P2' position such as T48P for the  $\zeta$ 46 cleavage[193]. This strongly suggests that the  $\beta$ 3-PS1-hybrid  $\beta$ -sheet is also indispensable in the subsequent cleavage of shorter A $\beta$  peptides.

#### 6.3.4. Simulations explain why P1' $\rightarrow$ F enhances while P2' $\rightarrow$ F mutation weakens the formation of the catalytically active geometry

Since the major products from C99 during  $\gamma$ -secretase sequential trimming follow either the A $\beta$ 49-A $\beta$ 46-A $\beta$ 43-A $\beta$ 40 or the A $\beta$ 48-A $\beta$ 45-A $\beta$ 42 product line, the initial endoproteolytic cleavage at the  $\epsilon$ 49 or  $\epsilon$ 48 sites largely controls also the final products. Among many C99 mutations altering the final A $\beta$ 42/A $\beta$ 40 ratio, the size of the initial substrate P2' position was suggested to have a direct impact on the  $\epsilon$ -cleavage site since it needs to fit into a proposed size-limiting S2' pocket in PS1, which is particularly problematic for bulky aromatic amino acids[66]. While  $\epsilon$ 49 cleavage is impaired for the C99 M51F mutant,  $\epsilon$ 48 cleavage is impaired in the V50F mutant. Furthermore, the conventional production lines can be altered by introducing large Phe residues at other substrate positions, e.g. T48F and I45F, hypothetically the P2' residue for the APP<sub>L52</sub> and  $\gamma$ 43 and cleavage, respectively[66, 156, 193] (putative cleavages inhibited by Phe are listed in Table 6.3).

In order to better mimic a general substrate-binding scenario for the subsequent cleavage poses, we used the A $\beta$ 52 system, where three residues are present beyond the scissile bond at the  $\epsilon$ 49 cleavage pose, as a model. Phe was introduced *in silico* in the A $\beta$ 52-bound  $\gamma$ -secretase complex at substrate positions P1' or P2', generating A $\beta$ 52<sub>V50F</sub>- and A $\beta$ 52<sub>M51F</sub>-bound complexes to represent general P1'  $\rightarrow$  F and P2'  $\rightarrow$  F models, respectively. The  $\beta$ 3-strand association processes of these two complexes were sampled using the HIREUS method. While the free energy difference  $\Delta\Delta G_{res}$  was increased in the V50F mutant, it is significantly decreased in the M51F mutant (Figure 6.11A). The PMF profile of  $\beta$ 3 association of the V50F

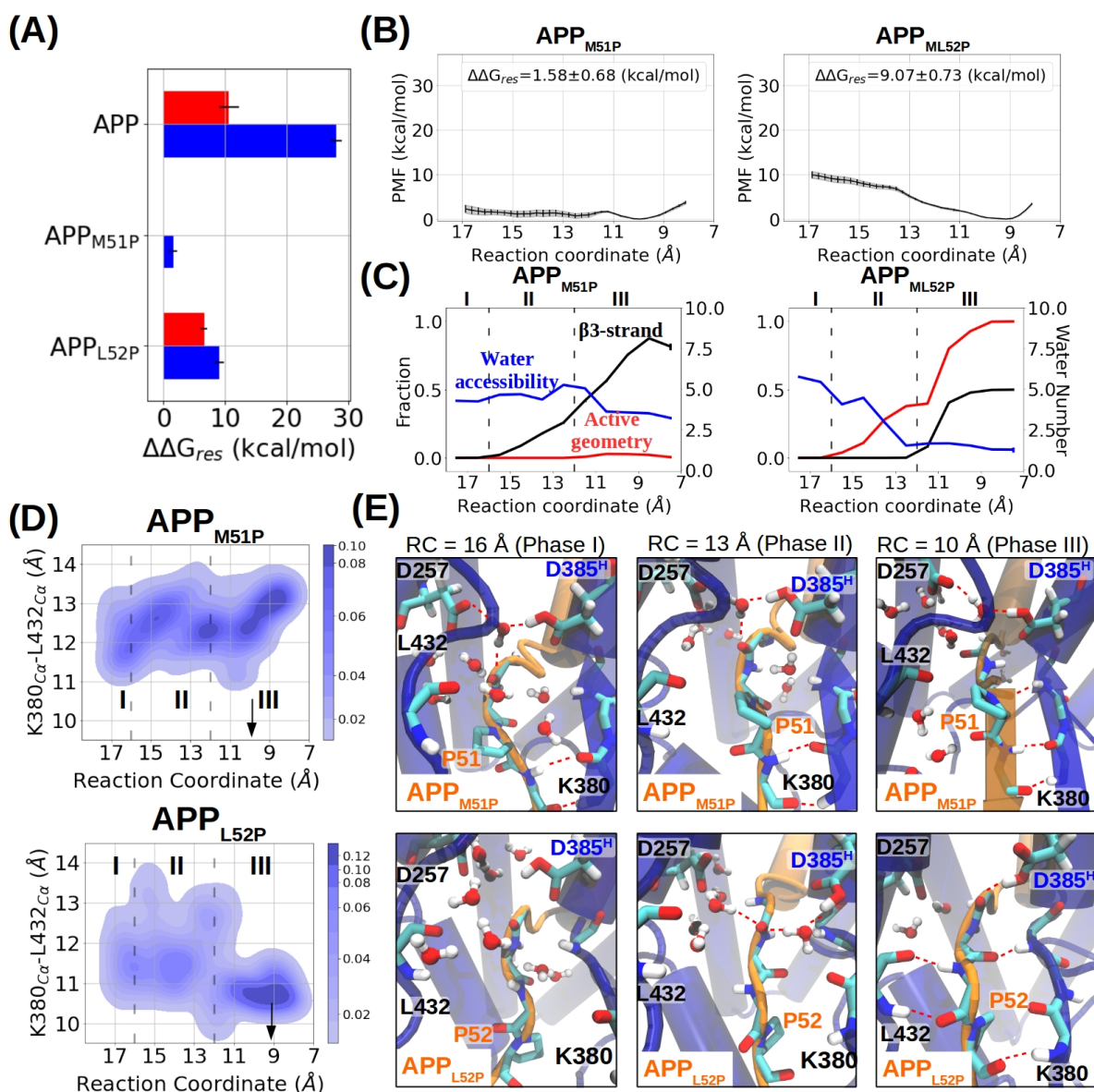


Figure 6.9.: Effect of APP substrate mutations M51P and L52P on binding to  $\gamma$ -secretase-PS1-D385<sup>H</sup> and APP catalysis (A) The free energy difference between the bound state and unbound state of substrate  $\beta$ 3-strand. The protonation state of PS1 is color-coded in red (D257<sup>H</sup>) and blue (D385<sup>H</sup>). (B) PMF profiles were calculated with the HREUS method along the substrate  $\beta$ 3-strand association reaction coordinate of APP with mutation M51P or L52P. (C) Fraction of sampled states that form an active site geometry compatible with cleavage (red),  $\beta$ 3-strand (black), and the number of water molecules around the catalytic center (blue), vs. the association reaction coordinate. (D) The change in the gap width between L432 and  $\beta$ 2 along the sampling pathway and the representative snapshots when APP<sub>M51P</sub> (upper panel) and APP<sub>L52P</sub> (lower panel) bind to  $\gamma$ -secretase. The atomic representations are shown in the same way as in Figure 6.2.

### 6.3. RESULT

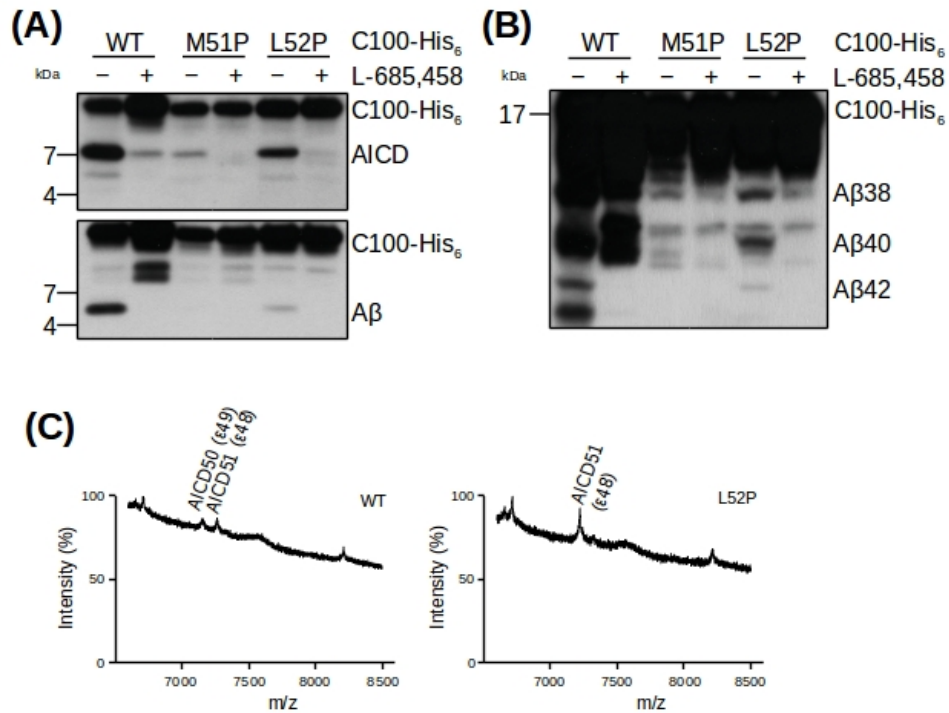


Figure 6.10.: Experimental verification on the cleavage efficiency of  $\gamma$ -secretase toward APP mutation M51P and L52P. (A)-(B) Analysis of WT and mutant C100-His<sub>6</sub> cleavage by  $\gamma$ -secretase after overnight incubation at 37 °C by immunoblotting for (A) AICD (Penta-His) and (B) A $\beta$  products (C) MALDI-TOF mass spectrometry analysis species of generated AICD species generated from WT (left) and mutant (right) C100-His<sub>6</sub> substrates. The intensity of the highest peak was set to 100%.

Table 6.3.: Putative inhibited cleavages in Phe mutations.

Mutation	Putative impaired cleavage	Reference
M51F	ε49	[66, 156]
V50F	ε48	[66]
T48F	ζ46	[66]
I47F	ζ45	[66]
I45F	ζ43	[66, 193]
V44F	ζ42	[66]

mutant is very similar to the WT and a flat free energy profile with the PMF minimum shifted toward dissociation is observed (Figure 6.7C, Figure 6.11B). In the  $A\beta_{52V50F}$  association process, water exclusion and  $\beta$ -strand formation are more completely executed than for WT (Figure 5C), and the gap between L432 and  $\beta 2$  is efficiently closed by the hydrogen bonds between PS1 and the substrate (Figure 6.11D, E). In contrast,  $A\beta_{52M51F}$  fails to prevent water molecules from reaching the catalytic center, and the active geometry is barely formed (Figure 56.11C). Although  $A\beta_{52M51F}$  is more capable of closing the gap between L432 and  $\beta 2$  than the other two impaired substrates  $A\beta_{51}$  and  $APP_{M51P}$  in the associated state, the gap is nonetheless not tight enough to drain out the water molecules (Figure 6.11D, E). Structural comparison between WT  $A\beta_{52}$  and the M51F mutant in the associated state shows that PS1 L432 deviated by around  $4\text{\AA}$  away from its original position to create enough space to accommodate substrate F51 (Figure B.8), supporting the idea of the “size-limiting” PS1 S2’ pocket suggested by previous work. The same effects of  $A\beta_{52V50F}$  and  $A\beta_{52M51F}$  were also found when forming the PS1-D257<sup>H</sup>  $\gamma$ -secretase complex (Figure B.9).

In Table 6.4, we list the residues of PS1 with a minimal residue-residue distance lower than  $5\text{\AA}$  from substrate P2’ in our simulations. We note that these contacting residues are fully consistent with the S2’ pocket described by the cryo-EM structures, inhibitor-bound model but the alternative S2’ pocket Bhattarai et al. indicated using the Gaussian accelerated molecular dynamics approach was not sampled[157]. Taken together, our results offer explanations in molecular detail why substrates with a large Phe at P2’ position shift the  $\epsilon$ -cleavage site positions. The large side chain at P2’ not only weakens the binding free energy of the hybrid  $\beta$ -sheet with PS1 but also expands the gap between L432 and  $\beta 2$  which results in water leakage to the catalytic center and destabilizes the active geometry. Our computational work also shows that mutating P1’ can increase the binding affinity of the substrate  $\beta 3$ -strand, forming a more stable hybrid  $\beta$ -sheet with PS1 and a more stable active geometry. These findings agree with and nicely explain the experimental result that the  $\epsilon 49/\epsilon 48$  ratio is increased in the C99 V50F mutation and decreased in the M51F as well as how  $\gamma$ -secretase skips cleavages when Phe is placed at the P2’ position in other processive cleavage steps.

## 6.4. Discussion and outlook

The reason of why  $\gamma$ -secretase cleaves the C99 substrate in steps of at least three residues remains largely elusive. In our recent computational work, we have related the three-residue-wise cleavage with the putative translational movement of the substrate  $3_{10}$ -helix at the PS1 internal docking site[146]. In this study, we used a combination of molecular simulations and biochemical experiments to elucidate the role of the  $\beta$ -strand cluster and water distribution for forming catalytically active geometries at the active site of  $\gamma$ -secretase. Despite its stability being called in question by computational work performed by Mehra et al[189], the hybrid  $\beta$ -sheet C-terminal to the  $\epsilon$ -cleavage sites, which consists of  $\beta 1$ ,  $\beta 2$ , and  $\beta 3$ , and PS1 L432 were shown to be critical for  $\gamma$ -secretase proteolysis[39, 58]. Our simulations indicate that  $\beta 3$  and consequently also the catalytically active geometry can only be stably formed when at least three substrate residues are present C-terminally of the scissile bond. The free energy

## 6.4. DISCUSSION AND OUTLOOK

Table 6.4.: PS1 residues found in within 5Å of any atom of substrate P2' using the last frame of the window with RC= 9.8Å in HREUS simulations.

Substrate	S2' in PS1-D257 <sup>H</sup>	S3' in PS1-D257 <sup>H</sup>
APP	L85,V379,K380,L381,G382,L418,T421, L422,L425,A431,L432,P433,A434	V82,L85,F86,V261,V379,K380,L381,G382, D385,L418,T421,L422,L425,K430, A431,L432,P433,A434
APP <sub>M51P</sub>	D257,V261,L268,V272,L381,G382, T421,L425,A431,L432,P433,A434	V379,K380,L381,G382,D385,T421,L422, L425,A431,L432,P433,A434
APP <sub>L52P</sub>	V261,V272,I287,V379,K380,L381,G382, L418,T421,L422,L425,K430,A431, L432,P433,A434	V379,K380,L381,G382,D385,T421,L422, L425,A431,L432,P433,A434
Aβ53	L85,V379,K380,L381,G382,D385,F386,Y389,L418,T421, L422,L425,A431, L432,P433,A434	L85,V261,I287,V379,K380,L381,G382, D385,Y389,L418,T421,L422,L425,K430, A431,L432,P433,A434
Aβ52	L85,V261,V272,V379,K380,L381,D385, Y389,L418,T421,L422,L425,K430, A431,L432,P433,A434	L85,V261,V379,K380,L381,G382,D385, L418,T421,L422,L425,K430,A431,L432, P433,A434
Aβ52 <sub>V50F</sub>	L85,I287,V379,K380,L381,G382,L418, T421,L422,L425,A431,L432,P433, A434	L85,V261,V379,K380,L381,G382,D385, L418,T421,L422,L425,K430,A431,L432,P433, A434
Aβ52 <sub>M51F</sub>	H81,L85,R377,V379,K380,L381,G382, L422,L425,L432,A434	H81,L85,V379,K380,L381,G382,T421, L422,L425,A431,L432,P433,A434
Aβ51	L85,V272,K380,L381,G382,D385,L418 ,T421,L422,L425,A431,L432,P433, A434	L85,Q276,K380,L381,G382,D385,Y389, L418,T421,L422,L425,A431,L432,A434

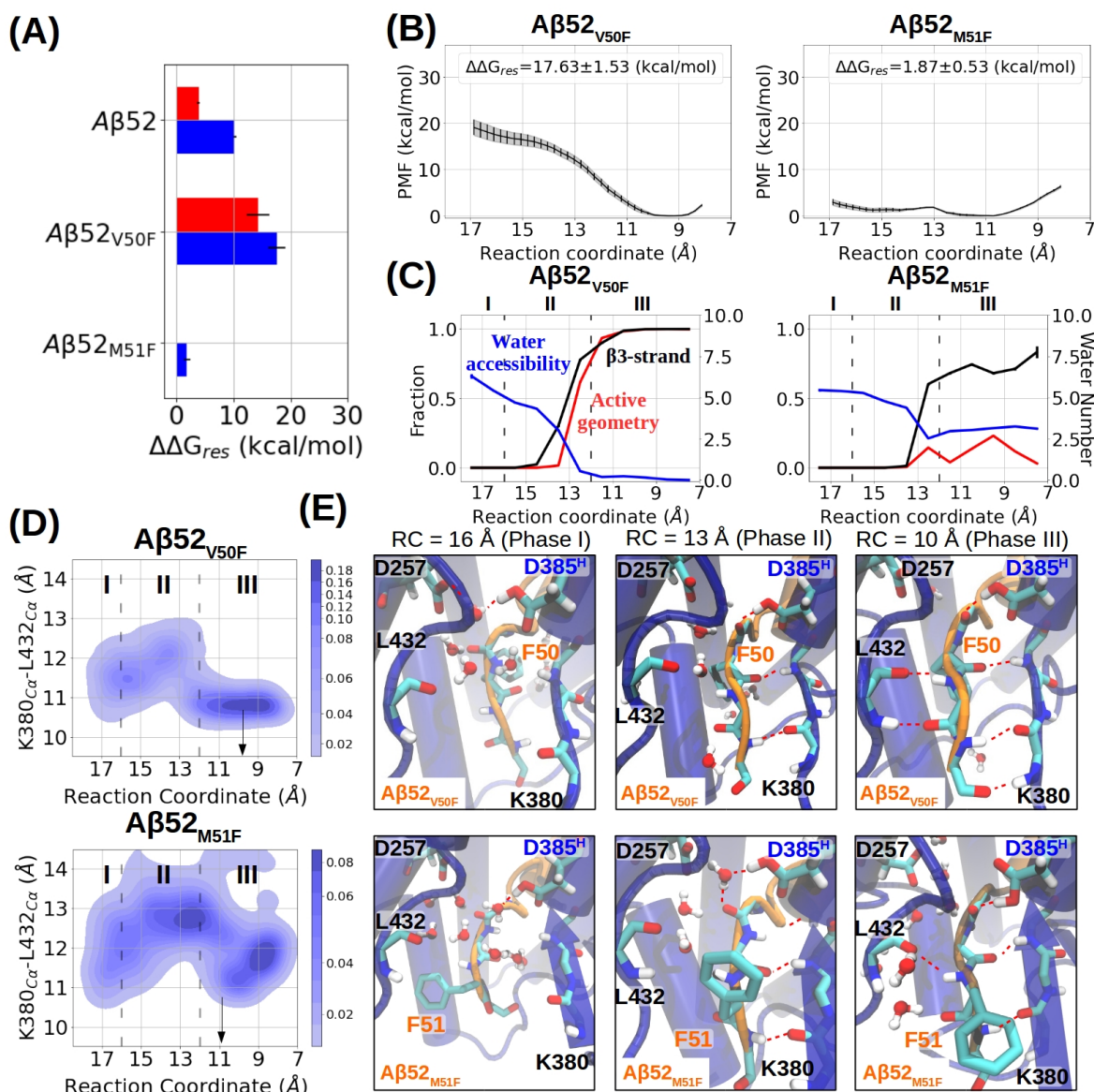


Figure 6.11.: Free energy calculations and unrestrained simulations on V50F and M51F substitutions in the  $A\beta_{52}$  substrate bound to  $\gamma$ -secretase-PS1-D385<sup>H</sup>. (A) The free energy difference between the bound state and unbound state of substrate  $\beta$ 3-strand. The protonation state of PS1 is color-coded in red (D257<sup>H</sup>) and blue (D385<sup>H</sup>). (B) PMF profiles were calculated with the HREUS method along the substrate 3-strand association reaction coordinate (RC) of  $A\beta_{52}$  with mutation V50F or M51F. (C) Fraction of sampled states that form an active site geometry compatible with cleavage (red),  $\beta$ 3-strand (black), and the number of water molecules around the catalytic center (blue), vs. RC. (D) The change in the gap width between L432 and  $\beta_2$  along the sampling pathway and the representative snapshots when  $A\beta_{52}$  mutants V50F (upper panel) and M51F (lower panel) bind to  $\gamma$ -secretase. The atomic representations are shown in the same way as in Figure 6.8.

profiles show for the longer substrates a strong  $\beta 3$  association affinity that diminishes for substrates with only 1 or 2 residues C-terminal of the scissile bond. Hence, this result offers a direct explanation of why  $\gamma$ -secretase cleaves APP in steps of three residues[39, 40].

Our simulations allow us also to draw a picture of the molecular details of the catalytic process: In the association process, the gap between  $\beta 2$  and L432 of PS1 serves as the bottleneck that recognizes and recruits the substrate P2' residue into the PS1 S2' sub-site, whereas  $\beta 3$  serves as a bottle plug which prevents redundant water molecules from flowing into the catalytic site. Before the  $\beta 3$  association, more than five water molecules from the intracellular side can access the catalytic center through the gap between  $\beta 2$  and PS1 L432 and prevent hydrolysis by disturbing the topology at the catalytic center. Upon association,  $\beta 3$  is attracted to the gap between  $\beta 2$  and PS1 L432 mediated by VDW interaction and drains the water from the catalytic center. In the associated regime,  $\beta 3$  zips up the gap between  $\beta 2$  and L432 by forming hydrogen bonds with them, and only less than two water molecules are trapped at the cleavage site allowing the formation of a stable catalytically active geometry.

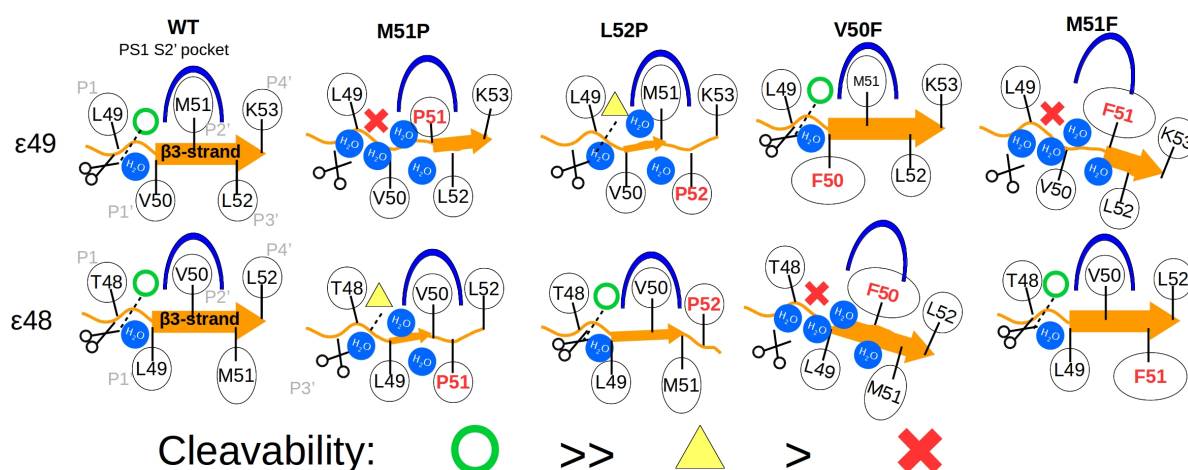


Figure 6.12.: Schematics of the APP mutations effect on  $\epsilon 49/\epsilon 48$  cleavage ratio according to our simulation model. The  $\epsilon 49$  cleavability of (A) WT, (B) M51P, (C) L52P, (D) V50F, and (E) M51F APP are shown according to our computational results. The cleavability at the  $\epsilon 48$  cleavage site is conjectured by assuming the same binding pose as in the  $\epsilon 49$  cleavage with one residue shift. The hypothetical cleavability of substrates at either the  $\epsilon 49$  or  $\epsilon 48$  cleavage is indicated with the green circle (unaffected), yellow triangle (weakened), and red cross (inhibited). The PS1 S2' pocket is indicated in blue and APP is represented in orange. The  $\gamma$ -secretase catalysis is represented by black scissors and water molecules are shown with blue circles. The  $\epsilon 49/\epsilon 48$  ratios shown in the bottom panel are predicted according to our model. Conformation of the  $\beta 3$ -strand is depicted based on the residue-wise secondary structure calculation shown in Figure B.10.

By assuming that the binding pose of the  $\epsilon 48$  cleavage is structurally similar to the  $\epsilon 49$  cleavage, our simulation model is able to explain the shift in  $\epsilon 49/\epsilon 48$  ratio observed

experimentally, as illustrated in Figure 6.12. In WT APP, M51 fits into the PS1 S2' pocket and the hybrid  $\beta$ -sheet can efficiently control limited water access to the catalytic center (Figure 6.12A). Introducing a Pro to either the substrate P2' or P3' positions weakens the hybrid  $\beta$ -sheet stability (Figure 6.12B, C). Placing a Phe at the substrate P2' position expands the gap between PS1 K380 and L432 because of the limited size of the PS1 S2' pocket (Figure 6.12B, C). In the M51P mutant, the hydrogen bond between substrate P2' and PS1 L432 is lost and this creates a water-accessible gateway. In the case of the L52P mutant, the substrate  $\beta$ 3-strand is shortened and the binding affinity is weaker than for the WT case, nevertheless, the simulations indicate that the active geometry was still sampled when the hybrid  $\beta$ -sheet is formed. However, the problem of weak  $\beta$ 3-strand binding affinity of the L52P mutant can be circumvented by adopting the binding pose for the  $\epsilon$ 48 cleavage (Figure 6.12C). Indeed, consistent with our model,  $\epsilon$ 48 cleavage in the L52P mutant was observed in our cell-free assay as a sole  $\epsilon$ -cleavage event. Exclusive or strongly preferred cleavage, respectively, at T48, has also been observed in previous experimental studies[193, 61]. Moreover, also consistent with our model, the M51P substrate mutation caused a very strong drop in total activity in our assay. In contrast to the Pro mutations, the M51F mutant does not alter the  $\beta$ 3-strand stability but expands the gap between the substrate and PS1 residue L432 and fails to block the water from accessing the catalytic center when associated in the  $\epsilon$ 49 cleavage pose although energetically unfavorable, leading to a decrease in the  $\epsilon$ 49/ $\epsilon$ 48 ratio[66, 156] (Figure 6.12E). Along these lines, the V50F mutation should also suffer from water leakage through the same mechanism and should fail to form a stable active geometry when binding in the  $\epsilon$ 48 cleavage pose, hence, it should lead to an increase in the  $\epsilon$ 49/ $\epsilon$ 48 ratio which has been observed experimentally [66] (Figure 6.12D). It is important to add that our simulation-derived model can also explain the observed effect of Pro and Phe substitutions on several subsequent APP substrate cleavages. This includes the inhibition of the  $\zeta$ 46 cleavage ( $A\beta$ 49  $\rightarrow$   $A\beta$ 46) by T48F[157, 193] and T48P, the observed inhibition of the  $\gamma$ 43 cleavage ( $A\beta$ 46  $\rightarrow$   $A\beta$ 43) by I45F[66, 193], and the inhibition of  $\gamma$ 40 cleavage ( $A\beta$ 43  $\rightarrow$   $A\beta$ 40) by A42F (Supplementary Figure S19), suggesting that the hybrid  $\beta$ -sheet conformation, especially at the P2' position of the substrate, is also indispensable for the cleavage of shorter  $A\beta$  peptides. Our study also gives important insights into the influence of the protonation states of the active site residues D257 and D385 in PS1. While qualitatively similar results were obtained for both protonation states in support of the role of the hybrid- $\beta$ -sheet, differences in active geometry, water recruitment, and  $\beta$ 3 association strength were observed. While the hydrophobic surface of TMD6a in contact with V50 and L52 of C99 makes the proximity of D257 more hydrophobic, the GxGD motif in TMD7 allows water to dwell in the cavity in the vicinity of D385. In the simulations, the D257-protonated PS1 on average recruited more water molecules to the catalytic center than the D385-protonated state. Meanwhile, in all the studied E-S complexes, D385-protonated PS1 binds substrate  $\beta$ 3 more firmly at the  $\beta$ -sheet region and traps the limited water molecules longer around the catalytic center. Taking into account that the peptide hydrolysis results in the net proton transfer from the protonated aspartate to the unprotonated, the functional discrepancies between two protonation states might suggest that  $\gamma$ -secretase captures the substrate in the D385<sup>H</sup> protonation state and releases the product



when the proton is transferred to D257. This could be the subject of a future simulation study that allows the sampling of different protonation states during MD simulations. Our work provides mechanistic insight into how the hybrid  $\beta$ -sheet facilitates  $\gamma$ -secretase proteolysis by excluding water from the catalytic site and the fundamental reason for three-residue-wise cleavage on C99. Notably, an enzyme-substrate  $\beta$ -sheet and the size-dependent modulation near the catalytic center are also found in rhomboid, an intramembrane serine protease[194], and the other wide range of proteases [195] (Figure B.11). It is of broader interest for further work to examine if the formation E-S hybrid  $\beta$ -strand is a necessary factor in intramembrane proteolysis using both computational and experimental methods.

## 7. An internal docking site stabilizes substrate binding to $\gamma$ -secretase: Analysis by Molecular Dynamics Simulations

Amyloid precursor protein (APP) is cleaved and processed sequentially by  $\gamma$ -secretase yielding amyloid  $\beta$ -peptides ( $A\beta$ ) of different lengths. Longer  $A\beta$  peptides are associated with the formation of neurotoxic plaques related to Alzheimer's disease. Based on the APP substrate-bound structure of  $\gamma$ -secretase we investigated the enzyme-substrate interaction using Molecular Dynamics simulations and generated model structures that represent the sequentially cleaved intermediates during the processing reaction. The simulations indicated an internal docking site providing strong enzyme-substrate packing interaction. In the enzyme-substrate complex, it is located close to the region where the helical conformation of the substrate is interrupted and continues towards the active site in an extended conformation. The internal docking site consists of two non-polar pockets that are preferentially filled by large hydrophobic or aromatic substrate side chains to stabilize binding. The placement of smaller residues such as glycine can trigger a shift in the cleavage pattern during the simulations or result in the destabilization of substrate binding. The reduced packing by smaller residues also influences the hydration of the active site and the formation of a catalytically active state. The simulations on processed substrate intermediates and a substrate G33I mutation offer an explanation of the experimentally observed relative increase of short  $A\beta$  fragments production with this mutation. In addition, studies on a substrate K28A mutation indicate that the internal docking site opposes the tendency of substrate dissociation due to a hydrophobic mismatch at the membrane boundary caused by K28 during processing and substrate movement towards the enzyme active site. The proposed internal docking site could also be useful for the specific design of new  $\gamma$ -secretase modulators. The work presented in this chapter are reproduced from our published work[146].

### 7.1. Introduction

$\gamma$ -Secretase is an intramembrane protease that cleaves over 140 type-I membrane peptides[22]. The catalytic center involves two aspartic acid residues located on transmembrane domains 6 and 7 of the catalytic subunit presenilin (PS1 or PS2)[29, 196, 22]. Among the known substrates the amyloid precursor protein (APP) is the most studied substrate because of its potential involvement in the pathology of Alzheimer's disease (AD)[27, 73]. According to the amyloid hypothesis, the cerebral plaques composed of amyloid  $\beta$  ( $A\beta$ ) are the neurotoxic substance inducing neuro-inflammation that eventually leads to cell death[4, 28]. Prior to

the  $\gamma$ -secretase cleavage, the extracellular domain of APP is first shedded by the soluble protease  $\beta$ -secretase to produce a transmembrane peptide termed APP CTF $\beta$  or C99[6]. C99 is subsequently recognized and cleaved in its transmembrane domain by  $\gamma$ -secretase, releasing N-terminal A $\beta$  peptides and the C-terminal APP intracellular domain (AICD)[197, 198, 199, 200]. The major A $\beta$  peptides contain between 37 (A $\beta$ 37) and 43 (A $\beta$ 43) amino acids[9]. The longer A $\beta$  peptides such as A $\beta$ 42 and A $\beta$ 43 are predominantly generated in case of familial AD (FAD) due to mutations in PS1 or APP, and are highly aggregation-prone and deposited in plaques[18, 19, 73, 201]. Secretion of A $\beta$  follows mainly two production lines starting by cleaving the amide bond between L49-V50 ( $\epsilon$ 49) or between T48-L49 ( $\epsilon$ 48). After the initial cleavage,  $\gamma$ -secretase trims A $\beta$  peptides in a three residue-stepwise proteolysis, following the sequence: A $\beta$ 49-A $\beta$ 46-A $\beta$ 43-A $\beta$ 40-A $\beta$ 37 or A $\beta$ 48-A $\beta$ 45-A $\beta$ 42-A $\beta$ 38[67]. In addition, A $\beta$ 38 can also be derived from A $\beta$ 43, generating the VVIAT peptide[67]. Modulators (GSM) and inhibitors (GSI) specific to desired substrates of  $\gamma$ -secretase are of significant medical interest because especially the latter compounds can selectively enhance the A $\beta$  processivity and may reduce the generation of long A $\beta$  that can form toxic aggregates[202, 21]. To understand how APP is cleaved into A $\beta$  peptides, biochemical photo-crosslinking assays have been used to study the process of C99 recruitment[60]. In addition, substrate-based chemical probes have been employed to understand how C99 binds to  $\gamma$ -secretase[203, 204, 86, 87]. According to these studies, following interactions with exosites in the nicastrin (NCT) and presenilin enhancer-2 (PEN-2) subunits of  $\gamma$ -secretase, C99 binds to the  $\gamma$ -secretase exosite in the N-terminal fragment of PS1 (PS1-NTF) before it reaches the catalytic center between PS1-NTF and PS1-CTF. It has also been found that FAD mutations in APP and PS1 weaken the binding between  $\gamma$ -secretase and A $\beta$  intermediates[76]. Moreover, PS1 FAD mutants misposition the C99 cleavage site domain by altering the interactions between substrate and enzyme. However, despite extensive research on PS1 mutations[77, 192, 71, 205, 78, 148, 193], A $\beta$  variants and AICD species, and binding of modulators and inhibitors, molecular mechanism of how FAD mutations and GSMs modulate the generation of secreted A $\beta$  species is still not completely resolved and fully understood[206, 207, 102]. The first three-dimensional (3D) structures of substrate-bound  $\gamma$ -secretase complexes were solved in early 2019[39, 40]. These structures show in molecular detail how C83 (a shortened C99), and Notch, another  $\gamma$ -secretase substrate regulating cell differentiation, bind to  $\gamma$ -secretase. Both substrates are bound in a very similar fashion in the space between TMD2 and TMD3 of PS1 (Figure 1A). In addition, both substrates adopt the same helix-loop-strand secondary structure profile. The cleavage site is located at the connection between the loop domain and the extended strand domain, which was predicted and later confirmed as a binding site of transition state analogs (TSA)[155, 58]. Notably, several FAD mutations are found at the connection between the helical and loop segment of the A $\beta$  substrate (over 12 FAD mutations were found within A42-V46, Table 4.3). The  $\beta$ -strand at the C-terminal side of the substrate scissile bond, termed  $\beta$ 3, is stabilized by forming a  $\beta$ -sheet with  $\beta$ 1,  $\beta$ 2-strands and L432 of PS1 (Figure 1A right panel). The disruption of the hybrid  $\beta$ -sheet has been shown to reduce the secretion of AICD, suggesting its crucial role in  $\gamma$ -secretase proteolysis. Another characteristic interaction feature shared by the two substrates is a hydrophobic well-packed contact between substrate residues

P5, P6 (measured relative to the initial cleavage site) and the TMD2-TMD5 of PS1 around the transition between the substrate loop segment and the helical part (Figure 1A left panel). Interestingly, in contrast to the helical part of the bound substrate, the Cryo-EM density of this substrate region is very well defined for the  $\gamma$ -secretase–substrate complexes[155, 58]. The interaction region contributes significantly to substrate binding since designed helical substrate-derived inhibitors of  $\gamma$ -secretase that include residues beyond the helical part (but not reaching beyond the active site to form a  $\beta$ -sheet), termed D-13, inhibit PS1 much more strongly than just the helical segment of the substrate (D-10)[208]. In the following, we term the contact site, distinct from the active site, the PS1 internal docking site to distinguish it from the previously mentioned external exosite[60]. In a recent inhibitor design study, the helix-loop-strand feature of the C83 and Notch1 substrates was adopted by connecting a helical peptide inhibitor (HPI) and a TSA motif with a  $\omega$ -aminoalkanoyl linker[86]. The resulting inhibitor achieved a very low subnanomolar IC<sub>50</sub> potency, five-fold more potent than the TSA alone, suggesting it binds both at the PS1 internal docking site and at the active site[209, 204, 86, 203]. In the present study, we used molecular dynamics (MD) and homology modeling approaches to investigate the non-polar contact between the substrate helical domain and the PS1 internal docking site and how APP mutations attenuate or promote the E-S stability. We constructed 23 substrate-bound  $\gamma$ -secretase complex variants by introducing APP mutations and modeling shorter A $\beta$  peptides *in silico* and studied the molecular dynamics in a POPC lipid bilayer. The simulations indicate that the packing at the internal docking site is an essential element for stabilizing the helical domain of the substrate. Moreover, occupying the cavity between PS1 TMD3-TMD5 prevents excessive water molecules from disturbing the catalytic geometry. We modeled the putative binding poses of C99, A $\beta$ 49, A $\beta$ 46, A $\beta$ 43, and A $\beta$ 40, and identified competing forces that determine the A $\beta$  peptide to stay or leave  $\gamma$ -secretase. As the substrate is trimmed shorter, the membrane-anchoring residue K28 which also interacts with the NCT domain[210] pulls down the peripheral lipid molecules as the helical part of the substrate proceeds to the next cleavage pose. It attenuates the local membrane thickness and imposes a force to pull the substrate out of the binding channel. The binding of especially hydrophobic residues at the internal docking site stabilizes binding and substrate processing.

To further verify the model, we simulated the Ab40- $\gamma$ -secretase complexes with the Ab37/Ab38-promoting APP mutation K28A [211, 212] and G33I [147, 213], which presumably weakens the membrane up-pulling force (meaning in the following toward the extracellular region) and enhances the substrate-binding affinity, respectively. Our model provides a molecular explanation of how mutations around V44 and I45 may alter the  $\epsilon$ -cleavage and how APP mutations at K28 and G33 facilitate the secretion of shorter Abproducts. We offer a framework to elucidate how Ab peptides are bound, trimmed, processed, and secreted by  $\gamma$ -secretase, and anticipate that our work will facilitate the structure-based design of more potent GSMs

## 7.2. Methods and Materials

Because this chapter is focusing on the processivity of the A $\beta$  substrates, we termed the APP substrate C99 in this chapter for better consistency with Figure 2.5. Simulations for the C99-bound  $\gamma$ -secretase complex are described as in Chapter 4 section 3.3. For the enzyme-substrate complexes, two different protonation states of the active site with either the D257 or D385 protonated were considered. Start structures for  $\gamma$ -secretase in complex with the C99-A $\beta$ 49-A $\beta$ 46-A $\beta$ 43-A $\beta$ 40-A $\beta$ 37 substrates representing consecutive cleavages are constructed using MODELLER [116] comparative modeling using the last frame of the simulation at the prior cleavage binding pose as the template. The aligned sequences are shown in Table 7.1. The modeled A $\beta$ -bound structures solvated in simulation boxes the same way as the C99-bound  $\gamma$ -secretase complex described in Chapter 4. Secondary structure analysis, water accessibilities, and binding free energy by MM/PBSA method followed the same procedure as described in Chapter 4

## 7.3. Results

### 7.3.1. Weakening the interaction at PS1 internal docking site leads to re-adjustment of substrate helical interface

In Chapter 4, we observed that the hydrophobic interaction between the PS1 internal docking site between substrate P5, P6 contributes a substantial E-S interaction (Figure 4.4A) and stabilizes TMD2-TMD5 of PS1 (Figure 4.4B) in the C99-bound and Notch1-bound simulations. When no substrate is presented at the internal docking site, this region is filled with water molecules and the water is excluded from the region when hydrophobic residues are placed in it (Figure A.7). In this chapter, we will focus on how the E-S interactions at the internal docking site are affected when smaller and less hydrophobic amino acids are placed inside the hydrophobic cavities.

Presumably, along the consecutive trimming process, C99 needs to move “downward” (from the extracellular towards the intracellular side of PS1) in order to expose the upstream scissile bonds to the catalytic center. Our simulations suggest that due to the  $3_{10}$ -helix formed within residues T43-I45, a translational movement without altering the substrate orientation is sufficient for the substrate proceeding to the next cleavage binding pose (Figure A.6).

However, sequence alignment along the A $\beta$ 49-A $\beta$ 46-A $\beta$ 43-A $\beta$ 40-A $\beta$ 37 production line shows that for this process sometimes small residues such as A42, G38, and G33 are placed in the internal docking site (see Table 7.1). According to our observation in the C99-bound and Notch1-bound complex such small residues may fail to stabilize the substrate at the PS1 internal docking site. Indeed, in support of our model, the experimental double mutation V44G-I45G on C99 was shown to significantly decrease the substrate cleavability and skewing the  $\epsilon$ -cleavage site to generate longer AICD such as AICD49-99, AICD46-99, and AICD45-99 (corresponding to the  $\epsilon$ 48,  $\epsilon$ 45, and  $\epsilon$ 44 cleavages, respectively)[214]. To study these scenarios, we introduced *in silico* glycine substitutions of V44 or/and I45 on C99 to generate

### 7.3. RESULTS

Table 7.1.: Residues of  $\gamma$ -secretase substrate at the certain cleavage pose denoted by the subscript. The membrane-anchoring residue K28 is highlighted in blue, P5 and P6 are highlighted in green.

Position	C99 <sub><math>\epsilon</math>49</sub>	C99 <sub><math>\epsilon</math>48</sub>	A $\beta$ 49 <sub><math>\zeta</math>46</sub>	A $\beta$ 46 <sub><math>\gamma</math>43</sub>	A $\beta$ 43 <sub><math>\gamma</math>40</sub>	A $\beta$ 43 <sub><math>\gamma</math>37</sub>
P22	K28	N27	G25	E22	F19	K16
P21	G29	K28	S26	D23	F20	L17
P20	A30	G29	N27	V24	A21	V18
P19	I31	A30	K28	G25	E22	F19
P18	I32	I31	G29	S26	D23	F20
P17	G33	I32	A30	N27	V24	A21
P16	L34	G33	I31	K28	G25	E22
P15	M35	L34	I32	G29	S26	D23
P14	V36	M35	G33	A30	N27	V24
P13	G37	V36	L34	I31	K28	G25
P12	G38	G37	M35	I32	G29	S26
P11	V39	G38	V36	G33	A30	N27
P10	V40	V39	G37	L34	I31	K28
P9	I41	V40	G38	M35	I32	G29
P8	A42	I41	V39	V36	G33	A30
P7	T43	A42	V40	G37	L34	I31
<b>P6</b>	<b>V44</b>	<b>T43</b>	<b>I41</b>	<b>G38</b>	<b>M35</b>	<b>I32</b>
<b>P5</b>	<b>I45</b>	<b>V44</b>	<b>A42</b>	<b>V39</b>	<b>V36</b>	<b>G33</b>
P4	V46	I45	T43	V40	G37	L34
P3	I47	V46	V44	I41	G38	M35
P2	T48	I47	I45	A42	V39	V36
P1	L49	T48	V46	T43	V40	G37
P1'	V50	L49	I47	V44	I41	G38
P2'	M51	V50	T48	I45	A42	V39
P3'	L52	M51	L49	V46	T43	V40
P4'	K53	L52	V50	I47	V44	I41

the complexes of C99<sub>V44G</sub>, C99<sub>I45G</sub>, and C99<sub>GG</sub> with  $\gamma$ -secretase and study the effect in MD simulations. Snapshots at the internal docking site of the mutants are shown in Figure 7.1A.

Among the three C99 mutants, only C99<sub>V44G</sub> was able to largely preserve the formation of the catalytic hydrogen bond (Figure 7.1B) and fluctuated the least (Figure 7.1C). In contrast, I45G and V44G-I45G mutants formed less frequent catalytic hydrogen bonds (Figure 7.1B) and showed respectively mildly and largely more fluctuations at around V44-V46 of the substrates (Figure 7.1C). By computing the residue-wise water accessibility, we found that while I41-I47 remain in a dry environment in C99<sub>V44G</sub> and C99<sub>wt</sub> models, they are more exposed to the aqueous environment in the case of bound I45G and double-glycine substrate variants (Figure 7.1D). As depicted in Figure 7.1C, in fact, in the simulations of substrate-free  $\gamma$ -secretase, the PS1 internal docking site is filled up with water molecules coming from the intracellular side of the membrane, in particular in the cavity between TMD3-TMD5 (Shown as the white surface of Figure 7.1A). When C99<sub>V44G</sub> or C99<sub>wt</sub> binds to the PS1 interior, the water-filling cavity is occupied by C99 I45 sidechain with a calculated binding energy contribution of  $\sim -4.95$  kcal/mol (Figure 2C, Figure 7.1A and Figure C.2A). However, when the water-blocking I45 is mutated to glycine, contributing a binding energy of only  $\sim -0.61$  kcal/mol, the cavity is again filled with water (Figure 7.1A and Figure C.2A). In case of the double substitution of V44G and I45G, the substrate helical domain rotates counterclockwise and V46 and I47 move into the PS1 internal docking site in one of the simulations (Figure C.3 and Video S2 from our published work). We note that none of the C99 glycine mutants perturbed the post-cleavage site hybrid  $\beta$ -sheet but they all altered the local helicity around V44-V46 (Figure C.2B). The C99<sub>I45G</sub> and C99<sub>GG</sub> variants enrich the helical occupation on V46, by which V46 is drawn spatially closer to the water-filling cavity formed between TMD3-TMD5 (similar results were observed with the D257<sup>H</sup> protonation state with dissociation of the post-cleavage site hybrid  $\beta$ -strand observed in the I45G mutant Figure C.4-C.5).

### 7.3.2. Modeling the $\epsilon 48$ binding pose of C99

Besides the dominant  $\epsilon 49$  cleavage,  $\gamma$ -secretase can cleave C99 alternatively also at the amide bond between T48 and L49, generating a C-terminal peptide (AICD 51-99) and N-terminal  $A\beta 48$ . According to the processive cleavage process of  $\gamma$ -secretase (Figure 2.6),  $A\beta 48$ . The  $\epsilon 48$  cleavage is the first cleavage of the production line producing the plaque-forming  $A\beta 42$ . Hence, it is both of biological and pathological interest to distinguish the structural difference between the  $\epsilon 49$  and  $\epsilon 48$  binding pose. We modeled two C99- $\gamma$ -secretase putative binding models at the  $\epsilon 48$  cleavage using comparative modeling with two different sequence alignment strategies (Figure 7.2A) The first model is generated by shifting the whole sequence by one residue toward the C-terminal side (C99 <sub>$\epsilon 48-1$</sub> ), whereas the second model keeps V44 and I45 at the same position and exposes the scissile bond by inserting an alignment gap between I45 and V46 (C99 <sub>$\epsilon 48-2$</sub> ).

During the MD simulations of both reconstructed models, the catalytic hydrogen bond between the protonated aspartate and the carbonyl group of C99 T48 was found in over 90% of the trajectory frames, except for the first few nanoseconds while the systems were still considered undergoing the equilibration (Figure 7.2B). In addition, the  $\beta$ -strand conformation

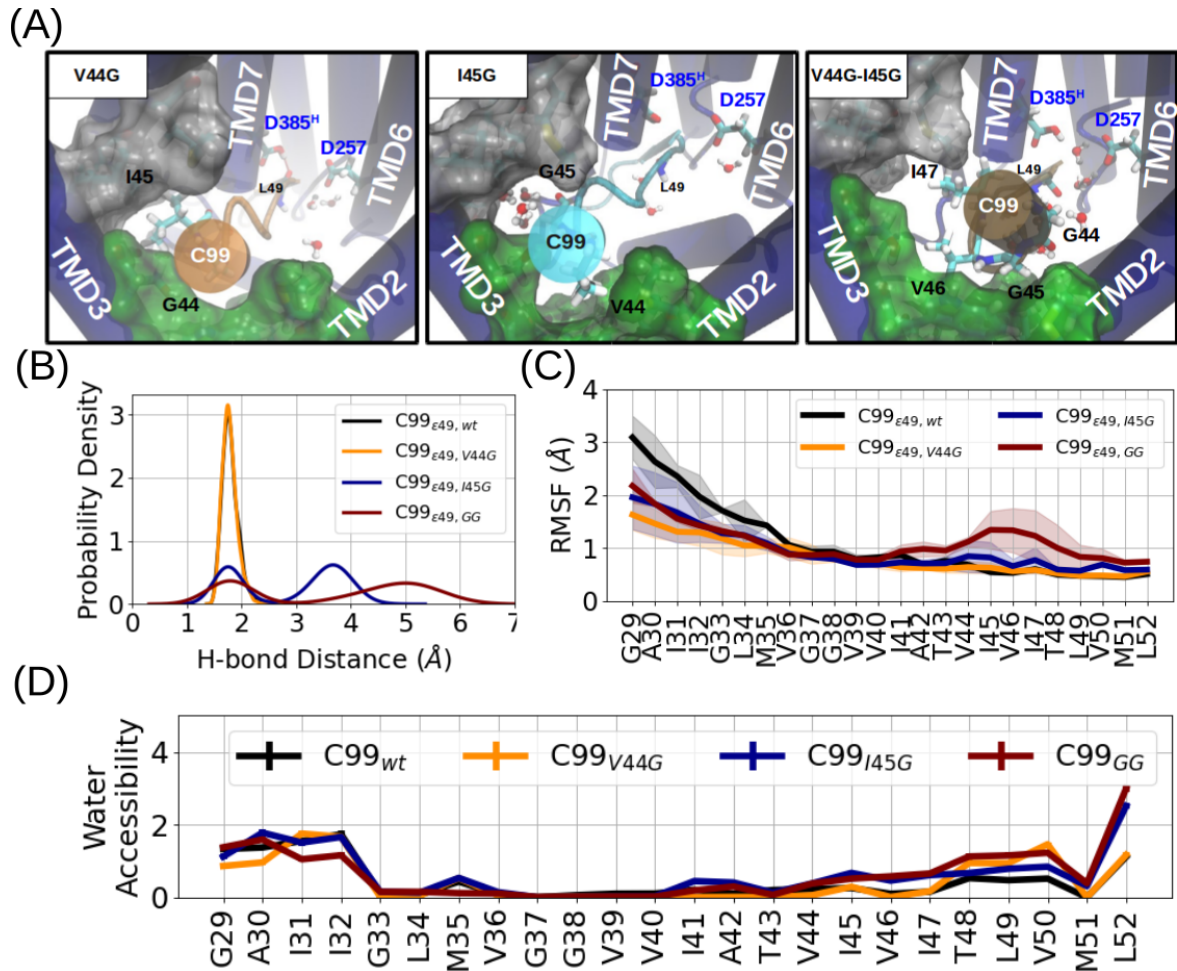


Figure 7.1.: Glycine mutations at the internal docking site disturb the E-S interaction. (A) Top-view at the PS1 internal docking site of C99<sub>V44G</sub> (left), C99<sub>I45G</sub> (middle), and C99<sub>GG</sub> bound γ-secretase complexes (representation same as in Figure 2C). (B) Probability density of the catalytic hydrogen bond distance during simulations. (C) RMSF of the substrate TMD and (D) residue-wise water accessibility of C99<sub>wt</sub> (black), C99<sub>V44G</sub> (orange), C99<sub>I45G</sub> (blue), and C99<sub>GG</sub> (brown) during simulations in complex with γ-secretase.



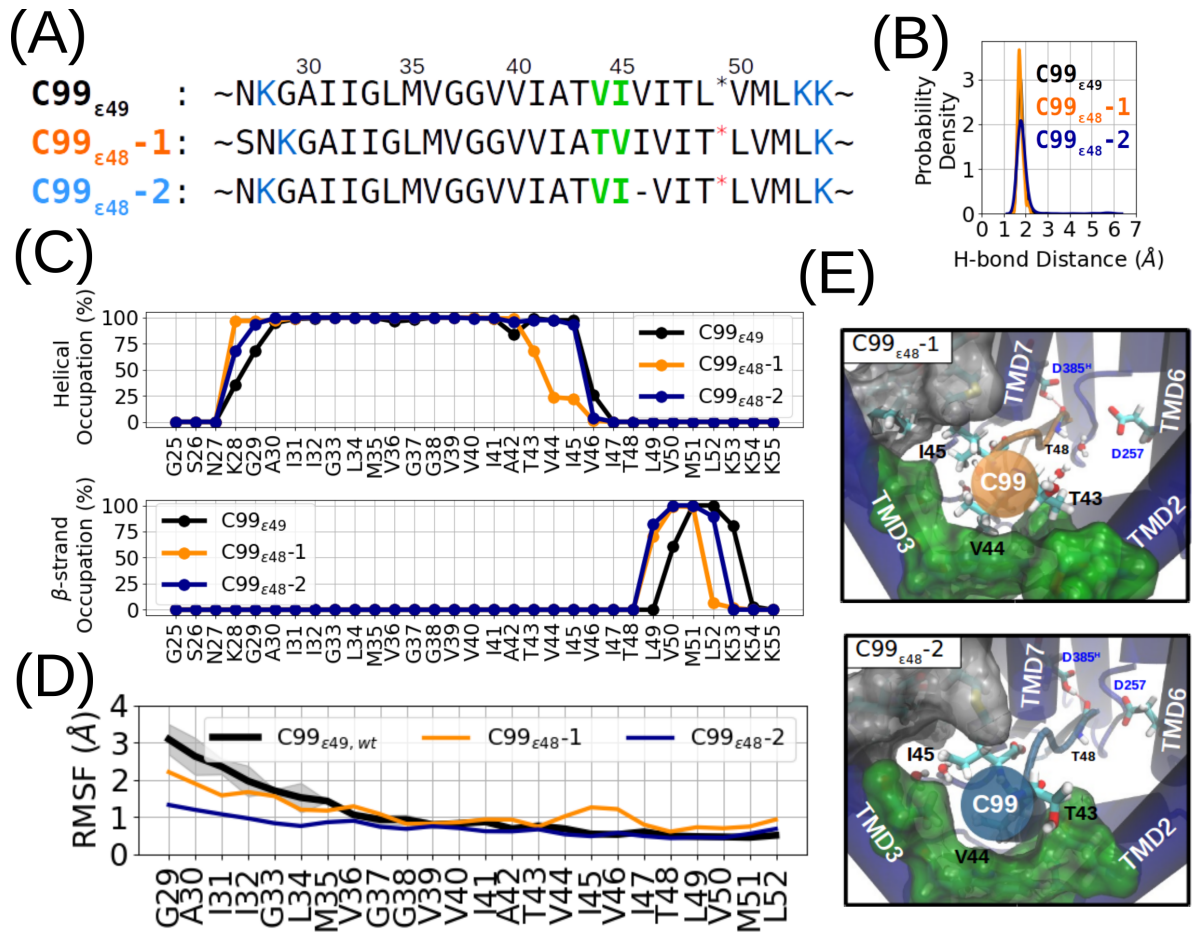


Figure 7.2.: MD-simulations of  $\gamma$ -secretase in complex with C99 shifted for cleavage at  $\epsilon 48$  site. (A) Sequence alignments of the two models termed C99<sub>ε48-1</sub> (orange), and C99<sub>ε48-2</sub> (dark blue) relative to the residue positions in C99<sub>ε49</sub> (black). Membrane anchoring residues K28, K53, and K54 are marked in blue, and residues residing in the hydrophobic pockets are marked in green. (B) Sampled probability density of the catalytic hydrogen bond distance in the MD simulation. (C) Secondary structure analysis of simulation on C99<sub>ε49</sub> (black), C99<sub>ε48-1</sub> (orange), C99<sub>ε48-2</sub> (dark blue) bound to  $\gamma$ -secretase complexes. (D) RMSF of the substrates. (E) View (from the extracellular side) into the PS1 internal docking site of the two models C99<sub>ε48-1</sub> (upper) and C99<sub>ε48-2</sub> (lower). (representation same as in Figure 2C)

is preserved at position P2' and P3' (Figure 7.2C). Despite the formed catalytic hydrogen bond, larger fluctuations were observed around substrate V44-V46 (Figure 7.2D), and a rotational substrate motion at the internal docking site was observed in the first reconstructed complex C99<sub>ε48-1</sub>. With the T43 rotated out of its starting position, the internal docking site was again filled up with V44 and I45 after  $\sim 100$ ns (Figure C.6 and Video S1 from our published work). As shown in the evolution of the secondary structure, the rotation was facilitated by the unwinding of the helix between T43-I45 (Figure 7.2C, Figure C.2). In contrast, V44 and I45 stayed firmly in the PS1 internal docking site in the second reconstructed complex C99<sub>ε48-2</sub>. Consistent with the C99<sub>ε49</sub> complex, residues T43-I45 exhibit a  $3_{10}$ -helix while A42 being slightly disordered (Figure 7.2C, Figure C.1). Notably, while V46 stays mostly in the 'Turn' and 'Bend' conformation, calculated by the DSSP method, in the C99<sub>ε49</sub> complex (Figure 4.2A), it is classified into the 'None' state in the C99<sub>ε48</sub> complex, where the peptide bond is stretched so that the carbonyl group of T48 can be exposed to the catalytic aspartates (Figure C.1).

Importantly, our simulation result suggests that the V44 and I45 are recognized by the PS1 internal docking site and sterically allow both a cleavage at  $\epsilon 49$  or  $\epsilon 48$  position (a complete shift of the entire substrate sequence is sterically not required!). Even in the model with T43 intentionally placed in the P6 pocket of the internal docking site resulted in a rapid readjustment with V44 and I45 bound to the P5, P6 binding sites (Figure 7.2E). Hence, which  $\epsilon$ -cleavage is executed depends mainly on the conformation of the residues of V46-T48. This explains why mutations on residues remotely N-terminal to V44 such as A21-D23 and K28-A30 do not alter the  $\epsilon$ -cleavage preference. However, the docking pose present in the study might be altered when residues around V44 and I45 are mutated. Indeed, some APP mutations around this region such as T43I, V44A, and V44F have been shown to decrease the  $\epsilon 49/\epsilon 48$  cleavage propensity, while other mutations such as I45F increase it.

### 7.3.3. Local hydrophobic mismatch induced by K28 in A $\beta$ n- $\gamma$ -secretase binding complexes.

The substrate processing by PS1 in  $\gamma$ -secretase leads to a consecutive cleavage typically along the A $\beta$ 49-A $\beta$ 46-A $\beta$ 43-A $\beta$ 40-A $\beta$ 37 production line. It leads to the placement of substrate residues of different sizes into the internal docking site (see Table 7.1). In addition, the stepwise cleavage requires the movement of the substrate toward the interior of the enzyme and toward the cytosolic side of the cell. The N-terminal side of the C99 substrate includes a K28 residue that is considered a membrane-anchoring residue for the substrate[211]. K28 was also shown to interact with residue 241 on the NCT domain of  $\gamma$ -secretase[210]. Mutation of K28 to a negatively charged residue or neutral residue enhances the processivity of cleavage leading to a relative increase of shorter secreted A $\beta$  products[210, 211, 147, 101]. It is of interest to investigate the dynamics of the intermediate states that have to form during substrate processing and how binding to the internal docking site correlates with the movement of the K28 anchor residue. We consecutively constructed A $\beta$ n- $\gamma$ -secretase complex models using the known full substrate complex as a template for shifting the substrate sequence (See Method and Schematics in Figure 7.3A). Each product is named by its sequence length and

cleavage site. For instance,  $A\beta_{49}_{\zeta_{46}}$  stands for the docking pose of  $A\beta_{49}$  with V46 placed at the catalytic active site and ready for the  $\zeta_{46}$  cleavage (Figure C.7).

In case of the simulations of the complexes with bound  $A\beta_{49}_{\zeta_{46}}$ ,  $A\beta_{46}_{\gamma_{43}}$ ,  $A\beta_{43}_{\gamma_{40}}$  models the corresponding P5 and P6 residues (Table 7.1) all remained at the initial placement in the PS1 internal docking site (Figure C.7). However, in the case of the model for the last cleavage pose  $A\beta_{40}_{\gamma_{37}}$  the residues P5 and P6 dissociated from the internal docking site during the simulations (Figure 7.3B and Video S3 from our published work). It is likely that besides the stabilization interaction at the PS1 internal docking site, the K28 anchor plays a role during this dissociation process. To study the influence of the membrane-anchoring effect of K28, we analyzed the dynamics of the lipid molecules as well as the  $A\beta$ - $\gamma$ -secretase complex at the juxtamembrane region. The electron density of a lipid bilayer shows typically two peaks at around  $z=+18.5\text{\AA}$  and  $z=-18.5\text{\AA}$ , representing the choline groups of POPC of the two membrane leaves. The K28 and the P6 residues are not always staying at the same position on the  $z$ -axis relative to the center of the membrane for the different  $A\beta$ -bound complexes (Figure 7.3C). In the case of the  $A\beta_{49}_{\zeta_{46}}$ ,  $A\beta_{46}_{\gamma_{43}}$ ,  $A\beta_{43}_{\gamma_{40}}$  models, the K28 moved deeper into the membrane as the substrate proceeds towards the shorter cleavage poses (Figure 7.3A, with residues at the internal docking site staying in place). Only in the case of the  $A\beta_{40}_{\gamma_{37}}$  complex residue P6 moved towards the exterior membrane side (Figure 7.3B). The time plot along the simulations shows that K28 of  $A\beta_{40}$  quickly approached the membrane upper leaf within the first 20ns followed by the “pulling-up” motion of P6 (Figure C.8). K28, carrying a positive charge, interacts with the negatively charged lipid phosphate groups and alters the peripheral membrane thickness (Figure 7.3A). With the strongest deformation close to K28, the membrane adapts slowly to its average thickness and reaches its equilibrium at the radial distance of  $22.5\text{\AA}$  away from K28 on the  $xy$  plane, denoted as  $\rho_{xy,K28}$  (Figure 7.3A,D). According to the mattress model, a hydrophobic mismatch in an uneven membrane may induce an elastic deformation-free energy against the membrane deformation[215]. We estimated the mismatch amplitude by fitting the curves with a simple function  $z_{POPC}=\alpha e^{-\beta\rho} \cos(\gamma\rho) + z(\rho=\infty)$ , with  $\alpha$  the mismatch amplitude,  $\beta$  the radial decaying rate,  $\gamma$  the membrane curvature, and  $z(\rho=\infty)$  the equilibrium height of the membrane upper leaf (Figure C.9). In the first two cleavages, namely  $C99_{\epsilon_{49}}$  and  $A\beta_{49}_{\zeta_{46}}$ , K28 stayed above the upper membrane leaf, inducing a positive mismatch which promotes the substrate to move downward towards the next binding pose ( $\alpha \sim +3\text{\AA}$ , Figure 7.3A and Figure 7.4A). When proceeding to the  $A\beta_{46}_{\gamma_{43}}$  binding pose, K28 stayed at a similar height as the membrane extracellular surface ( $\alpha \sim 0\text{\AA}$ , Figure 7.4B). However, K28 no longer matches the average membrane height and creates a negative mismatch in the  $A\beta_{43}_{\gamma_{40}}$  and  $A\beta_{40}_{\gamma_{37}}$  docking poses, hence, the substrates, in this case, experience a membrane up-pulling force ( $\alpha \sim -4.75\text{\AA}$ , Figure 7.3A, Figure 7.4C).

On the C-terminal side of the TMD helix are P5 and P6 residues sitting in the PS1 internal docking site (Figure C.7). Residue-wise binding energy analysis shows that the strongest binding affinity from the substrate helical domain is contributed by these two residues and their binding strengths correlate with their size and hydrophobicity (Figure C.10A, S21C). Notably, when a small residue is placed at position P5, such as A42 from the  $A\beta_{49}_{\zeta_{46}}$  binding pose or G33 from the  $A\beta_{40}_{\gamma_{37}}$  binding pose, larger fluctuations were observed in the substrate

TMD, in particularly around P5 (Figure C.11A). In contrast, when a small residue is placed at P6, such as G38 from the A $\beta$ 46 $_{\gamma 43}$  binding pose, only a mild increase in the fluctuation of substrate TMD was observed (Figure C.11A). We note that the different effects upon placing a small residue at substrate P5 or P6 display a consistent behavior as observed in the simulations with C99 mutants V44G and/or I45G (Figure 7.1C). With the negative hydrophobic mismatch inducing an up-pulling force at the juxtamembrane region and the weak helical binding force at the PS1 internal docking site, we hypothesize that the P6 movement observed for A $\beta$ 40 $_{\gamma 37}$  is due to an imbalance between the two competing forces. For example, in the case of A $\beta$ 43 $_{\gamma 40}$  residue M35 and V36 placed in the internal docking site contribute calculated binding energy of  $-6.5 \pm 0.7$  kcal/mol, whereas residues G33 and I32 placed in the internal docking site for A $\beta$ 40 $_{\gamma 37}$  contribute only  $-2.8 \pm 0.4$  kcal/mol (Figure C.10C). In the latter case, this small interaction energy cannot outbalance the hydrophobic mismatch and the residues moved out of the pocket even though they experience a similar membrane mismatch amplitude (Figure 7.3B). On the other hand, although G38 and V39 of A $\beta$ 46 $_{\gamma 43}$  contribute a relatively weak binding energy at the internal docking site ( $-4.05 \pm 0.4$  kcal/mol, Figure C.10C), the absence of the hydrophobic mismatch makes A $\beta$ 46 stays steadily with the enzyme. (Figure 7.3D and Video S3 from our published work) The gradual movement of the K28 during processing has also an effect on the helicity and water accessibility of the substrate TMD. With the local POPC molecules adjusting their positions to match K28, the local water-membrane interface moved towards the intracellular side and residues around K28 are thus surrounded by water molecules (Figure S22B). The water molecules weaken the intramolecular hydrogen bonds required in helix formation and the helical part of the substrate becomes shorter and shorter along the processive cleavage with residues at the N-terminal side of K28 turning into disordered conformations (Figure S22B-C, time evolution of the secondary structures are shown in Figure C.12). Notably, substrate P2-P5 are more exposed to water molecules in the A $\beta$ 40 $_{\gamma 37}$  docking pose than other A $\beta$  docking poses, agreeing with the C99 $_{I45G}$  simulation where a glycine was also present at the P5 position (Figure C.11B). The catalytic hydrogen bonds are stable in the first two cleavage poses, while they are seldom formed in the later three complexes, although still not totally vanished (Figure 7.3E). As a result of the local hydrophobic mismatch, P5 P6 binding energies and trend of catalytic hydrogen bond were similarly reproduced in the D257-protonated PS1 complexes, except that the substrate was pulled out from the internal docking site in both the A $\beta$ 43 $_{\gamma 40}$  and A $\beta$ 40 $_{\gamma 37}$  docking poses (Figure C.7, Figure C.13-C.16).

#### 7.3.4. APP mutations K28A and G33I stabilize substrate binding and catalytic geometry of A $\beta$ 40

In the study of putative A $\beta$ n- $\gamma$ -secretase complexes within the lipid bilayer, we observed a pair of competing forces that characterizes the decisive factors of  $\gamma$ -secretase holding or releasing the substrate, the membrane up-pulling force toward the extracellular membrane surface and the E-S stabilization force at the PS1 internal docking site. This model provides a straightforward way to explain why certain APP mutations on K28[211, 212] and G33[147, 213] generate shorter A $\beta$  products. It is due to reducing the membrane-anchoring feature

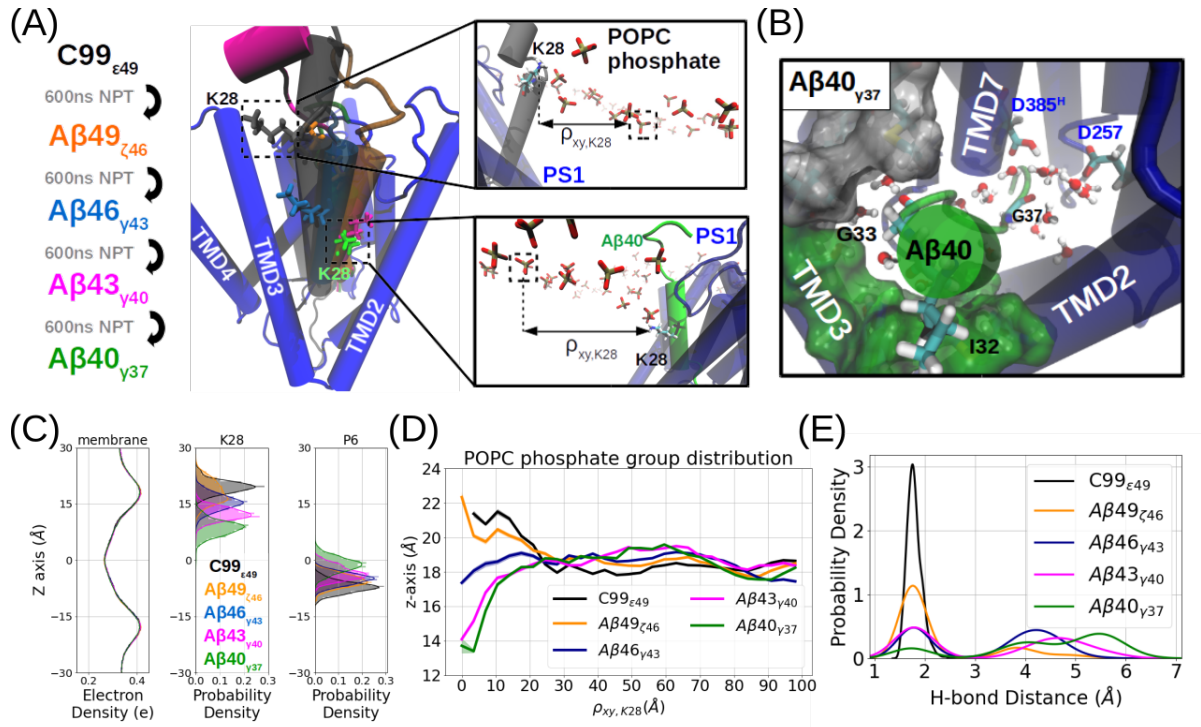


Figure 7.3.: Comparative modeling and simulations of Aβn-γ-secretase complexes. (A) Schematic views of the modeling workflow and the top views (from the extracellular domain) of the binding poses of C99<sub>ε49</sub> (black), Aβ49<sub>ζ46</sub> (orange), Aβ46<sub>γ43</sub> (blue), Aβ43<sub>γ40</sub> (magenta), and Aβ40<sub>γ37</sub> (green). Membrane anchoring residue K28 of each substrate is shown explicitly in the licorice representation. Zoom-in views of the C99<sub>ε49</sub> (upper) and Aβ40<sub>γ37</sub> (lower) show how the position of K28 influences the local distribution of the POPC phosphate groups with radial distance from K28 on the xy plane denoted by  $\rho_{xy,K28}$ . (B) View into the PS1 internal docking site in the Aβ40<sub>γ37</sub>-bound γ-secretase. (representation same as in Figure A.7) (C) Distribution of the membrane electron density (left), membrane-anchoring residue K28 (middle), and substrate P6 (right) along the z-axis in different Aβn-γ-secretase complexes. (D) The average z-axis of the POPC phosphate on the extracellular side is distributed along the radial distance  $\rho_{xy,K28}$ . (E) Probability density of the catalytic hydrogen bond distance.

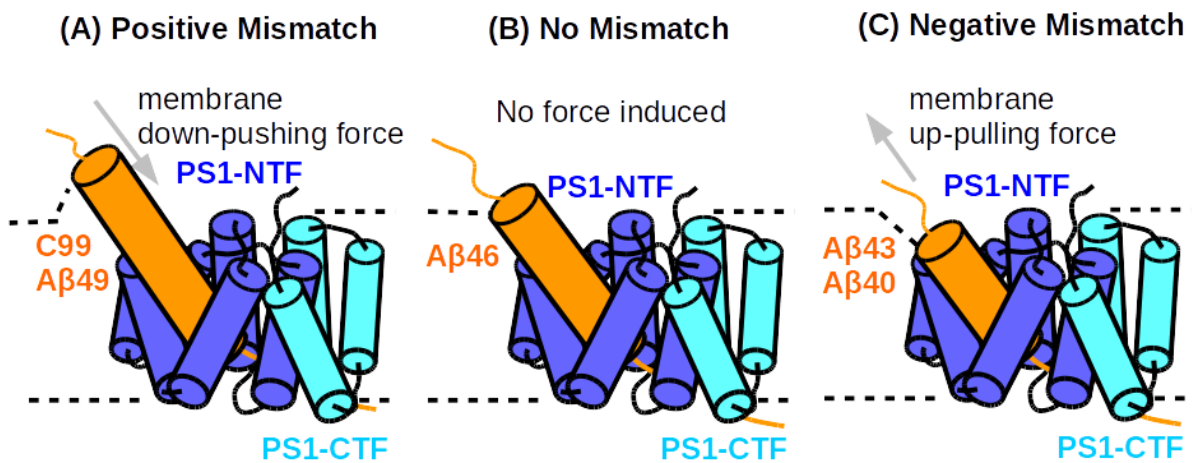


Figure 7.4.: Figure 7.4: Illustration of the membrane elastic force induced by local hydrophobic mismatch based on the mattress model. (A) Positive mismatch as observed in C99<sub>ε49</sub> and Aβ49<sub>ζ46</sub> substrate bound poses is induced when the substrate protrudes the membrane bilayer. The membrane elastic potential trying to restore the optimal membrane thickness causes a force pushing the substrate towards the PS1 active site (termed “down pushing force”). (B) No hydrophobic mismatch was observed in the Aβ46<sub>γ43</sub> binding pose (elastic potential is zero or close to zero). (C) Negative mismatch observed in the Aβ43<sub>γ40</sub> and Aβ40<sub>γ37</sub> substrate-bound cases is induced when the substrate tends to contract the membrane bilayer. The membrane elastic potential keeping the membrane thickness causes an “up-pulling” force on the substrate.

at position K28 and/or enhancing the E-S binding affinity at the internal docking site. To verify the hypothesis, simulations of A $\beta$ 40- $\gamma$ -secretase complexes with APP mutations K28A and G33I were conducted and analyzed. Both APP mutations G33I and K28A result in an increased fraction of sampled catalytic geometries and stable I32 placement in the PS1 internal docking site (Figure 7.5A-C and Video S4 from our published work). Notably, in the K28A mutant, water was observed accessing the pocket between TMD3-TMD5, similarly as was seen in the C99 I45G mutant at the  $\epsilon$ 49 cleavage pose (Figure 7.1A). G33I, in contrast, is expected to have a reverse influence compared to I45G in C99. Indeed, the introduced isoleucine with a larger hydrophobic side chain fills the pocket and prevents further water access (Figure 7.5B). In this case, side chains of P5 and P6 together contributed  $\Delta E_{I32} + \Delta E_{I33} = -8.2 \pm 0.6$  kcal/mol of E-S binding energy, around 5.4 kcal/mol stronger than estimated for the wild-type A $\beta$ 40 binding (Figure C.10B,C). Indeed, the more favorable binding energy at the PS1 internal docking site allowed in this case K28 to move deeper into the membrane and to maintain and balance a much more negative hydrophobic mismatch. ( $\alpha \sim -12.7\text{\AA}$ , Figure 7.5D, Figure C.9, Figure C.17). Mutation K28A, on the other hand, alleviates the difference between the local and global membrane thickness ( $\alpha \sim +1.02$ , Figure 7.5D, Figure C.9) and, thus, mitigates the elastic potential stored in the membrane layer. With the G33 and I32 remaining in contact with the PS1 internal docking site, the local binding energy was observed -2 kcal/mol more favorable than the wild-type A $\beta$ 40 mainly because of the enhanced interaction of I32 (Figure C.9B,C). Although less pronounced the same findings were also observed in the corresponding simulations with the D257-protonated  $\gamma$ -secretase complexes (Figure C.14-C.15 and Figure C.18-C.19).

## 7.4. Discussion and outlook

Intramembrane proteases (IMPs) are involved in several cell signaling pathways and a variety of diseases[216, 217, 218, 219, 220].  $\gamma$ -Secretase is an intramembrane aspartate protease cleaving its substrates within the lipid bilayer with the catalytic subunit presenilin. Once the APP substrate is bound to PS1  $\gamma$ -secretase, A $\beta$  peptides are sequentially produced and can dissociate. Longer products such as A $\beta$ 42 and A $\beta$ 43 are prone to aggregate into AD-related fibrils. Several FAD mutations located on PS1 and APP are found to decrease  $\gamma$ -secretase activity and promote the secretion of longer A $\beta$  peptides. We have previously suggested that PS1 mutations located on TMD6, TMD7, and TMD9 can directly influence the ability of the enzyme to recruit the scissile bond to the catalytic center by free-energy calculation[78]. Other groups have reported that mutations on PS1 and APP are consistently destabilizing the E-S complexes and, thus, tend to release longer A $\beta$  peptides rather than bringing them to the next cleavage site[221, 76, 222]. Since A $\beta$ 42 and A $\beta$ 43 are respectively generated from the A $\beta$ 48-A $\beta$ 45-A $\beta$ 42-A $\beta$ 38 and A $\beta$ 49-A $\beta$ 46-A $\beta$ 43-A $\beta$ 40-A $\beta$ 37 productions lines, understanding how  $\gamma$ -secretase decides the first cleavage site and releases A $\beta$  peptide at the certain step are of great biological and therapeutic interest. In the present study, we investigate the role of an internal docking site PS1 that shows a well-packed interaction with substrate residues P5 and P6 shared by both C99 and Notch1 (Table 1). Our simulations indicate that this

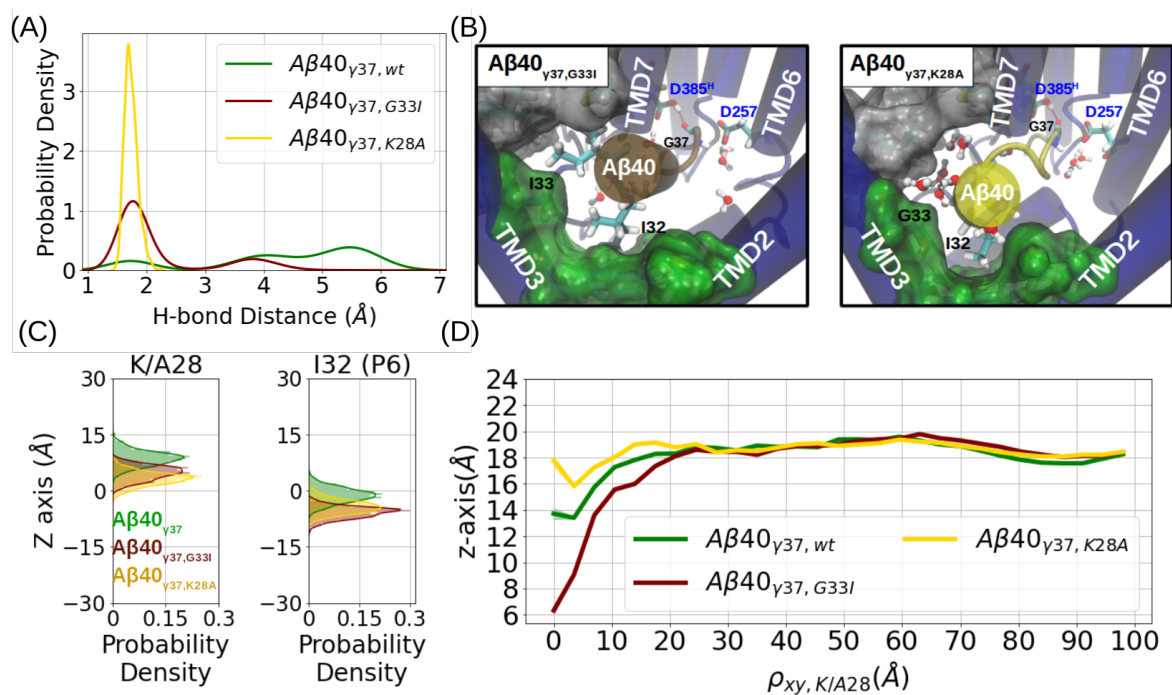


Figure 7.5.: Influence of the APP mutations G33I and K28A on the  $A\beta_{40}\gamma_{37}$  binding pose. (A) Probability density of the catalytic hydrogen bond distance. (B) Top-view at the PS1 internal docking site of  $A\beta_{40}\gamma_{37}$ , G33I- (left) and  $A\beta_{40}\gamma_{37}$ , K28A- (right) bound  $\gamma$ -secretase complexes. Atomic representation is described similarly as in Figure A.7. (C) Distribution of the membrane-anchoring residue K28 (middle) and substrate P6 (right) along the z-axis. (D) The average z-axis of the POPC phosphate on the extracellular side is distributed along the radial distance  $\rho_{xy, K/A28}$ .



region interacts strongly with the C-terminus of the substrate helical segment. In support of the critical role of the internal docking site, cross-linking experiments with the substituted cysteine accessibility method (SCAM) demonstrated that the pocket undergoes conformational changes in FAD-linked PS1 mutations and upon the binding of GSM-1. Furthermore, the photo-affinity labeling experiment showed that V44 is the strongest binding residue of C99. Compared to the substrate-free  $\gamma$ -secretase, we find that the water access to this pocket from the intracellular side is blocked upon substrate binding. During the simulations, the substrate P5-P7 residues situated at the C-terminus of the helical domain frequently adopted a 3-10 helix geometry, which well aligns the upstream  $i+3$  residues to the same interface. Hence, only a translational movement toward the next  $i+3$  cleavage pose is required during substrate processing (and no substrate rotation). The simulations on readjusted APP substrates support the view that both  $\epsilon 49$  and  $\epsilon 48$  cleavages are possible with the placement of V44 and I45 in the PS1 internal docking site (the cleavage then depends on the conformation between V46 to V50). During the modeling of the docking poses following the initial cleavage, we identified the gradual change in membrane thickness centered at the APP membrane-anchoring residue K28 to oppose the substrate movement towards the PS1 active site. K28 interacts with the membrane but in addition also with other residues in the PS1 and nicastrin subunits[210]. While the docking pose of  $A\beta_{46\gamma 43}$  induced the least membrane hydrophobic mismatch in POPC membrane, the positive mismatch was observed in the  $C99_{\epsilon 49}$  and  $A\beta_{49\zeta 46}$  complexes, and negative mismatches were observed in  $A\beta_{43\gamma 40}$ - and  $A\beta_{40\gamma 37}$ -bound complexes. The positive hydrophobic mismatch is beneficial for substrate processing because it causes a force to promote further substrate processing, whereas a negative mismatch promotes the release of the substrate to the extracellular space. Unlike  $A\beta_{43\gamma 40}$  having large hydrophobic residues M35 and V36 located in the PS1 internal docking site,  $A\beta_{40\gamma 37}$  with I32 and G33 in the internal docking site is indeed dragged out of the enzyme during simulations and the catalytic geometry at the active site was hardly formed. This indicates that the E-S interaction at the PS1 internal docking site on the one hand and the membrane up-pulling force caused by negative hydrophobic mismatch, on the other hand, are two critical factors determining whether the substrate can proceed to the next cleavage position. Our tug-of-war model between the membrane and  $\gamma$ -secretase provides a useful explanation of why many mutations around K28 and G33 have an impact on C99 processivity. Indeed, mutations G29K and A30K, which both interact more strongly with the membrane phosphate can induce an even stronger hydrophobic mismatch around K28 and have been shown to decrease the processivity[212]. On the contrary, K28E, which repels the membrane phosphate groups might further push the substrate to the next cleavage steps. Indeed, this mutation and K28A, K28Q and K28L, which all alleviate the K28-lipid interaction, are found to produce even shorter peptides such as  $A\beta_{33}$  and  $A\beta_{34}$ [211, 212, 223]. Furthermore, mutating G33 to larger amino acids such as G33A and G33I, which in our model enhance the enzyme binding force, also promote the  $A\beta_{40} \rightarrow A\beta_{37}$  processing in experiments[147, 213]. Among these mutations, we simulated the two well-reported mutations K28A and G33I, which address the balance between the forces in two different ways. Both mutations indeed stabilized the E-S geometry such that the correct formation of a stable catalytic geometry at the active site became possible.

To sum up, our work provides a mechanistic explanation of how the balance between the interaction at the proposed PS1 internal docking site and a negative hydrophobic mismatch at the substrate juxtamembrane domain plays an important role in the processivity of A $\beta$ . Furthermore, our simulations suggest that the PS1 internal docking site, spatially distinct from the catalytic center (but forming a FAD hot spot), is an essential element for substrate positioning and stabilization. The local hydrophobic mismatch induced by APP K28, on the other hand, imposes a force on the substrate against the substrate binding. Based on this model, we successfully demonstrated how A $\beta$  species can be modulated by enhancing the binding affinity at the PS1 internal docking site and weakening the hydrophobic mismatch. Our model paves a way for not only explaining how mutations on APP or PS1 alter the  $\epsilon 49/\epsilon 48$  or A $\beta 43/A\beta 40$  ratio but also provides a potential target site for the development of small molecules that promote the secretion of shorter A $\beta$  peptides.

## 8. A Putative Binding Mode of $\gamma$ -Secretase Modulator at the $A\beta$ - $\gamma$ -Secretase Interface Attenuates Membrane Negative Hydrophobic Mismatch

The intramembrane protease  $\gamma$ -secretase has been a drug target to cure Alzheimer's disease by inhibiting the secretion of plaque-forming Amyloid- $\beta$  ( $A\beta$ ) peptides. However, inhibitors that impair the cleavage activity of  $\gamma$ -secretase suffer from severe side effects because of their off-targeting nature. As an alternative,  $\gamma$ -secretase modulators (GSM) are designed to reduce the generation of toxic peptides by enhancing the cleavage processivity without impairing the enzyme activity. We performed MD simulations with an experimentally determined GSM binding mode with and without the presence of  $A\beta$ 43 peptide to study the interactions between enzyme, substrate, GSM, and lipid. Our result suggests that the binding of imidazole-based GSM E2012 can attenuate the membrane distortion caused by processive cleavage of  $\gamma$ -secretase through helix unwinding and direct interaction with the  $A\beta$  substrate. The critical interactions contributed by the fluorophenyl and 2-piperidine moieties explain the functions of the preserved hydrogen acceptor and the aromatic ring preserved also in many other imidazole-based GSMs.

### 8.1. Introduction

Alzheimer's disease (AD) is a chronic neurodegenerative disease characterized by memory loss, poor judgment, and impaired daily-life activity[224]. Based on the statistics from the world health organization (WHO), there are currently more than 55 million people living with dementia of which 60-70% of the cases are in the form of AD[225]. According to the amyloid hypothesis, the overproduced  $A\beta$  peptides in the brain initiate a cascade of events that leads to the formation of senile plaques and causes neural inflammation and oxidation[3, 4].  $A\beta$  peptides are generated upon consecutive cleavages of amyloid precursor protein (APP) by the soluble protease  $\beta$ -secretase (BACE1) and intramembrane protease  $\gamma$ -secretase[5, 6, 7, 8]. The C-terminal product of APP upon the initial cleavage of BACE1 contains 99 residues, termed C99, and remains in the membrane bilayer. C99 is later recognized by  $\gamma$ -secretase inside the membrane environment and cleaved into the C-terminal product amyloid intracellular domain (AICD) and the N-terminal product  $A\beta$  peptides inside the catalytic subunit presenilin (PS)[226, 197, 199, 200]. Importantly,  $\gamma$ -secretase cleaves C99 in a successive way, following either the  $A\beta$ 49 $\rightarrow$  $A\beta$ 46 $\rightarrow$ A43 $\rightarrow$  $A\beta$ 40 $\rightarrow$  $A\beta$ 37 or  $A\beta$ 48 $\rightarrow$  $A\beta$ 45 $\rightarrow$ A42 $\rightarrow$  $A\beta$ 38

production line[66, 67, 69, 226]. However, the  $\gamma$ -secretase does not always process C99 until the end of each production line and A $\beta$  peptide has the probability to escape from the enzyme during the successive cleavage process, leading to a secretion of A $\beta$  peptides in different lengths[226, 9]. The ability of  $\gamma$ -secretase to produce the shorter peptide, called processivity, is usually characterized by the A $\beta$ 37/A $\beta$ 40 or A $\beta$ 42/A $\beta$ 38 ratio. Among all A $\beta$  products, the A $\beta$ 40 is the major species, followed by A $\beta$ 43, A $\beta$ 42, A $\beta$ 38, and A $\beta$ 37[9]. Due to the hydrophobic nature of the longer species, A $\beta$ 42 and A $\beta$ 43 are more prone to form aggregations than the shorter peptides[3, 10, 11, 12]. Although inhibitors against BACE1 and  $\gamma$ -secretase were developed to lower the level of longer A $\beta$  peptides, none of these have been FDA-approved for clinical use in AD treatment[20, 21, 82, 83, 84, 79]. In particular, non-selective  $\gamma$ -secretase inhibitors (GSI) such as Semagacestat (LY450,139) can increase the risk for skin cancer because of the inhibition of Notch1 signaling[227]. To alleviate the off-targeting effects of GSI,  $\gamma$ -secretase modulators (GSM) are proposed as an alternative to increasing the processivity of  $\gamma$ -secretase without changing its total activity[25, 26].

Despite its potential therapeutic application, the underlying mechanism of how GSM modulates the processivity of  $\gamma$ -secretase against C99 and A $\beta$  peptides remains unclear. Several attempts have been reached to understand the binding sites and the conformational changes of PS upon the binding of GSMs[99, 102, 93, 98, 229, 230, 231, 232, 233, 228]. Also, computational approaches have been implemented to understand the structure-activity relationship (SAR) to the measured effective A $\beta$ 42 inhibition potency ( $EC_{50}$ )[160]. The exact binding mode of an imidazole-based GSM E2012 was confirmed only recently by cryo-EM technique together with transition state analog (TSA) GSI L458[58]. The structure of E2012 revealed an allosteric site located at the extracellular surface of PS1 and indicated a hydrogen bond with PS1 Y106, which was found to be critical for the binding of E2012 by several studies[161, 98, 233, 234]. Nonetheless, the location of E2012 determined by the cryo-EM structure partially overlaps with the APP-bound structure[39], indicating a possibility that APP may undergo a conformational change when coexisting with GSM. A recent work combining comprehensive experimental and computational work suggests another binding mode for a similar imidazole-based GSM (GSM III)[161] slightly differently from the cryo-EM resolved position but similar to what was proposed for another imidazole-based GSM (ST2038)[98]. Despite different binding modes of GSM being proposed, it is suggested that GSMs conduct their mode of action through active site allosteric modulation[102] and enhancing the enzyme-substrate (E-S) interaction[161].

It has been suggested that GSMs such as GSM-1 and E2012 mainly affect  $\gamma$ -secretase proteolysis at the later stage along the processive cleavage process and have little impact on initial cleavage[192], it is therefore of more mechanistic interest how GSM binds to the E-S complex when C99 is trimmed into the shorter A $\beta$  peptides. Our previous computational work suggests that the release of A $\beta$  peptides can be facilitated by a local hydrophobic mismatch induced by the C99 anchoring residues K28 at the  $\gamma$ 40 and  $\gamma$ 37 cleavages[146]. Interestingly, while C99 mutant K28R shows a positive modulation response against E2012, an inverse modulation, e.g. an increase in longer A $\beta$  species was observed in the E2012 treatment toward K28E mutant, implying an ionic interaction between K28 and GSM[101].

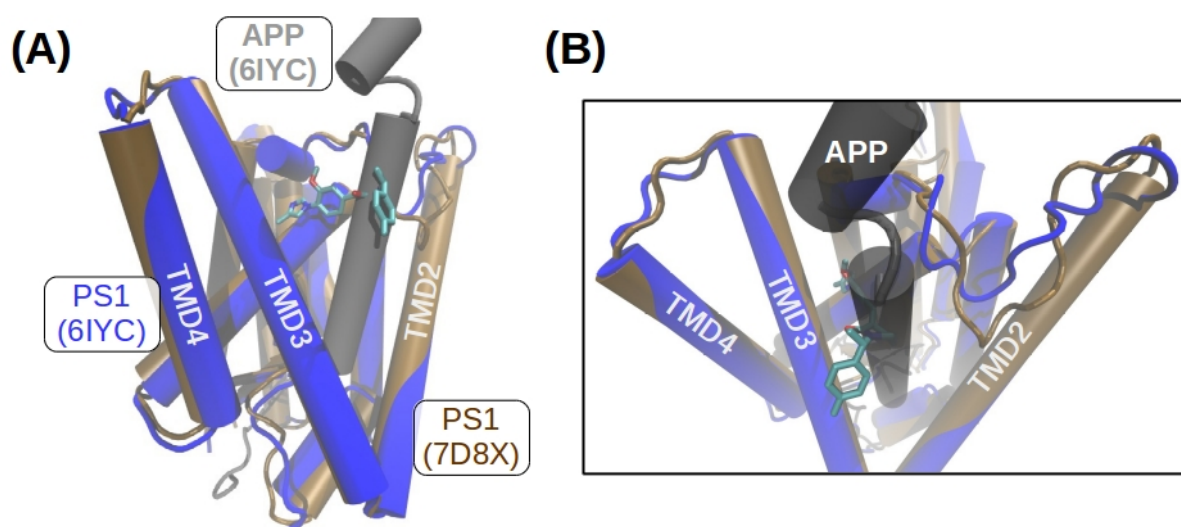


Figure 8.1.: Structural overlap between the cryo-EM structures of the E2012-bound (PDBID 7D8X, cyan E2012 and brown PS1) and APP-bound (PDBID 6IYC, gray APP and blue PS1)  $\gamma$ -secretase. (A) Side view of the structural overlap. (B) Top view of the structural overlap.

By comparing the size of the E2012 binding pocket determined in the cryo-EM structure between the binding modes of C99 and different A $\beta$  substrates from our previous work[146], we suggest that this E2012 can bind into the A $\beta$ 49-bound, A $\beta$ 46-bound, and A $\beta$ 43-bound  $\gamma$ -secretase complex without deviating from its binding position. Consistent with our previous observation, our study using molecular dynamics (MD) reveals that the A $\beta$ 43- $\gamma$ -secretase can induce a negative hydrophobic mismatch at the lipid-substrate interface that was suggested to facilitate substrate release[146]. We identified a putative binding mode of E2012 that attenuates the negative hydrophobic mismatch by two mechanisms. First, the binding of E2012 unwinds the helix of A $\beta$ 43 around residue K28, making K28 capable of forming ionic interaction with its upstream such as D23 or S26. Second, E2012 can also be directly from ionic interaction with K28 with the fluorophenyl and the 2-piperidine fragments. These two mechanisms of E2012 neutralize the electrostatic potential of K28 and are expected to reduce the membrane up-pulling force and extend the enzyme residence of short A $\beta$  peptides.

## 8.2. Methods

The pocket size distribution of each E-S complex is calculated based on the simulation trajectories generated in our previous work[146] through the following analysis processes. First, each trajectory was aligned to the E2012-bound cryo-EM structure (PDBID 7D8X[58]) by minimizing the RMSD between PS1-NTE, as depicted in Figure 8.1A, B, and Figure 8.2A-D). Then, an E2012 binding pocket was defined by the volume constructed by probes placed within 5 Å of E2012 by Fpocket[235] (Figure D.1). During the simulations, the volume of the binding

pocket in each simulation was calculated as the sum of the total available volume evaluated by MD pocket[236]. A volume of 346.192 Å<sup>3</sup> was calculated for E2012 at the conformation revealed by cryo-EM using Sanjeevini online server[237].

The E2012-bound  $\gamma$ -secretase structure was taken from the RCSB protein data bank PDBID 7D8X and neural capping groups were added at the C-terminus of PS1-NTF (after T291) and the N-terminus of PS1-CTF (before R377) by AmberTools22[**AMBER22**]. The A $\beta$ 43-bound complex from our previous homology work was used as a starting structure of the current study for a longer simulation time (1  $\mu$ s) and a different lipid force field lipid21[164] instead of lipid17). The A $\beta$ 43-E2012- $\gamma$ -secretase ternary structure was constructed by docking E2012 at the putative GSM binding site of the A $\beta$ 43-bound homology using Autodock vina[238]. The geometry of E2012 was optimized three a three-step optimization. The first 3D structure of E2012 was generated with Openbabel[239] and optimized using Avogadro[240] with the MMFF94 forcefield[241] until the energy difference  $dE$  is smaller than  $10^{-3}$  kJ/mol. Secondly, geometry optimization was carried out using the BP86 functional and def2-SVP, def2/J, and D3BJ basis sets under the resolution of identity (RI) approximation[242] with TIGHTSCF SCF convergence criteria in ORCA[243]. Lastly, the geometry is optimized with B3LYP functional and def2-TZVP basis sets with near SCF convergence criteria in ORCA[243]. The chosen docking pose adopts an orientation similar to the cryo-EM structure resolved E2012 and forms a hydrogen bond with PS1 Y106. Simulation boxes are prepared using CHARMM-GUI[179] and PPM online servers[187] and solvated with 500 POPC molecules, TIP3P water molecules[163], and 0.15 M KCl salt. All simulations were performed using the Amber20 package[111] and the ff14SB force field[163] for the proteins, and the lipid21 force field[164] for the membrane lipids. Charges of E2012 were calculated with Hatree-Fock single point calculation with 6-34+G(d,p) basis sets and AUTOAUX (Automatic Generation of Auxiliary Basis Sets [244]) under RIJDX approximation[245] using ORCA[243]. The calculated wave functions were analyzed by Multiwfn[246] to generate point charges on each atom. Finally, atom types and forcefield parameters were generated using GAFF2[103] and antechamber. The catalytic aspartic acid D385 is protonated and D257 is unprotonated based on a recent pKa calculation[145].

MD simulations were performed through an energy minimization, equilibration, and production run process. First, the energy of each simulation box was minimized with maximal 70,000 steps with  $10 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$  positional restraint on protein using the MPI version of the pmemd program in Amber20[111], followed by equilibration with gradually releasing positional restraint, from 10 to 0.1  $\text{kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$  on protein and 2.5 to 0  $\text{kcal} \cdot \text{mol}^{-1} \cdot \text{Å}^{-2}$  on the membrane, for 400 ps at 303.15 K using the cuda version of pmemd[**AMBER22**]. The equilibrated systems are submitted for 1 $\mu$ s production run at a temperature of 303.15 K using the Langevin thermostat[126] and a pressure of 1 bar by Berendsen barostat[125]. A time step of 4 fs was allowed with the SHAKE algorithm [123] hydrogen mass repartitioning[124]. A cut-off distance of 9 Å was chosen for the non-bonded interaction based on our previous calculation[146]. In total, 3 trajectories for the E2012-only system and 5 trajectories for A $\beta$ 43-bound and A $\beta$ 43-E2012-bound systems were generated. RMSD of PS1 along the simulation time is shown in Figure D.2 to ensure protein stability and the secondary structure

is calculated using the DSSP algorithm [165]. Hydrophobic mismatch amplitude is estimated by fitting the phosphate distribution in the z-axis along the radial distance from substrate K28 or E2012 in the E2012-bound complex, denoted as  $\rho_{xy,K28}$  or  $\rho_{xy,E2012}$ , with a function  $Z_{POPC} = \alpha e^{-\beta\rho} \cos(\gamma\rho) + Z_{POPC}(\rho = \infty)$ , with  $\alpha$  the mismatch amplitude,  $\beta$  the radial decaying rate, and  $\gamma$  the membrane curvature.  $Z_{POPC}(\rho = \infty)$  describes the equilibrium membrane height of the membrane upper leaf.

### 8.3. Results

#### ***In silico* pocket detection reveals GSM-compatible binding pockets in A $\beta$ -bound $\gamma$ -secretase complex**

The mode of action of a molecule can be usually explained by its binding pose toward its biological target. However, because of the substantial overlap between the cryo-EM structures of the APP-bound and the E2012-bound  $\gamma$ -secretase (Figure 8.1A, B)[58, 39], how E2012 binds to E-S complex remains under debate. Under the assumption that A $\beta$  peptides bind to  $\gamma$ -secretase in a similar manner as APP and Notch1, we have previously modeled the binding mode of A $\beta$  peptides at different stages along the  $\epsilon$ 49- $\zeta$ 46- $\gamma$ 43- $\gamma$ 40- $\gamma$ 37 production line[146]. Interestingly, no or little steric overlap was observed when aligning the E2012-bound structure with the constructed A $\beta$  homologies (Figure 8.2A-D).

Next, we measured the volume of the putative GSM binding pocket by placing probes 5 Å around the resolved E2012 structure (see Methods) to evaluate how well the GSM binding pocket can remain during the simulations. In the substrate-free simulations, the size of the pocket was found larger than the size of E2012 in 67% of the simulated configuration (Figure 8.2). However, this pocket was occupied when C99 or A $\beta$ 40 bound to  $\gamma$ -secretase and E2012 can only fit the putative binding pocket in 35% of the simulation (Figure 8.2). Interestingly, in the binding mode of A $\beta$ 49, A $\beta$ 46, and A $\beta$ 43 binding modes, more than 65% of the sampled conformations are E2012-compatible (Figure 8.2). In particular, the volume of the binding pocket in the A $\beta$ 46-bound and A $\beta$ 43-bound  $\gamma$ -secretase showed even more suitable for accommodating E2012 than in the substrate-free form. This indicates a probability that E2012 might not bind directly at the C99-bound form but rather in the later stage of the processive cleavage and be released after the  $\gamma$ 37 (A $\beta$ 40  $\rightarrow$  A $\beta$ 37) cleavage. This provides us an opportunity to investigate the dynamics of a potential substrate-GSM- $\gamma$ -secretase ternary complex inside the lipid bilayer and to understand how GSM improves the cleavage toward shorter A $\beta$  peptides.

#### **Binding of E2012 at PS1 is not stable without substrate or $\gamma$ -secretase inhibitor**

E2012 is an imidazole-based GSM containing four ring segments: an arylimidazol ring, a methoxyphenyl ring, a 2-piperidine ring, and a fluorophenyl ring. The first two rings are a conserved moiety preserved in many other GSMs and form contact with several residues of  $\gamma$ -secretase (Figure 8.3A). The arylimidazol group is located in the vicinity of PS1 Y106, which plays a decisive role in the binding of E2012 and other imidazole-based GSM[161, 98, 233, 234],

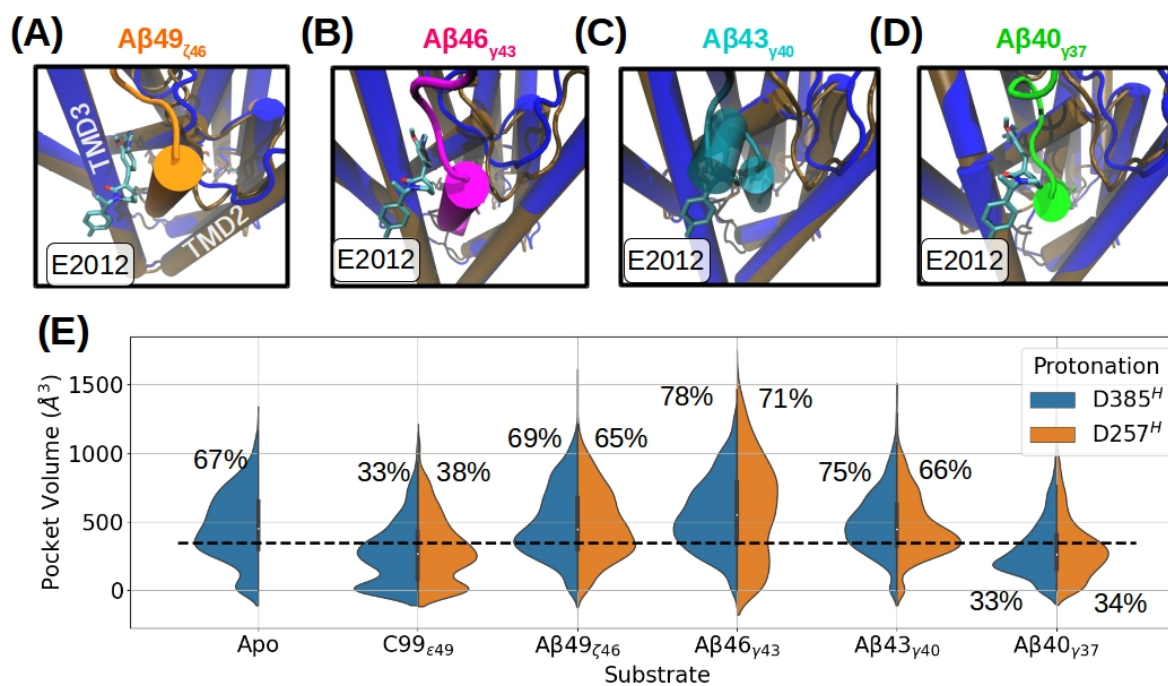


Figure 8.2.: Evaluation of the putative GSM binding pocket in the apo-form and substrate-bound  $\gamma$ -secretase simulations from our previous work. (A)-(D) Superposition of cryo-EM structure 7D8X (PS1 in brown helices and E2012 in licorice representation) and the constructed (A) A $\beta$ 49-bound, (B) A $\beta$ 46-bound, (C) A $\beta$ 43-bound, and (D) A $\beta$ 40-bound  $\gamma$ -secretase homologies from our previous work (PS1 shown in blue). Subscripts show the corresponding cleavage site for each substrate. (E) The volume of the putative binding pocket was measured during MD simulations from our previous work (see Methods and Material). The volume of E2012 (346.192 Å<sup>3</sup>) is indicated by the black dashed line.



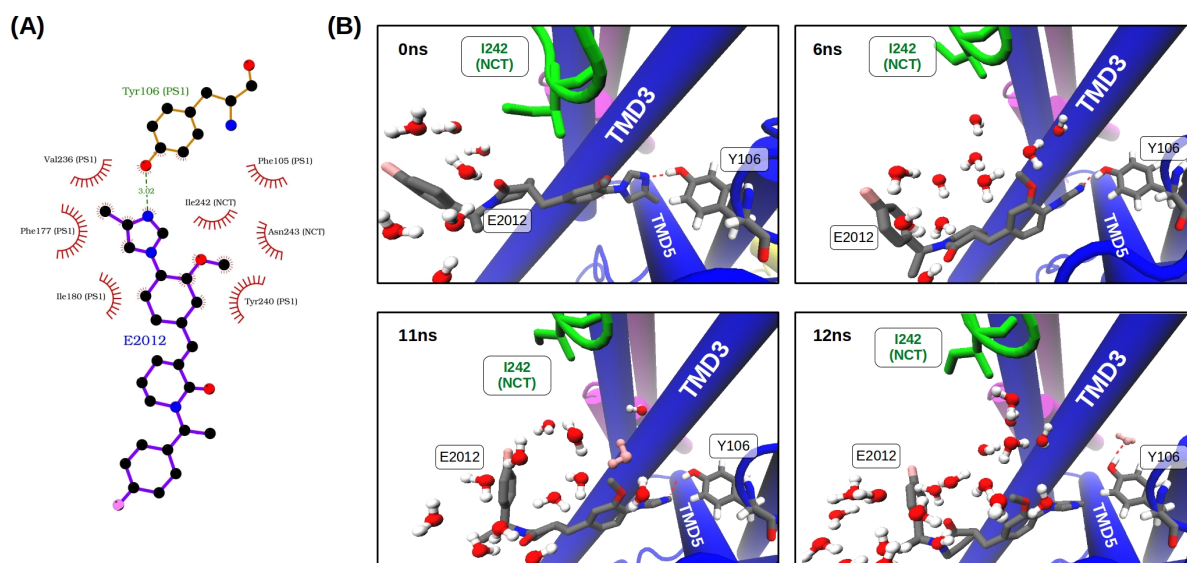


Figure 8.3.: Dynamics of E2012 in the  $\gamma$ -secretase bound state revealed by cryo-EM and MD simulations. (A) Interaction diagram of the cryo-EM resolved structure. The diagram is generated via LigPlus[247]. (B) Snapshots of the structure of E2012 surrounded by Y106, F177, and Y240 of PS1 in the MD simulation. A water molecule was recruited between TMD1, TMD3, and TMD6 of PS1. The hydrogen bond between PS1 Y106 and E2012 arylimidazole and the hydrogen bond between PS1 Y106 and the recruited water molecule is indicated by the dashed red line. Hydrogen atoms of E2012 are hidden for better visualization.

and the methoxyphenyl group might form  $\pi$ - $\pi$  interaction with F177 on TMD3 and Y240 on TMD5 of PS1, which was also reported to interact with E2012[161]. In contrast, the third and the fourth rings are usually modified in the development of other imidazole-based GSM and form only little contact with  $\gamma$ -secretase (Figure 8.3A). Notably, E2012 was only resolved with the presence of TSA GSI in the cryo-EM structure[58]. Furthermore, the photo-affinity labeling experiment has suggested that the binding of E2012 is improved by the coexistence of TSA GSI[98]. These findings together indicate an allosteric modulation between the active site and the GSM-binding site. To study the dynamics of E2012 without the modulation of TSA GSI, we performed three MD simulations starting with randomly assigned initial velocity distribution and studying its interaction with  $\gamma$ -secretase.

Surprisingly, I242 of NCT moved away from E2012 in the first few nanoseconds of all simulations (Figure D.3A). This leads to water leakage from the extracellular domain to the cavity between TMD1, TMD3, and TMD5 of PS1 and interferes with the hydrogen bond between E2012 and PS1 Y106 (Figure 8.3B and Figure D.3A). The hydrogen bond can be restored if E2012 stayed inside the binding pocket when the recruited water molecule left Y106. Nonetheless, E2012 was completely dissociated from  $\gamma$ -secretase in one of the simulations, indicated by a large root-mean-square-deviation value (Figure D.3C). The dissociation was initiated by the broken E2012-Y106 hydrogen bond mediated by a water molecule coming

from the extracellular side, leading to a slight deviation of Y106 (Figure D.4A, B). Without the hydrogen bond, E2012 resided for around 40ns before it completely left the binding cavity, and the GSM-binding cavity was filled with water molecules. (Figure D.4C, D). Notably, upon the dissociation of E2012, F177, and Y240 rotated toward the GSM-binding cavity and Y106 returned to its original position where it recognized E2012 and formed a hydrogen bond with the water molecules inside the GSM-binding cavity (Figure D.42, D).

While the arylimidazol ring and the methoxyphenyl ring of E2012 were mostly confined in the GSM-binding pocket during the simulation, the 2-piperidine ring and the fluorophenyl ring of E2012 did not form any specific interaction with  $\gamma$ -secretase and adopted several conformations (Figure D.5). E2012 was embedded inside the membrane environment throughout all simulations, even when it was dissociated from  $\gamma$ -secretase (Figure D.6A, B). We note that no specific interactions between E2012 and the lipid phosphate were identified and the membrane thickness at the E2012 periphery remained at the same thickness as the bulk membrane environment (see Methods, Figure D.6C, and Table 8.1).

Table 8.1.: Mismatch amplitude  $\alpha$  calculated by fitting the membrane curve to  $z_{POPC} = \alpha e^{\beta \rho} \cos(\gamma \rho) + z_{POPC}(\rho = \infty)$  (see Methods). Fitting of membrane curves is shown in Figure D.8.

Model	run 1	run 2	run 3	run 4	run 5
E2012-bound complex	1.33Å	-1.3Å	1.66Å	-	-
A $\beta$ 43-bound complex	-5.97Å	-6.14Å	-6.31Å	-6.79Å	-5.63Å
E2012-A $\beta$ 43-bound	-2.24Å	0.937Å	-0.99Å	-3.77Å	-1.09Å

### A $\beta$ 43-bound $\gamma$ -secretase induces negative hydrophobic mismatch

In Figure 8.2D we have demonstrated that E2012 can fit into the GSM-binding pocket when A $\beta$ 49, A $\beta$ 46, or A $\beta$ 43 is associated at the substrate binding site by pocket size analysis. Since the cleavage of A $\beta$ 49 and A $\beta$ 46 are usually not measured in the experiment and the secretion of A $\beta$ 43 is suggested as a pathogenic A $\beta$  species in AD-causing amyloid plaques[11, 12], it is of biological and therapeutic importance to understanding how A $\beta$ 43 is released from  $\gamma$ -secretase and how GSM reduces its production. According to our previous work, the balance between E-S interaction mostly stabilized at the PS1 internal docking site and the membrane-pulling force induced by hydrophobic mismatch were suggested to be the dominant factors determining whether an A $\beta$  peptide is to be released or to be delivered to the next cleavage[146]. Therefore, monitoring whether the binding of E2012 can skew the balance might provide a mechanistic explanation of its mode of action.

As a consistency check with a newly published lipid force field (lipid21[164]), we took the A $\beta$ 43-bound  $\gamma$ -secretase homology from the previous work[146] and performed MD simulations (5 replicas \* 1 $\mu$ s). The simulations show a consistent secondary structure across five replicas (Figure 8.4A and the time-evolution of the secondary structure are depicted in Figure D.9A). A  $\beta$ -strand was formed C-terminal to the scissile bond (I41-T43) of A $\beta$ 43 and

stabilized by the hydrogen bonds formed with K380, G382, and L432 of PS1, similar to the  $\beta$ 3-strand resolved in the APP-bound cryo-EM structure[39]. Furthermore, The catalytic hydrogen bond was observed between substrate V40 and PS1 D385 in most sampled configurations (Figure D.10A). Notably, the breaking of the hydrogen bond coincided with the deformation of  $\beta$ -strand conformation (Figure D.10A and Figure D.9A). When the hydrogen bonds between PS1 L432 and A $\beta$ 43 A42 are dissociated, water molecules can reach the active site and disrupt the catalytic hydrogen bond (Figure D.10B). Residues between A21 to D23 and residues between N27 to G37 formed into a helical conformation (Figure 8.4A). In particular, M35, V36, and G37 resided in the internal docking site and occasionally turned into a  $3_{10}$  helix, which we have suggested an essential feature for the  $\gamma$ -secretase tripeptide mechanism[146]. Importantly, the hydrophobic mismatch induced by the ionic interaction between lipid phosphate and substrate K28 ( Figure 8.4C, D) was reproduced in all simulations with a similar mismatch amplitude  $\alpha = -5.77\text{\AA}$  (see Methods and Material, Figure D.8B, and Table 8.1). According to the mattress model[215], the negative mismatch induced at the substrate-membrane interface induces an elastic up-pulling force on the substrate anchoring residue K28.

### **GSM E2012 induces helical unwinding and attenuates the negative hydrophobic mismatch**

As we have observed in the E2012-bound simulations, the fluorophenyl and the 2-piperidine fragments of E2012 are very flexible and showed no specific interactions with  $\gamma$ -secretase. Nonetheless, modifications at this segment are closely related to the modulation potency, usually denoted as the A $\beta$ 42 EC<sub>50</sub> value[25]. This indicates that besides  $\gamma$ -secretase, E2012 potentially also interacts either with the A $\beta$  substrate or its environment. When looking at the simulation of A $\beta$ 43-bound  $\gamma$ -secretase, we observed that the GSM-binding site was filled with water molecules and formed a hydrogen bond with PS1 Y106 (Figure D.12), resembling a water distribution highly similar to what was observed in the E2012-dissociated simulation (Figure D.5D). Therefore, we performed molecular docking to dock E2012 at the GSM-binding site (Figure D.13A, B). The docking pose preserves the same orientation of the cryo-EM resolved structure and does not collide with the A $\beta$ 43 substrate. Next, we conducted MD simulations to study the dynamics of the A $\beta$ 43-E2012- $\gamma$ -secretase ternary structure inside the membrane bilayer.

Unlike the consistent outcome in the A $\beta$ 43-bound complex simulation, simulations of A $\beta$ 43-E2012- $\gamma$ -secretase system showed a diverged result, depending on the binding mode of E2012. The binding of E2012 largely deviated from the docked pose in three out of five simulations (Figure D.14A). The conformations of the ternary complex at the end of each simulation are shown in Figure D.15 (sim1 to sim4) and Figure 8.5B, C (sim5). Interestingly, in the simulations where E2012 did not deviate from the initial docking pose (sim3 and sim5), the  $\beta$ -strand conformation C-terminal to the scissile bond was destroyed (Figure 8.5A and time-evolution of secondary structure is depicted in Figure D.8B). Consistent with the observation in the A $\beta$ 43 bound simulations, the catalytic hydrogen bond was also broken with the disrupted  $\beta$ -strand conformation (Figure D.14B). In contrast, E2012 relocated to a docking pose similar to what was revealed in the cryo-EM structure in the other three simulations

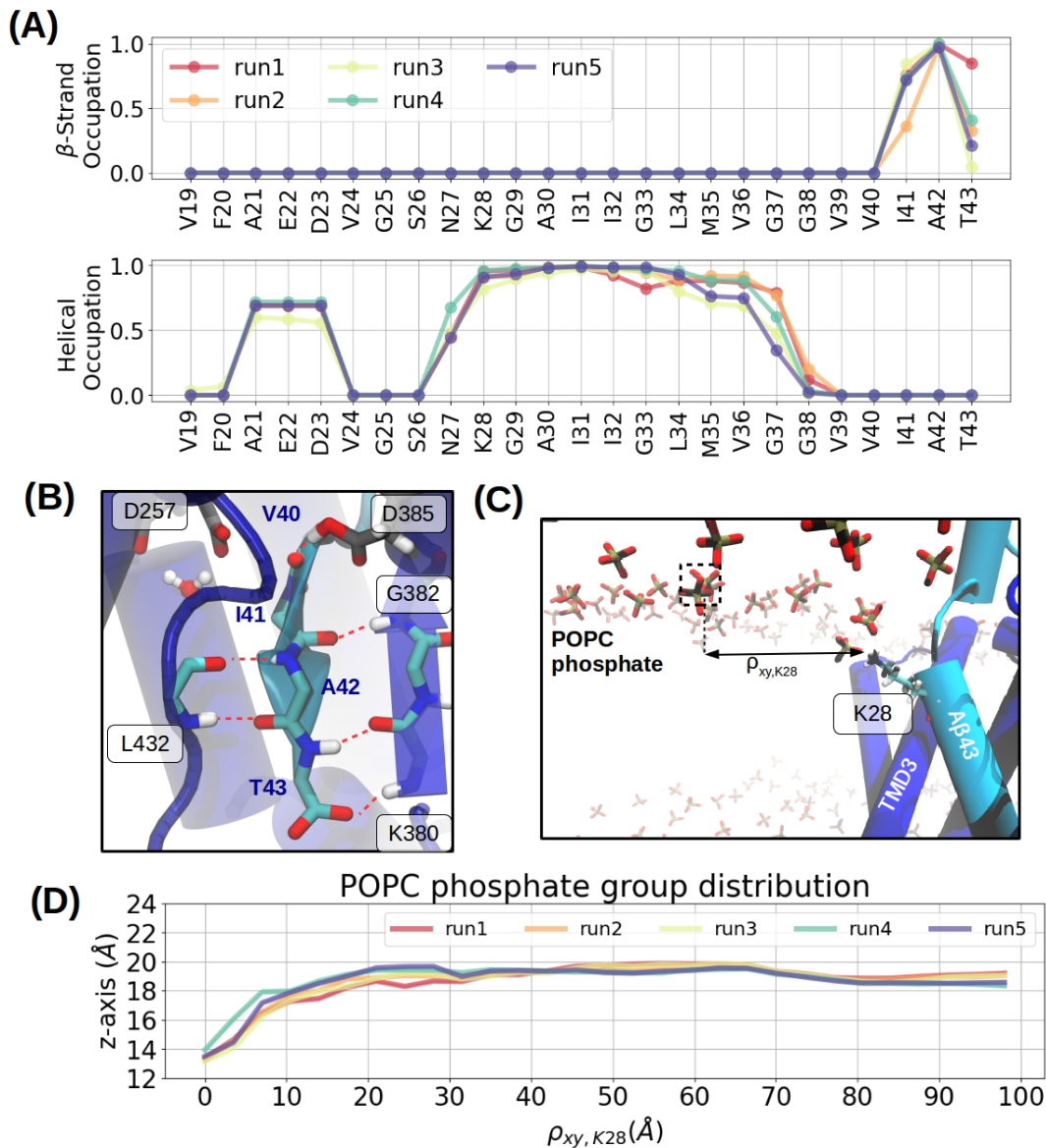


Figure 8.4.: Dynamics of A $\beta$ 43-bound  $\gamma$ -secretase revealed by MD simulations. (A) Averaged secondary structure of A $\beta$ 43 residues analyzed using the DSSP method. Time-evolution of the secondary structure in each simulation is depicted in Figure D.8A (B) Hybrid  $\beta$ -sheet between A $\beta$ 43 and PS1. Hydrogen bonds are indicated with the red dashed lines. (C) Hydrophobic mismatch induced by the ionic interaction between lipid phosphate and K28 of A $\beta$ 43. The radial distance between a phosphate group and A $\beta$ 43 K28 is projected on the x-y plane and denoted as  $\rho_{xy,K28}$ . (D) Z-axis of the phosphate groups at the upper leaf as a function of  $\rho_{xy,K28}$  in five independent 1 $\mu$ s simulations. Convergence is guaranteed as depicted in Figure D.11.

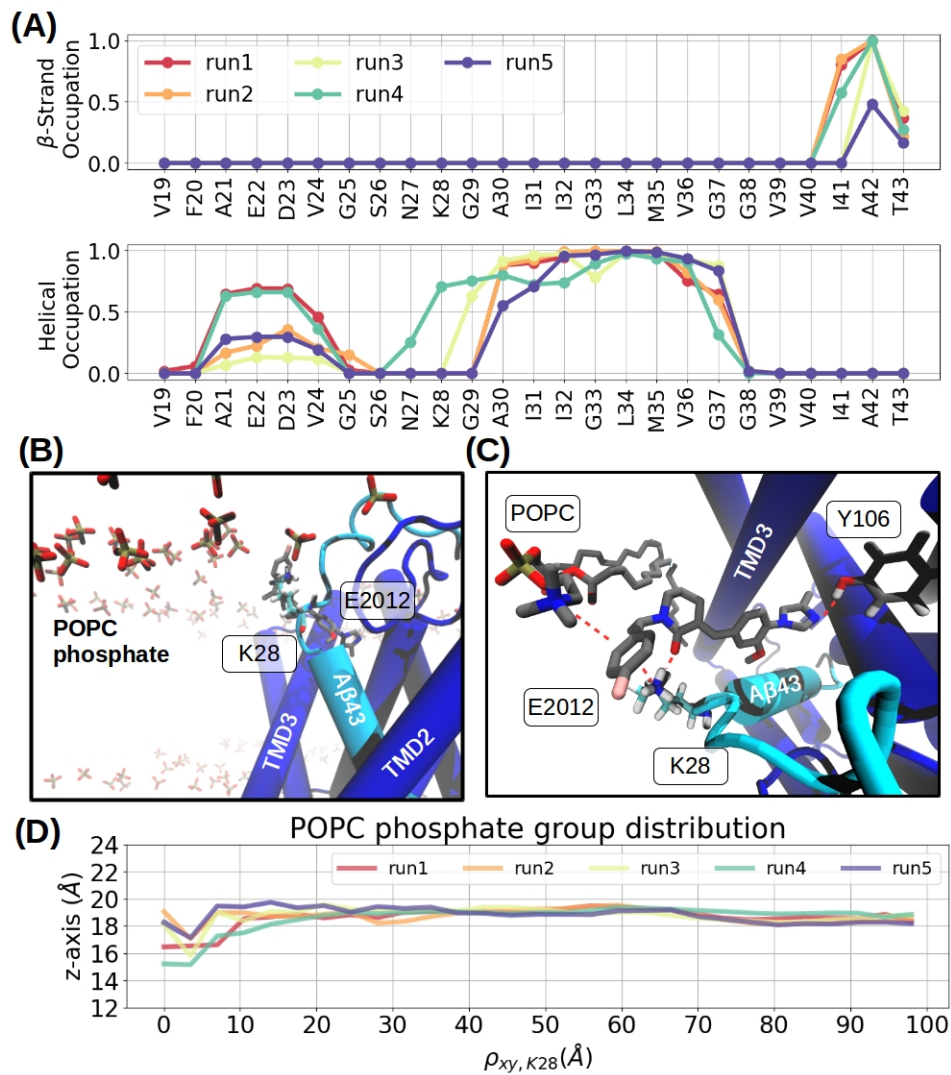


Figure 8.5.: Dynamics of E2012-A $\beta$ 43- $\gamma$ -secretase ternary complex revealed by MD simulations. (A) Averaged secondary structure of A $\beta$ 43 residues analyzed using the DSSP method. Time-evolution of the secondary structure in each simulation is depicted in Figure D.8B (B) Side view of an E2012-A $\beta$ 43- $\gamma$ -secretase ternary complex in side the membrane environment. (C) Top view of an E2012-A $\beta$ 43- $\gamma$ -secretase ternary complex around substrate K28. E2012 is sandwiched by the choline group from a POPC lipid and K28. Hydrogen bonds and  $\pi$ -cation interactions are indicated with the red dashed lines. (D) Z-axis of the phosphate groups at the upper leaf as a function of  $\rho_{xy, K28}$  in five independent 1 $\mu$ s simulations. Convergence is guaranteed as depicted in Figure D.17.

(Figure D.15A, B, D), and the  $\beta$ -strand and catalytic hydrogen bond were constantly formed (Figure D.14B). Besides one simulation (run 4) showing a similar helical domain as in the GSM-free simulation, the helical part of N27-G29 was unwound in the other four simulations. We note that the unwinding of the helix around this location enables K28 to reach upward and form ionic interaction with its upstream amino acids such as D23 (Figure D.15B) and S26 (Figure D.15C).

Next, we analyzed the interactions between E2012,  $\gamma$ -secretase, and A $\beta$ 43 substrate. In contrast to the E2012-only simulations, the hydrogen bond between E2012 and Y106 was formed stably in the majority of the sampled configurations (Figure D.14C). Notably, the fluorophenyl and 2-piperidine fragments, which were found mobile and non-interacting in the substrate-free simulation, were stabilized in one binding mode (sim5, depicted in Figure 8.5B, C). The amine group of K28 occasionally interacted with E2012 through a hydrogen bond with the 2-piperidine ring and a cation- $\pi$  interaction with the fluorophenyl ring. Interestingly, the choline group of a POPC molecule on the other side of E2012 also formed a cation- $\pi$  interaction with the fluorophenyl ring. When K28 was pulled by the 2-piperidine ring of E2012, I242 of NCT also moved down to the close vicinity of K28 (Figure D.16A), which was not observed when E2012 was not bound (Figure D.16B) or in the other binding modes (Figure D.15A-D).

Lastly, we measured whether the membrane is distorted when both E2012 and A $\beta$ 43 are bound to  $\gamma$ -secretase. Surprisingly, three out of five simulations were found to attenuate the local hydrophobic mismatch that was found in the A $\beta$ 43-bound simulation (Figure 8.5D and Table 8.1). This indicates that the membrane pulling force observed in the A $\beta$ 43-bound  $\gamma$ -secretase simulations was eliminated in the presence of E2012. Although different binding modes were discovered, the interaction between K28 and E2012 or other residues was observed in all simulations where the hydrophobic mismatch was alleviated. For example the K28-D23 interaction in sim2 (Figure D.15B), the K28-S26 interaction in sim3 (Figure D.15C), and the K28-E2012 interaction in sim5 (Figure 8.5C). These interactions neutralize the positive electric potential surface (EPS) around K28 and weaken the electrostatic attraction of the lipid head groups. By contrast, hydrophobic mismatch remained in sim1 and sim4 where K28 still interacted with the membrane phosphates.

## 8.4. Discussion

Rather than inhibiting the proteolysis activity of  $\gamma$ -secretase, GSM was designed to reduce the long A $\beta$  species by increasing the cleavage processivity[192]. Although it has been speculated that GSMs promote the generation of short A $\beta$  peptides by increasing the E-S interaction and prolonging its residence time inside  $\gamma$ -secretase[248], no structural data has been provided to explain its mode of action. Thus, uncovering the binding mode of GSM is important not only for the understanding of the mechanism of  $\gamma$ -secretase modulation but also for further structure-based drug development. In this study, we combine computational approaches to study how the imidazole-based GSM E2012 can bind to the substrate-bound  $\gamma$ -secretase complex and study its dynamics inside the membrane environment with computational

approaches.

Using the pocket detection technique, we identified that the putative GSM binding pocket was most of the time available for E2012 binding in the A $\beta$ 49-bound, A $\beta$ 46-bound, and A $\beta$ 43-bound complexes. Among these three substrates, A $\beta$ 43-bound is considered the most relevant species because it is the longest A $\beta$  species in which the modulation effect upon GSM binding can be detected[249, 161]. Therefore, we set up simulations for the E2012-bound, A $\beta$ 43-bound, and E2012-A $\beta$ 43-bound  $\gamma$ -secretase and compare their dynamics. In the E2012-bound complex, we found that the hydrogen bond between E2012 and PS1 Y106 was not stable because of water molecules flowing through the gap between E2012 and NCT I242, which was suggested to interact with the juxtamembrane domain of APP[210], and no specific interaction was identified between  $\gamma$ -secretase and the 2-piperidine ring and the fluorophenyl ring of E2012. When E2012 was dissociated from PS1, the GSM-binding pocket was filled with water molecules. When only A $\beta$ 43 was placed inside  $\gamma$ -secretase, a rigid  $\beta$ -strand C-terminal to the target scissile bond and stable catalytic hydrogen bond were observed. Interestingly, when the  $\beta$ -strand is slightly dissociated from PS1, water molecules from the intracellular domain can access the catalytic center through the gap between A $\beta$ 43 and L432 of PS1, indicating the critical role of the  $\beta$ -strand in stabilizing the catalytic site geometry. In consistence with our previous work, the negatively charged membrane phosphates were attracted by the positive EPS of the substrate K28, through which the membrane became thinner than its bulk environment[146]. The membrane-thinning effect around K28, termed negative hydrophobic mismatch, can contribute to an up-pulling force on the A $\beta$ 43 peptide and facilitate the product release[146]. Importantly, the hydrophobic mismatch was alleviated when both E2012 and  $\beta$ 43 bind simultaneously at the  $\gamma$ -secretase GSM binding site. The reduction in membrane distortion was highly associated with the ionic interaction with K28 facilitated by E2012. First, the binding of E2012 induced a helical unwinding around K28 and enabled K28 to interact with the upstream residues of A $\beta$ 43 such as D23 and S26. Second, the 2-piperidine moiety can directly form a hydrogen bond with K28 and the fluorophenyl ring can shield the interaction between K28 and lipid phosphate by forming  $\pi$ -cation interaction with both sides. Note, the  $\pi$ -cation interaction from lysine has been also observed in a wide range of protein-protein and protein-ligand structures[250, 251, 252, 253]. The attenuation effect of K28-lipid interaction also explains why NCT mutation I242E, which locates in close proximity to E2012 binding site, increases the  $\gamma$ -secretase processivity[210] and why an inverse modulation effect was found in APP mutation K28E[101]. Although the fluorophenyl and 2-piperidine moieties are largely modified in the development of other imidazole-based GSM, the hydrogen acceptor in the 2-piperidine ring and the aromatic ring in the fluorophenyl ring are highly conserved across many other potent GSMs[25].

To our knowledge, this is the first study presenting the molecular mechanism of how GSM binds and modulates the interaction between short A $\beta$  peptide and  $\gamma$ -secretase. However, this chapter only presents the preliminary result of the E2012 binding mode with low consistency in MD simulations, more efforts are still required to determine a realistic GSM binding mode.

## 9. What makes a good protein-protein interaction stabilizer: Analysis and Application of the Dual-Binding Mechanism

Protein-protein interactions (PPI) are essential for many biological processes including signal transduction, immune reactions and many diseases. Inhibition of protein-protein interactions by drug-like compounds is the most common basis for therapeutic approaches. However, due to the large and flat interface of most protein-protein complexes, it is often difficult to find compounds that bind specifically to a partner and inhibit binding. However, complex formation often results in new pockets at the interface, and binding of complex stabilizing compounds to such cavities is an alternative but so far much less explored strategy. In many cases, selective PPI stabilization is as desirable as inhibition for treatment of diseases. Herein, we employ molecular dynamics simulation and pocket detection techniques to investigate 18 known stabilizers and associated protein complexes. For most cases, we find that a dual-binding mechanism, meaning a similar interaction strength of the stabilizer to each protein partner, is an important prerequisite for effective stabilization. A few stabilizers follow an allosteric mechanism by stabilizing the bound structure of a protein partner and/or increase the PPI indirectly. On a large set of 226 protein-protein complexes, we find in > 75% of the cases interface cavities suitable for binding of drug-like compounds. Based on the analysis we propose a computational workflow to identify potential compounds that exploit new cavities formed at protein-protein interfaces and optimize the dual-binding mechanism and apply it to 5 protein-protein complexes. Our study demonstrates a great potential for *in silico* PPI stabilizers discovery that can be helpful in a wide range of therapeutic applications.

### 9.1. Introduction

Protein-protein (PP) complex formation plays a fundamental role in the majority of biological processes such as cell fate[254], immune reactions, and signal transduction[255, 256], and is involved in many human diseases[182, 183], offering substantial therapeutic potentials for modern drug design[182, 257, 258]. Approaches to protein-protein interaction (PPI) modulation are typically inhibition[259] but more recently also stabilization[260, 261, 262] of interactions. PPIs can be inhibited by allosteric effects (involving binding of compounds not necessarily at the interface) or by targeting the protein-protein interface directly such that protein partners can not form a complex. However, the design of PPI inhibitors suffers from the general feature of a typically flat and large buried interface area (> 1500Å<sup>2</sup>[263, 264]), and also often hydrophobic[264] surface, which is frequently considered as ‘undruggable’



for small molecules[265, 266]. Larger peptide-based inhibitors[267] can be adopted as a promising strategy with low toxicity and indeed several peptides have been approved as clinical drugs[268]. In particular, cyclic peptides with pre-defined conformation and good cell permeability have received increasing attention in PPI inhibitor design[269, 270, 271] and using *in silico* 'hot-spots' matching with cyclic peptides has been proposed for rapid rational design[272].

For many diseases, the promotion or stabilization can be more helpful than the inhibition of PPIs. Such stabilization can be achieved by small organic molecules through allosteric (not at the interface) or a direct interface binding mechanism[262]. The former mechanism describes the allosteric regulation of one protein partner upon stabilizer binding which increases the binding affinity to another binding partner while the later mechanism describes the direct binding of the stabilizer at the PPI interface with a potentially druggable cavity for small molecules[273, 274]. With several advantages compared to PPI inhibitors stabilization of PPI has become an emerging field in recent years. Zarzycka et. al.[275] and Andrei et. al.[261] surveyed the available protein-protein-stabilizer complexes and discussed different groups of stabilizers according to the stabilized PPI type and the origin of the stabilizers, respectively. So far only a few drug design efforts have been directed toward stabilizer design. For example, fragment-based drug discovery was used for the design of the adapter protein 14-3-3 with p65-derived protein[276], the tumor suppressor protein p53[277], oncogenic transcription factors TAZ[277], and estrogen receptor-derived peptide[278]. Sijbesma et. al. performed structure-based virtual screening over nearly 6 million compounds from the Molport database and tested 13 compounds *in vitro*, among which 2 are validated as potent stabilizers[279]. Tang et. al. combined molecular dynamics (MD) and molecular docking technique to investigate 22 14-3-3 protein-peptide-stabilizer complexes and concluded that a simulation time of 20-50ns combined with molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) and generalized-Born surface area (MM/GBSA) methods are useful to identify stabilizers from decoys[280].

With the rapid advance in protein structure prediction[281, 117], protein-protein docking[115, 282, 283], and protein complex structure prediction[118], the prediction of potential PPI stabilizers from merely the knowledge of just protein sequences can be a promising route at relatively low cost. To achieve such a goal, it is of fundamental importance to understand the functional mechanisms of the existing stabilizers.

In the first part of the present study, we discuss the theory of PPI stabilization and emphasize the importance of a dual-binding mechanism. Next, we analyze existing PP complexes with bound stabilizers using MD simulations combined with the calculation of interaction-free energies between protein partners and the stabilizer. We use the term stabilizer-induced PPI for complexes that do not form without a stabilizer (Figure 9.2A) whereas stabilizer-enhanced PPI includes those that form complexes already without stabilizers (Figure 9.2B). We find that the more potent stabilizers in most cases distribute the calculated interaction evenly between both protein partners regardless of the total binding free energy between the PP complex and the stabilizer. This follows the expectation based on a mathematical model for the equilibrium binding state whereas less potent stabilizers tend to bind more strongly to one protein than the

other. However, a few compounds mediate the stabilization through an allosteric mechanism by indirectly increasing the PP affinity (4 out of 18 cases). In addition, indirect positive but also negative (destabilization) effects on the direct protein partner interaction were observed. Importantly, we show that 80% of the stabilizer-binding pockets can be detected *in silico* by direct computational probing of the PP complex (without stabilizer) crystal structure. Those pockets hidden in the PP complex structure can be revealed by running short MD simulations and subsequent pocket detection. We showed that the dual-binding mechanism can be useful to identify potential PPI stabilizers. Based on the model, we propose a protocol for PPI stabilizer discovery combining pocket probing, molecular docking, and MD simulation. We also check a large set of known PP complexes for druggable pockets and find for 75% of the cases interface cavities useful for stabilizer binding and apply the suggested protocol to a subset of 5 PP complexes. It indicates promising results that could be implemented in the structure-based drug design of PPI stabilizers.

## 9.2. Methods and Materials

In this study, in total 23 protein complex structures are used from the RCSB PDB database, including 18 RLS complexes and 5 RL complexes (Table 9.1). Due to missing side chains and short backbone segments in some proteins (and some mutated residues), the comparative modeling software MODELLER[116], was used with the wild-type sequences listed in Table 9.1. This includes the restoration of the mutations N755S in PDBID 3bbr, T686A in PDBID 3b6q, L22Q, T51D, V52I in the ligand of PDBID 3o98, S13A, T30D, V31I in the ligand of PDBID 3m50 and 3m51, C6S in PDBID 3u15, and L9E and 3vbg. The RL complexes of set A and set B are generated by stripping the stabilizers from the corresponding RLS complex. Because of the high structural similarity, A1-a, A2-a, and A3-a were taken to represent the RL complexes of A1, A2, and A3, respectively. Besides the 14-3-3 and its ligands, where the terminal residues are as well interacting with each other and the stabilizers, the N-terminus and C-terminus are capped with neutral capping group ACE and NME with Ambergtools 20[111], respectively. To maintain the conformations of the complex, two calcium ions are kept in the S100A4 homodimer and CaM/CaMBD2a complex, four glutamic acids are kept in the iGluR2 homodimer, and a zinc ion is kept in the CK1 $\alpha$ /CRL4. Also, disulfide bonds in PD1L/PD1L and lambda-6A/lambda-6A complexes are preserved.

Each system was solvated in a box 12.5 Å extended from the nearest atom of the solute with OPC 4-point water molecules and 0.15M sodium chloride. The protonation state of the stabilizers is determined with openbabel [239] under pH=7.4. The atomic interactions are described by the ff19SB forcefield[284] for protein, OPC forcefield[285] for water, TIP4PEW[286] forcefield for ions, and GAFF2[103] for the stabilizers. The same procedure was also implemented to prepare the MD simulation of the 5 example complexes from protein docking benchmark dataset 5.0. 5 replicas of 50ns trajectories of each RL and RLS complex and 5 replicas of 200ns long trajectories of each RS complex and receptor-only system of PDBID 3m51, 2o98, 4mdk, and 1kkq are generated with MD simulation with the CUDA-accelerated version of PMEMD from AMBER20 package with randomly assigned initial velocities. Each simulation

box underwent the minimization process of maximal 3,5000 steps and was equilibrated for 75ps under constant volume and 225ps under constant pressure. Gradually decreasing positional restraints on protein atoms are applied during the minimization and equilibrium processes. The temperature was kept at 310K using a Langevin thermostat and the pressure was kept at 1 bar with a Berendsen barostat. The non-bonded cut-off distance was set to 9 Å and the simulation was integrated with a timestep of 2ps with the SHAKE algorithm. Interface residues are defined by the residue within 5 Å of the partner protein calculated using VMD-PYTHON. RMSD and iRMSD are calculated with the starting structure as the reference. RMSF, RMSD, and iRMSD are calculated using CPPTRAJ. The last 30ns, in a total of 75 frames, of each 50ns simulation in the RL and RLS simulation are taken to calculate the binding free energies between receptor, ligand, and stabilizer using molecular mechanics coupled with generalized-Born surface area (MM/GBSA) method with the internal dielectric constant of 1 and external dielectric constant of 80. The effective Born radii are calculated with the GBOBCII model (igb=5) in ref [287]. Pocket detection was performed with Fpocket4.0 on the crystal structure of the RL complexes in set C. Each pocket consists of polar and apolar probes, as shown in Figure 5C in red and gray spheres, respectively. An atom and a pocket probe are considered in contact if the distance between the atom and the pocket probe is smaller than the sum of their radii with the radii of atoms set to H:1.2 Å, C:1.7 Å, N:1.55 Å, O:1.52 Å, F:1.47 Å, B:1.92 Å, P:1.8 Å, S:1.8 Å, Cl:0.2Å. Ligand coverage fraction is calculated as  $\frac{N_{ligand} \cap N_{pocket}}{N_{pocket}}$ . Where  $N_{ligand} \cap N_{pocket}$  is the number of ligand atoms contacting the pocket probes, as illustrated in Figure E.1 The dynamics of the stabilizer binding pocket are analyzed using MDtraj by first creating probes within 2Å from the known stabilizers, as shown in Figure E.2.

Re-evaluation of the 13 compounds performed by Sigbesma et. al [279] was generated with Openbabel[239] and docked with Autodock vina[238]. The docking procedure was performed on the 3D structure from PDB 5F74. The preparation of 5\* 50ns trajectories and the binding energy calculation were performed with the same procedure as mentioned above.

Considering many stabilizers in the present study possess molecular weights between 375 Dalton and 425 Dalton, and LogP between 2 and 4.5 ( Figure E.3), we fetched 122,847,475 compounds fulfilling such molecular properties from the ZINC20 database[288] with standard highest reactivity. We further eliminated compounds with complex structures, which might cause difficulties for force field parameterization, by filtering the compounds with BertzCT complexity[289] > 700 using RDkit[290], leading to 15,072,167, from which we randomly selected 1 million compounds to perform molecular docking with AutoDock-GPU[291].

## 9.3. Results and Discussion

### Stabilization of PPI through the Dual-binding Mechanism

Arguably, the aim of a potent PPI stabilizer is to increase the formation of the RLS ternary complex. In the case where R and L do not interact, e.g. stabilizer-induced PPI, the system is similar to the bivalent scaffold discussed by Yang and Hlavacek[292] (Figure 9.2A). With the

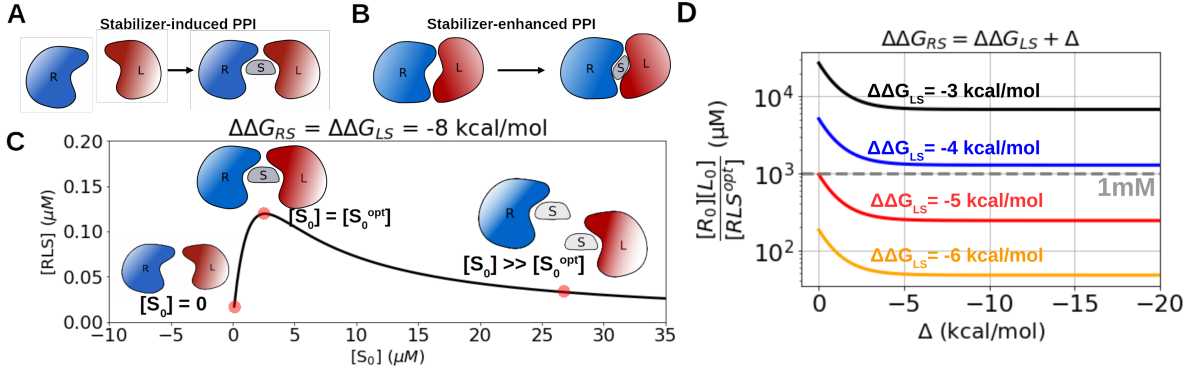


Figure 9.1.: Theoretical model of the dual binding mechanism. (A) Schematics of stabilizer-induced PPI where two non-interacting or weakly-interacting proteins R and L are stabilized by the presence of a stabilizer S. (B) Schematics of stabilizer-enhanced PPI where the interaction between a native RL complex is enhanced by the presence of a stabilizer S. (C) Schematics of the theoretical model according to Eq. 1.  $[R_0]$  and  $[L_0]$  are set to  $1\mu\text{M}$  and the cooperative factor  $\phi$  is set to 1 (non-cooperative binding). Schematics from left to right represent the ternary complexation with no stabilizer, optimal stabilizer, and over-saturated stabilizer. (D) The theoretical dependency of the effective dissociation constant  $K_{RL,eff} = [R_0][L_0]/[RLS^{opt}]$  with different set of  $(\Delta\Delta G_{RS}, \Delta\Delta G_{LS})$ . The gray dashed line shows the boundary between millimolar and micromolar ranges of  $K_{RL,eff}$ .

total concentration of receptor  $[R_0]$ , ligand  $[L_0]$ , and stabilizer  $[S_0]$ , the formation of the RLS complex  $[RLS]$  at the equilibrium state can be formulated as

$$[RLS] = \frac{C - \sqrt{C^2 - 4[R_0][L_0]}}{2} \quad (9.1)$$

where

$$C \equiv [R_0] + [L_0] + ([S_f] + K_{RS})([S_f] + K_{LS})/(\phi[S_f]). \quad (9.2)$$

$\phi$  is the cooperative factor,  $[S_f]$  indicates the concentration of the unbound stabilizer, and  $K_{RS}$  and  $K_{LS}$  are the dissociation equilibrium constants of the RS and LS complex, respectively. Eq. 1 describes that the RLS complex cannot be efficiently formed when  $[S_0]$  is too high because of the saturation of the individual RS and LS complexes and there exists an optimal concentration of the total stabilizer,  $[S_0^{opt}] = \sqrt{K_{RS}K_{LS}}$ , that maximizes the RLS formation to  $[RLS^{opt}]$  (Figure 9.2C).

With the relation  $K_D = e^{\Delta\Delta G/(RT)}$ , where  $K_D$  is the dissociation equilibrium constant and  $\Delta\Delta G$  is the Gibbs binding free energy,  $k_B$  is Boltzmann constant, T is the temperature, we can study the how  $[RLS^{opt}]$  changes with  $\Delta\Delta G_{RS}$  and  $\Delta\Delta G_{LS}$  according to Eq. 1. In the case where  $\Delta\Delta G_{RS} < \Delta\Delta G_{LS}$ , e.g. the stabilizer binds more strongly to the receptor, enhancing both  $\Delta\Delta G_{RS}$  and  $\Delta\Delta G_{LS}$  by a small amount improves the effective dissociation equilibrium constant  $K_{RL,eff} = [R_0][L_0]/[RLS^{opt}]$  much more efficiently than enhancing  $\Delta\Delta G_{RS}$  alone.

### 9.3. RESULTS AND DISCUSSION

	Name (R/L)	PDBID (RLS)	Stabilizer	PDBID (RL)	TM-score[293, 294]*	R residue	L residue
<b>Set A. Stabilizer-Induced PPI with 2 stabilizers</b>							
A1	14-3-3/PMA2	(a) 3m50[295] (b) 3m51[295]	Epibestatin Pyrrolidone1		0.9960	9-240	926-956
A2	PD-1L/PD-1L	(a) 5j89[296] (b) 5j8o[296]	BMS-202 BMS-8		0.9641	18-134	18-134
A3	BRD4/BRD4	(a) 5ad3[297] (b) 5ad2[297]	Compound 6 Compound 2		0.9634	42-168	42-168
<b>Set B. Other Stabilizer-Induced PPI</b>							
B1	14-3-3/ChREBP	6ygj[279]	Compound 3			4-232	117-136
B2	14-3-3/H <sup>+</sup> -ATPase	2o98[298]	Fusicoccin			2-241	905-956
B3	CK1 $\alpha$ /CRL4	5fqd[299]	Lenalidomide			47-436	14-303
B4	Cdc34/Ubiquitin 1 $\alpha$	4mdk[300]	CC0651			3-184	3-74
B5	PPAR $\alpha$ /SMRT	1kkq [301]	GW6471			200-269	682-700
B6	MDM4/MDM4	3u15[302]	2x RO-2443			26-107	26-107
B7	MDM2/MDM2	3vbg[302]	2x RO-2443			26-109	26-109
<b>Set C. Stabilizer-enhanced PPI</b>							
C1	TTR/TTR	3tet[303]	2x Tafamidis	4tlt[304]	0.9912	11-125	11-125
C2	S100A4/S100A4	3ko0[305]	4x Trifluoperazine	3cga[306]	0.9331	2-94	2-94
C3	iGluR2/iGluR2	3bbr[307]	(R,R)-2a	3b6q[308]	0.7748	390-775	390-775
C4	lambda-6A/lambda-6A	6mg5[309]	Coumarin	6mg4[309]	0.9867	1-214	1-214
C5	CaM/CaMBD2-a	4j9z[310]	2x NS309	4j9y[310]	0.9790	395-491	2-147

Table 9.1.: Protein-protein complexes stabilized by bound compounds investigated in the present study.

Note: R residue and L residue indicate the residue number of the receptor and the ligand taken in this study, respectively. \*TM-scores in set A are calculated between the RL conformations between the two crystal structures and TM-scores in set C are calculated between the RL conformations from the RL and RLS crystal structures.

This indicates that the protein that binds weaker to the stabilizer, namely  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$ , plays a more decisive role in determining the stabilization efficiency (Figure 9.2D). Hence, a dual-binding activity is effective for a stabilizer to efficiently induce RLS complex formation. At the extreme case where  $\Delta\Delta G_{RS}$  is strong enough that all receptors are in the stabilizer-bound form, it is apparent that the formation of RLS depends solely on ligand-stabilizer interaction. Note, the dual-binding mechanism is one mechanism to stabilize PPI, other allosteric mechanisms are discussed in the next paragraphs.

#### Investigation of known PP complexes with bound stabilizer

In the present study, we investigate 18 stabilizers binding to 15 different PPIs using MD simulations and pocket detection techniques. An overview of our data set is listed in Table 9.1 and visualized in Figure 9.1.

A stabilized PPI contains a receptor protein (R), a ligand protein (L, typically the smaller protein partner), and a stabilizer (S) and is abbreviated as an RLS complex. Similarly, a PPI without a stabilizer is abbreviated as an RL complex. The 15 PPIs with available three-dimensional (3D) structures are split into 3 data sets. Set A contains 3 stabilizer-induced PPIs where the stabilization effect of two stabilizers on the same R-L pair are experimentally

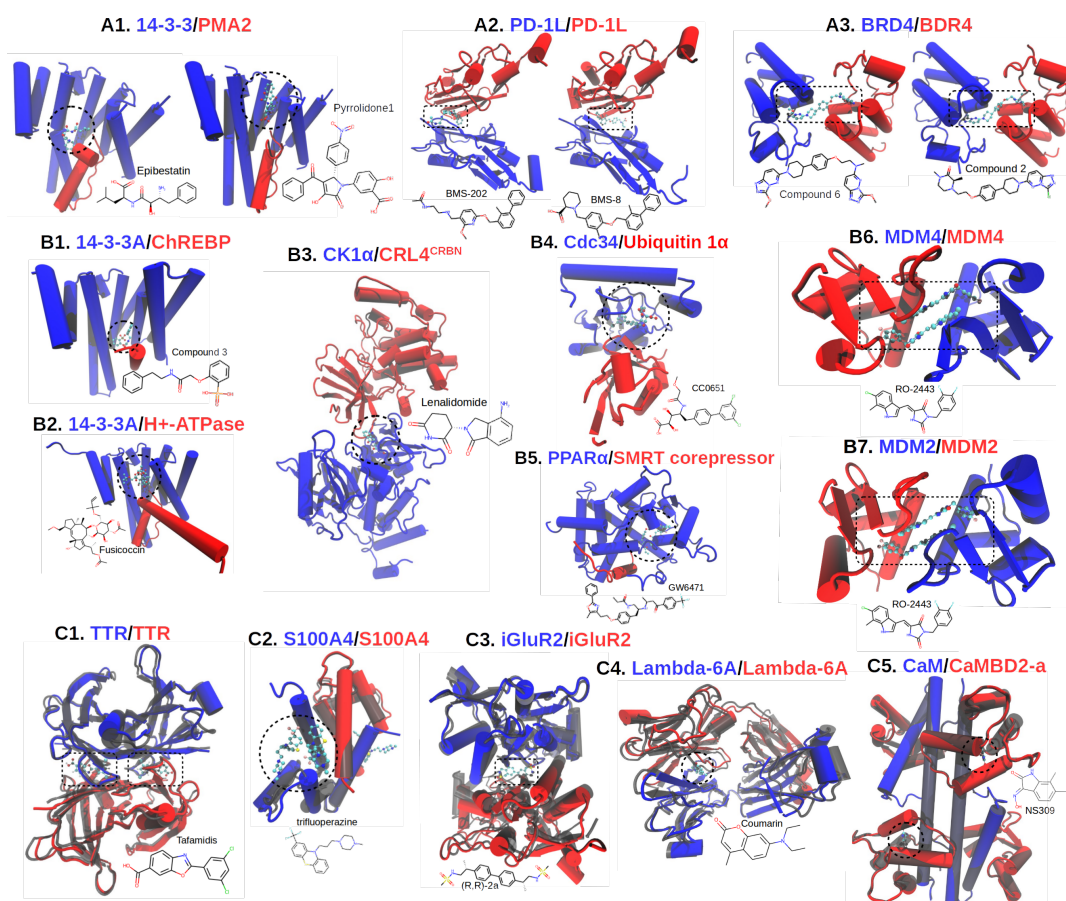


Figure 9.2.: 3D structures of the RLS ternary complexes investigated in the present study. The receptor and ligand are labeled blue and red, respectively. The corresponding PDBIDs of are listed in Table 9.1. The locations of stabilizers are encircled by black dashed lines. Set A (top row) contains 6 RLS complexes with common RL partners binding to a more potent stabilizer (left) and a less potent stabilizer (right). Set B (middle rows) contains 7 RLS complexes with only one stabilizer without an available RL structure. Set C (bottom row) contains 5 RLS complexes with the 3D structures of the RL complexes available (transparent gray).

measured and co-crystallized, were chosen to investigate how the experimentally measured stabilization efficiency is correlated with the structural, dynamical, and thermodynamic properties of the RLS complex. Set B is formed by additional 7 stabilizer-induced PPIs (structure known in complex with one stabilizer) and 5 stabilizer-enhanced PPIs in set C (for which 3D structures also in the absence of stabilizer are known). In example A1, the complex between 14-3-3 and the C-terminal of PMA2 can be induced by the dipeptide Epibestatin and (A1-a,  $K_{RL,eff} = 1.8\mu\text{M}$ ) and Pyrrolidone 1 (A1-b,  $K_{RL,eff} = 80\mu\text{M}$ )[298]. In example A2, the dimerization of the weakly dimerizing protein immunologic regulators programmed death 1 (PD-1) ligand (PD-L1, melting temperature  $T_m = 35.4^\circ\text{C}$ ) are enhanced by small molecules BMS-202 (A2-a,  $T_m = 48.4^\circ\text{C}$ ) and BMS-8 (A2-b,  $T_m = 44.8^\circ\text{C}$ )[299]. Similarly, the dimerization of bromodomain-containing protein 4 (BRD4) can be also induced by small triazolopyridazine-containing molecules compound 6 (A3-a,  $pK_{RL,eff} = 8.1$ ) and compound 6 (A3-b,  $pK_{RL,eff} = 7.2$ )[300] in example A3. Notably, although the different binding poses of the stabilizers in A1 and A2 are observed, the RL complexes adopt similar conformations upon stabilizer binding, as indicated by the high TM-scores in Table 9.1.

We first calculated the buried surface area (BSA) of RS and LS of the studied 18 RLS complexes (for a quick estimation of contacts between stabilizers and proteins, Figure 9.3A). For most stabilizers, a similar BSA is shared with both protein partners indicating a similar number of contacts between the stabilizer and both proteins (Table 9.2). However, surprisingly, although A2-a and A3-a are more potent stabilizers than A2-b and A3-b, respectively, they form less BSA with R and L. Furthermore, we observed four stabilizers in A1-b, B2, B4, and B5 having only a little or no contact with the ligand, which violates the requirement for the dual-binding mechanism discussed above.

To estimate the binding free energy more accurately, we perform MD simulations and calculated the interaction free energies  $\Delta\Delta G_{RS}$ ,  $\Delta\Delta G_{LS}$ , and the total interaction free energy between the protein partners and the stabilizer  $\Delta\Delta G_{(RL)S}$  (Table 9.2 and Figure 9.3A) using the MMGBSA (Molecular Mechanics Generalized Born surface area) endpoint method (see Methods section). Indeed, a comparison between the more potent and less potent stabilizers in set A shows that the weaker stabilizer-binding partner, namely  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$ , has a stronger binding affinity to the more potent stabilizers than to the less potent stabilizers. It agrees with the theoretical model described in Eq.1, despite weaker total stabilizer interaction free energy  $\Delta\Delta G_{(RL)S}$  (Table 9.2, Figure 9.3B).

However, stabilizers in complexes A1-b, B2, B4, and B5 do not follow this rule and bind to the ligand-protein partner with calculated interaction free energies  $\Delta\Delta G_{LS}$  less than  $-5$  kcal/mol, which is the minimum binding free energy required to induce a PPI complex with  $K_{RL,eff}$  at a  $\mu\text{M}$  level at  $[R_0] = [L_0] = 1 \mu\text{M}$ , according to the theoretical model (Figure 9.2D), implicating another possible PPI-inducing mechanism (Table 9.2, Figure 9.3B).

In order to understand how these four stabilizers induce the RLS complexation without bivalent binding to both protein partners, we removed the ligand proteins and performed 200ns MD simulations with (R+S) and without stabilizers (R-only). We found that the interface residues of the receptors, e.g. the residues in contact with the ligand protein in

the RLS complexes, remain stable only in the presence of stabilizers but become much more mobile when the stabilizers are removed, indicated by increased interface root-mean-square deviation (iRMSD, Figure 9.3C-D). Hence, in these cases, the stabilizers bind to one protein partner (receptor) and stabilize the bound conformation, and in turn, allosterically stabilize PP complex formation.

We next calculated the direct interaction between receptor and ligand with the stabilizer ( $\Delta\Delta G_{RL}^S$ ) and without ( $\Delta\Delta G_{RL}$ ) the stabilizer (see Methods). Figure 9.3E shows that the calculated RL interaction is weakened in most stabilizer-induced PPIs when the stabilizers are removed but is stronger or similar in most stabilizer-enhanced PPIs without stabilizers. In order to understand this effect, we performed residue-wise energy decomposition analysis on the RL complexes simulations and identified the favorable RL contacts (indicated by negative binding energies), and unfavorable RL contacts (indicated by positive binding energies, Figure E.4-E.5). Interestingly, the unfavorable RL contacts in the stabilizer-induced PPI complexes are often formed by the residues responsible for stabilizer binding (Figure E.4, and Figure E.5).

Thus, in addition to the dual-binding mechanism, stabilizers can also contribute to a more favorable binding by shielding the unfavorable RL interactions. This may explain why the conformations of receptor and ligand in set A tend to form a similar conformation upon stabilizer binding where favorable RL contacts and druggable pockets near the unfavorable RL contacts can be formed simultaneously (Table 9.2). When stabilizers are removed from these RL(S) complexes, these unfavorable contacts are exposed to their binding partner causing interface distortions. Such an effect was observed in complexes A3, B6, and B7 within 50ns simulations (Figure 9.3F, Figure E.6, and Figure E.7). In contrast, the interfaces of the stabilizer-enhanced PPIs remained stable without stabilizers (Figure E.8). The binding of these stabilizers involves many favorable RL contacts (Figure 9.3G and Figure E.9) and is therefore more likely to disturb the direct RL interactions.

### **MD simulations reveal hidden stabilizer-binding pockets**

In order to discover potential PPI stabilizers in the scheme of structure-based drug design (SBDD), it is crucial to identify useful potential ligand-binding pockets. Structural comparison has shown that PPI stabilizers can bind into an interface pocket of an RL complex without deforming the conformation (Table 9.1), offering an opportunity for virtual screening. We used Fpocket4.0[235] to probe druggable pockets of the natural RL complexes and detected at least 10 pockets in each RL complex (Figure 9.4A). Although the dual-binding mechanism already excludes the non-interface pockets, performing virtual screening through all possible interface pockets is still computationally expensive, and, therefore, a reliable approach to rank the interface pockets is desired to narrow down the sampling phase space. To examine whether the scores provided by Fpocket, e.g. drug score and pocket score, can accurately distinguish the stabilizer-binding pocket from other interface pockets, we calculated the ligand coverage fraction of each pocket (see Methods). Besides the C4 complex, all stabilizer-binding pockets are fully or partially detectable from the RL complex and well ranked by both pocket score and drug score, indicating that it is a reliable approach to identify potential stabilizer-



### 9.3. RESULTS AND DISCUSSION

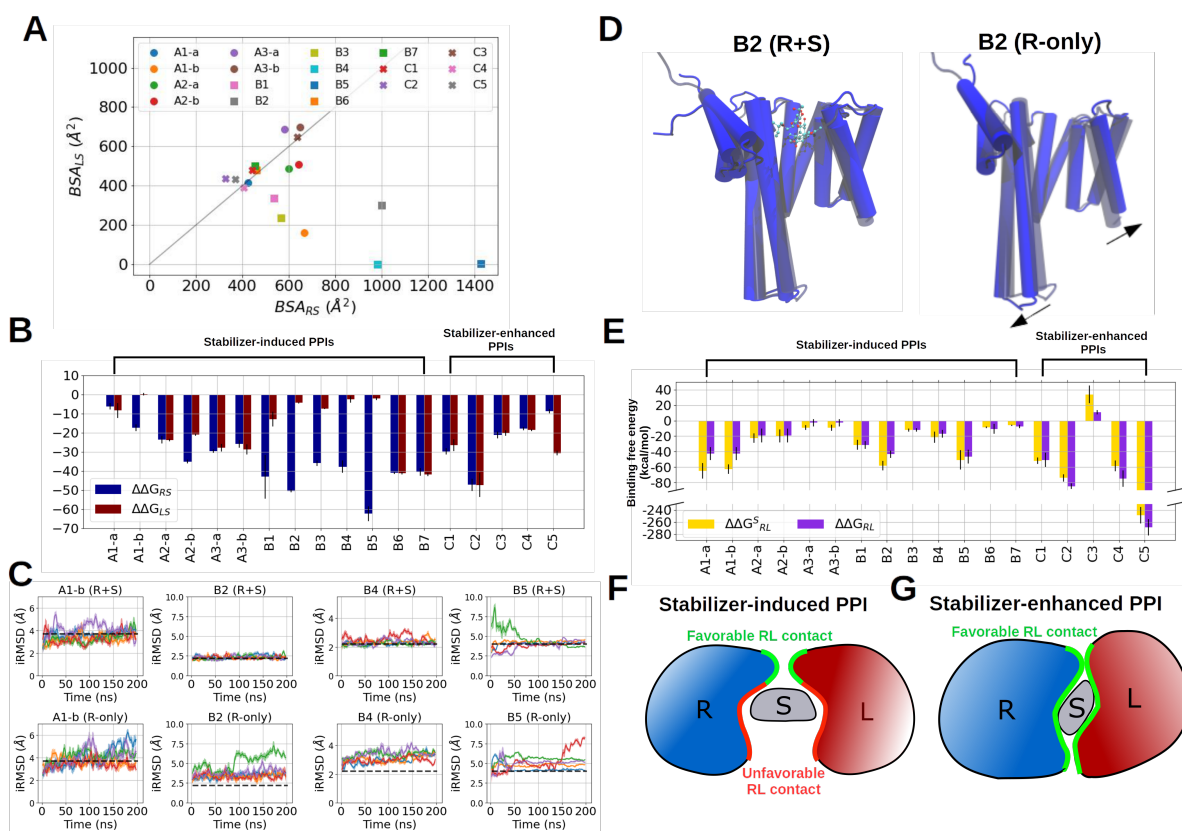


Figure 9.3.: Binding free energy and interface stability calculations with MD simulations. (A) Buried surface area between ligand and stabilizer  $BSA_{LS}$  and between receptor and stabilizer  $BSA_{RS}$  from the RLS crystal structures listed in Table 9.1. (B) Binding free energy between receptor and stabilizer (dark blue) and between ligand and stabilizer (dark red). (C) Interface root-mean-square deviation (iRMSD) of the receptor in the stabilizer-bound form (top) and the receptor-only form (bottom). Five different colors indicate five independent simulations. (D) Schematics of the allosteric effect of PPI stabilizer exemplified with B2 receptor with the stabilizer (left) and without the stabilizer (right). The receptor of the crystal structure is shown in transparent and the last frame of the 200 ns MD simulation is shown in solid representation. PPI stabilizer is shown in CPK representation and the movement of the receptor-ligand binding interface is indicated by the black arrows. (E) Binding free energy between receptor and ligand in the stabilizer-bound form (yellow) and the stabilizer-free form (purple). (F) Schematic of stabilizer shielding the unfavorable RL contacts of stabilizer-induced PPIs. (G) Schematic of stabilizer shielding the favorable RL contacts of stabilizer-enhanced PPIs. Data are shown as rolling averages over every 2 ns with standard deviation indicated by the shaded area.

### 9.3. RESULTS AND DISCUSSION

	PDBID	BSA <sub>LS</sub> /BSA <sub>RS</sub>	$\Delta\Delta G_{RS}$ (kcal/mol)	$\Delta\Delta G_{LS}$ (kcal/mol)	$ \Delta\Delta G_{RS} - \Delta\Delta G_{LS} $ (kcal/mol)	$\Delta\Delta G_{(RL)S}$ (kcal/mol)
<b>Set A. Stabilizer-Induced PPI with 2 stabilizers</b>						
A1	(a) 3m50	0.98	<b>-6.27 ± 1.61</b>	-8.43 ± 3.97	4.54 ± 3.36	-17.16 ± 3.52
	(b) 3m51	0.24	-17.45 ± 1.7	<b>0.11 ± 0.57</b>	17.56 ± 2.22	-24.46 ± 2.5
A2	(a) 5j89	0.81	<b>-23.68 ± 2.06</b>	-23.79 ± 0.66	1.78 ± 1.07	-54.89 ± 2.59
	(b) 5j8o	0.79	-35.16 ± 0.97	<b>-21.08 ± 0.52</b>	14.08 ± 0.79	-62.58 ± 1.53
A3	(a) 5ad3	1.18	-29.64 ± 0.87	<b>-27.95 ± 2.01</b>	2.3 ± 1.2	-61.94 ± 2.93
	(b) 5ad2	1.08	<b>-25.93 ± 2.02</b>	-28.7 ± 2.65	3.77 ± 2.97	-58.47 ± 3.24
<b>Set B. Other Stabilizer-Induced PPI</b>						
B1	6ygi	0.63	-43.03 ± 11.45	<b>-12.93 ± 3.94</b>	30.11 ± 15.11	-87.96 ± 4.4
B2	2o98	0.3	-50.51 ± 0.78	<b>-4.35 ± 0.34</b>	46.16 ± 0.72	-64.26 ± 2.36
B3	5fqd	0.42	-35.9 ± 1.43	<b>-7.33 ± 0.24</b>	28.57 ± 1.47	-48.65 ± 1.53
B4	4mdk	0	-37.84 ± 3.22	<b>-2.66 ± 1.67</b>	35.18 ± 3.51	-45.46 ± 3.68
B5	1kkq	0	-62.43 ± 3.84	<b>-2.17 ± 0.75</b>	60.26 ± 4.31	-66.32 ± 3.73
B6	3u15	1.04	<b>-41.06 ± 0.54</b>	-41.25 ± 0.6	0.78 ± 0.59	-93.54 ± 0.72
B7	3vbg	1.11	<b>-40.61 ± 1.99</b>	-41.84 ± 0.95	2.19 ± 1.79	-93.43 ± 2.25
<b>Set C. Stabilizer-enhanced PPI</b>						
C1	3tct	1.08	-29.84 ± 1.67	<b>-26.53 ± 3.12</b>	3.31 ± 2.7	-61.82 ± 2.93
C2	3ko0	1.34	<b>-47.25 ± 3.22</b>	-47.39 ± 6.29	8.81 ± 3.35	-104.13 ± 3.42
C3	3bbr	1.02	-21.14 ± 1.79	<b>-20.29 ± 1.41</b>	1.77 ± 1.57	-53.64 ± 2.43
C4	6mg5	0.96	<b>-17.93 ± 0.86</b>	-18.61 ± 0.61	1.25 ± 0.26	-44.36 ± 1.06
C5	4j9z	1.17	<b>-8.79 ± 1.12</b>	-30.69 ± 1.2	21.9 ± 2.16	-42.76 ± 0.65

Table 9.2.: Calculation of mean interaction energies in RL and RLS complexes.

Note: Buried surface area and interaction free energy were calculated with the MMGBSA endpoint method. The more potent stabilizers in set A are highlighted in grey. The calculated interaction free energies of the weaker interaction with the stabilizer are shown in bold font. The error shows the standard deviation of five values obtained from five independent simulations.

### 9.3. RESULTS AND DISCUSSION

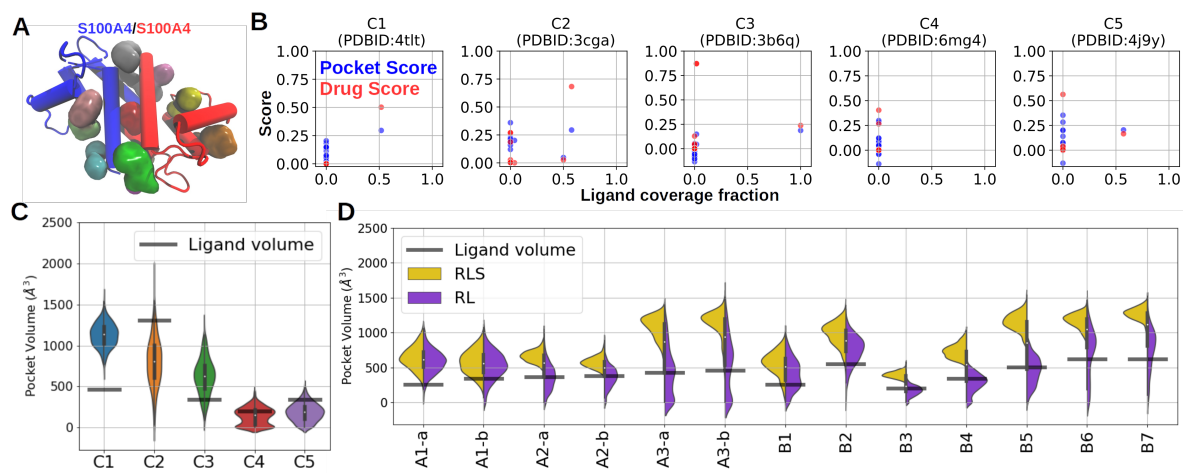


Figure 9.4.: Interface pocket detection and discovery with MD simulations. (A) A schematic of the pockets revealed by Fpocket from the C2 RL complex (PDBID: 3cga). (B) Correlation of the ligand coverage fraction and the scores provided by Fpocket. (C) Distribution of volume of the stabilizer-binding pocket in set C in MD simulations. (D) Distribution of volume of the stabilizer-binding pocket in sets A and B in MD simulations in the stabilizer-bound state (RLS complex, yellow) and the stabilizer-free state (RL complex, purple). The volume of ligands is calculated from <http://www.scfbio-iitd.res.in/Sanjeevini/Molecular-volume-calculator.php> and shown with transparent lines.

binding pockets based solely on the RL structure (Figure 9.4B). Further use of MDpocket to track the dynamics of stabilizer-binding pockets from the simulation (see Methods) reveals that all pockets exhibit a breathing motion with a considerable range of volume sampled in the simulations (Figure 9.4C and time series shown in Figure E.10). Importantly, the stabilizer-binding pocket hidden in the RL crystal structure of the C4 complex occasionally expands to a volume large enough to accommodate the stabilizer with a rapid fluctuation in pocket volume (Figure 9.4C and Figure E.10). Although the volume of the stabilizer-binding pocket remained in some cases of the stabilizer-induced PPI complexes when the stabilizers are removed from the RLS complexes, many of them exhibited a decreased pocket size smaller than the ligand volume (Figure 9.4D). Furthermore, the stabilizer-binding pocket became completely undetectable at the distorted RL interface such as in A3, B6, and B7 (Figure E.6, Figure E.7, Figure E.11, and Figure E.12). Note, interface distortion and RL complex dissociation are expected in all stabilizer-induced PPI if simulations were extended to much longer (e.g. microsecond) timescales due to the only weakly interacting nature.

### Application of the Dual-Binding Mechanism to the *In Silico* Identification of PPI Stabilizers

Our systematic study on 18 PPI stabilizers reveals PPI stabilization mechanisms and the dynamics of the stabilizer-binding pockets. In particular, the dual-binding mechanism derived from the starting theoretical model and supported by experimental results shows that the binding affinity between the stabilizer and the weakly interacting protein partner plays a decisive role in the RLS complexation process. Hence, it may provide a better ranking than the conventional structure-based drug design (SBDD) workflow which usually only considers the total calculated interaction energy between the compound and the target of interest.

We tested the dual-binding ranking approach on the 14-3-3/ChREBP complex with 13 compounds preselected by Sigbesma et. al.[279] through a virtual screening approach ( Figure E.13A). We performed molecular docking, MD simulations, and MM/GBSA calculations to check the dual-binding mechanism of these compounds (see Method, Figure E.13B). For all compounds, a stronger calculated interaction with the receptor compared to the ligand was observed (Figure 9.5A). According to the proposed dual-binding mechanism, the interaction of the compound and the weaker ligand partner plays then a more decisive role in stabilization efficiency. Indeed, an improvement in ranking performance (compared to the experiment) was found when the compounds are ranked using the binding affinity to the ligand instead of the receptor or vs. total binding energy (Figure 9.5A).

To this end, we propose an *in silico* protocol adapted for PPI stabilizers discovery (Figure 9.5B). A 3D structure of the desired RL pair needs to be available, e.g. experimental structure or based on protein-protein complex modeling. Next, Fpocket can be used to probe and rank druggable pockets at the PPI interface. One can further assess the stability of the pocket or explore hidden pockets by performing MD simulations. Once a suitable binding pocket is identified, virtual screening can be conducted with molecular docking programs to screen through a large chemical database (e.g. following the protocol described in the Methods section). Compounds with high docking scores can then be submitted to MD simulation to study the binding stability and interaction free energy. At this step, it is important to check if the compound-protein interaction is of similar magnitude with both protein partners (desired to optimize the stabilizing effect according to the dual binding mechanism). Finally, compounds that fit the criteria can be submitted to experimental validation and lead optimization.

As a demonstration, we implemented the protocol to a subset of protein docking benchmark set 5.0[311], which contains 226 high-quality protein complexes. Pocket detection revealed that among 226 tested complexes, 223 complexes (98.7%) have at least one pocket consisting of 5 atoms from each side of the interface, 173 complexes (76.5%) have at least 10 atoms from each side of the interface, and 57 complexes (25.2%) have at least 20 atoms from each side of the interface ( Figure E.14A-B). We further extracted the best-scored interface pockets composed of at least 5 atoms from each protein and several pockets were identified with pocket scores and drug scores higher than the stabilizer-binding pocket in the present study (Figure 9.4B, Figure E.14C-D).

Among these pockets, we chose 5 interface pockets with high drug scores, including PDBID

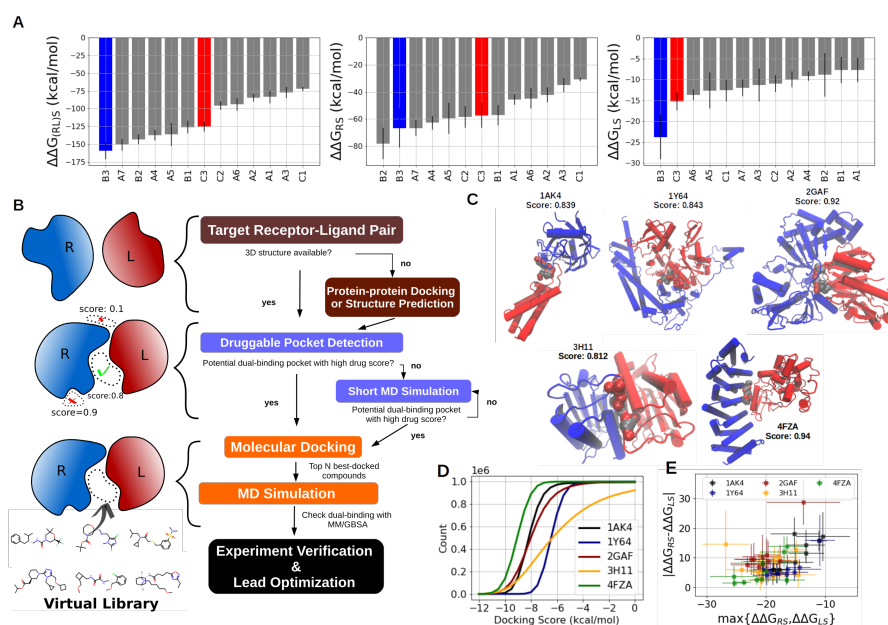


Figure 9.5.: Application of the dual-binding mechanism to PPI stabilizer discovery. (A) Reranking of 13 potential 13-4-4/ChREBP stabilizers using total binding energy (left,  $\Delta\Delta G_{(RL)S}$ ), binding energy to the receptor (middle,  $\Delta\Delta G_{RS}$ ), and binding energy to the ligand (middle,  $\Delta\Delta G_{LS}$ ). Compounds B3 and C3 are the most potent stabilizers with  $EC_{50}$  values of  $0.7\mu M$  and  $45\mu M$ [279], respectively. (B) *In silico* stabilizer discovery workflow. Brown boxes show the PPI structure collection. The blue boxes indicate the pocket detection phase and the orange boxes point to *in silico* binding free energy evaluation with molecular docking and MD simulations. (C) Structures of 5 PPI complexes with potential stabilizer pockets (receptor proteins in blue and ligands proteins in the red cartoon). The pocket is defined by the Van der Waals sphere at the interface with the red and gray spheres representing the polar and apolar pocket probes. (E) The population of the docking score of the 1 million compounds screened through the 5 example PPI complexes. (F) Binding free energy difference  $|\Delta\Delta G_{LS} - \Delta\Delta G_{RS}|$  and the weaker stabilizer-binding energy  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$  of the ten ligands with the best docking score at each PP interface. The numerical value of  $\Delta\Delta G_{RS}$ ,  $\Delta\Delta G_{LS}$ ,  $|\Delta\Delta G_{LS} - \Delta\Delta G_{RS}|$ , and  $\Delta\Delta G_{(RL)S}$  are listed in Supplementary Table E.1-E.5

1AK4, 1Y64, 2GAF, 3H11, and 4FZA, for virtual screening (Figure 9.5C). We performed molecular docking on these five interface pockets with 1 million compounds selected from the ZINC20 database[288] (See Method). Besides the pocket predicted with the poorest drug score (score = 0.812, PDBID 3H11), all pockets can accommodate over 75% of the screening compounds with a docking score over -6 kcal/mol, indicating good druggability preselected with drug scores (Figure 9.5D). The top-ranked 10 docked compounds of each protein complex and their docking score are listed in Figure E.15-E.19 and were submitted to MD simulation for interaction free energy calculation (Supplementary Table E.1-E.5). As demonstrated in the present study, an ideal PPI stabilizer usually possesses a strong binding affinity to the weaker stabilizer-binding partner, indicated by a low  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$ , and a similar binding affinity to both proteins, indicated by a low  $|\Delta\Delta G_{RS} - \Delta\Delta G_{LS}|$ . Among the 50 selected compounds, 34 compounds (68%) possess a  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$  lower than -15 kcal/mol and a  $|\Delta\Delta G_{RS} - \Delta\Delta G_{LS}|$  smaller than 10 kcal/mol (Figure 9.5E). In particular, the best three compounds at the PP interface of 4FZA (compound 5, compound 8, and compound 9, Supplementary Table E.5) all exhibited a  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$  lower than -24 kcal/mol and  $|\Delta\Delta G_{RS} - \Delta\Delta G_{LS}|$  smaller than 4 kcal/mol, are expected to be promising stabilizers. Together with the dual-binding mechanism of stabilizers, our results provide not only a prediction of potential PPI stabilizers but also a guide for further lead optimization, for example, by improving the RS interaction of compound 3 in PDBID 1AK4 or by improving the LS interaction of compound 4 in PDBID 3H11 (Supplementary Table E.1 and Supplementary Table E.4).

## Discussion

To facilitate the development of PPI stabilizer discovery, it is essential to understand how these compounds stabilize PP complexes. In the present study, we have case-studied the binding behavior of 18 stabilizers on their corresponding PP complexes using MD simulations, binding free energy calculations, and pocket detection techniques. Stabilizers can facilitate PPI complex formation through the dual-binding mechanism and/or through conformational stabilization of the bound receptor (or ligand) structure. The more common optimal dual-binding mechanism requires the stabilizer to interact approximately equally with the protein partners. The effective stabilization depends predominantly on the partner that binds weaker to the stabilizer. Furthermore, stabilizers can shield the unfavorable contacts between a non-interacting RL pair while not affecting the favorable contacts. We also found that a fraction of stabilizers that do not fulfill the dual-binding mechanism may also support PPI complexation by stabilizing the interface of one protein to a conformation ready for partner protein binding. The latter more complex allosteric mechanism could in principle be detected by extensive MD simulations of a protein partner in the presence and absence of the stabilizer. However, the result of such simulations can be sensitive to force field artifacts and due to the computational demand, it may not be useful for investigating many compounds. Hence, in practical design efforts, it might be most promising to focus on identifying putative stabilizer compounds that follow the dual-binding mechanism.

Indeed, our analysis of the stabilizer-binding pockets in experimental structures without stabilizers demonstrated that most stabilizer-binding pockets can be readily detected. Those pockets hidden in the RL complexes show up in short MD simulations. Furthermore, interface pockets that can accommodate ligands allowing stabilization by the dual-binding mechanism could be detected in the majority of cases in a large set of known PP complex structures. Based on our analysis we demonstrated that the dual-binding mechanism can be useful to identify potent stabilizers and propose a protocol for PPI stabilizer discovery. We applied the protocol to 5 test cases which can suggest potential PPI stabilizers for five protein-protein complexes. Combining the pocket-scoring technique and dual-binding criteria, our SBDD workflow can efficiently identify druggable interface pockets and estimate the stabilizing efficiency *in silico*.

Ideally, with the advances in structural prediction and protein-protein docking techniques, one can design molecules that induce the complexation of selected proteins. For a given interaction geometry our protocol could also be useful for the identification of such compounds that induce a PPI or support a very weak association. Our work is the first computational study systematically characterizing the binding behavior across different kinds of PPI stabilizers. We anticipate our study to facilitate future developments for the discovery of PPI stabilizers.

## Conclusion and Outlook

Understanding the mechanism of how neurotoxic A $\beta$  peptides are generated from the cleavage of  $\gamma$ -secretase is essential for the cure of Alzheimer's disease and for understanding the mechanism of intramembrane proteolysis. Combining MD simulations, homology modeling, pocket detection, and molecular docking techniques, we attempted to decipher how single point mutations on C99 or A $\beta$  peptides affect the enzyme-substrate interaction and provide a molecular hypothesis for the experimentally observed phenotype.

Starting by comparing the dynamics of  $\gamma$ -secretase between the substrate-bound and substrate-free states in Chapter 4, three main interacting sites of PS1 were identified, namely the TMD6a hydrophobic patch, the hybrid  $\beta$ -sheet, and the internal docking site. By either collaborating with experimental partners or comparing with the previously reported literature, the findings reported in each chapter are strongly supported by sufficient experimental evidence. In Chapter 5, we reported on the interaction at the TMD6a hydrophobic patch that was observed in the interactions of all GSI-bound or substrate-bound  $\gamma$ -secretase structures using the PSH model in membrane bilayer and detergent micelle environments. Both the catalytic active site and the stability of TMD6a were mildly disrupted by the detergent molecules and severely disrupted by the lysine mutations on TMD6a in MD simulations and validated by *in vitro* cleavage assay.

In Chapter 6 we studied how the length of a substrate determines its cleavability by  $\gamma$ -secretase by gradually truncating the C-terminal residues of C99. We observed that the stability of the catalytic active site and the  $\beta$ 3-strand of the substrate are highly correlated to each other and both started to deform partially when only 3 residues are present C-terminal to the active site and deform completely if less. With the use of the HREMD advanced sampling method, we concluded that the E-S hybrid  $\beta$ -sheet is essential to maintain the stability of the catalytic center by controlling the water access and is only sufficiently rigid with at least 3 amino acids. The molecular model not only explains the long-observed  $\gamma$ -secretase tripeptide cleavage but also explains why substrates with Phe and Pro at position P2' can not be cleaved.

In Chapter 4, we identified that E-S interaction at the PS1 internal docking site contributes the most to the substrate stabilization effect and continued the study in Chapter 7. As K28 is gradually brought into the membrane bilayer, the lipid was distorted because of the electrostatic interaction between the phospholipid headgroup and A $\beta$  K28. The interaction induces a positive mismatch at the  $\epsilon$ 49 and  $\zeta$ 46 cleavage and a negative hydrophobic mismatch at the  $\gamma$ 40 and  $\gamma$ 37 cleavage. The balance between the hydrophobic mismatch and the E-S interaction is suggested to largely determine the length of the released A $\beta$  products. The model also explains how  $\gamma$ 37 cleavage is promoted by increasing the E-S interaction by APP mutation G33I or eliminating the negative mismatch by APP mutation K28A.

Using the A $\beta$ -bound models we prepared in Chapter 7, we found that the GSM-binding



site suggested by the cryo-EM structure is exclusively available during the simulation of the A $\beta$ 49, A $\beta$ 46, and A $\beta$ 43 bound form and studied the modulation mechanism of an imidazole-based GSM E2012 in Chapter 8. Our preliminary result suggests a binding mode that elucidates both the functions of the conserved pharmacophores of the design of GSM as well as the mechanism of the enhanced processivity. As an associated project to the study of the GSM-A $\beta$ - $\gamma$ -secretase ternary complex, we investigated 19 PPI complexes stabilized by small molecules and demonstrated the fundamental requirement both from the theoretical mathematical model and from a real-world case study in Chapter 9. We reported that a druggable pocket can be usually found at a protein-protein interface and a dual-binding mechanism is an important requirement for many stabilizers.

The results obtained from each chapter are likely to be investigated deeper as part of additional stand-alone projects. First, based on the structural differences of PSH in different membrane environments in Chapter 5, it is worth investigating to study the dynamics of other membrane-embedded enzymes that also exhibit membrane-dependent activities. Second, we used the HREUS method to examine the functional role of the E-S hybrid  $\beta$ -sheet between  $\gamma$ -secretase and its substrates in Chapter 6). The findings that  $\beta$ -sheet controls the water access to the active site and the active site stability directly raise a question if the hybrid  $\beta$ -sheet, which is already widely found in multiple E-S complexes, is a universal requirement for also other proteases and can be validated by either *in vitro* or *in sicilo* proline mutations on the E-S hybrid  $\beta$ -sheet region. Third, in Chapter 7, we have qualitatively characterized how the balance between the "lipid pulling force" and "E-S interaction" decides whether the peptide product is to stay with the protease or to be released. In principle, the shortening of the substrate length is expected to occur also in other intramembrane proteolysis processes and therefore the hydrophobic mismatch between the cleavage product and the membrane bilayer is also expected to be a driving force for product release. Although computational studies have been done on studying the free energy profile of insertion of amino acid in membrane bilayer[312] and how the length of lipid modulates the conformation of KALP peptidies[313], no work has been published to quantitatively report how hydrophobic mismatch can facilitate the product release in intramembrane proteolysis. Fourth, the putative binding mode of E2012 we suggested in Chapter 8 can successfully elucidate the functions of the consistent chemical motifs in multiple imidazole-based GSMs. Hence, it might be very useful to investigate whether other imidazole-based GSM can also adopt a similar binding mode and correlate their binding free energies to their *in vitro* potency (A $\beta$ 42 IC<sub>50</sub>). Lastly, the dual-binding mechanism we deduced from the existing PPI stabilizers in Chapter 9 paves a promising way for future computational drug design by targeting the interface pocket. In principle, the stabilizer discovery pipeline can also be applied to enhance the interactions between two arbitrary biomolecules other than proteins such as DNAs, RNAs, and polysaccharides.

Overall, our study provides important insights into the molecular mechanism of intramembrane proteolysis by  $\gamma$ -secretase and the role of key interacting sites in substrate recognition and cleavage. Our findings also have implications for the development of therapeutic strategies for Alzheimer's disease by targeting  $\gamma$ -secretase processivity and for the design of PPI stabilizers.

# Acknowledgments

During the course of my Ph.D. study, I was really lucky to get enormous help both physically and mentally from friends, colleagues, and my supervisor. I thank my supervisor Martin for introducing me to the field of Molecular Dynamics simulation and giving me a lot of useful suggestions and directions. Martin always has countless brilliant ideas for computer experiments albeit never really putting pressure on me to do my research the way he wants. The way of being so supportive while leaving room for the student to do the research in their own way is also how I expect myself to be if I as well need to supervise someone in the future. Many thanks go to the very important people that supported me when I first came to Germany back in the winter of 2017, including Katrin Vu and her family, Benjiro, Peter, and Grace. Without their support, I am very likely not able to reach this far in my academic career.

During my 5-year stay in T38, I especially thank Julian for being so welcoming when I first joined the group and helping me out in countless hopeless moments dealing with the bureaucracy and Manuel for giving me so much good advice in the  $\gamma$ -secretase project. I thank Richard, Max, Brianda, Luis, Asha, Christian, Simone, and Patrick for the joyful moments we had together. In addition, I also thank Jonathan for taking care of our cluster, Sonja for managing so much paperwork for our group, and Maria for giving me precious advice in writing and research.

I would like to thank the financial support from DAAD Taiwan and Deutsche Forschungsgemeinschaft. I give many thanks to my collaboration partners professor Steiner, professor Sieber, Lukas, and Jan. Collaboration projects play a very important role in my doctoral study and it is very encouraging to see that the theory derived from computer simulations can really explain and even predict the experimental outcome.

I thank my Taiwanese friends Yi-Ning, Brian, Kimi, Christine, Tracy, Tzu-Chia, Bo-Hung, Chun-Te, Joe, and the 8+1 group for the warm accompaniment and for making my abroad life feel at home.

My biggest thanks go to my family. I thank my parents for being so unconditionally supportive of my decision, my sister for always taking care of the family matters in Taiwan for us, and my brother for being very encouraging on the way of scientific research.

## Bibliography

- [1] U. C. Müller, T. Deller, and M. Korte. “Not just amyloid: physiological functions of the amyloid precursor protein family”. eng. In: *Nature Reviews. Neuroscience* 18.5 (2017), pp. 281–298. DOI: 10.1038/nrn.2017.29.
- [2] G. Thinakaran and E. H. Koo. “Amyloid Precursor Protein Trafficking, Processing, and Function \*”. English. In: *Journal of Biological Chemistry* 283.44 (2008), pp. 29615–29619. DOI: 10.1074/jbc.R800019200.
- [3] D. J. Selkoe. “Alzheimer’s Disease: Genes, Proteins, and Therapy”. In: *Physiological Reviews* 81.2 (2001), pp. 741–766. DOI: 10.1152/physrev.2001.81.2.741.
- [4] J. A. Hardy and G. A. Higgins. “Alzheimer’s Disease: The Amyloid Cascade Hypothesis”. eng. In: *Science* 256.5054 (1992), pp. 184–185. DOI: 10.1126/science.1566067.
- [5] S. L. Cole and R. Vassar. “The Alzheimer’s disease  $\beta$ -secretase enzyme, BACE1”. In: *Molecular Neurodegeneration* 2.1 (2007), p. 22. DOI: 10.1186/1750-1326-2-22.
- [6] C. Venugopal, C. M. Demos, K. Jagannatha Rao, M. A. Pappolla, and K. Sambamurti. “Beta-secretase: structure, function, and evolution”. In: *CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders)* 7.3 (2008), pp. 278–294.
- [7] D. M. Skovronsky, D. B. Moore, M. E. Milla, R. W. Doms, and V. M.-Y. Lee. “Protein Kinase C-dependent  $\alpha$ -Secretase Competes with  $\beta$ -Secretase for Cleavage of Amyloid- $\beta$  Precursor Protein in the Trans-Golgi Network \*”. English. In: *Journal of Biological Chemistry* 275.4 (2000), pp. 2568–2575. DOI: 10.1074/jbc.275.4.2568.
- [8] S. J. Tyler, D. Dawbarn, G. K. Wilcock, and S. J. Allen. “alpha- and beta-secretase: profound changes in Alzheimer’s disease”. eng. In: *Biochemical and Biophysical Research Communications* 299.3 (2002), pp. 373–376. DOI: 10.1016/s0006-291x(02)02635-9.
- [9] H. Steiner, A. Fukumori, S. Tagami, and M. Okochi. “Making the final cut: pathogenic amyloid- $\beta$  peptide generation by  $\gamma$ -secretase”. en. In: *Cell Stress* 2.11 (2018), pp. 292–310. DOI: 10.15698/cst2018.11.162.
- [10] J. T. Jarrett, E. P. Berger, and P. T. Lansbury. “The carboxy terminus of the beta amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer’s disease”. eng. In: *Biochemistry* 32.18 (1993), pp. 4693–4697. DOI: 10.1021/bi00069a001.

- [11] H. Welander, J. Frånberg, C. Graff, E. Sundström, B. Winblad, and L. O. Tjernberg. "Abeta43 is more frequent than Abeta40 in amyloid plaque cores from Alzheimer disease brains". eng. In: *Journal of Neurochemistry* 110.2 (2009), pp. 697–706. DOI: 10.1111/j.1471-4159.2009.06170.x.
- [12] A. Sandebring, H. Welander, B. Winblad, C. Graff, and L. O. Tjernberg. "The Pathogenic A $\beta$ 43 Is Enriched in Familial and Sporadic Alzheimer Disease". en. In: *PLoS ONE* 8.2 (2013). Ed. by P. Fernandez-Funez, e55847. DOI: 10.1371/journal.pone.0055847.
- [13] G. G. Glenner and C. W. Wong. "Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein". eng. In: *Biochemical and Biophysical Research Communications* 120.3 (1984), pp. 885–890. DOI: 10.1016/s0006-291x(84)80190-4.
- [14] G. G. Glenner and C. W. Wong. "Alzheimer's disease and Down's syndrome: sharing of a unique cerebrovascular amyloid fibril protein". eng. In: *Biochemical and Biophysical Research Communications* 122.3 (1984), pp. 1131–1135. DOI: 10.1016/0006-291x(84)91209-9.
- [15] S. E. Hickman, E. K. Allison, and J. El Khoury. "Microglial Dysfunction and Defective  $\beta$ -Amyloid Clearance Pathways in Aging Alzheimer's Disease Mice". In: *The Journal of Neuroscience* 28.33 (2008), pp. 8354–8360. DOI: 10.1523/JNEUROSCI.0616-08.2008.
- [16] F. M. Laird, H. Cai, A. V. Savonenko, M. H. Farah, K. He, T. Melnikova, H. Wen, H.-C. Chiang, G. Xu, V. E. Koliatsos, D. R. Borchelt, D. L. Price, H.-K. Lee, and P. C. Wong. "BACE1, a Major Determinant of Selective Vulnerability of the Brain to Amyloid- $\beta$  Amyloidogenesis, is Essential for Cognitive, Emotional, and Synaptic Functions". en. In: *The Journal of Neuroscience* 25.50 (2005), pp. 11693–11709. DOI: 10.1523/JNEUROSCI.2766-05.2005.
- [17] L. Placanica, L. Zhu, and Y.-M. Li. "Gender- and Age-Dependent  $\gamma$ -Secretase Activity in Mouse Brain and Its Implication in Sporadic Alzheimer Disease". In: *PLoS ONE* 4.4 (2009), e5088. DOI: 10.1371/journal.pone.0005088.
- [18] R. E. Tanzi. "The genetics of Alzheimer disease". eng. In: *Cold Spring Harbor Perspectives in Medicine* 2.10 (2012), a006296. DOI: 10.1101/cshperspect.a006296.
- [19] L. M. Bekris, C.-E. Yu, T. D. Bird, and D. W. Tsuang. "Genetics of Alzheimer Disease". In: *Journal of geriatric psychiatry and neurology* 23.4 (2010), pp. 213–227. DOI: 10.1177/0891988710383571.
- [20] M. S. Wolfe. "Inhibition and modulation of  $\gamma$ -secretase for Alzheimer's disease". In: *Neurotherapeutics* 5.3 (2008), pp. 391–398. DOI: 10.1016/j.nurt.2008.05.010.
- [21] M. S. Wolfe. " $\gamma$ -SECRETASE INHIBITORS AND MODULATORS FOR ALZHEIMER'S DISEASE". In: *Journal of Neurochemistry* 120.Suppl 1 (2012), pp. 89–98. DOI: 10.1111/j.1471-4159.2011.07501.x.
- [22] G. Güner and S. F. Lichtenthaler. "The substrate repertoire of  $\gamma$ -secretase/presenilin". en. In: *Seminars in Cell & Developmental Biology*. Gamma Secretase 105 (2020), pp. 27–42. DOI: 10.1016/j.semcd.2020.05.019.

- [23] I.-M. Shih and T.-L. Wang. "Notch signaling, gamma-secretase inhibitors, and cancer therapy". eng. In: *Cancer Research* 67.5 (2007), pp. 1879–1882. doi: 10.1158/0008-5472.CAN-06-3958.
- [24] S. Artavanis-Tsakonas, M. D. Rand, and R. J. Lake. "Notch signaling: cell fate control and signal integration in development". eng. In: *Science (New York, N.Y.)* 284.5415 (1999), pp. 770–776. doi: 10.1126/science.284.5415.770.
- [25] M. G. Bursavich, B. A. Harrison, and J.-F. Blain. "Gamma Secretase Modulators: New Alzheimer's Drugs on the Horizon?" In: *Journal of Medicinal Chemistry* 59.16 (2016), pp. 7389–7409. doi: 10.1021/acs.jmedchem.5b01960.
- [26] C. J. Crump, D. S. Johnson, and Y.-M. Li. "Development and Mechanism of  $\gamma$ -Secretase Modulators for Alzheimer's Disease". In: *Biochemistry* 52.19 (2013), pp. 3197–3216. doi: 10.1021/bi400377p.
- [27] M. S. Wolfe. "Dysfunctional  $\gamma$ -secretase in familial Alzheimer's disease". In: *Neurochemical research* 44.1 (2019), pp. 5–11. doi: 10.1007/s11064-018-2511-1.
- [28] D. J. Selkoe and J. Hardy. "The amyloid hypothesis of Alzheimer's disease at 25 years". In: *EMBO Molecular Medicine* 8.6 (2016), pp. 595–608. doi: 10.15252/emmm.201606210.
- [29] M. S. Wolfe, W. Xia, B. L. Ostaszewski, T. S. Diehl, W. T. Kimberly, and D. J. Selkoe. "Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and  $\gamma$ -secretase activity". en. In: *Nature* 398.6727 (1999), pp. 513–517. doi: 10.1038/19077.
- [30] M. S. Brown, J. Ye, R. B. Rawson, and J. L. Goldstein. "Regulated Intramembrane Proteolysis: A Control Mechanism Conserved from Bacteria to Humans". English. In: *Cell* 100.4 (2000), pp. 391–398. doi: 10.1016/S0092-8674(00)80675-3.
- [31] W. T. Kimberly, M. J. LaVoie, B. L. Ostaszewski, W. Ye, M. S. Wolfe, and D. J. Selkoe. " $\gamma$ -Secretase is a membrane protein complex comprised of presenilin, nicastrin, Aph-1, and Pen-2". In: *Proceedings of the National Academy of Sciences* 100.11 (2003), pp. 6382–6387.
- [32] R. Francis, G. McGrath, J. Zhang, D. A. Ruddy, M. Sym, J. Apfeld, M. Nicoll, M. Maxwell, B. Hai, M. C. Ellis, A. L. Parks, W. Xu, J. Li, M. Gurney, R. L. Myers, C. S. Himes, R. Hiebsch, C. Ruble, J. S. Nye, and D. Curtis. "aph-1 and pen-2 Are Required for Notch Pathway Signaling,  $\gamma$ -Secretase Cleavage of  $\beta$ APP, and Presenilin Protein Accumulation". English. In: *Developmental Cell* 3.1 (2002), pp. 85–97. doi: 10.1016/S1534-5807(02)00189-2.
- [33] W.-j. Luo, H. Wang, H. Li, B. S. Kim, S. Shah, H.-J. Lee, G. Thinakaran, T.-W. Kim, G. Yu, and H. Xu. "PEN-2 and APH-1 Coordinately Regulate Proteolytic Processing of Presenilin 1". English. In: *Journal of Biological Chemistry* 278.10 (2003), pp. 7850–7854. doi: 10.1074/jbc.C200648200.

- [34] D. M. Bolduc, D. R. Montagna, Y. Gu, D. J. Selkoe, and M. S. Wolfe. “Nicastrin functions to sterically hinder  $\gamma$ -secretase–substrate interactions driven by substrate transmembrane domain”. In: *Proceedings of the National Academy of Sciences* 113.5 (2016), E509–E518.
- [35] A. Yamasaki, S. Eimer, M. Okochi, A. Smialowska, C. Kaether, R. Baumeister, C. Haass, and H. Steiner. “The GxGD motif of presenilin contributes to catalytic function and substrate identification of  $\gamma$ -secretase”. In: *Journal of Neuroscience* 26.14 (2006), pp. 3821–3828.
- [36] B. I. Pérez-Revuelta, A. Fukumori, S. Lammich, A. Yamasaki, C. Haass, and H. Steiner. “Requirement for small side chain residues within the GxGD-motif of presenilin for  $\gamma$ -secretase substrate cleavage”. en. In: *Journal of Neurochemistry* 112.4 (2010), pp. 940–950. DOI: 10.1111/j.1471-4159.2009.06510.x.
- [37] B. Kretner, A. Fukumori, P.-H. Kuhn, B. I. Pérez-Revuelta, S. F. Lichtenthaler, C. Haass, and H. Steiner. “Important functional role of residue x of the presenilin GxGD protease active site motif for APP substrate cleavage specificity and substrate selectivity of  $\gamma$ -secretase”. en. In: *Journal of Neurochemistry* 125.1 (2013), pp. 144–156. DOI: 10.1111/jnc.12124.
- [38] T. Tomita, T. Watabiki, R. Takikawa, Y. Morohashi, N. Takasugi, R. Kopan, B. De Strooper, and T. Iwatsubo. “The first proline of PALP motif at the C terminus of presenilins is obligatory for stabilization, complex formation, and  $\gamma$ -secretase activities of presenilins”. In: *Journal of Biological Chemistry* 276.35 (2001), pp. 33273–33281.
- [39] R. Zhou, G. Yang, X. Guo, Q. Zhou, J. Lei, and Y. Shi. “Recognition of the amyloid precursor protein by human  $\gamma$ -secretase”. In: *Science* 363.6428 (2019), eaaw0930.
- [40] G. Yang, R. Zhou, Q. Zhou, X. Guo, C. Yan, M. Ke, J. Lei, and Y. Shi. “Structural basis of Notch recognition by human  $\gamma$ -secretase”. In: *Nature* 565.7738 (2019), p. 192.
- [41] A. A. Kermani. “A guide to membrane protein X-ray crystallography”. en. In: *The FEBS Journal* 288.20 (2021), pp. 5788–5804. DOI: 10.1111/febs.15676.
- [42] E. P. Carpenter, K. Beis, A. D. Cameron, and S. Iwata. “Overcoming the challenges of membrane protein crystallography”. In: *Current opinion in structural biology* 18.5 (2008), pp. 581–586.
- [43] X. Li, S. Dang, C. Yan, X. Gong, J. Wang, and Y. Shi. “Structure of a presenilin family intramembrane aspartate protease”. In: *Nature* 493.7430 (2013), p. 56.
- [44] D. Spasic, A. Tolia, K. Dillen, V. Baert, B. De Strooper, S. Vrijens, and W. Annaert. “Presenilin-1 Maintains a Nine-Transmembrane Topology throughout the Secretory Pathway”. en. In: *Journal of Biological Chemistry* 281.36 (2006), pp. 26569–26577. DOI: 10.1074/jbc.M600592200.
- [45] B. De Strooper, T. Iwatsubo, and M. S. Wolfe. “Presenilins and  $\gamma$ -secretase: structure, function, and role in Alzheimer disease”. In: *Cold Spring Harbor perspectives in medicine* 2.1 (2012), a006304.

- [46] S. Dang, S. Wu, J. Wang, H. Li, M. Huang, W. He, Y.-M. Li, C. C. Wong, and Y. Shi. "Cleavage of amyloid precursor protein by an archaeal presenilin homologue PSH". In: *Proceedings of the National Academy of Sciences* 112.11 (2015), pp. 3344–3349. doi: 10.1073/pnas.1502150112.
- [47] W. Kühlbrandt. "Forty years in cryoEM of membrane proteins". In: *Microscopy* 71.Supplement\_1 (2022), pp. i30–i50. doi: 10.1093/jmicro/dfab041.
- [48] V. K. Lazarov, P. C. Fraering, W. Ye, M. S. Wolfe, D. J. Selkoe, and H. Li. "Electron microscopic structure of purified, active  $\gamma$ -secretase reveals an aqueous intramembrane chamber and two pores". In: *Proceedings of the National Academy of Sciences* 103.18 (2006), pp. 6889–6894. doi: 10.1073/pnas.0602321103.
- [49] T. Ogura, K. Mio, I. Hayashi, H. Miyashita, R. Fukuda, R. Kopan, T. Kodama, T. Hamakubo, T. Iwastubo, T. Tomita, and C. Sato. "Three-dimensional structure of the  $\gamma$ -secretase complex". en. In: *Biochemical and Biophysical Research Communications* 343.2 (2006), pp. 525–534. doi: 10.1016/j.bbrc.2006.02.158.
- [50] P. Osenkowski, H. Li, W. Ye, D. Li, L. Aeschbach, P. C. Fraering, M. S. Wolfe, D. J. Selkoe, and H. Li. "Cryoelectron Microscopy Structure of Purified  $\gamma$ -Secretase at 12 Å Resolution". en. In: *Journal of Molecular Biology* 385.2 (2009), pp. 642–652. doi: 10.1016/j.jmb.2008.10.078.
- [51] F. Renzi, X. Zhang, W. J. Rice, C. Torres-Arancivia, Y. Gomez-Llorente, R. Diaz, K. Ahn, C. Yu, Y.-M. Li, S. S. Sisodia, and I. Ubarretxena-Belandia. "Structure of  $\gamma$ -Secretase and Its Trimeric Pre-activation Intermediate by Single-particle Electron Microscopy". en. In: *The Journal of Biological Chemistry* 286.24 (2011), p. 21440. doi: 10.1074/jbc.M110.193326.
- [52] Y. Li, S. H.-J. Lu, C.-J. Tsai, C. Bohm, S. Qamar, R. B. Dodd, W. Meadows, A. Jeon, A. McLeod, F. Chen, M. Arimon, O. Berezovska, B. T. Hyman, T. Tomita, T. Iwatsubo, C. M. Johnson, L. A. Farrer, G. Schmitt-Ulms, P. E. Fraser, and P. H. St George-Hyslop. "Structural Interactions between Inhibitor and Substrate Docking Sites Give Insight into Mechanisms of Human PS1 Complexes". In: *Structure(London, England:1993)* 22.1 (2014), pp. 125–135. doi: 10.1016/j.str.2013.09.018.
- [53] X.-C. Bai, C. Yan, G. Yang, P. Lu, D. Ma, L. Sun, R. Zhou, S. H. W. Scheres, and Y. Shi. "An atomic structure of human  $\gamma$ -secretase". eng. In: *Nature* 525.7568 (2015), pp. 212–217. doi: 10.1038/nature14892.
- [54] L. Sun, L. Zhao, G. Yang, C. Yan, R. Zhou, X. Zhou, T. Xie, Y. Zhao, S. Wu, X. Li, et al. "Structural basis of human  $\gamma$ -secretase assembly". In: *Proceedings of the National Academy of Sciences* 112.19 (2015), pp. 6003–6008. doi: 10.1073/pnas.1506242112.
- [55] X.-c. Bai, E. Rajendra, G. Yang, Y. Shi, and S. H. Scheres. "Sampling the conformational space of the catalytic subunit of human  $\gamma$ -secretase". In: *Elife* 4 (2015), e11182.
- [56] K. Nadezhdin, O. Bocharova, E. Bocharov, and A. Arseniev. "Structural and Dynamic Study of the Transmembrane Domain of the Amyloid Precursor Protein". In: *Acta Naturae* 3.1 (2011), pp. 69–76.

- [57] C. L. Deatherage, Z. Lu, B. M. Kroncke, S. Ma, J. A. Smith, M. W. Voehler, R. L. McFeeters, and C. R. Sanders. "Structural and biochemical differences between the Notch and the amyloid precursor protein transmembrane domains". In: *Science Advances* 3.4 (2017), e1602794. DOI: 10.1126/sciadv.1602794.
- [58] G. Yang, R. Zhou, X. Guo, C. Yan, J. Lei, and Y. Shi. "Structural basis of  $\gamma$ -secretase inhibition and modulation by small molecule drugs". en. In: *Cell* 184.2 (2021), 521–533.e14. DOI: 10.1016/j.cell.2020.11.049.
- [59] X. Guo, Y. Wang, J. Zhou, C. Jin, J. Wang, B. Jia, D. Jing, C. Yan, J. Lei, R. Zhou, and Y. Shi. "Molecular basis for isoform-selective inhibition of presenilin-1 by MRK-560". In: *Nature Communications* 13 (2022), p. 6299. DOI: 10.1038/s41467-022-33817-5.
- [60] A. Fukumori and H. Steiner. "Substrate recruitment of  $\gamma$ -secretase and mechanism of clinical presenilin mutations revealed by photoaffinity mapping". In: *The EMBO journal* 35.15 (2016), pp. 1628–1643.
- [61] A. Götz, N. Mylonas, P. Högel, M. Silber, H. Heinel, S. Menig, A. Vogel, H. Feyrer, D. Huster, B. Luy, D. Langosch, C. Scharnagl, C. Muhle-Goll, F. Kamp, and H. Steiner. "Modulating Hinge Flexibility in the APP Transmembrane Domain Alters  $\gamma$ -Secretase Cleavage". English. In: *Biophysical Journal* 116.11 (2019), pp. 2103–2120. DOI: 10.1016/j.bpj.2019.04.030.
- [62] M. Hitzengerber and M. Zacharias. "Structural Modeling of  $\gamma$ -Secretase A $\beta$ n Complex Formation and Substrate Processing". In: *ACS Chemical Neuroscience* 10.3 (2019), pp. 1826–1840. DOI: 10.1021/acscchemneuro.8b00725.
- [63] W. P. Esler, W. T. Kimberly, B. L. Ostaszewski, W. Ye, T. S. Diehl, D. J. Selkoe, and M. S. Wolfe. "Activity-dependent isolation of the presenilin- $\gamma$ -secretase complex reveals nicastrin and a  $\gamma$  substrate". In: *Proceedings of the National Academy of Sciences* 99.5 (2002), pp. 2720–2725. DOI: 10.1073/pnas.052436599.
- [64] S. Takagi-Niidome, T. Sasaki, S. Osawa, T. Sato, K. Morishima, T. Cai, T. Iwatsubo, and T. Tomita. "Cooperative Roles of Hydrophilic Loop 1 and the C-Terminus of Presenilin 1 in the Substrate-Gating Mechanism of  $\gamma$ -Secretase". en. In: *Journal of Neuroscience* 35.6 (2015), pp. 2646–2656. DOI: 10.1523/JNEUROSCI.3164-14.2015.
- [65] I. Schechter and A. Berger. "On the size of the active site in proteases. I. Papain". en. In: *Biochemical and Biophysical Research Communications* 27.2 (1967), pp. 157–162. DOI: 10.1016/S0006-291X(67)80055-X.
- [66] D. M. Bolduc, D. R. Montagna, M. C. Seghers, M. S. Wolfe, and D. J. Selkoe. "The amyloid-beta forming tripeptide cleavage mechanism of  $\gamma$ -secretase". In: *eLife* 5 (2016). Ed. by B. F. Cravatt, e17578. DOI: 10.7554/eLife.17578.
- [67] M. Takami, Y. Nagashima, Y. Sano, S. Ishihara, M. Morishima-Kawashima, S. Funamoto, and Y. Ihara. " $\gamma$ -Secretase: successive tripeptide and tetrapeptide release from the transmembrane domain of  $\beta$ -carboxyl terminal fragment". In: *Journal of Neuroscience* 29.41 (2009), pp. 13042–13052.



- [68] T. Sato, N. Dohmae, Y. Qi, N. Kakuda, H. Misonou, R. Mitsumori, H. Maruyama, E. H. Koo, C. Haass, K. Takio, M. Morishima-Kawashima, S. Ishiura, and Y. Ihara. "Potential Link between Amyloid  $\beta$ -Protein 42 and C-terminal Fragment  $\gamma$  49–99 of  $\beta$ -Amyloid Precursor Protein \*". English. In: *Journal of Biological Chemistry* 278.27 (2003), pp. 24294–24301. doi: 10.1074/jbc.M211161200.
- [69] N. Kakuda, S. Funamoto, S. Yagishita, M. Takami, S. Osawa, N. Dohmae, and Y. Ihara. "Equimolar Production of Amyloid  $\beta$ -Protein and Amyloid Precursor Protein Intracellular Domain from  $\beta$ -Carboxyl-terminal Fragment by  $\gamma$ -Secretase \*". English. In: *Journal of Biological Chemistry* 281.21 (2006), pp. 14776–14786. doi: 10.1074/jbc.M513453200.
- [70] M. A. Fernandez, J. A. Klutkowski, T. Freret, and M. S. Wolfe. "Alzheimer Presenilin-1 Mutations Dramatically Reduce Trimming of Long Amyloid  $\beta$ -Peptides ( $A\beta$ ) by  $\gamma$ -Secretase to Increase 42-to-40-Residue  $A\beta$  \*". English. In: *Journal of Biological Chemistry* 289.45 (2014), pp. 31043–31052. doi: 10.1074/jbc.M114.581165.
- [71] L. Sun, R. Zhou, G. Yang, and Y. Shi. "Analysis of 138 pathogenic mutations in presenilin-1 on the in vitro production of  $A\beta$ 42 and  $A\beta$ 40 peptides by  $\gamma$ -secretase". In: *Proceedings of the National Academy of Sciences* 114.4 (2017), E476–E485.
- [72] O. Quintero-Monzon, M. M. Martin, M. A. Fernandez, C. A. Cappello, A. J. Krzysiak, P. Osenkowski, and M. S. Wolfe. "Dissociation between Processivity and Total Activity of  $\gamma$ -Secretase: Implications for the Mechanism of Alzheimer-Causing Presenilin Mutations". In: *Biochemistry* 50.42 (2011), pp. 9023–9035. doi: 10.1021/bi2007146.
- [73] D. Scheuner, C. Eckman, M. Jensen, X. Song, M. Citron, N. Suzuki, T. D. Bird, J. Hardy, M. Hutton, W. Kukull, E. Larson, E. Levy-Lahad, M. Viitanen, E. Peskind, P. Poorkaj, G. Schellenberg, R. Tanzi, W. Wasco, L. Lannfelt, D. Selkoe, and S. Younkin. "Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease". eng. In: *Nature Medicine* 2.8 (1996), pp. 864–870. doi: 10.1038/nm0896-864.
- [74] D. R. Borchelt, G. Thinakaran, C. B. Eckman, M. K. Lee, F. Davenport, T. Ratovitsky, C. M. Prada, G. Kim, S. Seekins, D. Yager, H. H. Slunt, R. Wang, M. Seeger, A. I. Levey, S. E. Gandy, N. G. Copeland, N. A. Jenkins, D. L. Price, S. G. Younkin, and S. S. Sisodia. "Familial Alzheimer's disease-linked presenilin 1 variants elevate  $A\beta$ 1-42/1-40 ratio in vitro and in vivo". eng. In: *Neuron* 17.5 (1996), pp. 1005–1013. doi: 10.1016/s0896-6273(00)80230-5.
- [75] W. Chen, E. Gamache, D. J. Rosenman, J. Xie, M. M. Lopez, Y.-M. Li, and C. Wang. "Familial Alzheimer's mutations within APPTM increase  $A\beta$ 42 production by enhancing accessibility of  $\epsilon$ -cleavage site". In: *Nature communications* 5 (2014), p. 3037. doi: 10.1038/ncomms4037.

- [76] M. Szaruga, B. Munteanu, S. Lismont, S. Veugelen, K. Horr , M. Mercken, T. C. Saido, N. S. Ryan, T. D. Vos, S. N. Savvides, R. Gallardo, J. Schymkowitz, F. Rousseau, N. C. Fox, C. Hopf, B. D. Strooper, and L. Ch vez-Guti rrez. "Alzheimer's-Causing Mutations Shift A $\beta$  Length by Destabilizing  $\gamma$ -Secretase-A $\beta$ n Interactions". In: *Cell* 170.3 (2017), 443–456.e14. doi: <https://doi.org/10.1016/j.cell.2017.07.004>.
- [77] D.-M. Chau, C. J. Crump, J. C. Villa, D. A. Scheinberg, and Y.-M. Li. "Familial Alzheimer disease presenilin-1 mutations alter the active site conformation of  $\gamma$ -secretase". In: *Journal of Biological Chemistry* 287.21 (2012), pp. 17288–17296.
- [78] S.-Y. Chen and M. Zacharias. "How Mutations Perturb  $\gamma$ -Secretase Active Site Studied by Free Energy Simulations". en. In: *ACS Chemical Neuroscience* 11.20 (2020), pp. 3321–3332. doi: 10.1021/acscchemneuro.0c00440.
- [79] D. Leung, G. Abbenante, and D. P. Fairlie. "Protease Inhibitors: Current Status and Future Prospects". In: *Journal of Medicinal Chemistry* 43.3 (2000), pp. 305–341. doi: 10.1021/jm990412m.
- [80] M. S. Shearman, D. Beher, E. E. Clarke, H. D. Lewis, T. Harrison, P. Hunt, A. Nadin, A. L. Smith, G. Stevenson, and J. L. Castro. "L-685,458, an Aspartyl Protease Transition State Mimic, Is a Potent Inhibitor of Amyloid  $\beta$ -Protein Precursor  $\gamma$ -Secretase Activity". In: *Biochemistry* 39.30 (2000), pp. 8698–8704. doi: 10.1021/bi0005456.
- [81] W. P. Esler, W. T. Kimberly, B. L. Ostaszewski, T. S. Diehl, C. L. Moore, J.-Y. Tsai, T. Rahmati, W. Xia, D. J. Selkoe, and M. S. Wolfe. "Transition-state analogue inhibitors of  $\gamma$ -secretase bind directly to presenilin-1". en. In: *Nature Cell Biology* 2.7 (2000), pp. 428–434. doi: 10.1038/35017062.
- [82] R. S. Doody, R. Raman, M. Farlow, T. Iwatsubo, B. Vellas, S. Joffe, K. Kieburtz, F. He, X. Sun, R. G. Thomas, P. S. Aisen, E. Siemers, G. Sethuraman, and R. Mohs. "A Phase 3 Trial of Semagacestat for Treatment of Alzheimer's Disease". In: *New England Journal of Medicine* 369.4 (2013), pp. 341–350. doi: 10.1056/NEJMoa1210951.
- [83] T. R. McCaw, E. Inga, H. Chen, R. Jaskula-Sztul, V. Dudeja, J. A. Bibb, B. Ren, and J. B. Rose. "Gamma Secretase Inhibitors in Cancer: A Current Perspective on Clinical Performance". en. In: *The Oncologist* 26.4 (2021), e608–e621. doi: 10.1002/onco.13627.
- [84] B. D. Strooper. "Lessons from a Failed  $\gamma$ -Secretase Alzheimer Trial". English. In: *Cell* 159.4 (2014), pp. 721–726. doi: 10.1016/j.cell.2014.10.016.
- [85] V. Coric, S. Salloway, C. H. van Dyck, B. Dubois, N. Andreasen, M. Brody, C. Curtis, H. Soininen, S. Thein, T. Shiovitz, G. Pilcher, S. Ferris, S. Colby, W. Kerselaers, R. Dockens, H. Soares, S. Kaplita, F. Luo, C. Pachai, L. Bracoud, M. Mintun, J. D. Grill, K. Marek, J. Seibyl, J. M. Cedarbaum, C. Albright, H. H. Feldman, and R. M. Berman. "Targeting Prodromal Alzheimer Disease With Avagacestat: A Randomized Clinical Trial". In: *JAMA Neurology* 72.11 (2015), pp. 1324–1333. doi: 10.1001/jamaneuro.2015.0607.

- [86] S. Bhattarai, S. Devkota, K. M. Meneely, M. Xing, J. T. Douglas, and M. S. Wolfe. "Design of Substrate Transmembrane Mimetics as Structural Probes for  $\gamma$ -Secretase". In: *Journal of the American Chemical Society* 142.7 (2020), pp. 3351–3355. doi: 10.1021/jacs.9b13405.
- [87] C. Das, O. Berezovska, T. S. Diehl, C. Genet, I. Buldyrev, J.-Y. Tsai, B. T. Hyman, and M. S. Wolfe. "Designed Helical Peptides Inhibit an Intramembrane Protease". en. In: *Journal of the American Chemical Society* 125.39 (2003), pp. 11794–11795. doi: 10.1021/ja037131v.
- [88] L. Zhang and J. Hermans. "310 Helix Versus  $\alpha$ -Helix: A Molecular Dynamics Study of Conformational Preferences of Aib and Alanine". en. In: *Journal of the American Chemical Society* 116.26 (1994), pp. 11915–11921. doi: 10.1021/ja00105a034.
- [89] A. Bavoso, E. Benedetti, B. D. Blasio, V. Pavone, C. Pedone, C. Toniolo, and G. M. Bonora. "Long polypeptide 310-helices at atomic resolution". en. In: *Proceedings of the National Academy of Sciences* 83.7 (1986), pp. 1988–1992. doi: 10.1073/pnas.83.7.1988.
- [90] S. Weggen, J. L. Eriksen, P. Das, S. A. Sagi, R. Wang, C. U. Pietrzik, K. A. Findlay, T. E. Smith, M. P. Murphy, T. Bulter, D. E. Kang, N. Marquez-Sterling, T. E. Golde, and E. H. Koo. "A subset of NSAIDs lower amyloidogenic A $\beta$ 42 independently of cyclooxygenase activity". en. In: *Nature* 414.6860 (2001), pp. 212–216. doi: 10.1038/35102591.
- [91] G. K. Wilcock, S. E. Black, S. B. Hendrix, K. H. Zavitz, E. A. Swabb, M. A. Laughlin, and Tarenflurbil Phase II Study investigators. "Efficacy and safety of tarenflurbil in mild to moderate Alzheimer's disease: a randomised phase II trial". eng. In: *The Lancet. Neurology* 7.6 (2008), pp. 483–493. doi: 10.1016/S1474-4422(08)70090-5.
- [92] J. L. Eriksen, S. A. Sagi, T. E. Smith, S. Weggen, P. Das, D. C. McLendon, V. V. Ozols, K. W. Jessing, K. H. Zavitz, E. H. Koo, and T. E. Golde. "NSAIDs and enantiomers of flurbiprofen target  $\gamma$ -secretase and lower A $\beta$ 42 in vivo". en. In: *The Journal of Clinical Investigation* 112.3 (2003), pp. 440–449. doi: 10.1172/JCI18162.
- [93] Y. Ohki, T. Higo, K. Uemura, N. Shimada, S. Osawa, O. Berezovska, S. Yokoshima, T. Fukuyama, T. Tomita, and T. Iwatsubo. "Phenylpiperidine-type  $\gamma$ -secretase modulators target the transmembrane domain 1 of presenilin 1". In: *The EMBO Journal* 30.23 (2011), pp. 4815–4824. doi: 10.1038/emboj.2011.372.
- [94] Y. Mitani, J. Yarimizu, K. Saita, H. Uchino, H. Akashiba, Y. Shitaka, K. Ni, and N. Matsuoka. "Differential Effects between  $\gamma$ -Secretase Inhibitors and Modulators on Cognitive Function in Amyloid Precursor Protein-Transgenic and Nontransgenic Mice". en. In: *Journal of Neuroscience* 32.6 (2012), pp. 2037–2050. doi: 10.1523/JNEUROSCI.4264-11.2012.
- [95] T. L. Kukar, T. B. Ladd, M. A. Bann, P. C. Fraering, R. Narlawar, G. M. Maharvi, B. Healy, R. Chapman, A. T. Welzel, R. W. Price, B. Moore, V. Rangachari, B. Cusack, J. Eriksen, K. Jansen-West, C. Verbeeck, D. Yager, C. Eckman, W. Ye, S. Sagi, B. A. Cottrell, J. Torpey, T. L. Rosenberry, A. Fauq, M. S. Wolfe, B. Schmidt, D. M. Walsh,

- E. H. Koo, and T. E. Golde. "Substrate-targeting  $\gamma$ -secretase modulators". en. In: *Nature* 453.7197 (2008), pp. 925–929. doi: 10.1038/nature07055.
- [96] B. Kretner, A. Fukumori, A. Gutsmedl, R. M. Page, T. Luebbbers, G. Galley, K. Baumann, C. Haass, and H. Steiner. "Attenuated A $\beta$ 42 Responses to Low Potency  $\gamma$ -Secretase Modulators Can Be Overcome for Many Pathogenic Presenilin Mutants by Second-generation Compounds \*". English. In: *Journal of Biological Chemistry* 286.17 (2011), pp. 15240–15251. doi: 10.1074/jbc.M110.213587.
- [97] R. M. Page, K. Baumann, M. Tomioka, B. I. Pérez-Revuelta, A. Fukumori, H. Jacobsen, A. Flohr, T. Luebbbers, L. Ozmen, H. Steiner, and C. Haass. "Generation of A $\beta$ 38 and A $\beta$ 42 Is Independently and Differentially Affected by Familial Alzheimer Disease-associated Presenilin Mutations and  $\gamma$ -Secretase Modulation". en. In: *Journal of Biological Chemistry* 283.2 (2008), pp. 677–683. doi: 10.1074/jbc.M708754200.
- [98] K. Takeo, S. Tanimura, T. Shinoda, S. Osawa, I. K. Zahariev, N. Takegami, Y. Ishizuka-Katsura, N. Shinya, S. Takagi-Niidome, A. Tominaga, N. Ohsawa, T. Kimura-Someya, M. Shirouzu, S. Yokoshima, S. Yokoyama, T. Fukuyama, T. Tomita, and T. Iwatsubo. "Allosteric regulation of  $\gamma$ -secretase activity by a phenylimidazole-type  $\gamma$ -secretase modulator". In: *Proceedings of the National Academy of Sciences* 111.29 (2014), pp. 10544–10549. doi: 10.1073/pnas.1402171111.
- [99] M. Z. Kounnas, A. M. Danks, S. Cheng, C. Tyree, E. Ackerman, X. Zhang, K. Ahn, P. Nguyen, D. Comer, L. Mao, C. Yu, D. Pleyntet, P. J. Digregorio, G. Velicelebi, K. A. Stauderman, W. T. Comer, W. C. Mobley, Y.-M. Li, S. S. Sisodia, R. E. Tanzi, and S. L. Wagner. "Modulation of gamma-secretase reduces beta-amyloid deposition in a transgenic mouse model of Alzheimer's disease". eng. In: *Neuron* 67.5 (2010), pp. 769–780. doi: 10.1016/j.neuron.2010.08.018.
- [100] C. Nagy, E. Schuck, A. Ishibashi, Y. Nakatani, B. Rege, and V. Logovinsky. "P3-415: E2012, a novel gamma-secretase modulator, decreases plasma amyloid-beta (A $\beta$ ) levels in humans". In: *Alzheimer's & Dementia* 6.4S\_Part\_19 (2010), S574–S574. doi: 10.1016/j.jalz.2010.05.1958.
- [101] S. Ousson, A. Saric, A. Baguet, C. Losberger, S. Genoud, F. Vilbois, B. Permanne, I. Hussain, and D. Beher. "Substrate determinants in the C99 juxtamembrane domains differentially affect  $\gamma$ -secretase cleavage specificity and modulator pharmacology". en. In: *Journal of Neurochemistry* 125.4 (2013), pp. 610–619. doi: 10.1111/jnc.12129.
- [102] N. Pozdnyakov, H. E. Murrey, C. J. Crump, M. Pettersson, T. E. Ballard, C. W. am Ende, K. Ahn, Y.-M. Li, K. R. Bales, and D. S. Johnson. " $\gamma$ -Secretase Modulator (GSM) Photoaffinity Probes Reveal Distinct Allosteric Binding Sites on Presenilin". In: *The Journal of Biological Chemistry* 288.14 (2013), pp. 9710–9720. doi: 10.1074/jbc.M112.398602.
- [103] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. "Development and testing of a general amber force field". en. In: *Journal of Computational Chemistry* 25.9 (2004), pp. 1157–1174. doi: 10.1002/jcc.20035.

- 
- [104] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". de. In: *Annalen der Physik* 389.20 (1927), pp. 457–484. DOI: 10.1002/andp.19273892002.
- [105] D. R. Hartree. "The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods". en. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 24.1 (1928), pp. 89–110. DOI: 10.1017/S0305004100011919.
- [106] V. Fock. "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems". de. In: *Zeitschrift für Physik* 61.1 (1930), pp. 126–148. DOI: 10.1007/BF01340294.
- [107] J. C. Slater. "Note on Hartree's Method". In: *Physical Review* 35.2 (1930), pp. 210–211. DOI: 10.1103/PhysRev.35.210.2.
- [108] E. R. Johnson and A. D. Becke. "A post-Hartree-Fock model of intermolecular interactions". In: *The Journal of Chemical Physics* 123.2 (2005), p. 024101. DOI: 10.1063/1.1949201.
- [109] W. Koch and M. C. Holthausen. *A Chemist's Guide to Density Functional Theory*. en. Google-Books-ID: qm5cCwAAQBAJ. John Wiley & Sons, 2015.
- [110] V. W. D. Cruzeiro, M. Manathunga, K. M. Merz, and A. W. Götz. "Open-Source Multi-GPU-Accelerated QM/MM Simulations with AMBER and QUICK". en. In: *Journal of Chemical Information and Modeling* 61.5 (2021), pp. 2109–2115. DOI: 10.1021/acs.jcim.1c00169.
- [111] D. Case, K. Belfon, I. Ben-Shalom, S. Brozell, D. Cerutti, I. T.E. Cheatham, V. Cruzeiro, T. Darden, R. Duke, G. Giambasu, M. Gilson, H. Gohlke, A. Goetz, S. I. R Harris, S. Izmailov, K. Kasavajhala, A. Kovalenko, R. Krasny, T. Kurtzman, T. Lee, S. LeGrand, C. L. P. Li, J. Liu, T. Luchko, R. Luo, V. Man, K. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, A. Onufriev, F. Pan, S. Pantano, R. Qi, D. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. Simmerling, N. Skrynnikov, J. Smith, J. Swails, R. Walker, J. Wang, L. Wilson, R. Wolf, X. Wu, Y. Xiong, Y. Xue, D. York, and P. Kollman. "AMBER 20". In: *University of California, San Francisco* (2020).
- [112] T. A. Halgren. "The representation of van der Waals (vdW) interactions in molecular mechanics force fields: potential form, combination rules, and vdW parameters". In: *Journal of the American Chemical Society* 114.20 (1992), pp. 7827–7843. DOI: 10.1021/ja00046a032.
- [113] T. A. Halgren and W. Damm. "Polarizable force fields". eng. In: *Current Opinion in Structural Biology* 11.2 (2001), pp. 236–242. DOI: 10.1016/s0959-440x(00)00196-2.
- [114] S. Riniker, J. R. Allison, and W. F. van Gunsteren. "On developing coarse-grained models for biomolecular simulation: a review". en. In: *Physical Chemistry Chemical Physics* 14.36 (2012), p. 12423. DOI: 10.1039/c2cp40934h.
- [115] G. Glashagen, S. de Vries, U. Uciechowska-Kaczmarzyk, S. A. Samsonov, S. Murail, P. Tuffery, and M. Zacharias. "Coarse-grained and atomic resolution biomolecular docking with the ATTRACT approach". en. In: *Proteins: Structure, Function, and Bioinformatics* 88.8 (2020), pp. 1018–1028. DOI: 10.1002/prot.25860.
-

- [116] N. Eswar, B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M.-Y. Shen, U. Pieper, and A. Sali. "Comparative protein structure modeling using Modeller". eng. In: *Current Protocols in Bioinformatics* Chapter 5 (2006), Unit-5.6. doi: 10.1002/0471250953.bi0506s15.
- [117] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis. "Highly accurate protein structure prediction with AlphaFold". en. In: *Nature* 596.7873 (2021), pp. 583–589. doi: 10.1038/s41586-021-03819-2.
- [118] P. Bryant, G. Pozzati, and A. Elofsson. "Improved prediction of protein-protein interactions using AlphaFold2". en. In: *Nature Communications* 13.1 (2022), p. 1265. doi: 10.1038/s41467-022-28865-w.
- [119] A. Y. Toukmaji and J. A. Board. "Ewald summation techniques in perspective: a survey". en. In: *Computer Physics Communications* 95.2 (1996), pp. 73–92. doi: 10.1016/0010-4655(96)00016-1.
- [120] T. Darden, D. York, and L. Pedersen. "Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems". In: *The Journal of Chemical Physics* 98.12 (1993), pp. 10089–10092. doi: 10.1063/1.464397.
- [121] L. Verlet. "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules". In: *Physical Review* 159.1 (1967), pp. 98–103. doi: 10.1103/PhysRev.159.98.
- [122] W. F. Van Gunsteren and H. J. C. Berendsen. "A Leap-frog Algorithm for Stochastic Dynamics". In: *Molecular Simulation* 1.3 (1988), pp. 173–185. doi: 10.1080/08927028808080941.
- [123] H. C. Andersen. "Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations". en. In: *Journal of Computational Physics* 52.1 (1983), pp. 24–34. doi: 10.1016/0021-9991(83)90014-1.
- [124] C. W. Hopkins, S. Le Grand, R. C. Walker, and A. E. Roitberg. "Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning". In: *Journal of Chemical Theory and Computation* 11.4 (2015), pp. 1864–1874. doi: 10.1021/ct5010406.
- [125] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. "Molecular dynamics with coupling to an external bath". en. In: *The Journal of Chemical Physics* 81.8 (1984), pp. 3684–3690. doi: 10.1063/1.448118.
- [126] N. Goga, A. J. Rzepiela, A. H. de Vries, S. J. Marrink, and H. J. C. Berendsen. "Efficient Algorithms for Langevin and DPD Dynamics". In: *Journal of Chemical Theory and Computation* 8.10 (2012), pp. 3637–3649. doi: 10.1021/ct3000876.
- [127] S. Genheden and U. Ryde. "The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities". In: *Expert Opinion on Drug Discovery* 10.5 (2015), pp. 449–461. doi: 10.1517/17460441.2015.1032936.

- [128] E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. H. Zhang, and T. Hou. "End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design". In: *Chemical Reviews* 119.16 (2019), pp. 9478–9508. DOI: 10.1021/acs.chemrev.9b00055.
- [129] S. Genheden, O. Kuhn, P. Mikulskis, D. Hoffmann, and U. Ryde. "The normal-mode entropy in the MM/GBSA method: effect of system truncation, buffer region, and dielectric constant". eng. In: *Journal of Chemical Information and Modeling* 52.8 (2012), pp. 2079–2088. DOI: 10.1021/ci3001919.
- [130] L. Duan, X. Liu, and J. Z. Zhang. "Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein–Ligand Binding Free Energy". In: *Journal of the American Chemical Society* 138.17 (2016), pp. 5722–5728. DOI: 10.1021/jacs.6b02682.
- [131] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson. "Semianalytical treatment of solvation for molecular mechanics and dynamics". In: *Journal of the American Chemical Society* 112.16 (1990), pp. 6127–6129. DOI: 10.1021/ja00172a038.
- [132] D. Bashford and D. Case. "Generalized Born Models of Macromolecular Solvation Effects". In: *Annual review of physical chemistry* 51 (2000), pp. 129–52. DOI: 10.1146/annurev.physchem.51.1.129.
- [133] J. Kästner. "Umbrella sampling". en. In: *WIREs Computational Molecular Science* 1.6 (2011), pp. 932–942. DOI: 10.1002/wcms.66.
- [134] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman. "THE weighted histogram analysis method for free-energy calculations on biomolecules." In: *Journal of Computational Chemistry* 13.8 (1992), pp. 1011–1021. DOI: 10.1002/jcc.540130812.
- [135] J. S. Hub, B. L. de Groot, and D. van der Spoel. "g\_wham—A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates". In: *Journal of Chemical Theory and Computation* 6.12 (2010), pp. 3713–3720. DOI: 10.1021/ct100494z.
- [136] H. Flyvbjerg and H. G. Petersen. "Error estimates on averages of correlated data". In: *The Journal of Chemical Physics* 91.1 (1989), pp. 461–466. DOI: 10.1063/1.457480.
- [137] Y. Sugita and Y. Okamoto. "Replica-exchange molecular dynamics method for protein folding". en. In: *Chemical Physics Letters* 314.1-2 (1999), pp. 141–151. DOI: 10.1016/S0009-2614(99)01123-9.
- [138] P.-C. Li, N. Miyashita, W. Im, S. Ishido, and Y. Sugita. "Multidimensional Umbrella Sampling and Replica-Exchange Molecular Dynamics Simulations for Structure Prediction of Transmembrane Helix Dimers". In: *Journal of computational chemistry* 35.4 (2014), pp. 300–308. DOI: 10.1002/jcc.23494.
- [139] R. Rajamani and C. H. Reynolds. "Modeling the Protonation States of the Catalytic Aspartates in  $\beta$ -Secretase". In: *Journal of Medicinal Chemistry* 47.21 (2004), pp. 5159–5166. DOI: 10.1021/jm049817j.

- [140] L. Coates, P. T. Erskine, S. Mall, R. Gill, S. P. Wood, D. A. A. Myles, and J. B. Cooper. "X-ray, neutron and NMR studies of the catalytic mechanism of aspartic proteinases". en. In: *European Biophysics Journal* 35.7 (2006), pp. 559–566. doi: 10.1007/s00249-006-0065-7.
- [141] L. Coates, H.-F. Tuan, S. Tomanicek, A. Kovalevsky, M. Mustyakimov, P. Erskine, and J. Cooper. "The Catalytic Mechanism of an Aspartic Proteinase Explored with Neutron and X-ray Diffraction". en. In: *Journal of the American Chemical Society* 130.23 (2008), pp. 7235–7237. doi: 10.1021/ja801269x.
- [142] N. Yu, S. A. Hayik, B. Wang, N. Liao, C. H. Reynolds, and K. M. Merz. "Assigning the Protonation States of the Key Aspartates in  $\beta$ -Secretase Using QM/MM X-ray Structure Refinement". In: *Journal of Chemical Theory and Computation* 2.4 (2006), pp. 1057–1069. doi: 10.1021/ct0600060.
- [143] T. J. Paul, A. Barman, M. Ozbil, R. P. Bora, T. Zhang, G. Sharma, Z. Hoffmann, and R. Prabhakar. "Mechanisms of peptide hydrolysis by aspartyl and metalloproteases". en. In: *Physical Chemistry Chemical Physics* 18.36 (2016), pp. 24790–24801. doi: 10.1039/C6CP02097F.
- [144] C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski, and J. H. Jensen. "Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values". In: *Journal of Chemical Theory and Computation* 7.7 (2011), pp. 2284–2295. doi: 10.1021/ct200133y.
- [145] D. C. Guzmán-Ocampo, R. Aguayo-Ortiz, J.-L. Velasco-Bolom, P. L. Gupta, A. E. Roitberg, and L. Dominguez. "Elucidating the Protonation State of the  $\gamma$ -Secretase Catalytic Dyad". In: *ACS Chemical Neuroscience* 14.2 (2023), pp. 261–269. doi: 10.1021/acscchemneuro.2c00563.
- [146] S.-Y. Chen and M. Zacharias. "An internal docking site stabilizes substrate binding to  $\gamma$ -secretase: Analysis by molecular dynamics simulations". en. In: *Biophysical Journal* 121.12 (2022), pp. 2330–2344. doi: 10.1016/j.bpj.2022.05.023.
- [147] R. M. Page, A. Gutsmedl, A. Fukumori, E. Winkler, C. Haass, and H. Steiner. " $\beta$ -Amyloid Precursor Protein Mutants Respond to  $\gamma$ -Secretase Modulators". English. In: *Journal of Biological Chemistry* 285.23 (2010). Publisher: Elsevier, pp. 17798–17810. doi: 10.1074/jbc.M110.103283.
- [148] T.-H. Xu, Y. Yan, Y. Kang, Y. Jiang, K. Melcher, and H. E. Xu. "Alzheimer's disease-associated mutations increase amyloid precursor protein resistance to  $\gamma$ -secretase cleavage and the A $\beta$ 42/A $\beta$ 40 ratio". en. In: *Cell Discovery* 2.1 (2016). Bandiera\_abtest: a Cc\_license\_type: cc\_by Cg\_type: Nature Research Journals Number: 1 Primary\_atype: Research Publisher: Nature Publishing Group Subject\_term: Mechanisms of disease;Molecular biology Subject\_term\_id: mechanisms-of-disease;molecular-biology, pp. 1–14. doi: 10.1038/celldisc.2016.26.



- [149] C. A. Lemere, F. Lopera, K. S. Kosik, C. L. Lendon, J. Ossa, T. C. Saido, H. Yamaguchi, A. Ruiz, A. Martinez, L. Madrigal, L. Hincapie, J. C. Arango L., D. C. Anthony, E. H. Koo, A. M. Goate, D. J. Selkoe, and J. C. Arango V. "The E280A presenilin 1 Alzheimer mutation produces increased A $\beta$ 42 deposition and severe cerebellar pathology". en. In: *Nature Medicine* 2.10 (1996), pp. 1146–1150. DOI: 10.1038/nm1096-1146.
- [150] A. K. Somavarapu and K. P. Kepp. "Membrane Dynamics of  $\gamma$ -Secretase Provides a Molecular Basis for  $\beta$ -Amyloid Binding and Processing". In: *ACS Chemical Neuroscience* 8.11 (2017), pp. 2424–2436. DOI: 10.1021/acscchemneuro.7b00208.
- [151] J. Y. Lee, Z. Feng, X.-Q. Xie, and I. Bahar. "Allosteric Modulation of Intact  $\gamma$ -Secretase Structural Dynamics". In: *Biophysical Journal* 113.12 (2017), pp. 2634–2649. DOI: 10.1016/j.bpj.2017.10.012.
- [152] R. Aguayo-Ortiz, C. Chávez-García, J. E. Straub, and L. Dominguez. "Characterizing the structural ensemble of  $\gamma$ -secretase using a multiscale molecular dynamics approach". en. In: *Chemical Science* 8.8 (2017), pp. 5576–5584. DOI: 10.1039/C7SC00980A.
- [153] M. Hitzenberger and M. Zacharias. " $\gamma$ -Secretase Studied by Atomistic Molecular Dynamics Simulations: Global Dynamics, Enzyme Activation, Water Distribution and Lipid Binding". In: *Frontiers in Chemistry* 6 (2019).
- [154] R. Mehra and K. P. Kepp. "Understanding familial Alzheimer's disease: The fit-stay-trim mechanism of  $\gamma$ -secretase". en. In: *WIREs Computational Molecular Science* 12.1 (2022), e1556. DOI: 10.1002/wcms.1556.
- [155] M. Hitzenberger and M. Zacharias. "Uncovering the Binding Mode of  $\gamma$ -Secretase Inhibitors". In: *ACS Chemical Neuroscience* 10.8 (2019), pp. 3398–3403. DOI: 10.1021/acscchemneuro.9b00272.
- [156] A. Bhattarai, S. Devkota, S. Bhattarai, M. S. Wolfe, and Y. Miao. "Mechanisms of  $\gamma$ -Secretase Activation and Substrate Processing". en. In: *ACS Central Science* 6.6 (2020), pp. 969–983. DOI: 10.1021/acscentsci.0c00296.
- [157] A. Bhattarai, S. Devkota, H. N. Do, J. Wang, S. Bhattarai, M. S. Wolfe, and Y. Miao. "Mechanism of Tripeptide Trimming of Amyloid  $\beta$ -Peptide 49 by  $\gamma$ -Secretase". en. In: *Journal of the American Chemical Society* (2022). DOI: 10.1021/jacs.1c10533.
- [158] B. Dehury, N. Tang, R. Mehra, T. L. Blundell, and K. P. Kepp. "Side-by-side comparison of Notch- and C83 binding to  $\gamma$ -secretase in a complete membrane model at physiological temperature". In: *RSC Advances* 10.52 (2020), pp. 31215–31232. DOI: 10.1039/d0ra04683c.
- [159] R. Mehra and K. P. Kepp. "Computational prediction and molecular mechanism of  $\gamma$ -secretase modulators". en. In: *European Journal of Pharmaceutical Sciences* 157 (2021), p. 105626. DOI: 10.1016/j.ejps.2020.105626.
- [160] N. Tang, A. K. Somavarapu, and K. P. Kepp. "Molecular Recipe for  $\gamma$ -Secretase Modulation from Computational Analysis of 60 Active Compounds". en. In: *ACS Omega* 3.12 (2018), pp. 18078–18088. DOI: 10.1021/acsomega.8b02196.

- [161] D. Petit, M. Hitzenberger, M. Koch, S. Lismont, K. M. Zoltowska, T. Enzlein, C. Hopf, M. Zacharias, and L. Chávez-Gutiérrez. “Enzyme–substrate interface targeting by imidazole-based  $\gamma$ -secretase modulators activates  $\gamma$ -secretase and stabilizes its interaction with APP”. In: *The EMBO Journal* 41.21 (2022), e111084. DOI: 10.15252/embj.2022111084.
- [162] B. Kav and M. S. Miettinen. *Amber Lipid17 Simulations of POPC/POPS Membranes with CaCl2*. eng. type: dataset. 2018. DOI: 10.5281/zenodo.1438848.
- [163] P. Mark and L. Nilsson. “Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K”. In: *The Journal of Physical Chemistry A* 105.43 (2001). Publisher: American Chemical Society, pp. 9954–9960. DOI: 10.1021/jp003020w.
- [164] C. J. Dickson, R. C. Walker, and I. R. Gould. “Lipid21: Complex Lipid Membrane Simulations with AMBER”. In: *Journal of Chemical Theory and Computation* 18.3 (2022), pp. 1726–1736. DOI: 10.1021/acs.jctc.1c01217.
- [165] W. Kabsch and C. Sander. “Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features”. en. In: *Biopolymers* 22.12 (1983), pp. 2577–2637. DOI: 10.1002/bip.360221211.
- [166] L. P. Feilen, S.-Y. Chen, A. Fukumori, R. Feederle, M. Zacharias, and H. Steiner. “Active site geometry stabilization of a presenilin homolog by the lipid bilayer promotes intramembrane proteolysis”. In: *eLife* 11 (2022). Ed. by M. J. Lemieux, J. D. Faraldo-Gómez, M. J. Lemieux, and S. Workman, e76090. DOI: 10.7554/eLife.76090.
- [167] E. Wallin and G. V. Heijne. “Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms”. en. In: *Protein Science* 7.4 (1998), pp. 1029–1038. DOI: 10.1002/pro.5560070420.
- [168] R. S. Hegde and R. J. Keenan. “The mechanisms of integral membrane protein biogenesis”. en. In: *Nature Reviews Molecular Cell Biology* 23.2 (2022), pp. 107–124. DOI: 10.1038/s41580-021-00413-2.
- [169] A. M. Seddon, P. Curnow, and P. J. Booth. “Membrane proteins, lipids and detergents: not just a soap opera”. en. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes*. Lipid-Protein Interactions 1666.1 (2004), pp. 105–117. DOI: 10.1016/j.bbamem.2004.04.011.
- [170] M. Orwick-Rydmark, T. Arnold, and D. Linke. “The Use of Detergents to Purify Membrane Proteins”. eng. In: *Current Protocols in Protein Science* 84 (2016), pp. 4.8.1–4.8.35. DOI: 10.1002/0471140864.ps0408s84.
- [171] M. D. Womack, D. A. Kendall, and R. C. MacDonald. “Detergent effects on enzyme activity and solubilization of lipid bilayer membranes”. en. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 733.2 (1983), pp. 210–215. DOI: 10.1016/0005-2736(83)90524-2.

- [172] V. Kurauskas, A. Hessel, P. Ma, P. Lunetti, K. Weinhäupl, L. Imbert, B. Brutscher, M. S. King, R. Sounier, V. Dolce, E. R. S. Kunji, L. Capobianco, C. Chipot, F. Dehez, B. Bersch, and P. Schanda. "How Detergent Impacts Membrane Proteins: Atomic-Level Views of Mitochondrial Carriers in Dodecylphosphocholine". In: *The Journal of Physical Chemistry Letters* 9.5 (2018), pp. 933–938. doi: 10.1021/acs.jpcllett.8b00269.
- [173] S. Paschkowsky, F. Oestereich, and L. M. Munter. "Embedded in the Membrane: How Lipids Confer Activity and Specificity to Intramembrane Proteases". en. In: *The Journal of Membrane Biology* 251.3 (2018), pp. 369–378. doi: 10.1007/s00232-017-0008-5.
- [174] H. Zhou, S. Zhou, P. J. Walian, and B. K. Jap. "Dependency of  $\gamma$ -secretase complex activity on the structural integrity of the bilayer". en. In: *Biochemical and Biophysical Research Communications* 402.2 (2010), pp. 291–296. doi: 10.1016/j.bbrc.2010.10.017.
- [175] P. Osenkowski, W. Ye, R. Wang, M. S. Wolfe, and D. J. Selkoe. "Direct and Potent Regulation of  $\gamma$ -Secretase by Its Lipid Microenvironment \*". English. In: *Journal of Biological Chemistry* 283.33 (2008), pp. 22529–22540. doi: 10.1074/jbc.M801925200.
- [176] O. Holmes, S. Paturi, W. Ye, M. S. Wolfe, and D. J. Selkoe. "Effects of Membrane Lipids on the Activity and Processivity of Purified  $\gamma$ -Secretase". In: *Biochemistry* 51.17 (2012), pp. 3565–3575. doi: 10.1021/bi300303g.
- [177] E. Winkler, F. Kamp, J. Scheuring, A. Ebke, A. Fukumori, and H. Steiner. "Generation of Alzheimer Disease-associated Amyloid  $\beta$ 42/43 Peptide by  $\gamma$ -Secretase Can Be Inhibited Directly by Modulation of Membrane Thickness \*". English. In: *Journal of Biological Chemistry* 287.25 (2012), pp. 21326–21334. doi: 10.1074/jbc.M112.356659.
- [178] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. de Beer, C. Rempfer, L. Bordoli, R. Lepore, and T. Schwede. "SWISS-MODEL: homology modelling of protein structures and complexes". In: *Nucleic Acids Research* 46.W1 (2018), W296–W303. doi: 10.1093/nar/gky427.
- [179] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D. MacKerell, J. B. Klauda, and W. Im. "CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field". eng. In: *Journal of Chemical Theory and Computation* 12.1 (2016), pp. 405–413. doi: 10.1021/acs.jctc.5b00935.
- [180] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, and A. D. MacKerell. "CHARMM36m: an improved force field for folded and intrinsically disordered proteins". en. In: *Nature Methods* 14.1 (2017), pp. 71–73. doi: 10.1038/nmeth.4067.
- [181] J. Hu, Y. Xue, S. Lee, and Y. Ha. "The crystal structure of GXGD membrane protease FlaK". en. In: *Nature* 475.7357 (2011), pp. 528–531. doi: 10.1038/nature10218.
- [182] A. W. White, A. D. Westwell, and G. Braheimi. "Protein-protein interactions as targets for small-molecule therapeutics in cancer". eng. In: *Expert Reviews in Molecular Medicine* 10 (2008), e8. doi: 10.1017/S1462399408000641.

- [183] L. L. Blazer and R. R. Neubig. "Small Molecule Protein–Protein Interaction Inhibitors as CNS Therapeutic Agents: Current Progress and Future Hurdles". en. In: *Neuropsychopharmacology* 34.1 (2009), pp. 126–141. doi: 10.1038/npp.2008.151.
- [184] T. E. Creighton. *Proteins : structures and molecular properties*. eng. New York : W.H. Freeman, 1993.
- [185] W. T. Kimberly, W. Xia, T. Rahmati, M. S. Wolfe, and D. J. Selkoe. "The transmembrane aspartates in presenilin 1 and 2 are obligatory for gamma-secretase activity and amyloid beta-protein generation". eng. In: *The Journal of Biological Chemistry* 275.5 (2000), pp. 3173–3178. doi: 10.1074/jbc.275.5.3173.
- [186] G.-f. Chen, T.-h. Xu, Y. Yan, Y.-r. Zhou, Y. Jiang, K. Melcher, and H. E. Xu. "Amyloid beta: structure, biology and structure-based therapeutic development". en. In: *Acta Pharmacologica Sinica* 38.9 (2017), pp. 1205–1235. doi: 10.1038/aps.2017.28.
- [187] M. A. Lomize, I. D. Pogozheva, H. Joo, H. I. Mosberg, and A. L. Lomize. "OPM database and PPM web server: resources for positioning of proteins in membranes". eng. In: *Nucleic Acids Research* 40.Database issue (2012), pp. D370–376. doi: 10.1093/nar/gkr703.
- [188] T.-S. Lee, D. S. Cerutti, D. Mermelstein, C. Lin, S. LeGrand, T. J. Giese, A. Roitberg, D. A. Case, R. C. Walker, and D. M. York. "GPU-accelerated molecular dynamics and free energy methods in Amber18: performance enhancements and new features". In: *Journal of chemical information and modeling* 58.10 (2018), pp. 2043–2050. doi: 10.1021/acs.jcim.8b00462.
- [189] R. Mehra, B. Dehury, and K. P. Kepp. "Cryo-temperature effects on membrane protein structure and dynamics". en. In: *Physical Chemistry Chemical Physics* 22.10 (2020). Publisher: The Royal Society of Chemistry, pp. 5427–5438. doi: 10.1039/C9CP06723J.
- [190] D. Edbauer, E. Winkler, J. T. Regula, B. Pesold, H. Steiner, and C. Haass. "Reconstitution of  $\gamma$ -secretase activity". en. In: *Nature Cell Biology* 5.5 (2003), pp. 486–488. doi: 10.1038/ncb960.
- [191] H. Acx, L. Chávez-Gutiérrez, L. Serneels, S. Lismont, M. Benurwar, N. Elad, and B. De Strooper. "Signature Amyloid  $\beta$  Profiles Are Produced by Different  $\gamma$ -Secretase Complexes". In: *The Journal of Biological Chemistry* 289.7 (2014), pp. 4346–4355. doi: 10.1074/jbc.M113.530907.
- [192] M. Dimitrov, J.-R. Alattia, T. Lemmin, R. Lehal, A. Fligier, J. Houacine, I. Hussain, F. Radtke, M. Dal Peraro, D. Beher, and P. C. Fraering. "Alzheimer's disease mutations in APP but not  $\gamma$ -secretase modulators affect epsilon-cleavage-dependent AICD production". en. In: *Nature Communications* 4.1 (2013), p. 2246. doi: 10.1038/ncomms3246.
- [193] S. Devkota, T. D. Williams, and M. S. Wolfe. "Familial Alzheimer's disease mutations in amyloid protein precursor alter proteolysis by  $\gamma$ -secretase to increase amyloid  $\beta$ -peptides of  $\geq 45$  residues". en. In: *Journal of Biological Chemistry* 296 (2021), p. 100281. doi: 10.1016/j.jbc.2021.100281.

- [194] S. Zoll, S. Stanchev, J. Began, J. Škerle, M. Lepšík, L. Peclinovská, P. Majer, and K. Strisovsky. "Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate–peptide complex structures". In: *The EMBO Journal* 33.20 (2014), pp. 2408–2421. doi: 10.15252/embj.201489367.
- [195] J. D. A. Tyndall, T. Nall, and D. P. Fairlie. "Proteases Universally Recognize Beta Strands In Their Active Sites". In: *Chemical Reviews* 105.3 (2005), pp. 973–1000. doi: 10.1021/cr040669e.
- [196] S. F. Lichtenthaler, C. Haass, and H. Steiner. "Regulated intramembrane proteolysis – lessons from amyloid precursor protein processing". en. In: *Journal of Neurochemistry* 117.5 (2011), pp. 779–796. doi: 10.1111/j.1471-4159.2011.07248.x.
- [197] Y. Gu, H. Misonou, T. Sato, N. Dohmae, K. Takio, and Y. Ihara. "Distinct intramembrane cleavage of the beta-amyloid precursor protein family resembling gamma-secretase-like cleavage of Notch". eng. In: *The Journal of Biological Chemistry* 276.38 (2001), pp. 35235–35238. doi: 10.1074/jbc.C100357200.
- [198] M. Sastre, H. Steiner, K. Fuchs, A. Capell, G. Multhaup, M. M. Condron, D. B. Teplow, and C. Haass. "Presenilin-dependent  $\gamma$ -secretase processing of  $\beta$ -amyloid precursor protein at a site corresponding to the S3 cleavage of Notch". In: *EMBO reports* 2.9 (2001), pp. 835–841. doi: 10.1093/embo-reports/kve180.
- [199] C. Yu, S.-H. Kim, T. Ikeuchi, H. Xu, L. Gasparini, R. Wang, and S. S. Sisodia. "Characterization of a Presenilin-mediated Amyloid Precursor Protein Carboxyl-terminal Fragment  $\gamma$  EVIDENCE FOR DISTINCT MECHANISMS INVOLVED IN  $\gamma$ -SECRETASE PROCESSING OF THE APP AND Notch1 TRANSMEMBRANE DOMAINS". en. In: *Journal of Biological Chemistry* 276.47 (2001), pp. 43756–43760. doi: 10.1074/jbc.C100410200.
- [200] A. Weidemann, S. Eggert, F. B. M. Reinhard, M. Vogel, K. Paliga, G. Baier, C. L. Masters, K. Beyreuther, and G. Evin. "A Novel  $\epsilon$ -Cleavage within the Transmembrane Domain of the Alzheimer Amyloid Precursor Protein Demonstrates Homology with Notch Processing". In: *Biochemistry* 41.8 (2002), pp. 2825–2835. doi: 10.1021/bi015794o.
- [201] T. Iwatsubo, A. Odaka, N. Suzuki, H. Mizusawa, N. Nukina, and Y. Ihara. "Visualization of A beta 42(43) and A beta 40 in senile plaques with end-specific A beta monoclonals: evidence that an initially deposited species is A beta 42(43)". eng. In: *Neuron* 13.1 (1994), pp. 45–53. doi: 10.1016/0896-6273(94)90458-8.
- [202] R. J. Bateman, E. R. Siemers, K. G. Mawuenyega, G. Wen, K. R. Browning, W. C. Sigurdson, K. E. Yarasheski, S. W. Friedrich, R. B. DeMattos, P. C. May, S. M. Paul, and D. M. Holtzman. "A gamma-secretase inhibitor decreases amyloid-beta production in the central nervous system". In: *Annals of neurology* 66.1 (2009), pp. 48–54. doi: 10.1002/ana.21623.
- [203] A. T. Philip, S. Devkota, S. Malvankar, S. Bhattarai, K. M. Meneely, T. D. Williams, and M. S. Wolfe. "Designed helical peptides as functional probes for  $\gamma$ -secretase". In: *Biochemistry* 58.44 (2019), pp. 4398–4407. doi: 10.1021/acs.biochem.9b00639.

- [204] F. Bihel, C. Das, M. J. Bowman, and M. S. Wolfe. "Discovery of a Subnanomolar helical D-tridecapeptide inhibitor of gamma-secretase". eng. In: *Journal of Medicinal Chemistry* 47.16 (2004), pp. 3931–3933. doi: 10.1021/jm049788c.
- [205] L. Hanbouch, B. Schaack, A. Kasri, G. Fontaine, E. Gkanatsiou, G. Brinkmalm, E. Camporesi, E. Portelius, K. Blennow, G. Mourier, N. Gilles, M. J. Millan, C. Marquer, H. Zetterberg, L. Boussicault, and M.-C. Potier. "Specific Mutations in the Cholesterol-Binding Site of APP Alter Its Processing and Favor the Production of Shorter, Less Toxic A $\beta$  Peptides". In: *Molecular Neurobiology* 59.11 (2022), pp. 7056–7073. doi: 10.1007/s12035-022-03025-9.
- [206] A. J. Beel, P. Barrett, P. D. Schnier, S. A. Hitchcock, D. Bagal, C. R. Sanders, and J. B. Jordan. "Nonspecificity of Binding of  $\gamma$ -Secretase Modulators to the Amyloid Precursor Protein". In: *Biochemistry* 48.50 (2009), pp. 11837–11839. doi: 10.1021/bi901839d.
- [207] N. Gertsik, C. W. am Ende, K. F. Geoghegan, C. Nguyen, P. Mukherjee, S. Mente, U. Seneviratne, D. S. Johnson, and Y.-M. Li. "Mapping the Binding Site of BMS-708163 on  $\gamma$ -Secretase with Cleavable Photoprobes". en. In: *Cell Chemical Biology* 24.1 (2017), pp. 3–8. doi: 10.1016/j.chembiol.2016.12.006.
- [208] A. Y. Kornilova, F. Bihel, C. Das, and M. S. Wolfe. "The initial substrate-binding site of  $\gamma$ -secretase is located on presenilin near the active site". eng. In: *Proceedings of the National Academy of Sciences* 102.9 (2005), pp. 3230–3235. doi: 10.1073/pnas.0407640102.
- [209] M. S. Wolfe. "Probing Mechanisms and Therapeutic Potential of  $\gamma$ -Secretase in Alzheimer's Disease". In: *Molecules* 26.2 (2021), p. 388. doi: 10.3390/molecules26020388.
- [210] D. Petit, M. Hitzenberger, S. Lismont, K. M. Zoltowska, N. S. Ryan, M. Mercken, F. Bischoff, M. Zacharias, and L. Chávez-Gutiérrez. "Extracellular interface between APP and Nicastrin regulates A $\beta$  length and response to  $\gamma$ -secretase modulators". In: *The EMBO Journal* 38.12 (2019), e101494. doi: 10.15252/embj.2019101494.
- [211] T. L. Kukar, T. B. Ladd, P. Robertson, S. A. Pintchovski, B. Moore, M. A. Bann, Z. Ren, K. Jansen-West, K. Malphrus, S. Eggert, H. Maruyama, B. A. Cottrell, P. Das, G. S. Basi, E. H. Koo, and T. E. Golde. "Lysine 624 of the Amyloid Precursor Protein (APP) Is a Critical Determinant of Amyloid  $\beta$  Peptide Length: SUPPORT FOR A SEQUENTIAL MODEL OF  $\gamma$ -SECRETASE INTRAMEMBRANE PROTEOLYSIS AND REGULATION BY THE AMYLOID  $\beta$  PRECURSOR PROTEIN (APP) JUXTAMEMBRANE REGION \*". English. In: *Journal of Biological Chemistry* 286.46 (2011). Publisher: Elsevier, pp. 39804–39812. doi: 10.1074/jbc.M111.274696.
- [212] J. I. Jung, Y. Ran, P. E. Cruz, A. M. Rosario, T. B. Ladd, T. L. Kukar, E. H. Koo, K. M. Felsenstein, and T. E. Golde. "Complex Relationships between Substrate Sequence and Sensitivity to Alterations in  $\gamma$ -Secretase Processivity Induced by  $\gamma$ -Secretase Modulators". In: *Biochemistry* 53.12 (2014). Publisher: American Chemical Society, pp. 1947–1957. doi: 10.1021/bi401521t.

- [213] L.-M. Munter, P. Voigt, A. Harmeier, D. Kaden, K. E. Gottschalk, C. Weise, R. Pipkorn, M. Schaefer, D. Langosch, and G. Multhaup. "GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of A $\beta$ 42". In: *The EMBO Journal* 26.6 (2007), pp. 1702–1712. doi: 10.1038/sj.emboj.7601616.
- [214] M. A. Fernandez, K. Biette, G. Dolios, D. Seth, R. Wang, and M. S. Wolfe. "Transmembrane substrate determinants for  $\gamma$ -secretase processing of APP CTF $\beta$ ". In: *Biochemistry* 55.40 (2016), pp. 5675–5688. doi: 10.1021/acs.biochem.6b00718.
- [215] O. G. Mouritsen and M. Bloom. "Mattress model of lipid-protein interactions in membranes." In: *Biophysical Journal* 46.2 (1984), pp. 141–153.
- [216] N. Kühnle, V. Dederer, and M. K. Lemberg. "Intramembrane proteolysis at a glance: from signalling to protein degradation". In: *Journal of Cell Science* 132.16 (2019). doi: 10.1242/jcs.217745.
- [217] S. Düsterhöft, U. Künzel, and M. Freeman. "Rhomboid proteases in human disease: Mechanisms and future prospects". en. In: *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. Proteolysis as a Regulatory Event in Pathophysiology 1864.11, Part B (2017), pp. 2200–2209. doi: 10.1016/j.bbamcr.2017.04.016.
- [218] S. H. L. Verhelst. "Intramembrane proteases as drug targets". en. In: *The FEBS Journal* 284.10 (2017), pp. 1489–1502. doi: 10.1111/febs.13979.
- [219] M. Drag and G. S. Salvesen. "Emerging principles in protease-based drug discovery". en. In: *Nature Reviews Drug Discovery* 9.9 (2010), pp. 690–701. doi: 10.1038/nrd3053.
- [220] B. Turk. "Targeting proteases: successes, failures and future prospects". en. In: *Nature Reviews Drug Discovery* 5.9 (2006), pp. 785–799. doi: 10.1038/nrd2092.
- [221] M. Okochi, S. Tagami, K. Yanagida, M. Takami, T. S. Kodama, K. Mori, T. Nakayama, Y. Ihara, and M. Takeda. " $\gamma$ -Secretase Modulators and Presenilin 1 Mutants Act Differently on Presenilin/ $\gamma$ -Secretase Function to Cleave A $\beta$ 42 and A $\beta$ 43". English. In: *Cell Reports* 3.1 (2013), pp. 42–51. doi: 10.1016/j.celrep.2012.11.028.
- [222] T. Ratovitski, H. H. Slunt, G. Thinakaran, D. L. Price, S. S. Sisodia, and D. R. Borchelt. "Endoproteolytic Processing and Stabilization of Wild-type and Mutant Presenilin \*". English. In: *Journal of Biological Chemistry* 272.39 (1997), pp. 24536–24541. doi: 10.1074/jbc.272.39.24536.
- [223] J. I. Jung, S. Premraj, P. E. Cruz, T. B. Ladd, Y. Kwak, E. H. Koo, K. M. Felsenstein, T. E. Golde, and Y. Ran. "Independent Relationship between Amyloid Precursor Protein (APP) Dimerization and  $\gamma$ -Secretase Processivity". en. In: *PLOS ONE* 9.10 (2014), e111553. doi: 10.1371/journal.pone.0111553.
- [224] P. Scheltens, K. Blennow, M. M. B. Breteler, B. de Strooper, G. B. Frisoni, S. Salloway, and W. M. Van der Flier. "Alzheimer's disease". en. In: *The Lancet* 388.10043 (2016), pp. 505–517. doi: 10.1016/S0140-6736(15)01124-1.
- [225] W. H. Organization. *Dementia*. 2022. URL: <https://www.who.int/news-room/factsheets/detail/dementia> (visited on 09/30/2010).

- [226] Y. Qi-Takahara, M. Morishima-Kawashima, Y. Tanimura, G. Dolios, N. Hirotsu, Y. Horikoshi, F. Kametani, M. Maeda, T. C. Saido, R. Wang, and Y. Ihara. "Longer Forms of Amyloid  $\beta$  Protein: Implications for the Mechanism of Intramembrane Cleavage by  $\gamma$ -Secretase". In: *Journal of Neuroscience* 25.2 (2005), pp. 436–445.
- [227] V. Coric, C. H. van Dyck, S. Salloway, N. Andreasen, M. Brody, R. W. Richter, H. Soininen, S. Thein, T. Shiovitz, G. Pilcher, S. Colby, L. Rollin, R. Dockens, C. Pachai, E. Portelius, U. Andreasson, K. Blennow, H. Soares, C. Albright, H. H. Feldman, and R. M. Berman. "Safety and tolerability of the  $\gamma$ -secretase inhibitor avagacestat in a phase 2 study of mild to moderate Alzheimer disease". eng. In: *Archives of Neurology* 69.11 (2012), pp. 1430–1440. doi: 10.1001/archneuro.1.2012.2194.
- [228] T. Cai, K. Morishima, S. Takagi-Niidome, A. Tominaga, and T. Tomita. "Conformational Dynamics of Transmembrane Domain 3 of Presenilin 1 Is Associated with the Trimming Activity of  $\gamma$ -Secretase". en. In: *Journal of Neuroscience* 39.43 (2019), pp. 8600–8610. doi: 10.1523/JNEUROSCI.0838-19.2019.
- [229] A. Tominaga, T. Cai, S. Takagi-Niidome, T. Iwatsubo, and T. Tomita. "Conformational Changes in Transmembrane Domain 4 of Presenilin 1 Are Associated with Altered Amyloid-42 Production". en. In: *Journal of Neuroscience* 36.4 (2016), pp. 1362–1372. doi: 10.1523/JNEUROSCI.5090-14.2016.
- [230] K. Uemura, K. C. Farner, T. Hashimoto, N. Nasser-Ghodsi, M. S. Wolfe, E. H. Koo, B. T. Hyman, and O. Berezovska. "Substrate docking to  $\gamma$ -secretase allows access of  $\gamma$ -secretase modulators to an allosteric site". en. In: *Nature Communications* 1.1 (2010), p. 130. doi: 10.1038/ncomms1129.
- [231] K. Uemura, C. M. Lill, X. Li, J. A. Peters, A. Ivanov, Z. Fan, B. DeStrooper, B. J. Bacskai, B. T. Hyman, and O. Berezovska. "Allosteric Modulation of PS1/ $\gamma$ -Secretase Conformation Correlates with Amyloid  $\beta$ 42/40 Ratio". en. In: *PLOS ONE* 4.11 (2009), e7893. doi: 10.1371/journal.pone.0007893.
- [232] F. Raven, J. F. Ward, K. M. Zoltowska, Y. Wan, E. Bylykbashi, S. J. Miller, X. Shen, S. H. Choi, K. D. Ryneerson, O. Berezovska, S. L. Wagner, R. E. Tanzi, and C. Zhang. "Soluble Gamma-secretase Modulators Attenuate Alzheimer's  $\beta$ -amyloid Pathology and Induce Conformational Changes in Presenilin 1". In: *EBioMedicine* 24 (2017), pp. 93–101. doi: 10.1016/j.ebiom.2017.08.028.
- [233] T. Cai, M. Yonaga, and T. Tomita. "Activation of  $\gamma$ -Secretase Trimming Activity by Topological Changes of Transmembrane Domain 1 of Presenilin 1". en. In: *Journal of Neuroscience* 37.50 (2017), pp. 12272–12280. doi: 10.1523/JNEUROSCI.1628-17.2017.
- [234] L. Liu, B. M. Lauro, M. S. Wolfe, and D. J. Selkoe. "Hydrophilic loop 1 of Presenilin-1 and the APP GxxxG transmembrane motif regulate  $\gamma$ -secretase function in generating Alzheimer-causing A $\beta$  peptides". English. In: *Journal of Biological Chemistry* 296 (2021). doi: 10.1016/j.jbc.2021.100393.



- [235] V. Le Guilloux, P. Schmidtke, and P. Tuffery. "Fpocket: An open source platform for ligand pocket detection". In: *BMC Bioinformatics* 10.1 (2009), p. 168. DOI: 10.1186/1471-2105-10-168.
- [236] P. Schmidtke, A. Bidon-Chanal, F. Luque, and X. Barril. "MDpocket : Open Source Cavity Detection and Characterization on Molecular Dynamics Trajectories". In: *Bioinformatics (Oxford, England)* 27 (2011), pp. 3276–85. DOI: 10.1093/bioinformatics/btr550.
- [237] B. Jayaram, T. Singh, G. Mukherjee, A. Mathur, S. Shekhar, and V. Shekhar. "Sanjeevini: a freely accessible web-server for target directed lead molecule discovery". In: *BMC Bioinformatics* 13.17 (2012), S7. DOI: 10.1186/1471-2105-13-S17-S7.
- [238] O. Trott and A. J. Olson. "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading". In: *Journal of computational chemistry* 31.2 (2010), pp. 455–461. DOI: 10.1002/jcc.21334.
- [239] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison. "Open Babel: An open chemical toolbox". In: *Journal of Cheminformatics* 3.1 (2011), p. 33. DOI: 10.1186/1758-2946-3-33.
- [240] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison. "Avogadro: an advanced semantic chemical editor, visualization, and analysis platform". In: *Journal of Cheminformatics* 4.1 (2012), p. 17. DOI: 10.1186/1758-2946-4-17.
- [241] T. A. Halgren. "Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules". en. In: *Journal of Computational Chemistry* 17.5-6 (1996), pp. 616–641. DOI: 10.1002/(SICI)1096-987X(199604)17:5/6<616::AID-JCC5>3.0.CO;2-X.
- [242] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler. "Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions". en. In: *New Journal of Physics* 14.5 (2012), p. 053020. DOI: 10.1088/1367-2630/14/5/053020.
- [243] F. Neese, F. Wennmohs, U. Becker, and C. Riplinger. "The ORCA quantum chemistry program package". In: *The Journal of Chemical Physics* 152.22 (2020), p. 224108. DOI: 10.1063/5.0004608.
- [244] G. L. Stoychev, A. A. Auer, and F. Neese. "Automatic Generation of Auxiliary Basis Sets". In: *Journal of Chemical Theory and Computation* 13.2 (2017), pp. 554–562. DOI: 10.1021/acs.jctc.6b01041.
- [245] F. Neese. "An improvement of the resolution of the identity approximation for the formation of the Coulomb matrix". en. In: *Journal of Computational Chemistry* 24.14 (2003), pp. 1740–1747. DOI: 10.1002/jcc.10318.
- [246] T. Lu and F. Chen. "Multiwfn: A multifunctional wavefunction analyzer". en. In: *Journal of Computational Chemistry* 33.5 (2012), pp. 580–592. DOI: 10.1002/jcc.22885.

- [247] R. A. Laskowski and M. B. Swindells. "LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery". In: *Journal of Chemical Information and Modeling* 51.10 (2011). Publisher: American Chemical Society, pp. 2778–2786. DOI: 10.1021/ci200227u.
- [248] T. E. Golde, E. H. Koo, K. M. Felsenstein, B. A. Osborne, and L. Miele. "γ-Secretase inhibitors and modulators". en. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes. Intramembrane Proteases* 1828.12 (2013), pp. 2898–2907. DOI: 10.1016/j.bbamem.2013.06.005.
- [249] J. Trambauer, R. M. R. Sarmiento, A. Fukumori, R. Feederle, K. Baumann, and H. Steiner. "Aβ43-producing PS1 FAD mutants cause altered substrate interactions and respond to γ-secretase modulation". en. In: *EMBO reports* (2019). DOI: 10.15252/embr.201947996.
- [250] D. T. Infield, A. Rasouli, G. D. Galles, C. Chipot, E. Tajkhorshid, and C. A. Ahern. "Cation-π interactions and their functional roles in membrane proteins". In: *Journal of molecular biology* 433.17 (2021), p. 167035. DOI: 10.1016/j.jmb.2021.167035.
- [251] A. S. Mahadevi and G. N. Sastry. "Cation-π Interaction: Its Role and Relevance in Chemistry, Biology, and Material Science". en. In: *Chemical Reviews* 113.3 (2013), pp. 2100–2138. DOI: 10.1021/cr300222d.
- [252] J. P. Gallivan and D. A. Dougherty. "Cation-π interactions in structural biology". In: *Proceedings of the National Academy of Sciences* 96.17 (1999), pp. 9459–9464. DOI: 10.1073/pnas.96.17.9459.
- [253] D. A. DOUGHERTY. "The Cation-π Interaction". In: *Accounts of chemical research* 46.4 (2013), pp. 885–893. DOI: 10.1021/ar300265y.
- [254] M. R. Arkin and A. Whitty. "The road less traveled: modulating signal transduction enzymes by inhibiting their protein-protein interactions". en. In: *Current Opinion in Chemical Biology. Next Generation Therapeutics* 13.3 (2009), pp. 284–290. DOI: 10.1016/j.cbpa.2009.05.125.
- [255] B. J. Mayer. "Protein-Protein Interactions in Signaling Cascades". In: *Transmembrane Signaling Protocols*. Ed. by H. Ali and B. Haribabu. Totowa, NJ: Humana Press, 2006, pp. 79–99. DOI: 10.1385/1-59745-048-0:79.
- [256] T. Pawson and P. Nash. "Protein-protein interactions define specificity in signal transduction". en. In: *Genes & Development* 14.9 (2000), pp. 1027–1047. DOI: 10.1101/gad.14.9.1027.
- [257] D. E. Scott, A. R. Bayly, C. Abell, and J. Skidmore. "Small molecules, big targets: drug discovery faces the protein-protein interaction challenge". en. In: *Nature Reviews Drug Discovery* 15.8 (2016), pp. 533–550. DOI: 10.1038/nrd.2016.29.
- [258] J. A. Wells and C. L. McClendon. "Reaching for high-hanging fruit in drug discovery at protein-protein interfaces". en. In: *Nature* 450.7172 (2007), pp. 1001–1009. DOI: 10.1038/nature06526.

- [259] A. Loregian and G. Palù. “Disruption of protein–protein interactions: Towards new targets for chemotherapy”. en. In: *Journal of Cellular Physiology* 204.3 (2005), pp. 750–762. doi: 10.1002/jcp.20356.
- [260] I. Petta, S. Lievens, C. Libert, J. Tavernier, and K. De Bosscher. “Modulation of Protein-Protein Interactions for the Development of Novel Therapeutics”. eng. In: *Molecular Therapy: The Journal of the American Society of Gene Therapy* 24.4 (2016), pp. 707–718. doi: 10.1038/mt.2015.214.
- [261] S. A. Andrei, E. Sijbesma, M. Hann, J. Davis, G. O’Mahony, M. W. D. Perry, A. Karawajczyk, J. Eickhoff, L. Brunsveld, R. G. Doveston, L.-G. Milroy, and C. Ottmann. “Stabilization of protein-protein interactions in drug discovery”. eng. In: *Expert Opinion on Drug Discovery* 12.9 (2017), pp. 925–940. doi: 10.1080/17460441.2017.1346608.
- [262] P. Thiel, M. Kaiser, and C. Ottmann. “Small-Molecule Stabilization of Protein–Protein Interactions: An Underestimated Concept in Drug Discovery?” en. In: *Angewandte Chemie International Edition* 51.9 (2012), pp. 2012–2018. doi: 10.1002/anie.201107616.
- [263] L. L. Conte, C. Chothia, and J. Janin. “The atomic structure of protein-protein recognition”. en. In: *Journal of Molecular Biology* 285.5 (1999), pp. 2177–2198. doi: 10.1006/jmbi.1998.2439.
- [264] M. C. Smith and J. E. Gestwicki. “Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity”. eng. In: *Expert Reviews in Molecular Medicine* 14 (2012), e16. doi: 10.1017/erm.2012.10.
- [265] P. Buchwald. “Small-molecule protein-protein interaction inhibitors: therapeutic potential in light of molecular size, chemical space, and ligand binding efficiency considerations”. eng. In: *IUBMB life* 62.10 (2010), pp. 724–731. doi: 10.1002/iub.383.
- [266] M. R. Arkin and J. A. Wells. “Small-molecule inhibitors of protein–protein interactions: progressing towards the dream”. en. In: *Nature Reviews Drug Discovery* 3.4 (2004), pp. 301–317. doi: 10.1038/nrd1343.
- [267] L. Nevola and E. Giralt. “Modulating protein–protein interactions: the potential of peptides”. en. In: *Chemical Communications* 51.16 (2015), pp. 3302–3315. doi: 10.1039/C4CC08565E.
- [268] W. Cabri, P. Cantelmi, D. Corbisiero, T. Fantoni, L. Ferrazzano, G. Martelli, A. Mattellone, and A. Tolomelli. “Therapeutic Peptides Targeting PPI in Clinical Development: Overview, Mechanism of Action and Perspectives”. In: *Frontiers in Molecular Biosciences* 8 (2021).
- [269] E. M. Driggers, S. P. Hale, J. Lee, and N. K. Terrett. “The exploration of macrocycles for drug discovery — an underexploited structural class”. en. In: *Nature Reviews Drug Discovery* 7.7 (2008), pp. 608–624. doi: 10.1038/nrd2590.
- [270] Z. Qian, P. G. Dougherty, and D. Pei. “Targeting intracellular protein-protein interactions with cell-permeable cyclic peptides”. In: *Current opinion in chemical biology* 38 (2017), pp. 80–86. doi: 10.1016/j.cbpa.2017.03.011.

- [271] S. Rubin and N. Qvit. "Cyclic Peptides for Protein-Protein Interaction Targets: Applications to Human Disease". English. In: *Critical Reviews & Trade; in Eukaryotic Gene Expression* 26.3 (2016). doi: 10.1615/CritRevEukaryotGeneExpr.2016016525.
- [272] B. L. Santini and M. Zacharias. "Rapid in silico Design of Potential Cyclic Peptide Binders Targeting Protein-Protein Interfaces". In: *Frontiers in Chemistry* 8 (2020).
- [273] F. Da Silva, G. Bret, L. Teixeira, C. F. Gonzalez, and D. Rognan. "Exhaustive Repertoire of Druggable Cavities at Protein-Protein Interfaces of Known Three-Dimensional Structure". en. In: *Journal of Medicinal Chemistry* 62.21 (2019), pp. 9732–9742. doi: 10.1021/acs.jmedchem.9b01184.
- [274] M. Gao and J. Skolnick. "The distribution of ligand-binding pockets around protein-protein interfaces suggests a general mechanism for pocket formation". In: *Proceedings of the National Academy of Sciences* 109.10 (2012), pp. 3784–3789. doi: 10.1073/pnas.1117768109.
- [275] B. Zarzycka, M. A. Kuenemann, M. A. Miteva, G. A. F. Nicolaes, G. Vriend, and O. Sperandio. "Stabilization of protein-protein interaction complexes through small molecules". en. In: *Drug Discovery Today* 21.1 (2016), pp. 48–57. doi: 10.1016/j.drudis.2015.09.011.
- [276] M. Wolter, D. Valenti, P. J. Cossar, L. M. Levy, S. Hristeva, T. Genski, T. Hoffmann, L. Brunsveld, D. Tzalis, and C. Ottmann. "Fragment-Based Stabilizers of Protein-Protein Interactions through Imine-Based Tethering". en. In: *Angewandte Chemie International Edition* 59.48 (2020), pp. 21520–21524. doi: 10.1002/anie.202008585.
- [277] X. Guillory, M. Wolter, S. Leysen, J. F. Neves, A. Kuusk, S. Genet, B. Somsen, J. K. Morrow, E. Rivers, L. van Beek, J. Patel, R. Goodnow, H. Schoenherr, N. Fuller, Q. Cao, R. G. Doveston, L. Brunsveld, M. R. Arkin, P. Castaldi, H. Boyd, I. Landrieu, H. Chen, and C. Ottmann. "Fragment-based Differential Targeting of PPI Stabilizer Interfaces". In: *Journal of Medicinal Chemistry* 63.13 (2020), pp. 6694–6707. doi: 10.1021/acs.jmedchem.9b01942.
- [278] E. Sijbesma, K. K. Hallenbeck, S. Leysen, P. J. de Vink, L. Skóra, W. Jahnke, L. Brunsveld, M. R. Arkin, and C. Ottmann. "Site-Directed Fragment-Based Screening for the Discovery of Protein-Protein Interaction Stabilizers". In: *Journal of the American Chemical Society* 141.8 (2019), pp. 3524–3531. doi: 10.1021/jacs.8b11658.
- [279] E. Sijbesma, E. Visser, K. Plitzko, P. Thiel, L.-G. Milroy, M. Kaiser, L. Brunsveld, and C. Ottmann. "Structure-based evolution of a promiscuous inhibitor to a selective stabilizer of protein-protein interactions". en. In: *Nature Communications* 11.1 (2020), p. 3954. doi: 10.1038/s41467-020-17741-0.
- [280] R. Tang, P. Chen, Z. Wang, L. Wang, H. Hao, T. Hou, and H. Sun. "Characterizing the stabilization effects of stabilizers in protein-protein systems with end-point binding free energy calculations". In: *Briefings in Bioinformatics* 23.3 (2022), bbac127. doi: 10.1093/bib/bbac127.

- [281] M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. D. Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J. H. Pereira, A. V. Rodrigues, A. A. van Dijk, A. C. Ebrecht, D. J. Opperman, T. Sagmeister, C. Buhheller, T. Pavkov-Keller, M. K. Rathinaswamy, U. Dalwadi, C. K. Yip, J. E. Burke, K. C. Garcia, N. V. Grishin, P. D. Adams, R. J. Read, and D. Baker. "Accurate prediction of protein structures and interactions using a three-track neural network". In: *Science* 373.6557 (2021), pp. 871–876. doi: 10.1126/science.abj8754.
- [282] J. I. Garzon, J. R. Lopéz-Blanco, C. Pons, J. Kovacs, R. Abagyan, J. Fernandez-Recio, and P. Chacon. "FRODOCK: a new approach for fast rotational protein–protein docking". In: *Bioinformatics* 25.19 (2009), pp. 2544–2551. doi: 10.1093/bioinformatics/btp447.
- [283] B. G. Pierce, K. Wiehe, H. Hwang, B.-H. Kim, T. Vreven, and Z. Weng. "ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers". In: *Bioinformatics* 30.12 (2014), pp. 1771–1773. doi: 10.1093/bioinformatics/btu097.
- [284] C. Tian, K. Kasavajhala, K. A. A. Belfon, L. Raguette, H. Huang, A. N. Miguez, J. Bickel, Y. Wang, J. Pincay, Q. Wu, and C. Simmerling. "ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution". In: *Journal of Chemical Theory and Computation* 16.1 (2020), pp. 528–552. doi: 10.1021/acs.jctc.9b00591.
- [285] S. Izadi, R. Anandakrishnan, and A. V. Onufriev. "Building Water Models: A Different Approach". In: *The Journal of Physical Chemistry Letters* 5.21 (2014), pp. 3863–3871. doi: 10.1021/jz501780a.
- [286] I. S. Joung and T. E. I. Cheatham. "Molecular Dynamics Simulations of the Dynamic and Energetic Properties of Alkali and Halide Ions Using Water-Model-Specific Ion Parameters". In: *The Journal of Physical Chemistry B* 113.40 (2009), pp. 13279–13290. doi: 10.1021/jp902584c.
- [287] A. Onufriev, D. Bashford, and D. A. Case. "Exploring protein native states and large-scale conformational changes with a modified generalized born model". en. In: *Proteins: Structure, Function, and Bioinformatics* 55.2 (2004), pp. 383–394. doi: 10.1002/prot.20033.
- [288] J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield, and R. A. Sayle. "ZINC20—A Free Ultralarge-Scale Chemical Database for Ligand Discovery". In: *Journal of Chemical Information and Modeling* 60.12 (2020), pp. 6065–6073. doi: 10.1021/acs.jcim.0c00675.
- [289] S. H. Bertz. "Convergence, molecular complexity, and synthetic analysis". In: *Journal of the American Chemical Society* 104.21 (1982), pp. 5801–5803. doi: 10.1021/ja00385a049.
- [290] G. Landrum. "RDKit: Open-Source Cheminformatics Software". In: (2016).

- [291] D. Santos-Martins, L. Solis-Vasquez, A. F. Tillack, M. F. Sanner, A. Koch, and S. Forli. "Accelerating AutoDock4 with GPUs and Gradient-Based Local Search". In: *Journal of Chemical Theory and Computation* 17.2 (2021), pp. 1060–1073. DOI: 10.1021/acs.jctc.0c01006.
- [292] J. Yang and W. S. Hlavacek. "Scaffold-mediated nucleation of protein signaling complexes: Elementary principles". en. In: *Mathematical Biosciences* 232.2 (2011), pp. 164–173. DOI: 10.1016/j.mbs.2011.06.003.
- [293] J. Xu and Y. Zhang. "How significant is a protein structure similarity with TM-score = 0.5?" In: *Bioinformatics* 26.7 (2010), pp. 889–895. DOI: 10.1093/bioinformatics/btq066.
- [294] Y. Zhang and J. Skolnick. "Scoring function for automated assessment of protein structure template quality". en. In: *Proteins: Structure, Function, and Bioinformatics* 57.4 (2004), pp. 702–710. DOI: 10.1002/prot.20264.
- [295] R. Rose, S. Erdmann, S. Bovens, A. Wolf, M. Rose, S. Hennig, H. Waldmann, and C. Ottmann. "Identification and Structure of Small-Molecule Stabilizers of 14–3–3 Protein–Protein Interactions". In: *Angewandte Chemie International Edition* 49.24 (2010), pp. 4129–4132. DOI: 10.1002/anie.200907203.
- [296] K. M. Zak, P. Grudnik, K. Guzik, B. J. Zieba, B. Musielak, A. Dömling, G. Dubin, and T. A. Holak. "Structural basis for small molecule targeting of the programmed death ligand 1 (PD-L1)". en. In: *Oncotarget* 7.21 (2016), pp. 30323–30335. DOI: 10.18632/oncotarget.8730.
- [297] M. J. Waring, H. Chen, A. A. Rabow, G. Walker, R. Bobby, S. Boiko, R. H. Bradbury, R. Callis, E. Clark, I. Dale, D. L. Daniels, A. Dulak, L. Flavell, G. Holdgate, T. A. Jowitt, A. Kikhney, M. McAlister, J. Méndez, D. Ogg, J. Patel, P. Petteruti, G. R. Robb, M. B. Robers, S. Saif, N. Stratton, D. I. Svergun, W. Wang, D. Whittaker, D. M. Wilson, and Y. Yao. "Potent and selective bivalent inhibitors of BET bromodomains". en. In: *Nature Chemical Biology* 12.12 (2016), pp. 1097–1104. DOI: 10.1038/nchembio.2210.
- [298] C. Ottmann, S. Marco, N. Jaspert, C. Marcon, N. Schauer, M. Weyand, C. Vandermeeren, G. Duby, M. Boutry, A. Wittinghofer, J.-L. Rigaud, and C. Oecking. "Structure of a 14-3-3 Coordinated Hexamer of the Plant Plasma Membrane H<sup>+</sup>-ATPase by Combining X-Ray Crystallography and Electron Cryomicroscopy". English. In: *Molecular Cell* 25.3 (2007), pp. 427–440. DOI: 10.1016/j.molcel.2006.12.017.
- [299] G. Petzold, E. S. Fischer, and N. H. Thomä. "Structural basis of lenalidomide-induced CK1 $\alpha$  degradation by the CRL4CRBN ubiquitin ligase". en. In: *Nature* 532.7597 (2016), pp. 127–130. DOI: 10.1038/nature16979.
- [300] H. Huang, D. F. Ceccarelli, S. Orlicky, D. J. St-Cyr, A. Ziemba, P. Garg, S. Plamondon, M. Auer, S. Sidhu, A. Marinier, G. Kleiger, M. Tyers, and F. Sicheri. "E2 enzyme inhibition by stabilization of a low-affinity interface with ubiquitin". en. In: *Nature Chemical Biology* 10.2 (2014), pp. 156–163. DOI: 10.1038/nchembio.1412.

- [301] H. E. Xu, T. B. Stanley, V. G. Montana, M. H. Lambert, B. G. Shearer, J. E. Cobb, D. D. McKee, C. M. Galardi, K. D. Plunket, R. T. Nolte, D. J. Parks, J. T. Moore, S. A. Kliewer, T. M. Willson, and J. B. Stimmel. "Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPAR $\alpha$ ". en. In: *Nature* 415.6873 (2002), pp. 813–817. doi: 10.1038/415813a.
- [302] B. Graves, T. Thompson, M. Xia, C. Janson, C. Lukacs, D. Deo, P. Di Lello, D. Fry, C. Garvie, K.-S. Huang, L. Gao, C. Tovar, A. Lovey, J. Wanner, and L. T. Vassilev. "Activation of the p53 pathway by small-molecule-induced MDM2 and MDMX dimerization". In: *Proceedings of the National Academy of Sciences* 109.29 (2012), pp. 11788–11793. doi: 10.1073/pnas.1203789109.
- [303] C. E. Bulawa, S. Connelly, M. DeVit, L. Wang, C. Weigel, J. A. Fleming, J. Packman, E. T. Powers, R. L. Wiseman, T. R. Foss, I. A. Wilson, J. W. Kelly, and R. Labaudinière. "Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade". In: *Proceedings of the National Academy of Sciences* 109.24 (2012), pp. 9629–9634. doi: 10.1073/pnas.1121005109.
- [304] L. Saelices, L. M. Johnson, W. Y. Liang, M. R. Sawaya, D. Cascio, P. Ruchala, J. Whitelegge, L. Jiang, R. Riek, and D. S. Eisenberg. "Uncovering the Mechanism of Aggregation of Human Transthyretin". English. In: *Journal of Biological Chemistry* 290.48 (2015), pp. 28932–28943. doi: 10.1074/jbc.M115.659912.
- [305] V. N. Malashkevich, N. G. Dulyaninova, U. A. Ramagopal, M. A. Liriano, K. M. Varney, D. Knight, M. Brenowitz, D. J. Weber, S. C. Almo, and A. R. Bresnick. "Phenothiazines inhibit S100A4 function by inducing protein oligomerization". In: *Proceedings of the National Academy of Sciences* 107.19 (2010), pp. 8605–8610. doi: 10.1073/pnas.0913660107.
- [306] P. Pathuri, L. Vogeley, and H. Luecke. "Crystal Structure of Metastasis-Associated Protein S100A4 in the Active Calcium-Bound Form". en. In: *Journal of Molecular Biology* 383.1 (2008), pp. 62–77. doi: 10.1016/j.jmb.2008.04.076.
- [307] B. H. Kaae, K. Harpsøe, J. S. Kastrup, A. C. Sanz, D. S. Pickering, B. Metzler, R. P. Clausen, M. Gajhede, P. Sauerberg, T. Liljefors, and U. Madsen. "Structural Proof of a Dimeric Positive Modulator Bridging Two Identical AMPA Receptor-Binding Sites". en. In: *Chemistry & Biology* 14.11 (2007), pp. 1294–1303. doi: 10.1016/j.chembiol.2007.10.012.
- [308] W. Zhang, Y. Cho, E. Lolis, and J. R. Howe. "Structural and Single-Channel Results Indicate That the Rates of Ligand Binding Domain Closing and Opening Directly Impact AMPA Receptor Gating". en. In: *Journal of Neuroscience* 28.4 (2008), pp. 932–943. doi: 10.1523/JNEUROSCI.3309-07.2008.
- [309] G. J. Morgan, N. L. Yan, D. E. Mortenson, E. Rennella, J. M. Blundon, R. M. Gwin, C.-Y. Lin, R. L. Stanfield, S. J. Brown, H. Rosen, T. P. Spicer, V. Fernandez-Vega, G. Merlini, L. E. Kay, I. A. Wilson, and J. W. Kelly. "Stabilization of amyloidogenic immunoglobulin light chains by small molecules". In: *Proceedings of the National Academy of Sciences* 116.17 (2019), pp. 8360–8369. doi: 10.1073/pnas.1817567116.

- [310] M. Zhang, J. M. Pascal, and J.-F. Zhang. “Unstructured to structured transition of an intrinsically disordered protein peptide in coupling Ca<sup>2+</sup>-sensing and SK channel activation”. In: *Proceedings of the National Academy of Sciences* 110.12 (2013), pp. 4828–4833. DOI: 10.1073/pnas.1220253110.
- [311] T. Vreven, I. H. Moal, A. Vangone, B. G. Pierce, P. L. Kastritis, M. Torchala, R. Chaleil, B. Jiménez-García, P. A. Bates, J. Fernandez-Recio, A. M. Bonvin, and Z. Weng. “Updates to the integrated protein-protein interaction benchmarks: Docking benchmark version 5 and affinity benchmark version 2”. In: *Journal of molecular biology* 427.19 (2015), pp. 3031–3041. DOI: 10.1016/j.jmb.2015.07.016.
- [312] Z. Cao, Y. Bian, G. Hu, L. Zhao, Z. Kong, Y. Yang, J. Wang, and Y. Zhou. “Bias-Exchange Metadynamics Simulation of Membrane Permeation of 20 Amino Acids”. en. In: *International Journal of Molecular Sciences* 19.3 (2018), p. 885. DOI: 10.3390/ijms19030885.
- [313] S. K. Kandasamy and R. G. Larson. “Molecular Dynamics Simulations of Model Trans-Membrane Peptides in Lipid Bilayers: A Systematic Investigation of Hydrophobic Mismatch”. en. In: *Biophysical Journal* 90.7 (2006), pp. 2326–2343. DOI: 10.1529/biophysj.105.073395.



## A. Supplementary Information for Chapter 4: Dynamics of wild type $\gamma$ -secretase

### A.1. Reconstruction of the Notch1-bound form

When taking the Notch1-bound cryo-EM structure (PDBID 6IDF) as the initial structure for simulation, the hydrogen bond was found between the protonated D385 and V1754 (Figure A.1A, B). However, it has been reported that only the amine bond between G1753 and V1754 is cleaved by  $\gamma$ -secretase and the hydrogen bond sampled in our simulation indicates that the cleavage is executed at the wrong site. This can be due to the wrong assignment of amino acid from the cryo-EM density or the artifact from the experiment and in either case a new Notch1-bound model needs to be reconstructed. Since F1748 and F1749 are not only well resolved from the cyro-EM density, their positions are located at the hydrophobic cavity which we termed the "internal docking site". Therefore, we considered that in the natural scenario, these two phenylalanines should stay in the same position but the loop domain needs to be elongated so that the experimentally determined scissile bond, between G1753 and V1754, can be exposed to the catalytic center. Therefore, we performed homology modeling by inserting a sequence gap at the Notch1 loop position between G1751 and C1752 (Table A.1 with MODELLER. The new structure allowed us to sample the more realistic catalytic hydrogen bond (Figure 4.5B) and is therefore taken for the computational study.

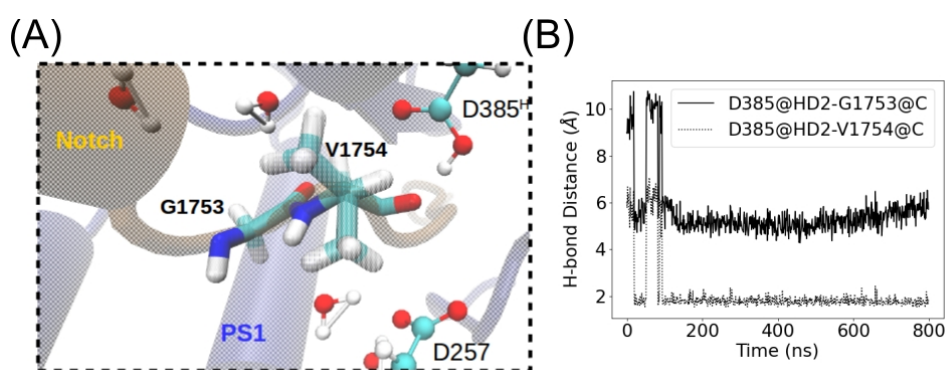


Figure A.1.: Catalytic hydrogen bonds from the simulation starting with the Notch1-bound  $\gamma$ -secretase structure determined by cryo-EM (PDBID 6IDF). (A) Schematic of the hydrogen bond sampled in the simulation. (B) Time-plot of the distances between the catalytic proton of D385 (D385@HD2) to the carbonyl group of G1753 (solid) and V1754 (dashed).

Table A.1.: Sequence alignment strategy used to reconstruct a more realistic Notch1-bound  $\gamma$ -secretase structure. The target scissile bond is indicated by the double lines between two residues.

	P10	P9	P8	P7	P6	P5	P4	P3	P2	P1	P1'	P2'	P3'
Notch1-old	V1745	L1746	L1747	F1748	F1749	V1750	G1751	C1752	G1753	V1754	L1755	L1756	S1757
Notch1-new	V1745	L1746	L1747	F1748	F1749	V1750	G1751	-	C1752	G1753	V1754	L1755	L1756

## A.2. Supplementary Figures

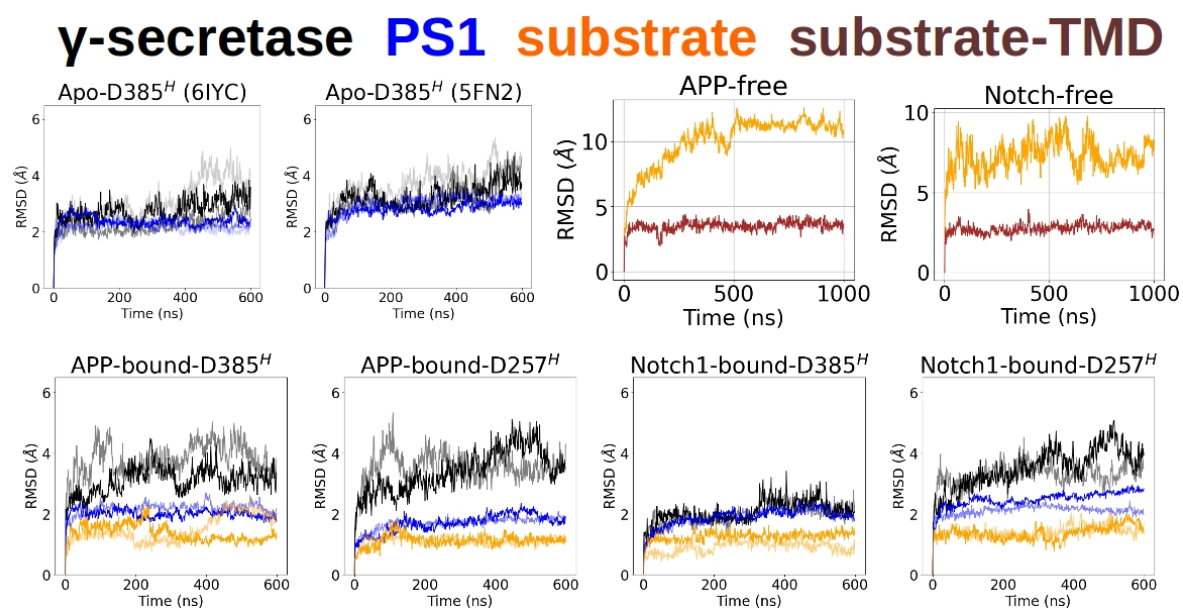


Figure A.2.: RMSD of proteins in POPC membrane. Proteins are color-coded by black ( $\gamma$ -secretase), blue (PS1) and orange (Substrate, APP, or Notch1). Opaque and transparent curves represent RMSD from the first and the second simulations. In addition, the RMSD of substrate TMDs is calculated in the substrate-only simulations to show their stability.

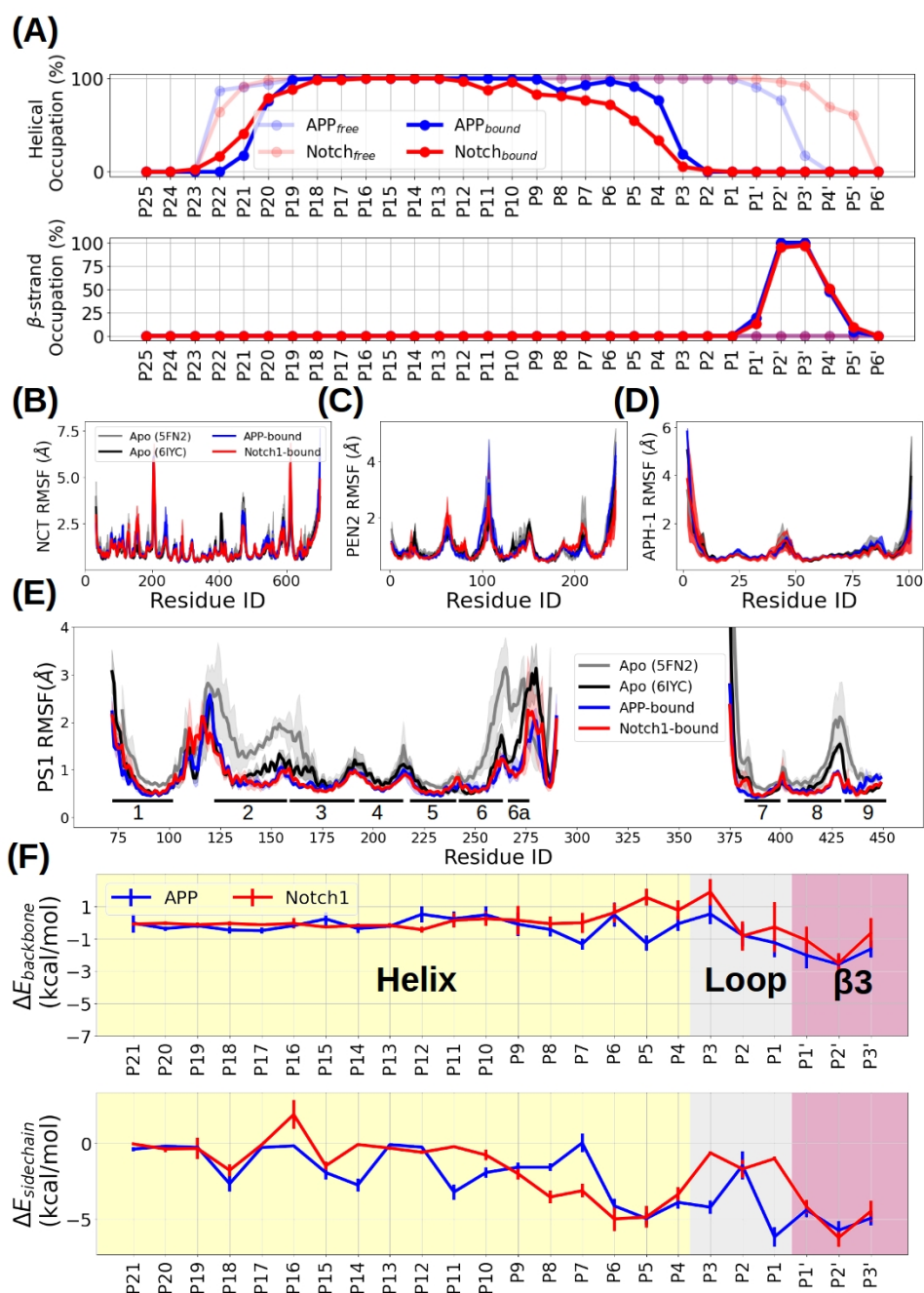


Figure A.3.: Change in the secondary structure of APP and Notch1 and  $\gamma$ -secretase subunits upon substrate binding and the binding energy analysis of substrate TMD. (A) Average occupation of helical (up) and  $\beta$ -strand (down) state of substrates over time and simulation replicas. (B) RMSF of NCT. (C) RMSF of PEN2. (D) RMSF of APH-1. (E) RMSF of PS1. (F) Residue-wise interaction energy of substrate TMD contributed by backbone (up) and side chain (down) atoms. Helical, loop, and  $\beta$ -3 domains are determined from the secondary structure analysis from (A).

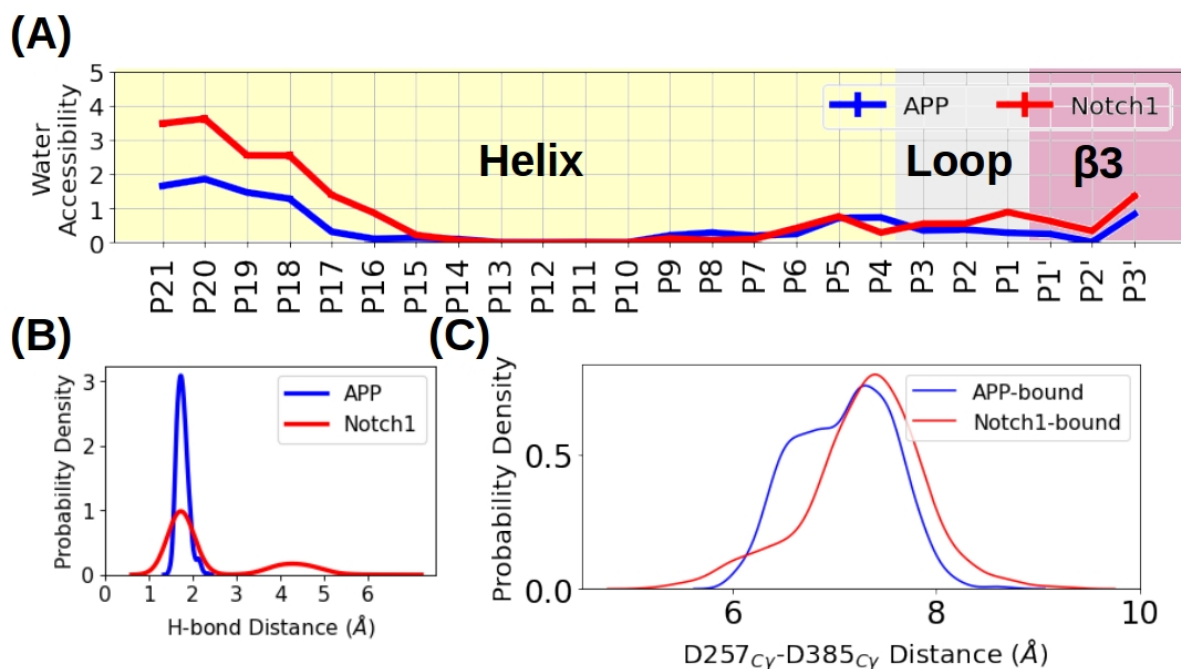


Figure A.4.: Water accessibility and geometry at the catalytic center. (A) Residue-wise water accessibility of the TMDs of APP (blue) and Notch1 (red). (B) Probability density of the distance of the catalytic hydrogen bond distance in APP-bound (blue) and Notch1-bound (red) simulations. (C) Probability density of the C $\gamma$ -C $\gamma$  distance between D257 and D385 sampled the simulations.

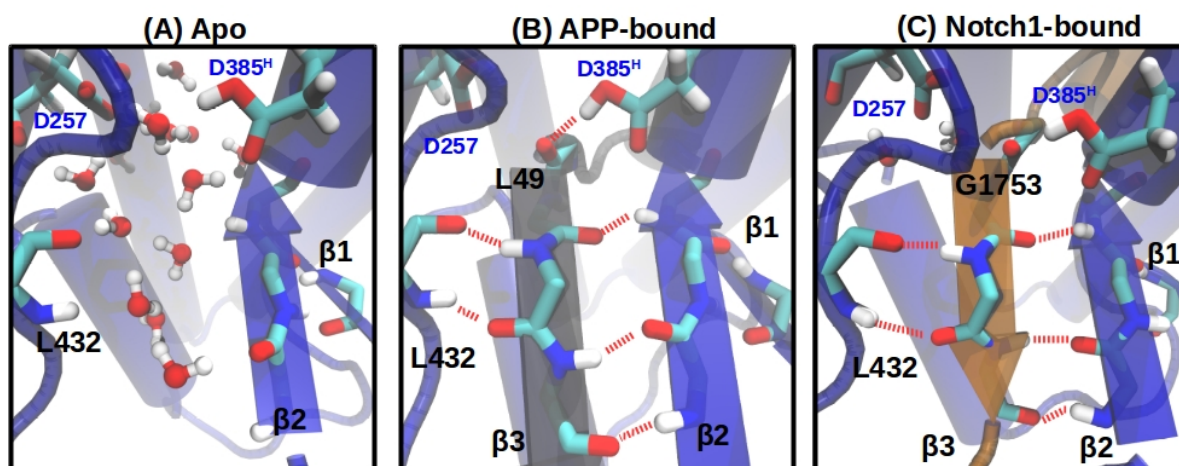


Figure A.5.: Schematics of the hybrid  $\beta$ -strand region in the (A) apo-state, (B) APP-bound state, and (C) Notch-1 bound state. PS1 is colored blue, APP is colored gray and Notch-1 is colored orange. The color codes are designed for comparison with Chapter 7 and are not consistent with Chapter 4.

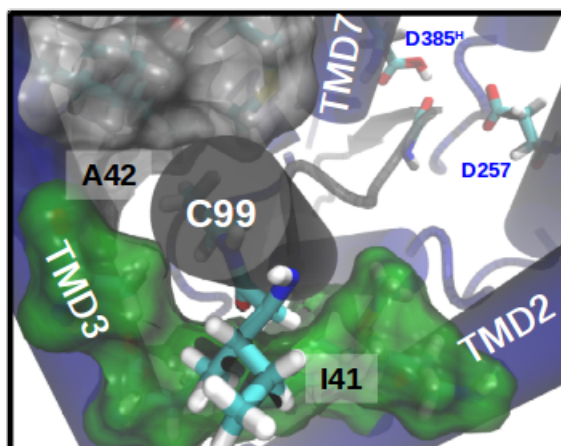


Figure A.6.: Top view at the PS1 internal docking site at the C99<sub>ε49</sub> binding pose with A42 and I41 well aligned to the PS1 internal docking site.. The color-codes are designed for comparison with Chapter 7 and is not inconsistent with Chapter 4.

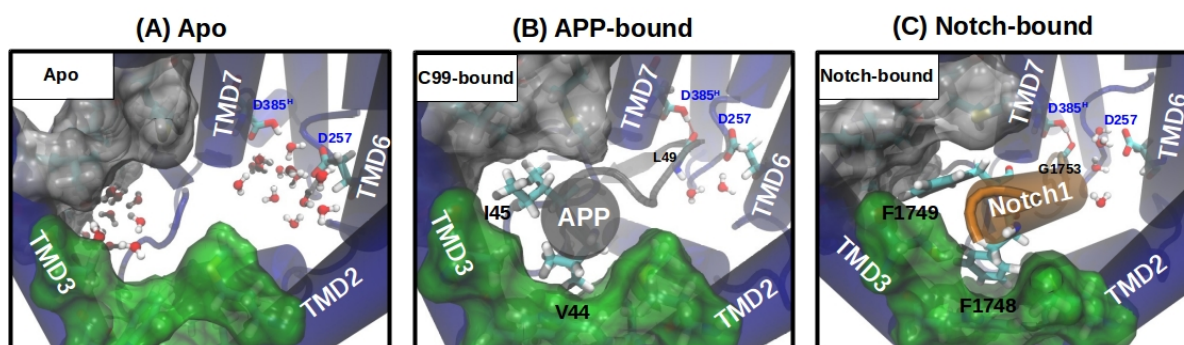


Figure A.7.: View into the PS1 internal docking site in the unbound-form (A), C99-bound (B), and Notch-bound (C)  $\gamma$ -secretase complexes. PS1 is shown in blue cartoon representation, C99 in grey, and Notch in orange. The sub-pocket formed between TMD2 and TMD3 is shown as a green surface, and the sub-pocket formed by TMD3-TMD5 and TMD7 is indicated as a light grey surface. All residues defining these two pockets are listed in Table 4.2. Water molecules are shown in the vdw+bond representation. P5, P6, and the backbone of P1 of C99 and Notch are shown in the licorice representation. The catalytic hydrogen bond is indicated as a red dashed line between the substrate scissile bond and the protonated aspartic acid. The color-codes are designed for comparison with Chapter 7 and is not inconsistent with Chapter 4.

## B. Supplementary Information for Chapter 6: Enzyme-Substrate Hybrid $\beta$ -Sheet Controls Geometry and Water Access to the $\gamma$ -Secretase Active Site

### B.1. Supplementary Figures

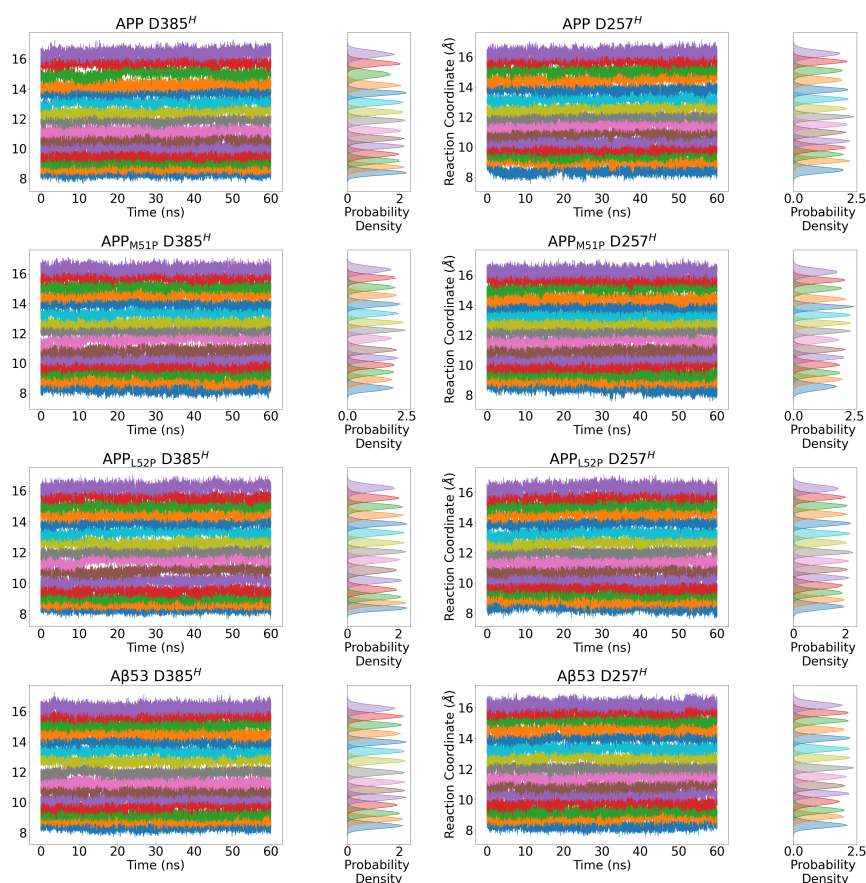


Figure B.1.: The sampled RC distance during the 60 ns H-REMD simulation (from top to bottom) APP, APP<sub>M51P</sub>, APP<sub>L52P</sub>, A $\beta$ 53, A $\beta$ 52, A $\beta$ 52<sub>V50F</sub>, A $\beta$ 52<sub>M51F</sub> and A $\beta$ 51 binding to  $\gamma$ -secretase-PS1-D385<sup>H</sup> (left) and  $\gamma$ -secretase-PS1-D257<sup>H</sup> (right).

## B.1. SUPPLEMENTARY FIGURES

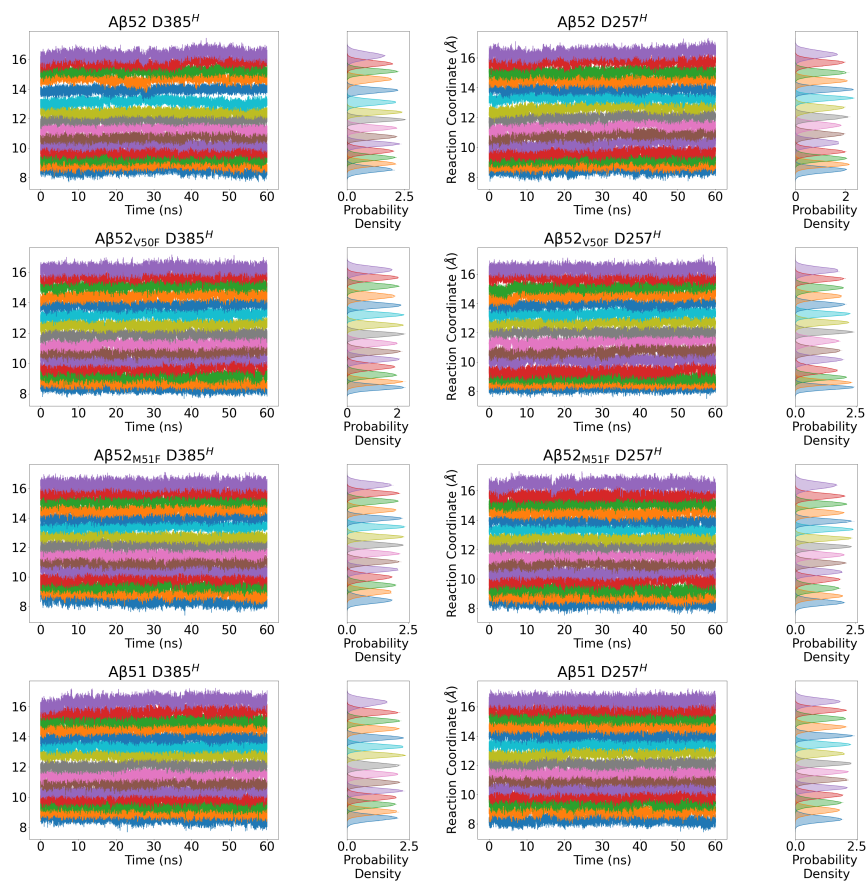


Figure B.2.: The sampled RC distance during the 60 ns H-REMD simulation (from top to bottom) of A $\beta$ 52, A $\beta$ 52<sub>V50F</sub>, A $\beta$ 52<sub>M51F</sub> and A $\beta$ 51 binding to  $\gamma$ -secretase-PS1-D385<sup>H</sup> (left) and  $\gamma$ -secretase-PS1-D257<sup>H</sup> (right).

## B.1. SUPPLEMENTARY FIGURES

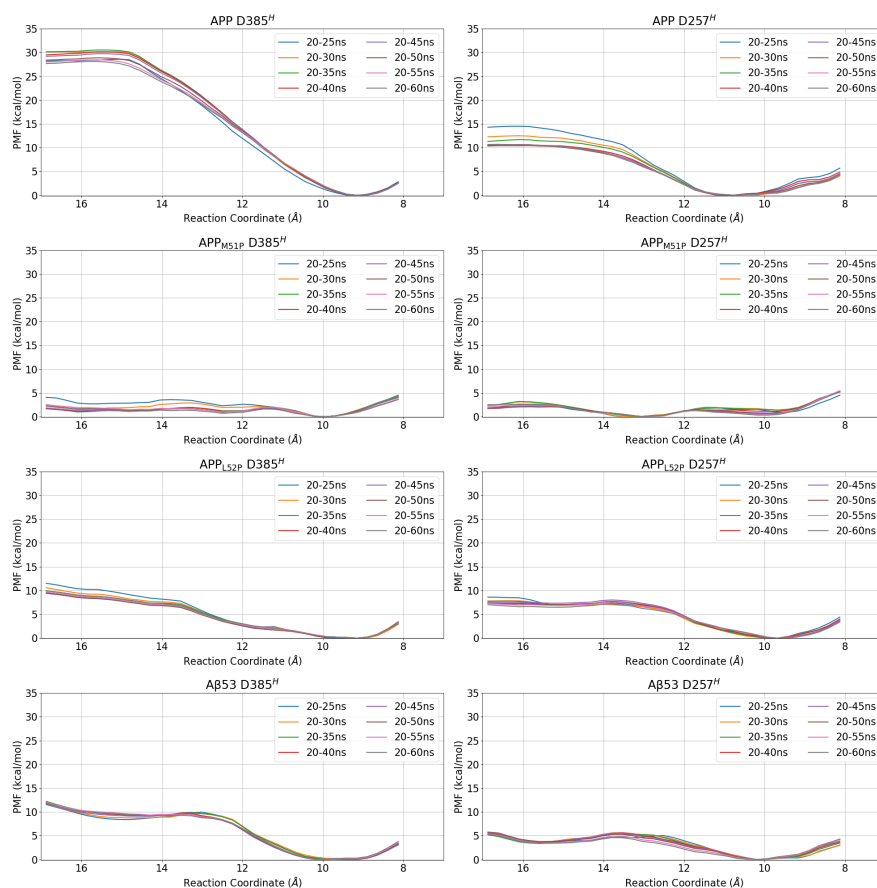


Figure B.3.: Convergence of PMF profiles along the H-REMD protocol (from top to bottom) of APP, APP<sub>M51P</sub>, APP<sub>L52P</sub>, and Aβ53 binding to γ-secretase-PS1-D385<sup>H</sup> (left) and γ-secretase-PS1-D257<sup>H</sup> (right).



## B.1. SUPPLEMENTARY FIGURES

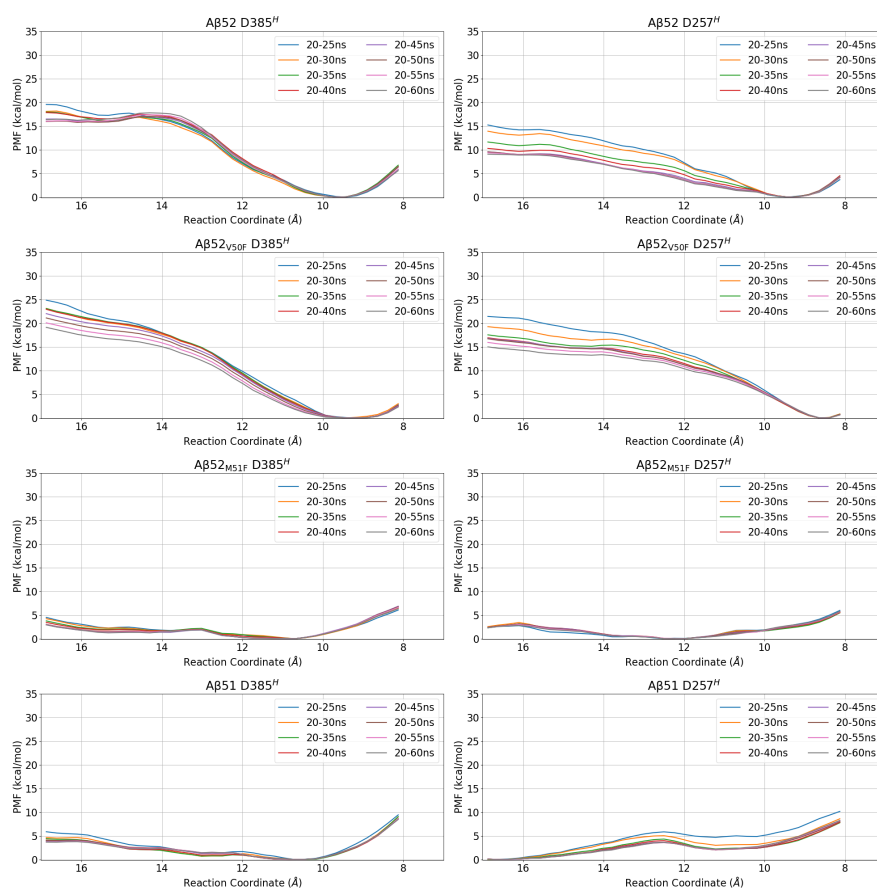


Figure B.4.: Convergence of PMF profiles along the H-REMD protocol (from top to bottom) of Aβ52, Aβ52<sub>V50F</sub>, Aβ52<sub>M51F</sub> and Aβ51 binding to γ-secretase-PS1-D385<sup>H</sup> (left) and γ-secretase-PS1-D257<sup>H</sup> (right).

B.1. SUPPLEMENTARY FIGURES

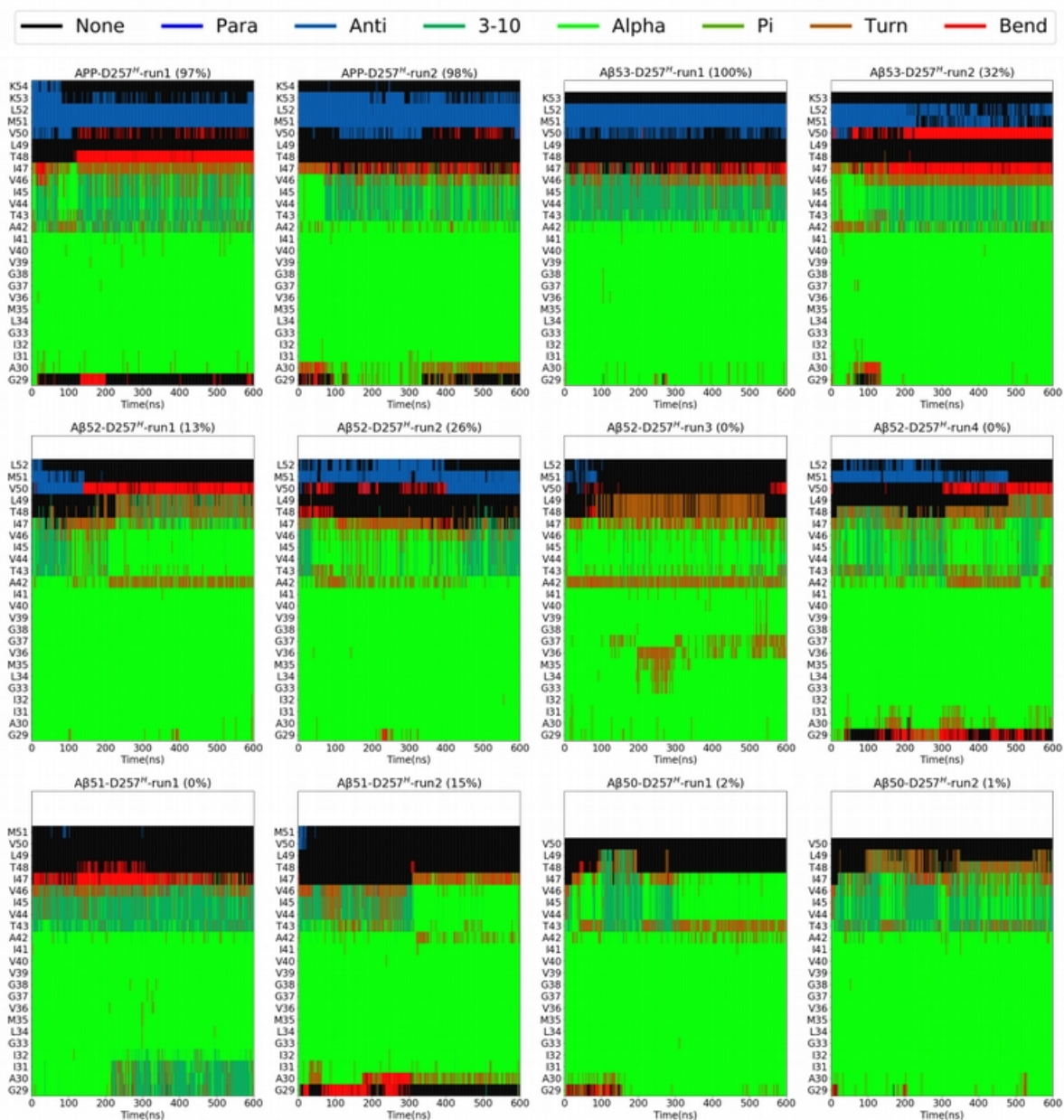


Figure B.5.: Secondary structures over time calculated by DSSP method in APP and its truncated variants binding to D257<sup>H</sup>  $\gamma$ -secretase. Each 600ns-long simulation is shown with the sampled fraction of active geometry written in the parenthesis.

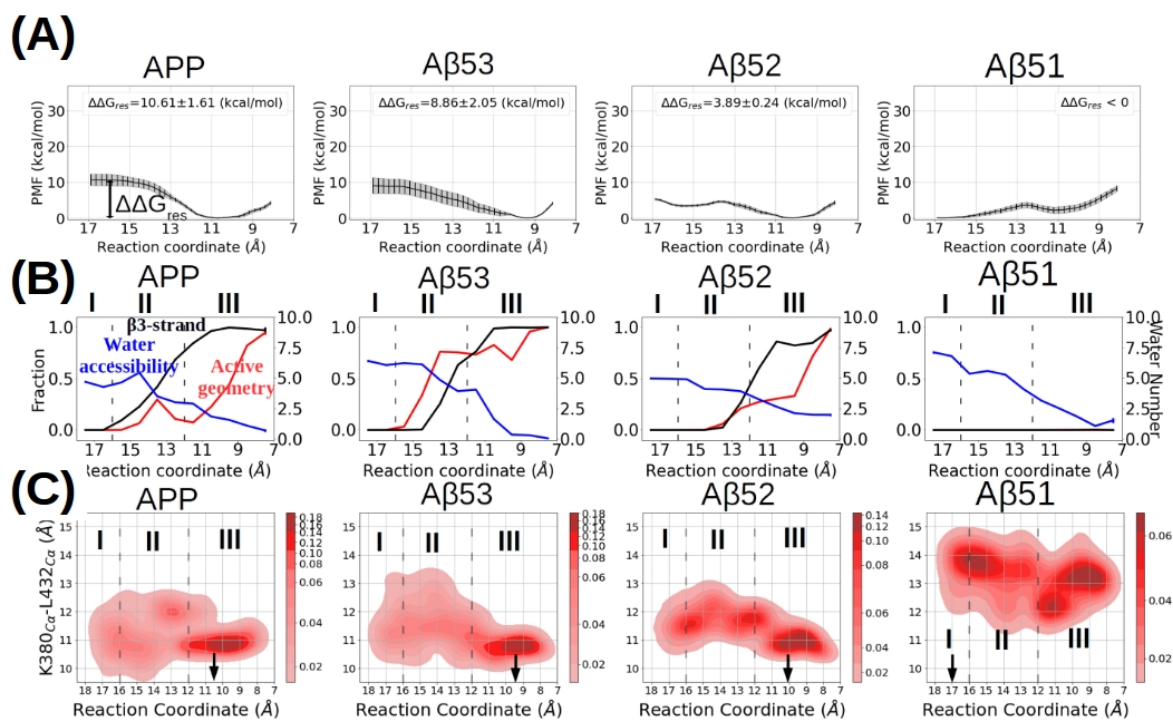


Figure B.6.: Hamiltonian Replica exchange MD along the  $\beta$ 3-strand association pathway in APP and its truncated variants binding to  $\gamma$ -secretase-PS1-D257<sup>H</sup>. (A) Potential-of-mean-force (PMF) profiles calculated with the HREUS method along the substrate  $\beta$ 3-strand association reaction of APP and its truncated substrates. (B) Fraction of sampled states that form an active site geometry compatible with cleavage (red),  $\beta$ 3-strand (black), and the number of water molecules around the catalytic center (blue), along the sampling pathway. (C) The change in the gap width between L432 and  $\beta$ 2, indicated by the C $\alpha$ -C $\alpha$  distance between L432 and K380, along the sampling pathway. The black arrows point to the RC that corresponds to the PMF minimum. Three regimes are distinguished by the transparent dashed lines with regime I corresponding to the dissociated regime, regime II the transition regime, and regime III forming the associated regime.

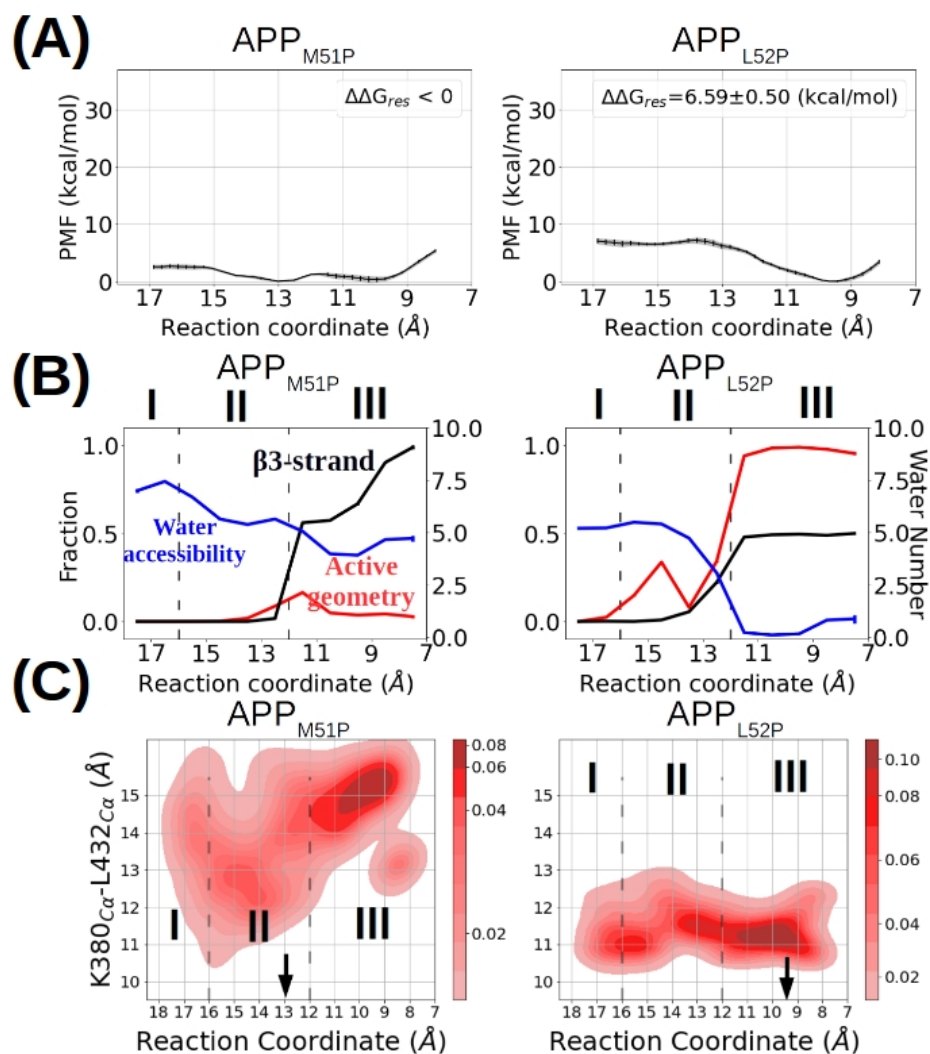


Figure B.7.: Hamiltonian Replica exchange MD along the  $\beta$ 3-strand association pathway in APP mutants M51P (left) and L52P (right) binding to  $\gamma$ -secretase-PS1-D257<sup>H</sup>. (A) Potential-of-mean-force (PMF) profiles calculated with the HREUS method along the substrate  $\beta$ 3-strand association reaction of APP mutants. (B) Fraction of sampled states that form an active site geometry compatible with cleavage (red),  $\beta$ 3-strand (black), and the number of water molecules around the catalytic center (blue), along the sampling pathway. (C) The change in the gap width between L432 and  $\beta$ 2, indicated by the C $\alpha$ -C $\alpha$  distance between L432 and K380, along the sampling pathway. The black arrows point to the RC that corresponds to the PMF minimum. Three regimes are distinguished by the black dashed lines with regime I corresponding to the dissociated regime, regime II the transition regime, and regime III forming the associated regime.

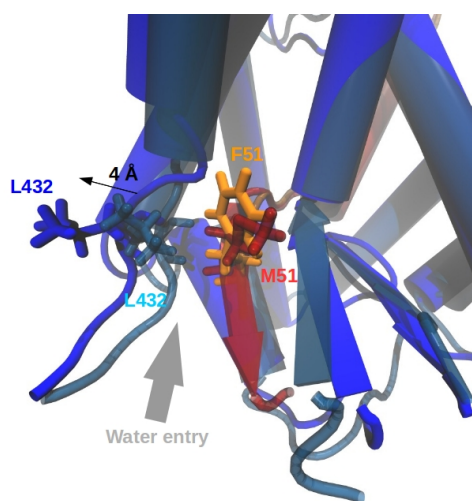


Figure B.8.: The effect of M51F mutant in the associated state ( $RC = 9.2\text{\AA}$ ). The overlap of snapshot shows the conformations of PS1 when binding to WT A $\beta$ 52 (light blue and red) and when binding to the M51F mutant (dark blue and orange). The large Phe sidechain of the M51F mutant causes a 4 $\text{\AA}$  deviation of PS1 L432 (measured by the C $\alpha$  atoms), and creates an entry for water access to the active site.

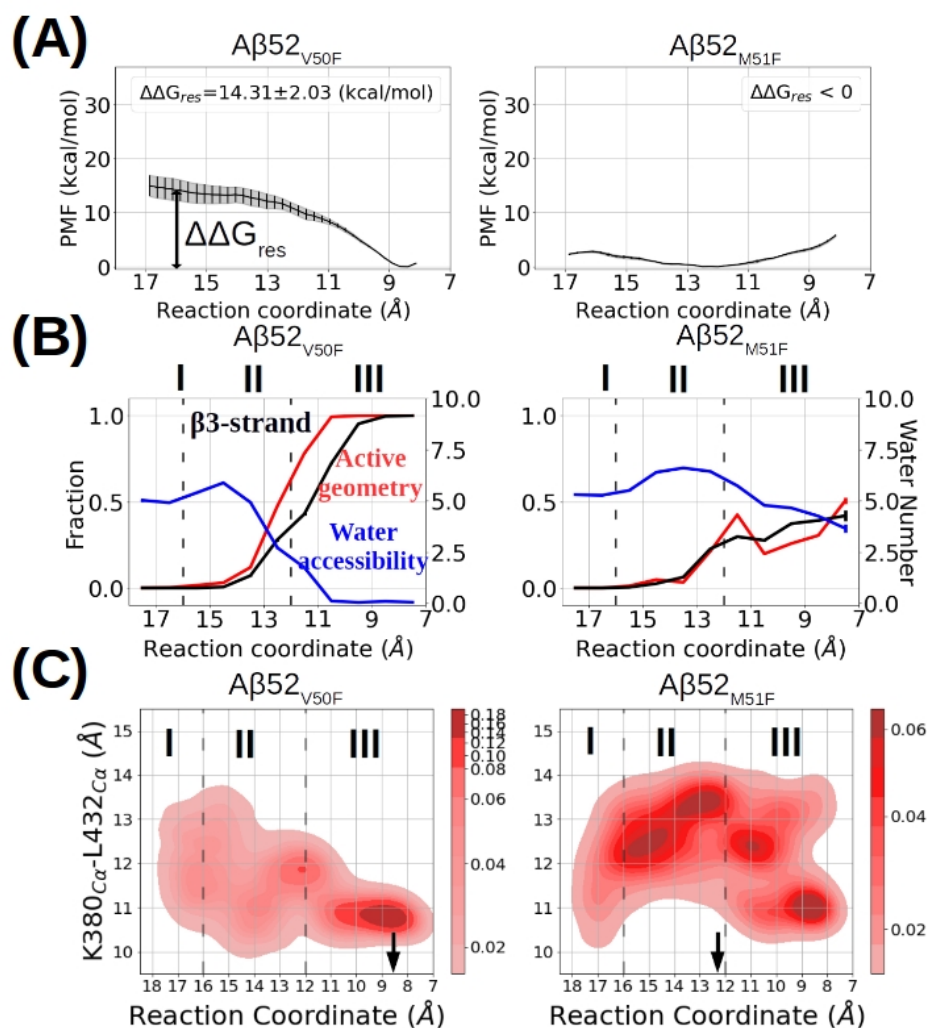


Figure B.9.: Hamiltonian Replica exchange MD along the  $\beta$ 3-strand association pathway in A $\beta$ 52 mutants V50F (left) and M51F (right) binding to  $\gamma$ -secretase-PS1-D257<sup>H</sup>. (A) Potential-of-mean-force (PMF) profiles were calculated with the HREUS method along the substrate  $\beta$ 3-strand association reaction of A $\beta$ 52 mutants V50F (left) and M51F (right). (B) Fraction of sampled states that form an active site geometry compatible with cleavage (red),  $\beta$ 3-strand (black), and the number of water molecules around the catalytic center (blue), along the sampling pathway. (C) The change in the gap width between L432 and  $\beta$ 2, indicated by the C $\alpha$ -C $\alpha$  distance between L432 and K380, along the sampling pathway. The black arrows point to the RC that corresponds to the PMF minimum. Three regimes are distinguished by the transparent dashed lines with regime I corresponding to the dissociated regime, regime II the transition regime, and regime III forming the associated regime.

B.1. SUPPLEMENTARY FIGURES

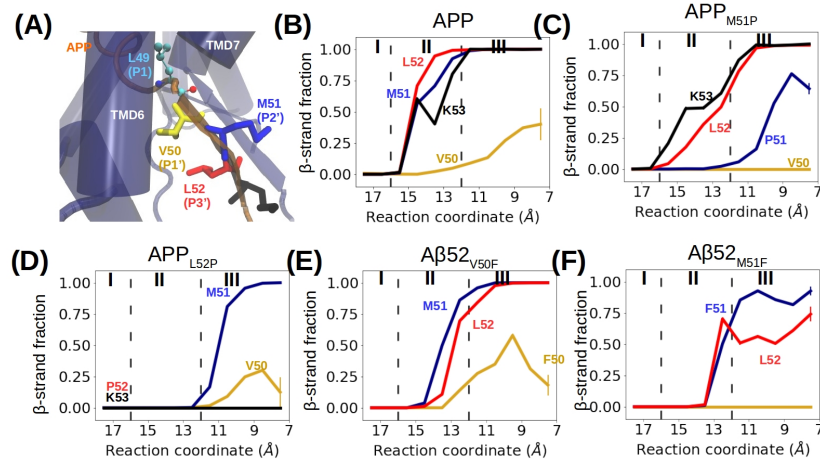


Figure B.10.:  $\beta$ -strand fraction of P1' to P4' residues when binding to D385H  $\gamma$ -secretase identified by DSSP in the HREUS simulations. (A) Schematic of the locations of P1 (cyan), P1' (yellow), P2' (blue), P3' (red), and P4' (yellow) of wild type APP (orange) binding to  $\gamma$ -secretase (blue). (B)-(F) The fraction of frames sampled in the  $\beta$ -strand conformation of P1'-P4' of (B) wildtype APP, (C)  $APP_{M51P}$  (D)  $APP_{L52P}$  (E)  $A\beta_{52V50F}$ , and (F)  $A\beta_{52M51F}$  in the HREUS simulations. Phase I (left, dissociated state), Phase II (middle, transition state), and Phase III (right, associated state) are separated with black dashed lines. The color of curves corresponds to the color coding in (A).

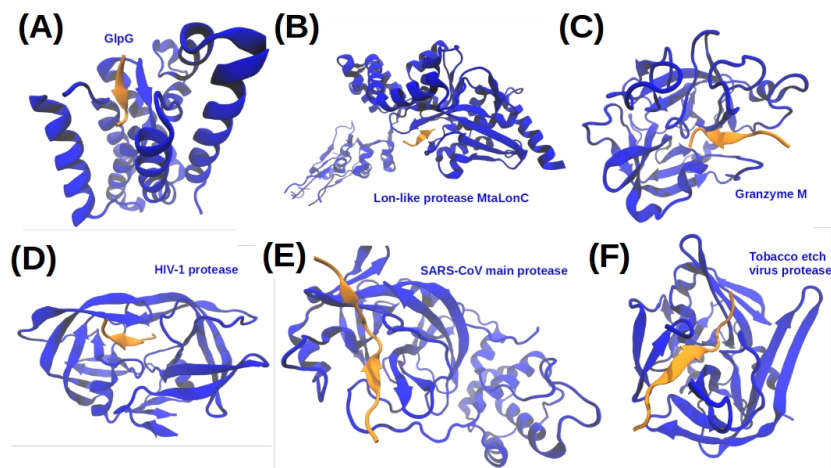


Figure B.11.: Examples of enzyme-substrate  $\beta$ -sheets formed in other proteases. (A) Intramembrane protease GlpG protease (PDBID 6PJA) (B) Lon-like protease MtaLonC. (PDBID 7EUX) (C) HIV-1 protease (PDBID 3D3T) (D) Granzyme M (PDBID 2ZGJ) (E) SARS-CoV main protease (PDBID 2Q6G) (F) Tobacco Etch Virus Protease (PDBID 1LVB). The enzymes are colored in blue and the substrates are colored in orange.

## C. Supplementary Information for Chapter 7: An internal docking site stabilizes substrate binding to $\gamma$ -secretase: Analysis by Molecular Dynamics Simulations

### C.1. Supplementary Figures

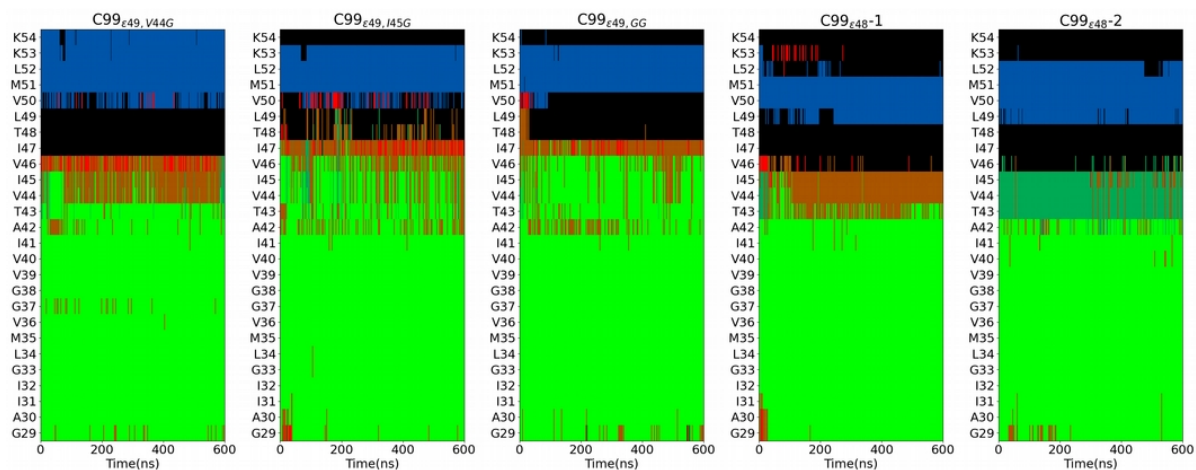


Figure C.1.: Evolution of the secondary structure of the substrates in time. (From left to right)  $C99_{\epsilon 48-1}$ ,  $C99_{\epsilon 48-2}$ ,  $C99_{V44G}$ ,  $C99_{I45G}$ ,  $C99_{GG}$ .



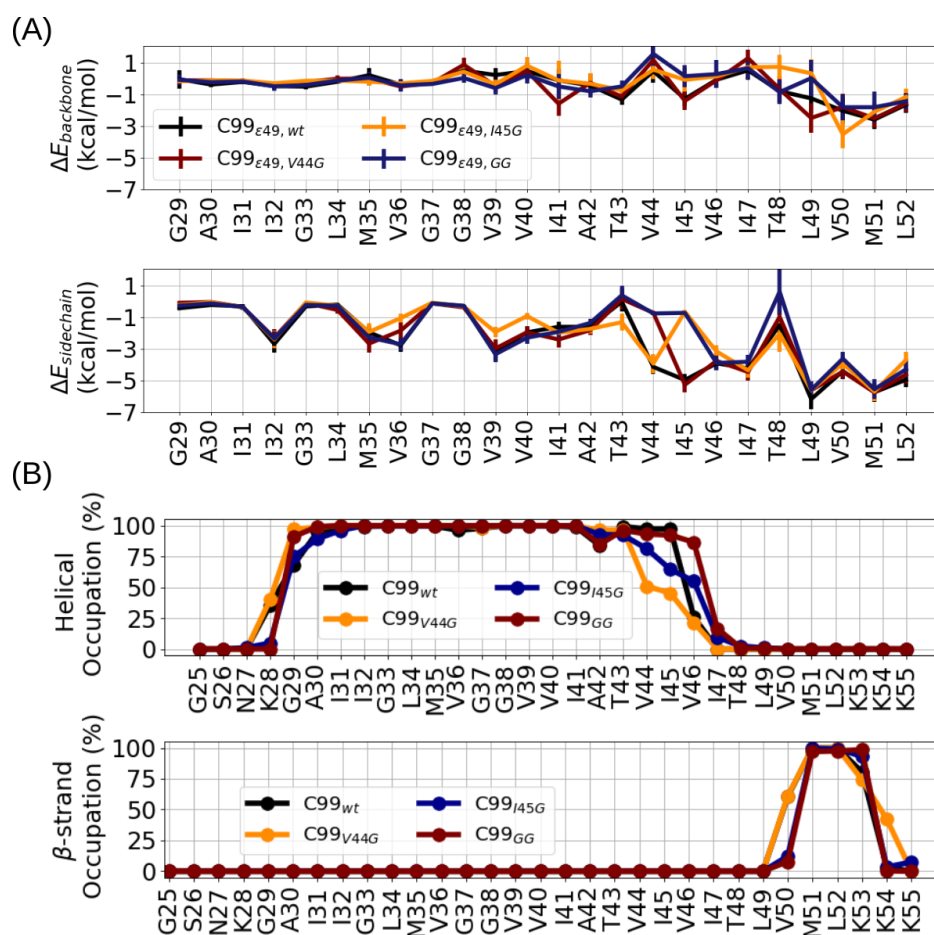


Figure C.2.: Analysis of the (A) residue-wise binding energy decomposition and (B) secondary structures in the V44G (orange), I45G (dark blue), and V44G+I45G (GG, brown) mutated complexes in comparison to the wildtype C99 (black) with D385-protonated PS1.

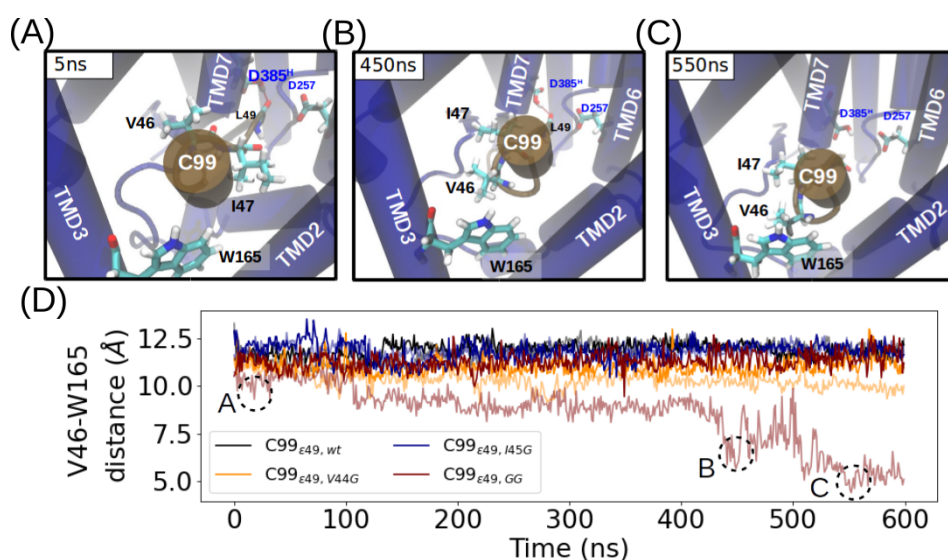


Figure C.3.: Rotation of the C99 helical domain observed in one of the C99<sub>GG</sub> simulation. Snapshots of the rotational movement at (A) 5ns, (B) 450ns, and (C) 550ns. PS1 is shown in blue and C99<sub>GG</sub> is shown in brown. (D) The center-of-mass (COM)-COM distance between V46 of C99 and W165 of PS1 along each 600ns simulation of C99<sub>e49</sub> (black and gray), C99<sub>e48-1</sub> (orange), and C99<sub>e48-2</sub> (darkblue) binding to  $\gamma$ -secretase. The representative time frames of C99<sub>GG</sub> are circled and shown in (A)-(C).

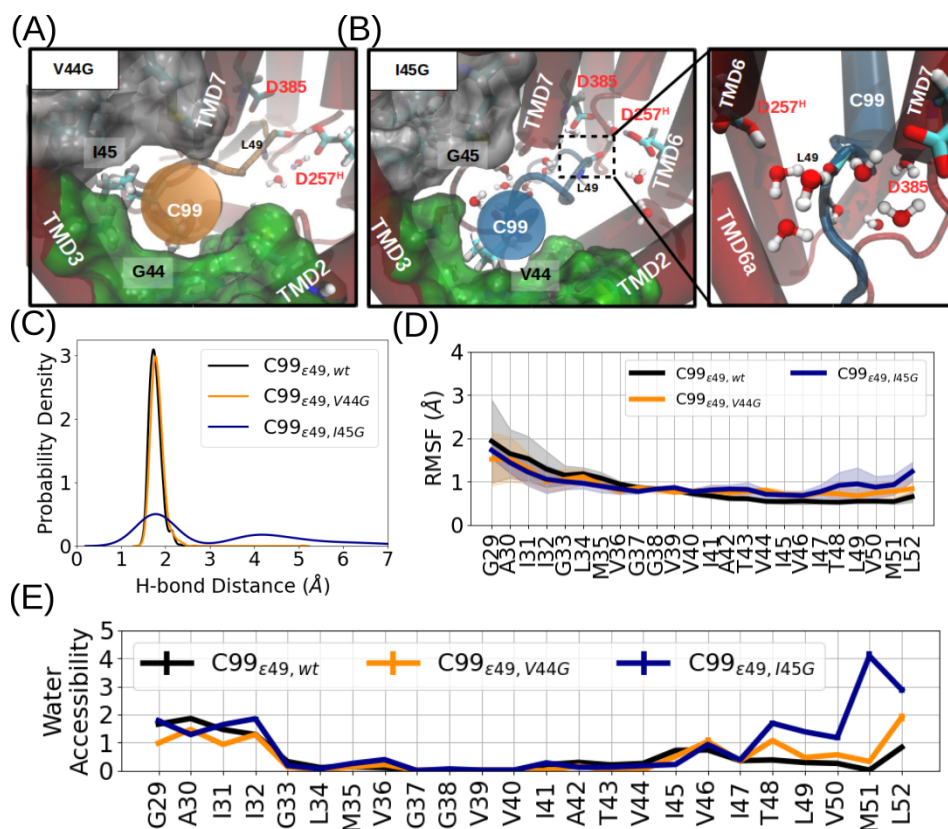


Figure C.4.: Glycine mutations at the internal docking site disturb the E-S interaction with D257-protonated PS1. (A) Top-view at the PS1 internal docking site of C99<sub>V44G</sub> bound  $\gamma$ -secretase complex. (B) Top-view at the PS1 internal docking site C99<sub>I45G</sub> bound  $\gamma$ -secretase complex with a zoom-in view shows the dissociated  $\beta$ -strand. Atomic representations are similar as described in Figure A.7. (C) Probability density of the catalytic hydrogen bond distance, (D) RMSF of the substrate TMD, and (E) secondary structure analysis. of C99<sub>wt</sub> (black), C99<sub>V44G</sub> (orange), and C99<sub>I45G</sub> (dark blue) in  $\gamma$ -secretase bound form.

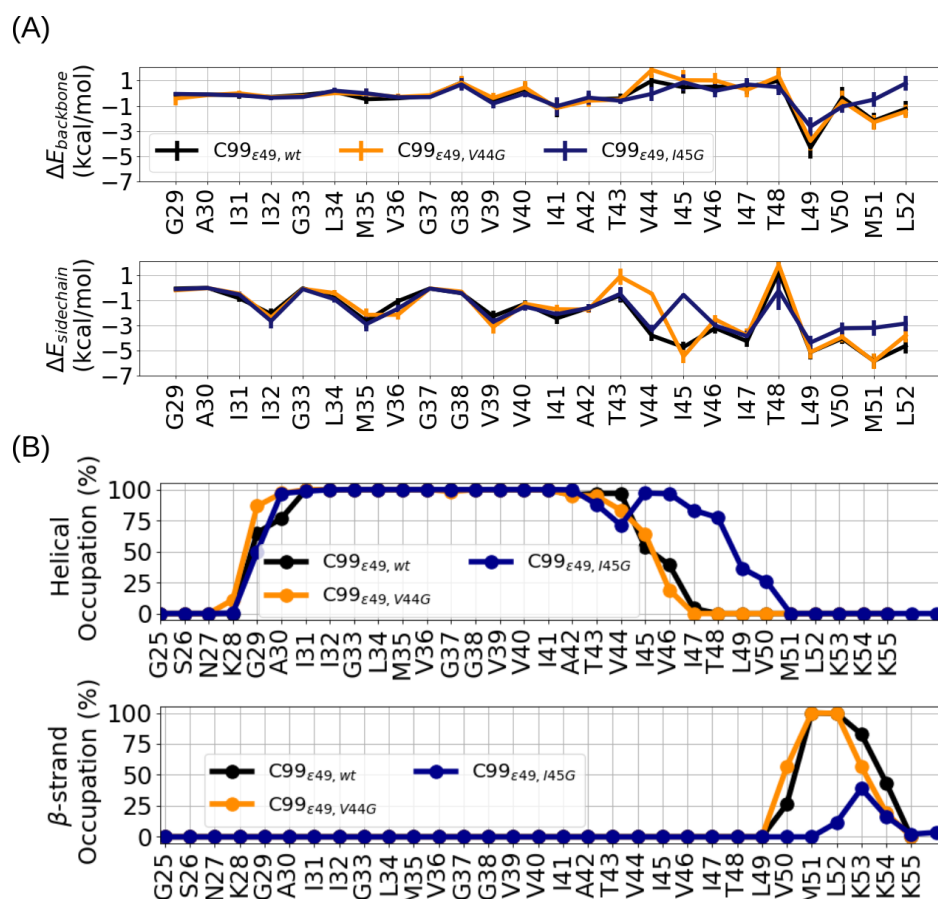


Figure C.5.: Analysis of the (A) residue-wise binding energy decomposition and (B) secondary structures in the V44G (orange) and I45G (dark blue) mutated complexes in comparison to the wildtype C99 (black) with D257-protonated PS1.

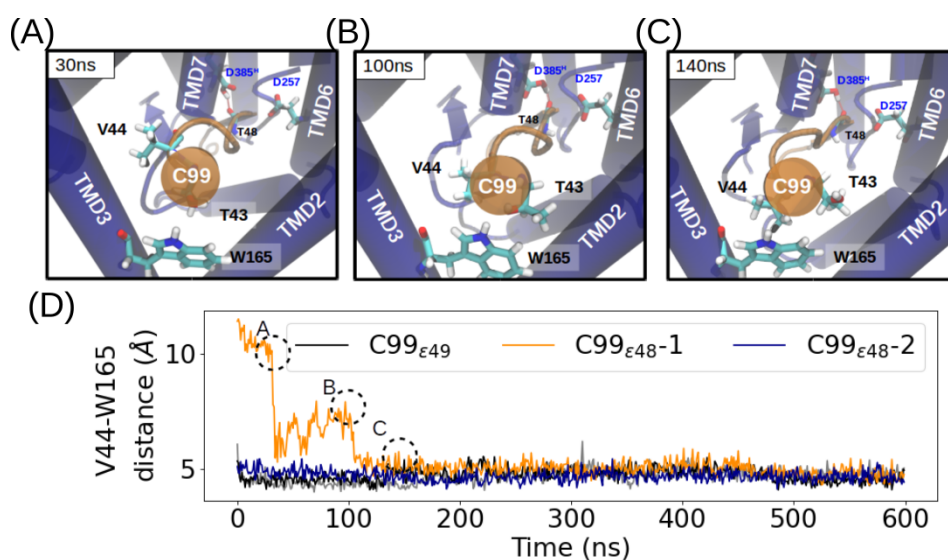


Figure C.6.: Rotation of the C99 helical domain observed in the  $C99_{\epsilon48-1}$  simulation. Snapshots of the rotational movement at (A) 30ns, (B) 100ns, and (C) 140ns. PS1 is shown in blue and C99 is shown in orange. (D) The center-of-mass (COM)-COM distance between V44 of C99 and W165 of PS1 along each 600ns simulation of  $C99_{\epsilon49}$  (black and gray),  $C99_{\epsilon48-1}$  (orange), and  $C99_{\epsilon48-2}$  (darkblue) binding to  $\gamma$ -secretase. The representative time frames of  $C99_{\epsilon48-1}$  are circled and shown in (A)-(C). The animated process is shown in Video S1 from our published work.

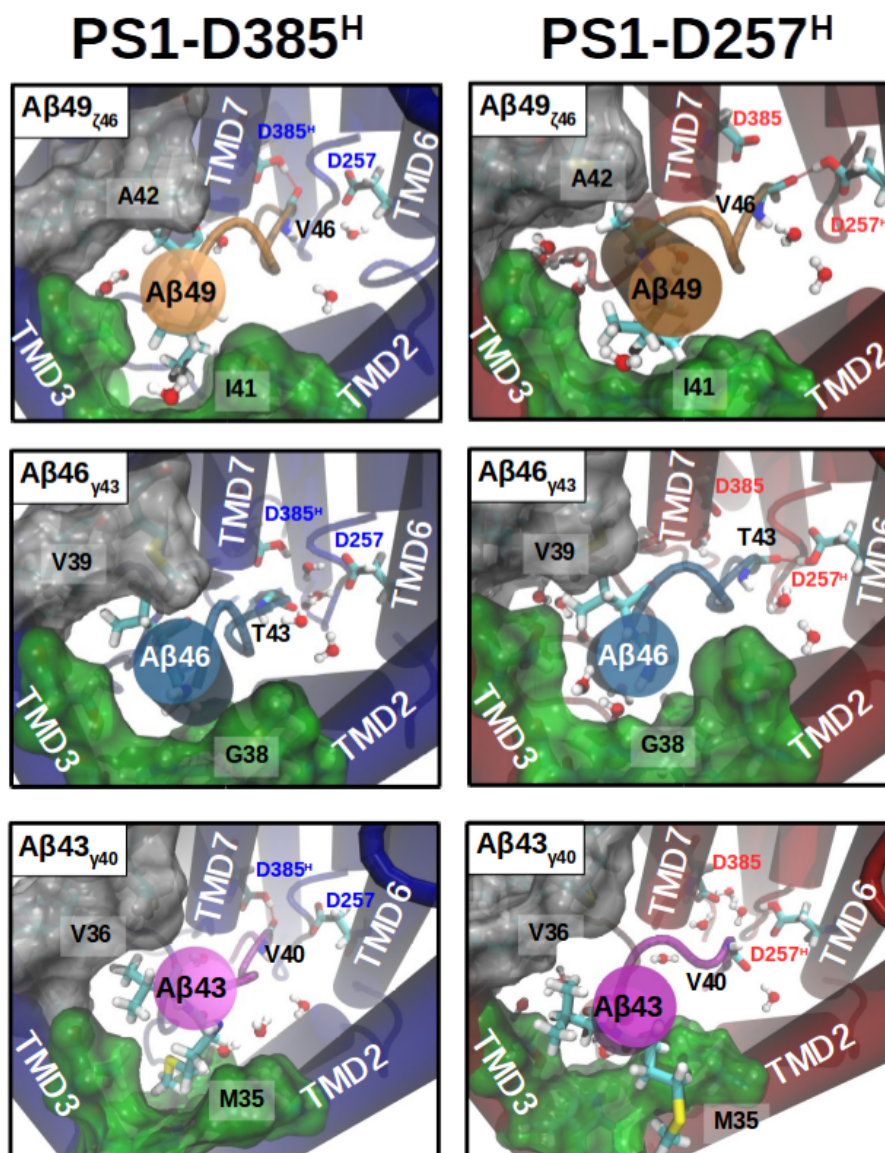


Figure C.7.: Top view of different A $\beta$  peptides binding to the PS1 internal docking site with (left) D385-protonated or (right) D257-protonated. From top to bottom: A $\beta$ 49 $_{\zeta 46}$ , A $\beta$ 46 $_{\gamma 43}$ , A $\beta$ 43 $_{\gamma 40}$ . Atomic representations are similar as described in Figure A.7.

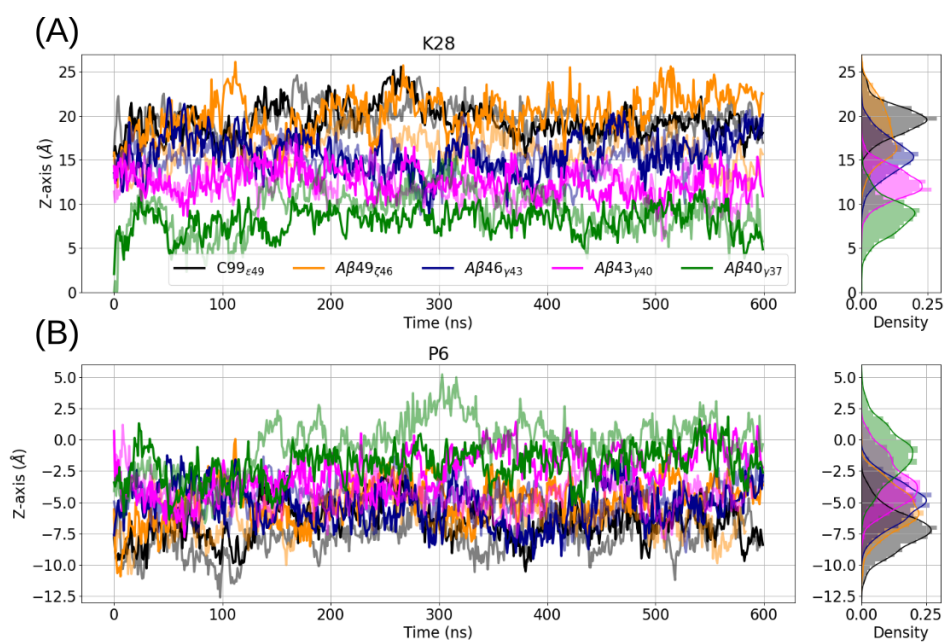


Figure C.8.: Z-axis position of substrate (A) K28 and (B) P6 in complex with  $\gamma$ -secretase with D385-protonated along the simulation time. Two replicas are represented by solid and transparent lines in the same color. The averaged phosphate groups are located at  $z=18\text{\AA}$  plane (see Figure 5D in the main text).

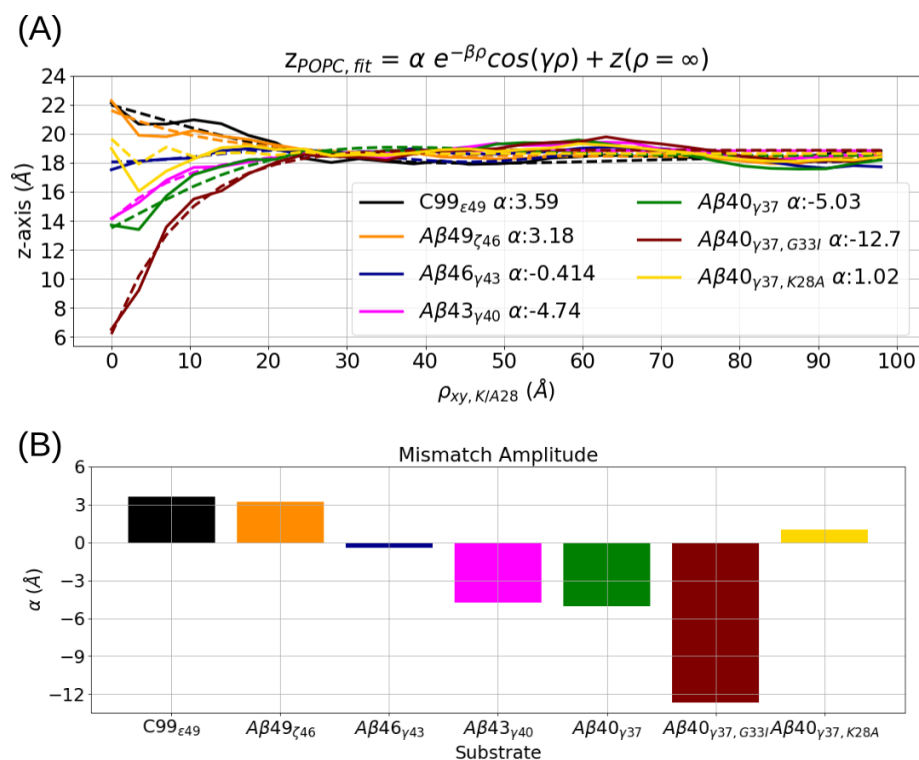


Figure C.9.: Fitting the membrane thickness distributed along the radial distance on the xy plane from K/A28 of substrates with D385-protonated PS1. (A) Fitting the hydrophobic mismatch profile with hydrophobic mismatch amplitude  $\alpha$ , radial decaying rate  $\beta$ , and harmonic oscillation  $\gamma$ . (B) Comparison of the hydrophobic mismatch amplitude  $\alpha$  calculated from (A) in different A $\beta$ -bound  $\gamma$ -secretase structures.



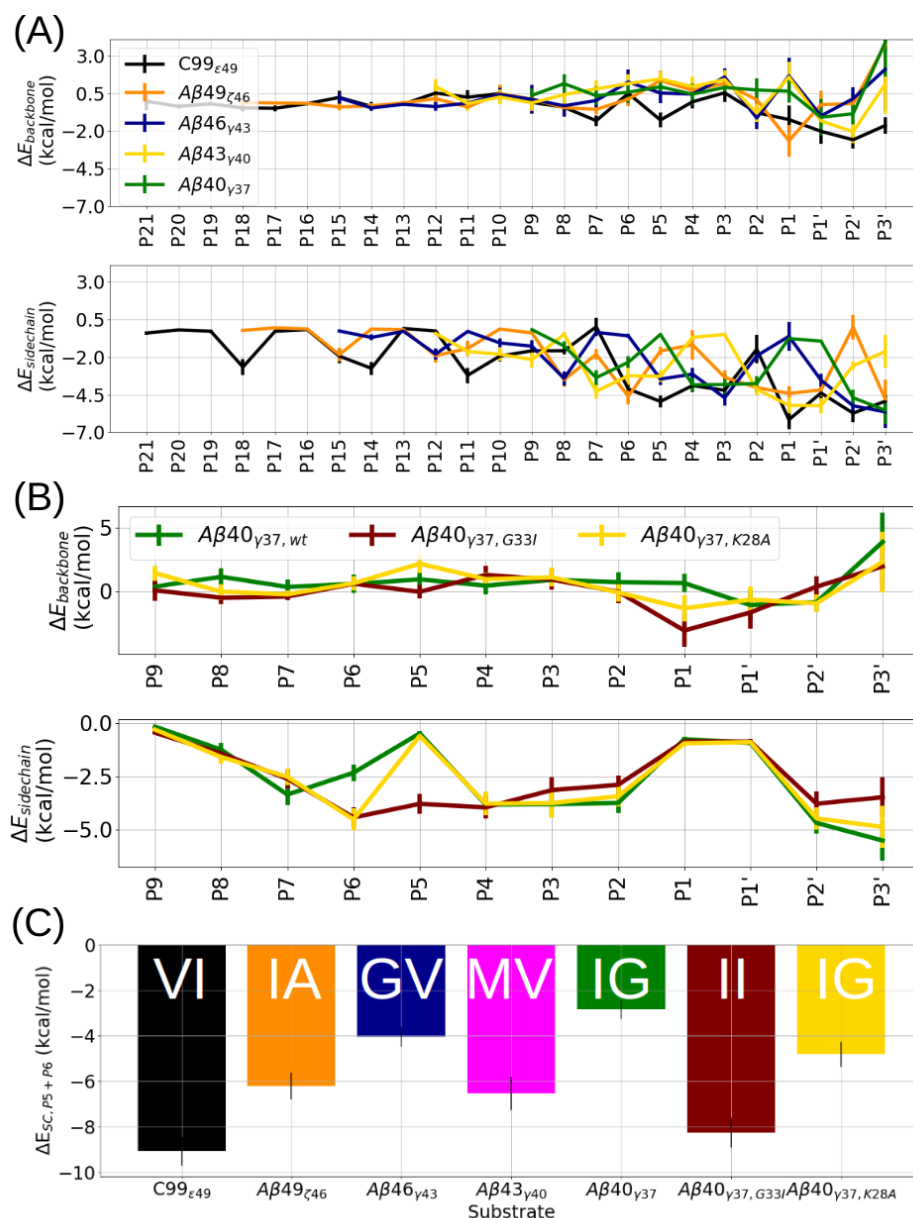


Figure C.10.: Binding energy between different A $\beta$  substrates and  $\gamma$ -secretase with the D385-protonated PS1. (A) Residue-wise binding energy decomposition between  $\gamma$ -secretase and C99 $_{\epsilon 49}$  (black), A $\beta$ 49 $_{\zeta 46}$  (orange), A $\beta$ 46 $_{\gamma 43}$  (dark blue), A $\beta$ 43 $_{\gamma 40}$  (magenta), A $\beta$ 40 $_{\gamma 37}$  (green). Backbone (top) and sidechain (bottom) contributions are averaged through two replicas. (B) Residue-wise binding energy decomposition between  $\gamma$ -secretase and A $\beta$ 40 $_{\gamma 37}$  with wild-type (solid) G33I (brown), and K28A (yellow) mutated sequences. Backbone (top) and sidechain (bottom) contributions are averaged through two replicas. (C) Summation of substrate P5 and P6 sidechain binding energy contribution. The corresponding amino acids at P6 and P5 are annotated at top of each bar.

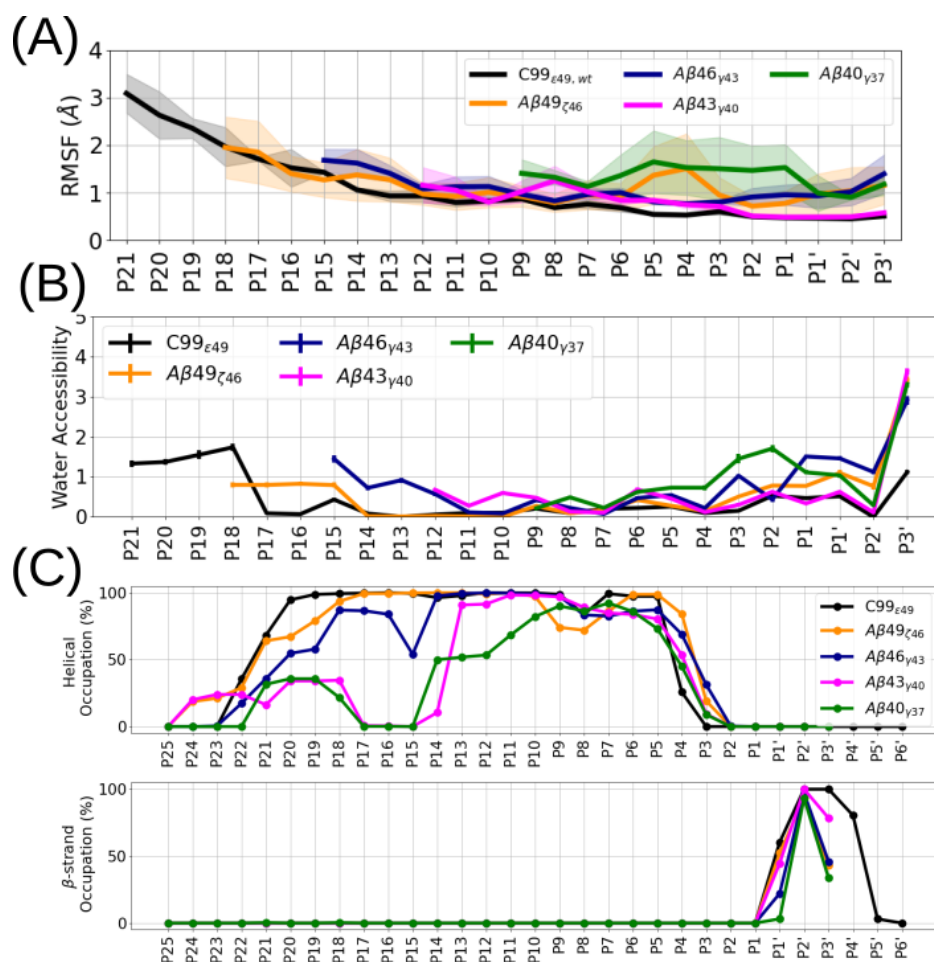


Figure C.11.: Properties of Aβ substrate bound with D385-protonated γ-secretase. (A) RMSF of the substrate TMD. (B) Residue-wise water accessibility. (C) Secondary structure analysis of C99 and Aβ peptides.

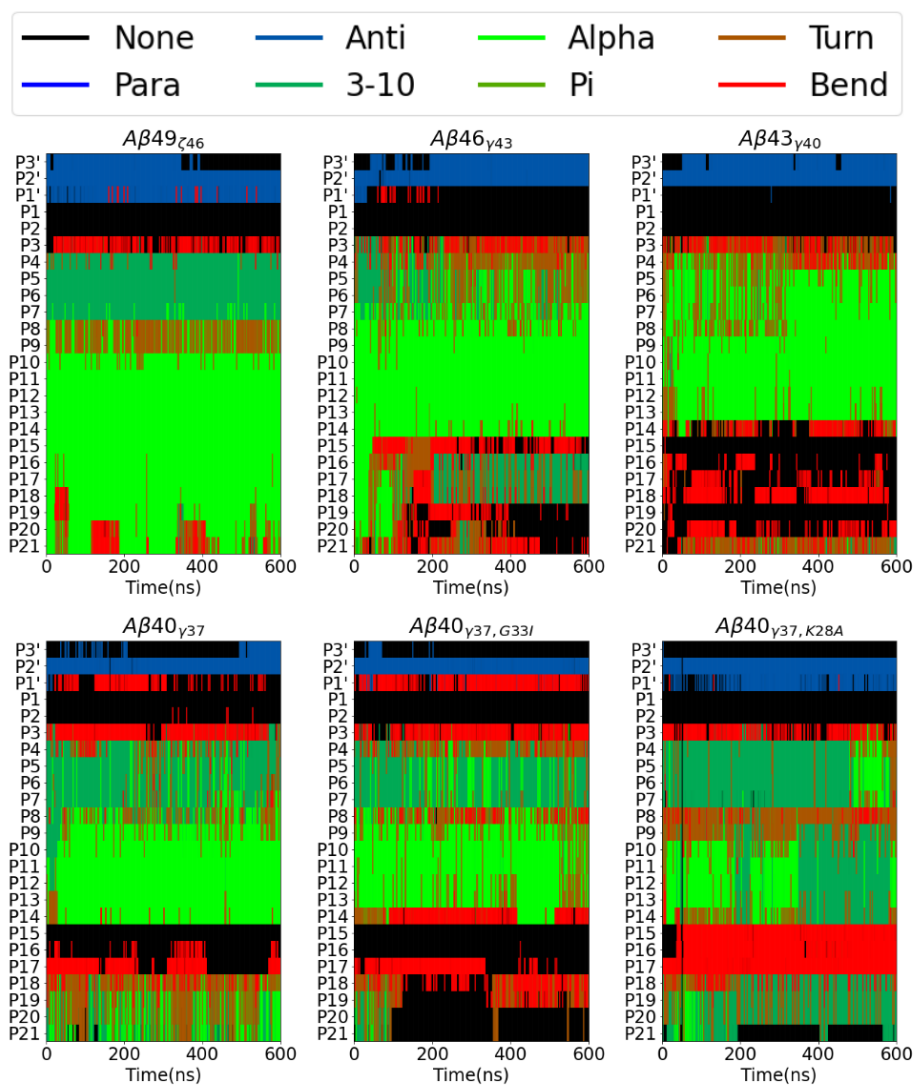


Figure C.12.: Evolution of the secondary structure of the substrates in time. From left to right: (Top)  $A\beta_{49\zeta_{46}}$ ,  $A\beta_{46\gamma_{43}}$ ,  $A\beta_{43\gamma_{40}}$ , (Bottom)  $A\beta_{40\gamma_{37}}$ ,  $A\beta_{40\gamma_{37},G33I}$ ,  $A\beta_{40\gamma_{37},K28A}$ .

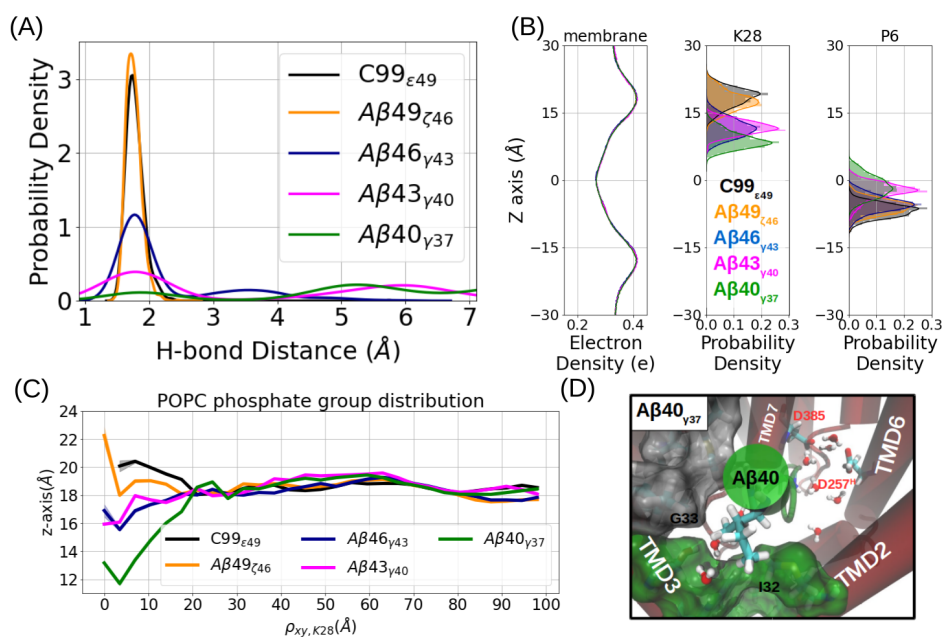


Figure C.13.: Comparative modeling and simulations of Aβn-γ-secretase complexes with D257-protonated PS1. (A) Probability density of the catalytic hydrogen bond distance. (B) Distribution of the membrane electron density (left), membrane-anchoring residue K28 (middle), and substrate P6 (right) along the z-axis in different Aβn-γ-secretase complexes. (C) The average z-axis of the POPC phosphate on the extracellular side is distributed along the radial distance  $\rho_{xy, K28}$ . (D) View into the PS1 internal docking site in the Aβ40<sub>γ37</sub>-bound γ-secretase. (representation same as in Figure A.7)

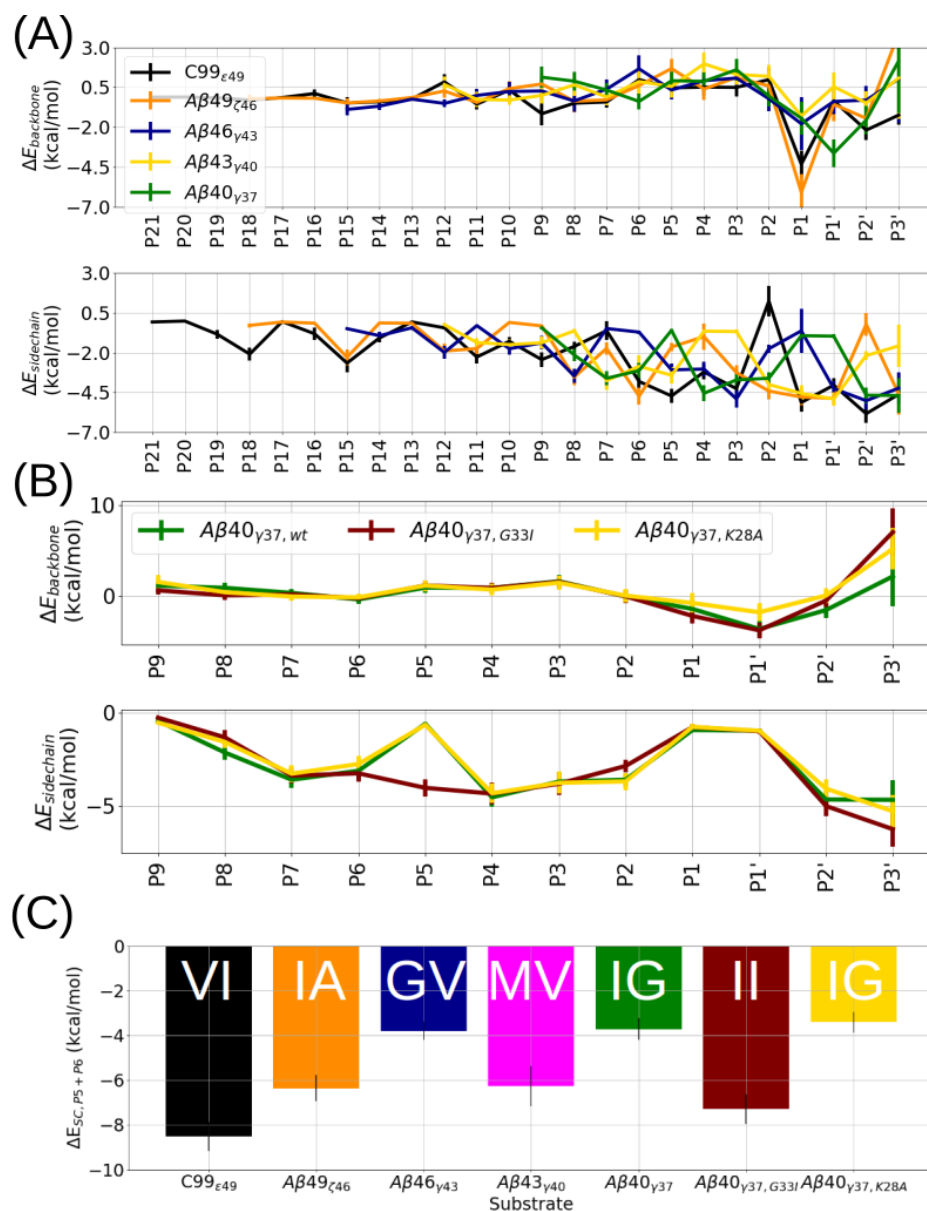


Figure C.14.: Binding energy between different substrates and  $\gamma$ -secretase with the D257-protonated PS1. (A) Residue-wise binding energy decomposition between  $\gamma$ -secretase and C99 $_{\epsilon 49}$  (black), A $\beta$ 49 $_{\zeta 46}$  (orange), A $\beta$ 46 $_{\gamma 43}$  (dark blue), A $\beta$ 43 $_{\gamma 40}$  (magenta), A $\beta$ 40 $_{\gamma 37}$  (green). Backbone (top) and sidechain (bottom) contributions are averaged through two replicas. (B) Residue-wise binding energy decomposition between  $\gamma$ -secretase and A $\beta$ 40 $_{\gamma 37}$  with wild-type (solid) G33I (brown), and K28A (yellow) mutated sequences. Backbone (top) and sidechain (bottom) contributions are averaged through two replicas. (C) Summation of substrate P5 and P6 sidechain binding energy contribution. The corresponding amino acids at P6 and P5 are annotated at top of each bar.

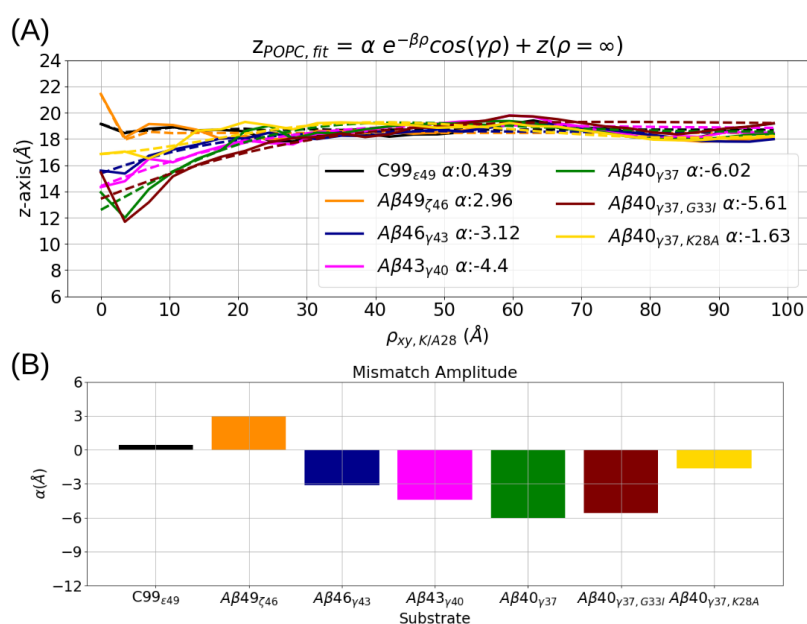


Figure C.15.: Fitting the membrane thickness distributed along the radial distance on the  $xy$  plane from K/A28 of substrates with D257-protonated PS1. (A) Fitting the hydrophobic mismatch profile with hydrophobic mismatch amplitude  $\alpha$ , radial decaying rate  $\beta$ , and harmonic oscillation  $\gamma$ . (B) Comparison of the hydrophobic mismatch amplitude  $\alpha$  calculated from (A) in different Aβ-bound γ-secretase structures.

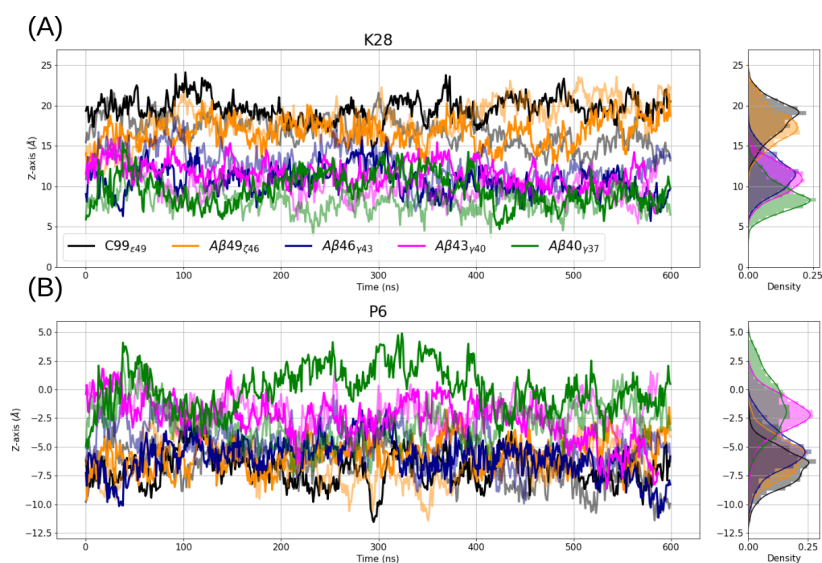


Figure C.16.: Z-axis position of substrate (A) K28 and (B) P6 in complex with  $\gamma$ -secretase with D257-protonated along the simulation time. Two replicas are represented by solid and transparent lines in the same color. The averaged phosphate groups are located at  $z=18\text{\AA}$  plane (see Figure S23C).

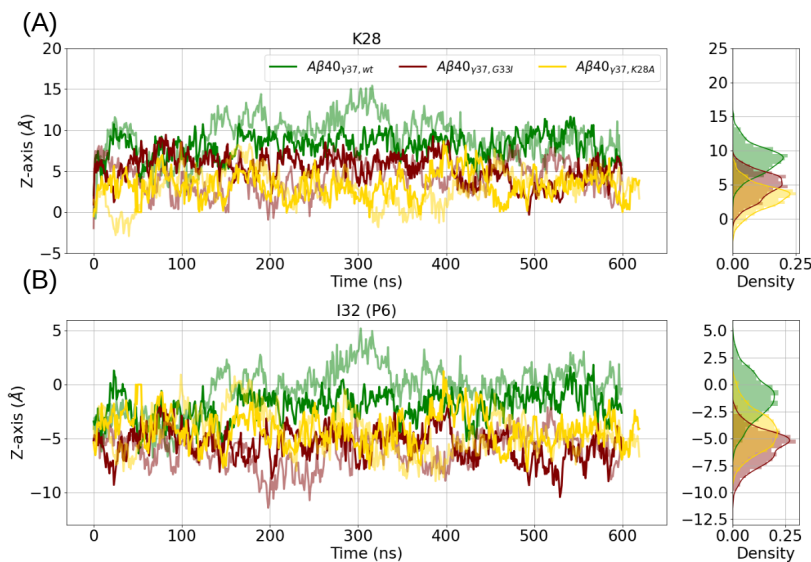


Figure C.17.: Z-axis position of substrate (A) K28 and (B) I32 in complex with  $\gamma$ -secretase with D385-protonated along the simulation time. Two replicas are represented by solid and transparent lines in the same color. The averaged phosphate groups are located at  $z=18\text{\AA}$  plane (see Figure 7C from the main text).

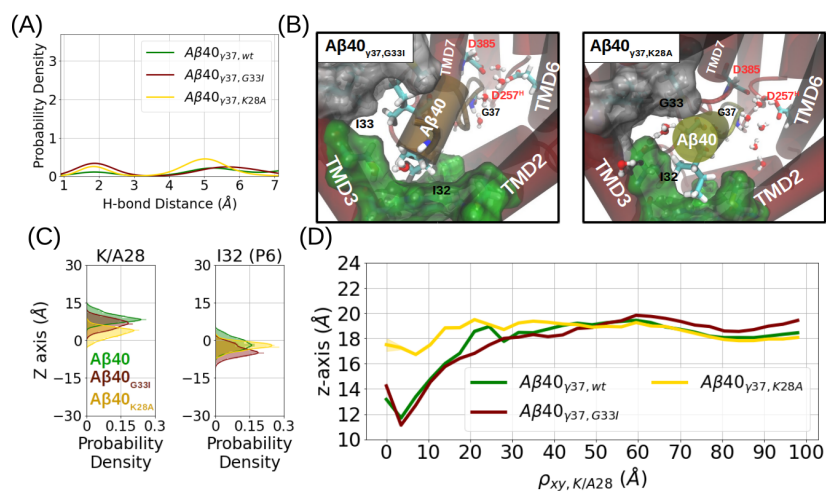


Figure C.18.: Influence of the APP mutations G33I and K28A on the A $\beta$ 40<sub>γ37</sub> binding pose with D257-protonated PS1. (A) Probability density of the catalytic hydrogen bond distance. (C) Distribution of the substrate K/A28 (middle) and substrate P6 (right) along the z-axis in different A $\beta$ n-γ-secretase complexes. (D) The average z-axis of the POPC phosphate on the extracellular side is distributed along the radial distance  $\rho_{xy, K28}$ .

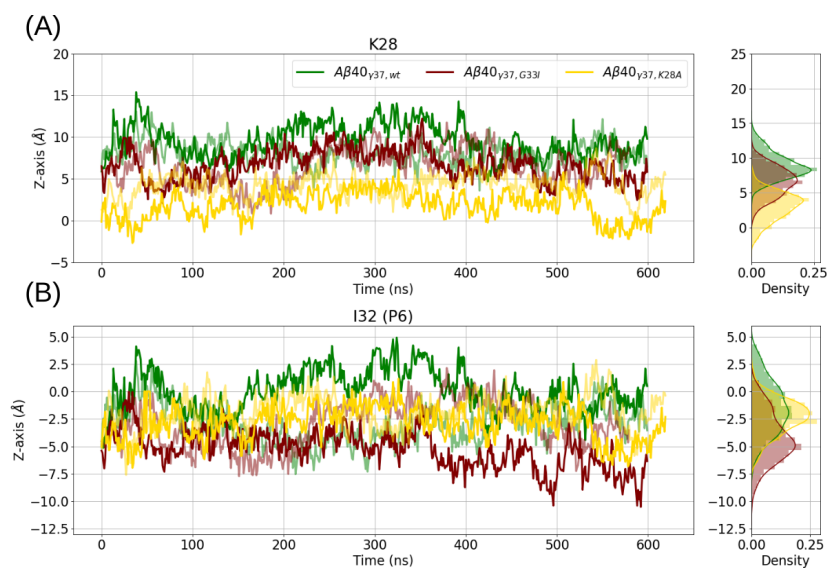


Figure C.19.: Z-axis position of substrate (A) K28 and (B) P6 in complex with γ-secretase with D257-protonated along the simulation time. Two replicas are represented by solid and transparent lines in the same color. The averaged phosphate groups are located at  $z=18\text{Å}$  plane (see Figure S23C).



## D. Supplementary Information for Chapter 8: A Putative Binding Mode of $\gamma$ -Secretase Modulator at the A $\beta$ -Membrane Interface Attenuates Negative Hydrophobic Mismatch

### D.1. Supplementary Figures

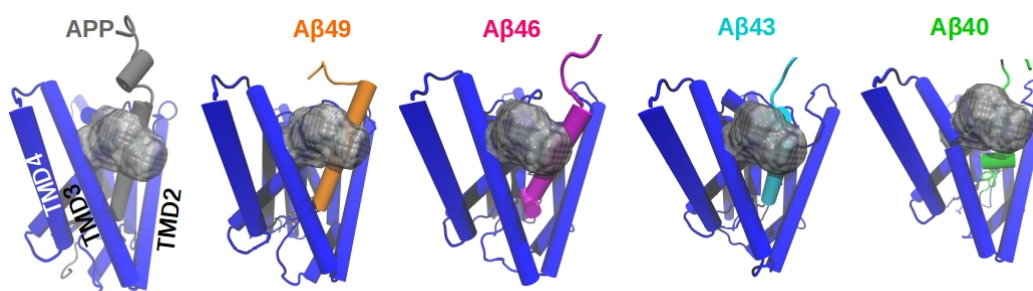


Figure D.1.: Probes placed 5 Å around the GSM binding site for pocket analysis during MD simulations (see Methods and Materials). Schematics show the location of the probes in the APP-bound (gray), A $\beta$ 49-bound (orange), A $\beta$ 46-bound (magenta), A $\beta$ 43-bound, and A $\beta$ 40-bound forms of  $\gamma$ -secretase.

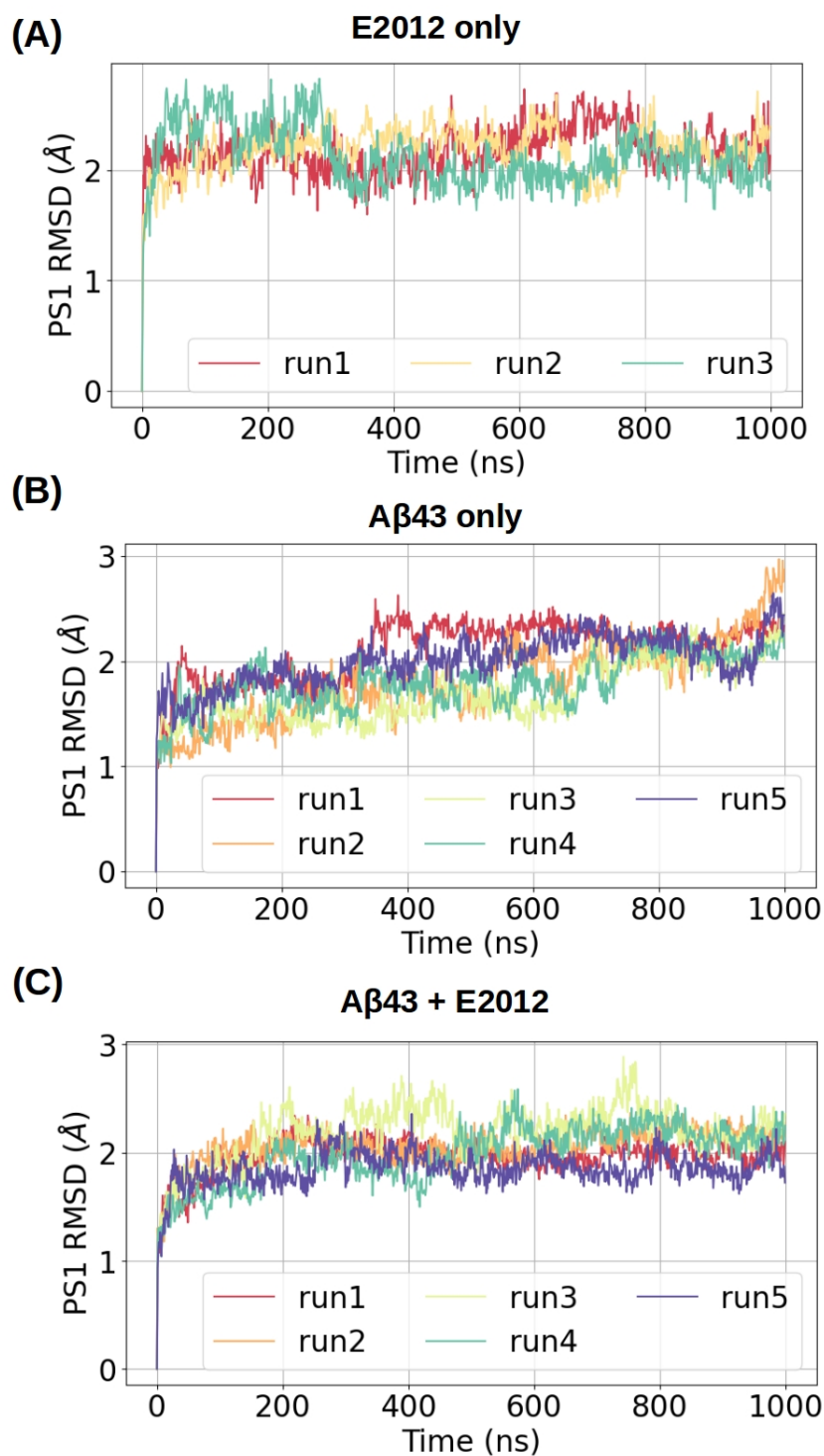


Figure D.2.: Stability assessment of PS1 in terms of root-mean-square deviation (RMSD) in the (A) E2012-bound, (B) A $\beta$ 43-bound, and (C) E2012-A $\beta$ 43-bound  $\gamma$ -secretase.

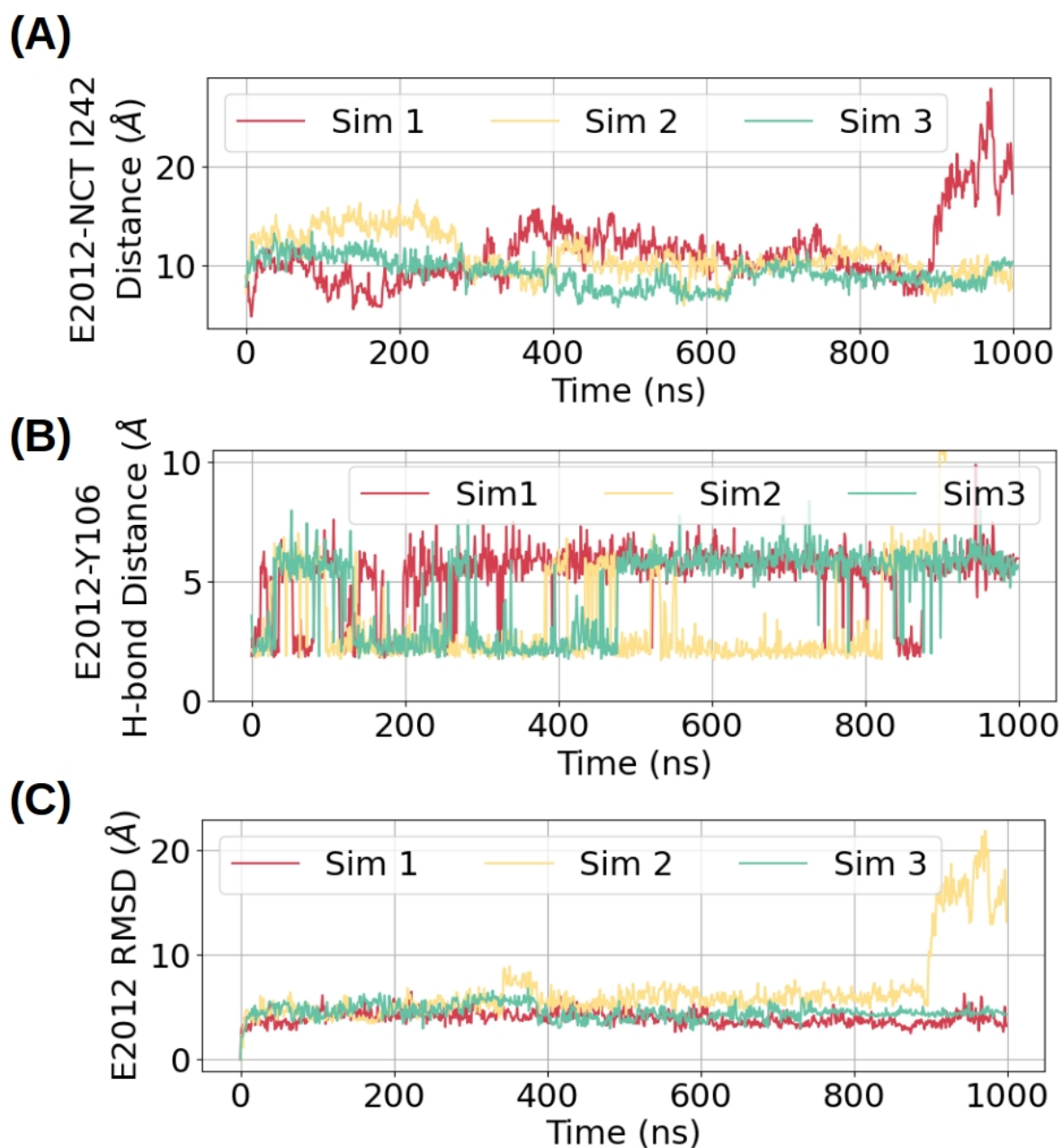


Figure D.3.: Dynamics and stability of E2012 during three 1 $\mu$ s simulations. (A) Time evolution of the center-of-mass distance between E2012 and I242 of NCT. (B) Time evolution of the hydrogen bond distance between the E2012 and PS1 Y106. (C) Time evolution of the RMSD of E2012 against the starting structure.

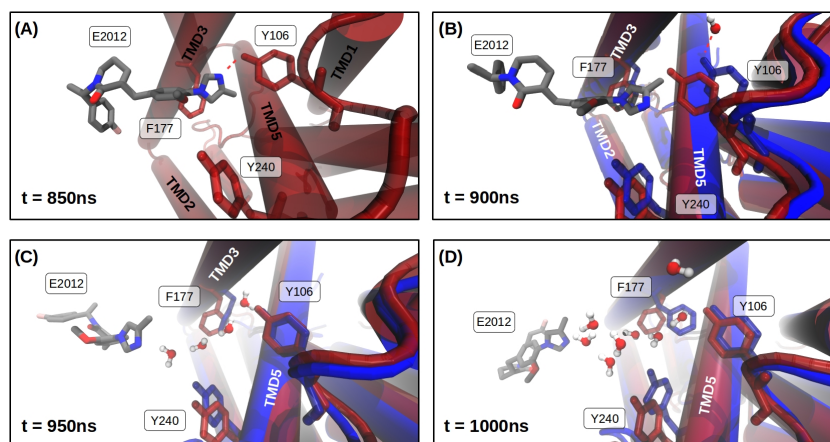


Figure D.4.: Dissociation pathway in the second simulation of E2012-bound  $\gamma$ -secretase. (A) Snapshot of the associated state at  $t = 850$ ns. (B) Snapshot of the associated state with the Y106 slightly deviated away from E2012 and formed a hydrogen bond with a water molecule at  $t = 900$ ns. (C) Snapshot of the dissociated state with Y106 moved back and formed a hydrogen bond with another water molecule in the GSM-binding site at  $t=950$ ns. (D) Snapshot of the dissociated state where F177 and Y240 moved toward the GSM-binding site. at  $t=1000$ ns (end of the simulation).

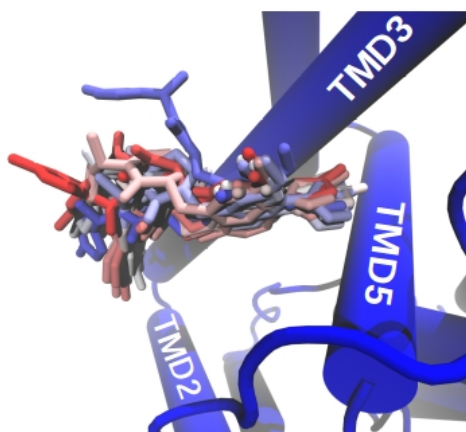


Figure D.5.: Different conformations sampled in the first simulation of E2012-bound  $\gamma$ -secretase complex. Conformations E2012 are shown for every 10ns between 0ns to 1000ns and color-coded according to their simulation time (from red at 0ns to blue at 1000ns).

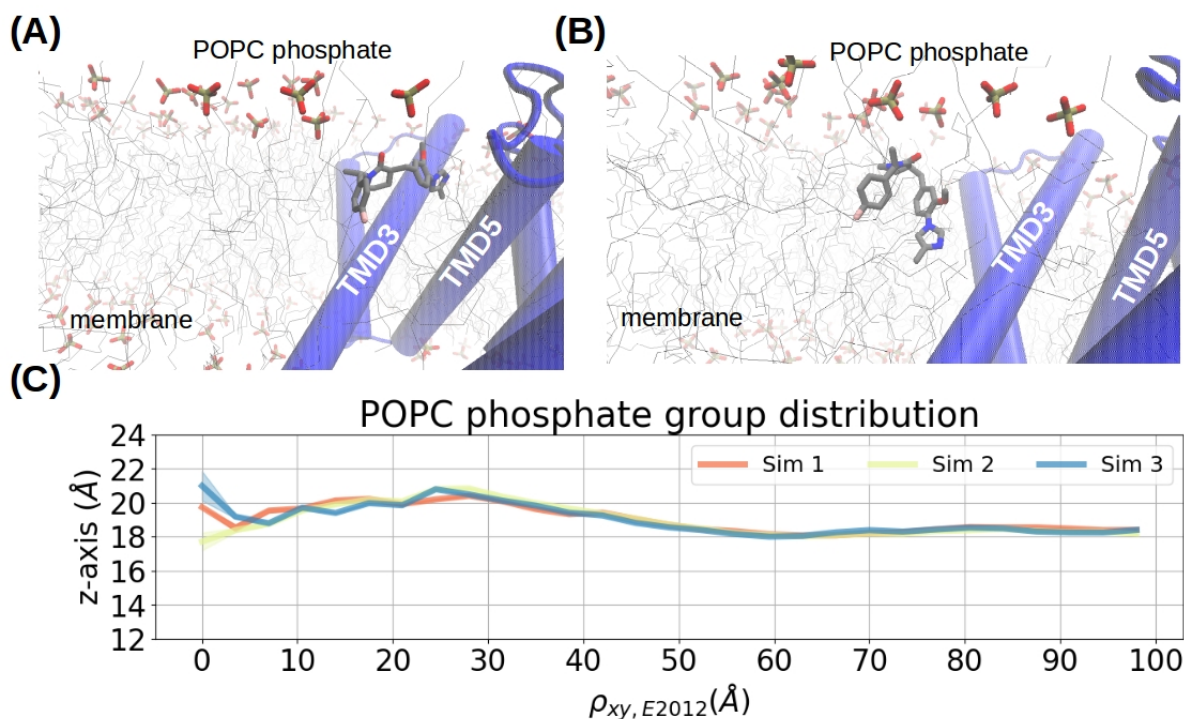


Figure D.6.: Relative location of E2012 and interaction with the POPC membrane. (A) Snapshot of the simulation when E2012 was associated with PS1. (B) Snapshot of the simulation when E2012 was dissociated from PS1. Phosphate groups of the POPC membrane are shown with the licorice representation and the hydrophobic tails are shown with gray lines. PS1 is shown as the blue helices and E2012 is shown in the gray licorice representation. (C) Membrane thickness fluctuation is represented by the z-axis of the phosphate groups as a function of the radial distance on the xy plane away from the center of mass of E2012. Data shown are the average phosphate height in each simulation with a 1ns data collection interval and the standard deviations are shown with the shaded area.

D.1. SUPPLEMENTARY FIGURES

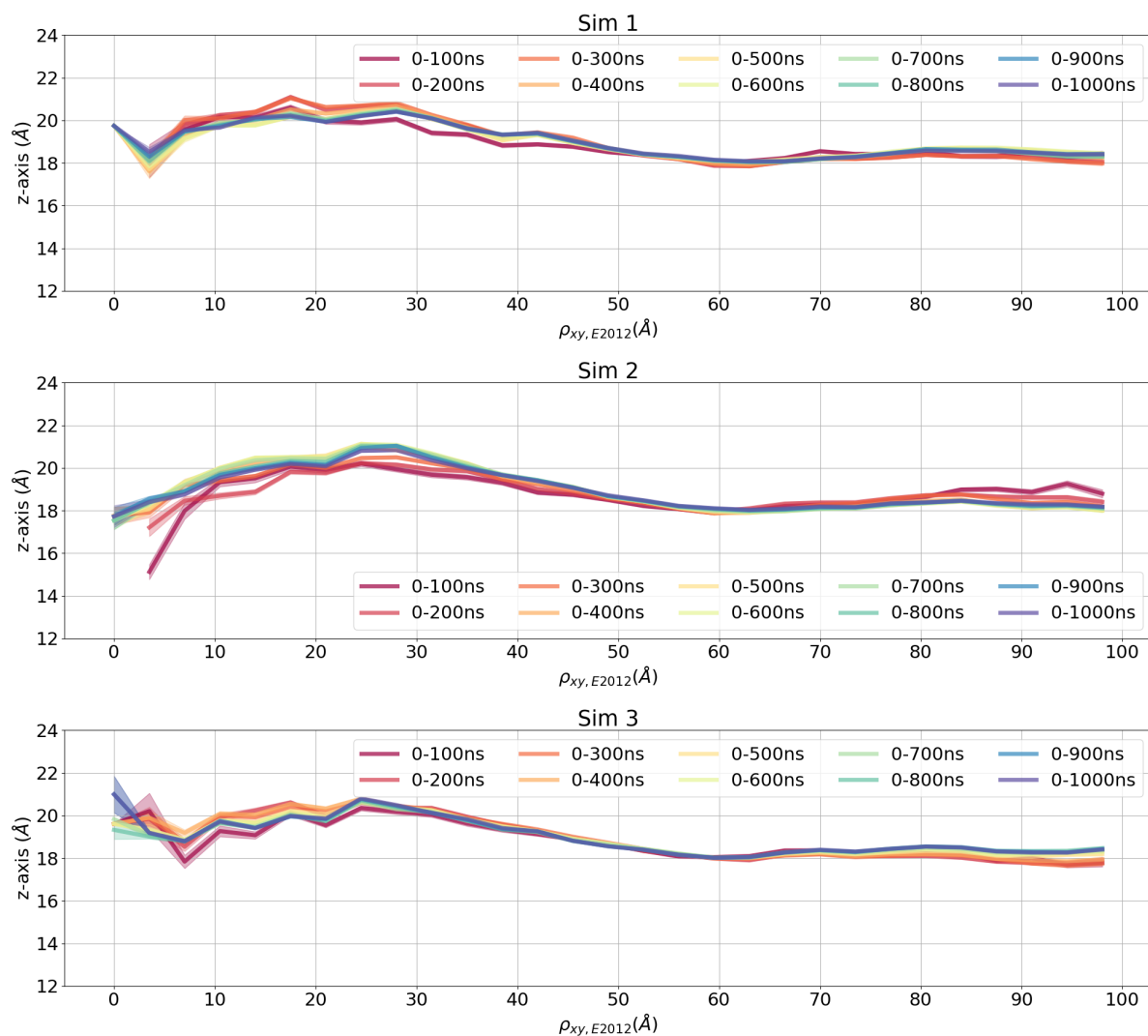


Figure D.7.: Convergence check of the membrane curvature in each E2012-bound  $\gamma$ -secretase simulation with an time increment of 100ns.

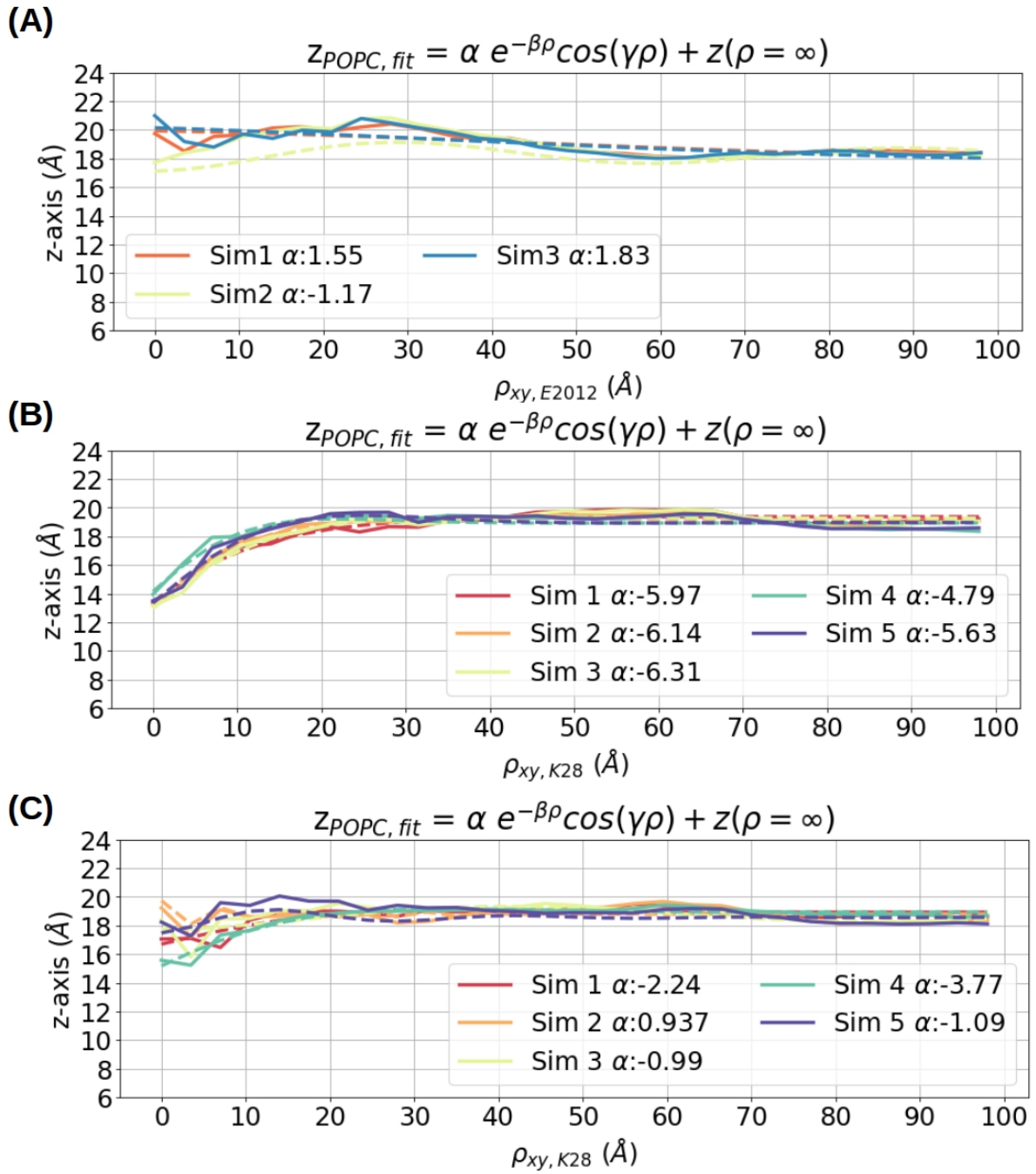


Figure D.8.: Membrane curvatures in the (A) E2012-bound, (B) A $\beta$ 43-bound, and (C) E2012-A $\beta$ 43-bound  $\gamma$ -secretase simulations fitted to the function  $z_{POPC} = \alpha e^{-\beta\rho} \cos(\gamma\rho) + z_{POPC}(\rho = \infty)$ . The mismatch amplitude  $\alpha$  is especially shown to indicate the mismatch amplitude.

D.1. SUPPLEMENTARY FIGURES

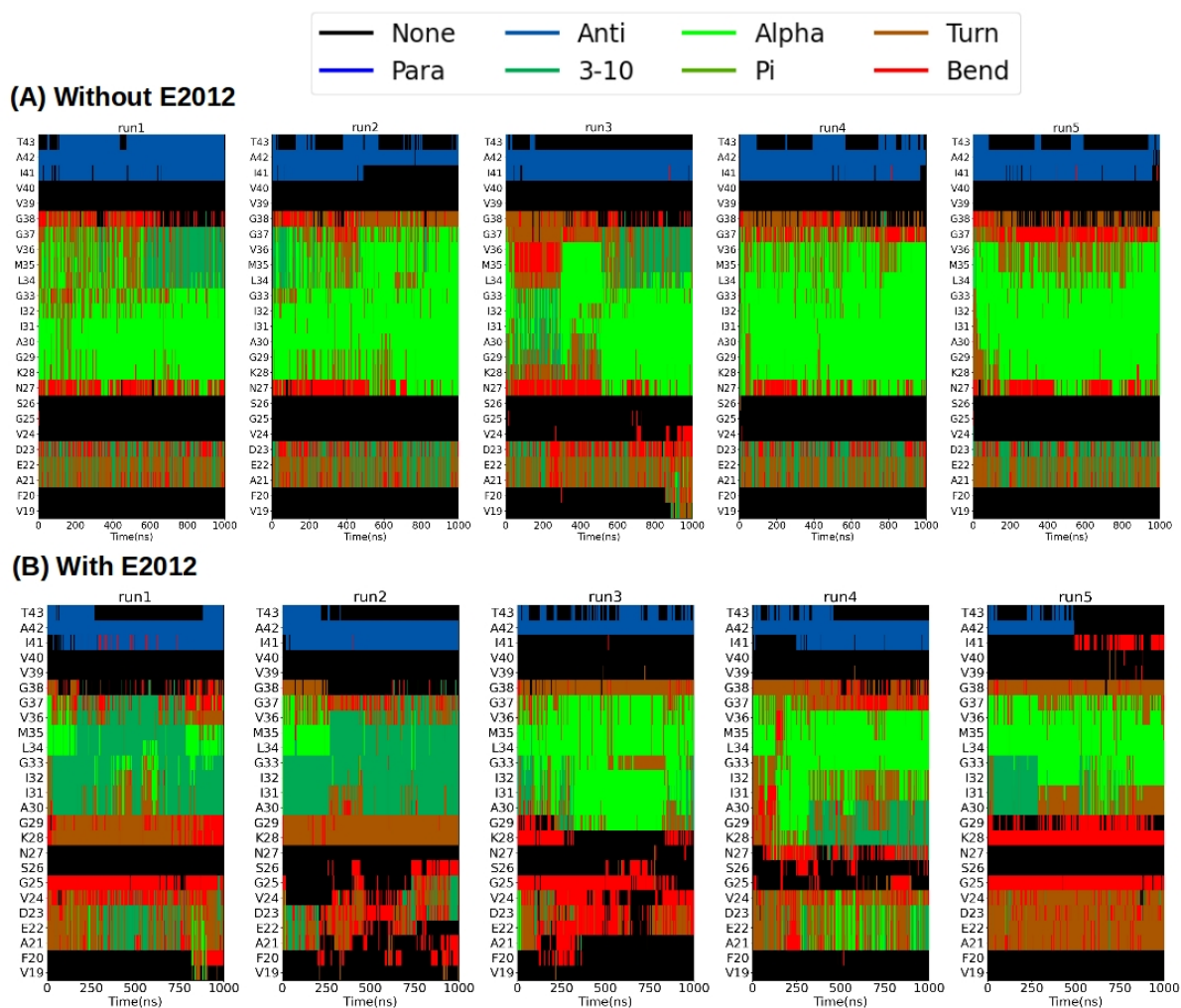


Figure D.9.: Secondary structure along simulations time during the simulation of (A) A $\beta$ 43-bound and (B) E2012-A $\beta$ 43-bound  $\gamma$ -secretase analyzed using the DSSP method.



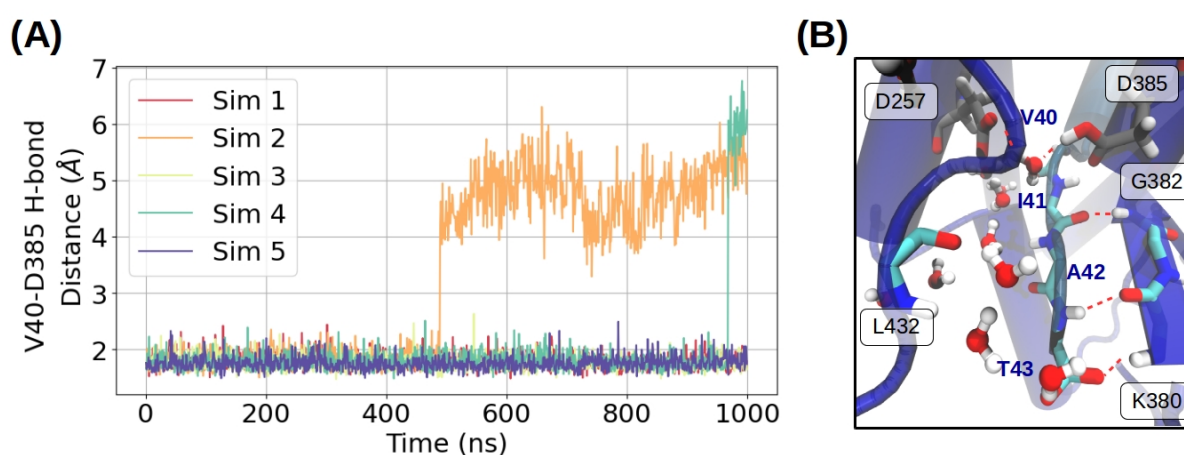


Figure D.10.: Dynamics of the catalytic hydrogen bond in the A $\beta$ 43-bound  $\gamma$ -secretase simulations. (A) Time evolution of the catalytic hydrogen bond formed between A $\beta$ 43 V40 and the protonated D385 of PS1. (B) Geometry at the hybrid  $\beta$ -strand area with the disrupted catalytic hydrogen bond because of the water molecules coming from the intracellular side. A $\beta$ 43 substrate is colored in light blue and PS1 in dark blue. Hydrogen bonds between A $\beta$ 43 and PS1 are indicated with the red dashed lines.

## D.1. SUPPLEMENTARY FIGURES

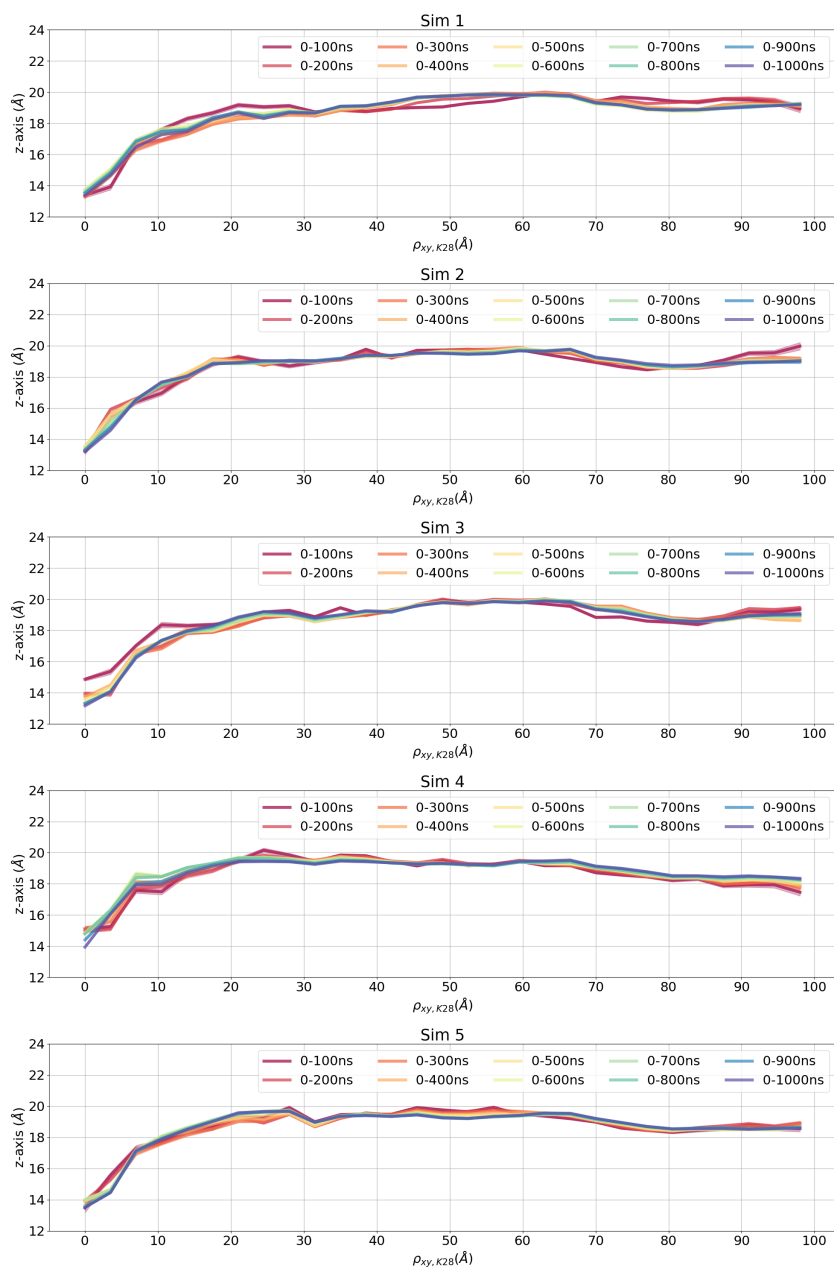


Figure D.11.: Convergence check of the membrane curvature in each A $\beta$ 43-bound  $\gamma$ -secretase simulation with an time increment of 100ns.

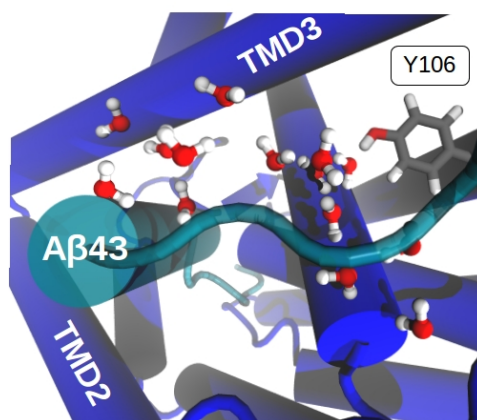


Figure D.12.: Water recruited into the GSM-binding site during the simulation of  $\beta$ 43-bound  $\gamma$ -secretase simulation.

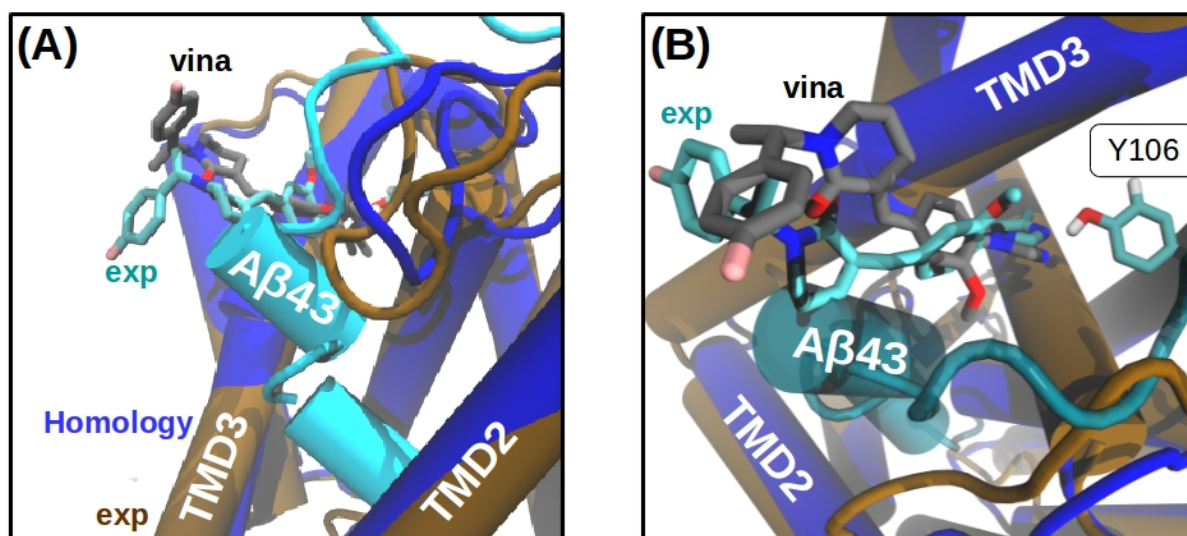


Figure D.13.: Starting structure of the E2012-A $\beta$ 43-bound  $\gamma$ -secretase simulation generated by molecular docking in the (A) side-view and (B) top-view. Carbons of the experimentally determined docking pose (PDBID: 7D8X) are shown with the cyan licorice representation and carbons of the docked E2012 generated by Autodock vina is shown with the gray representation.

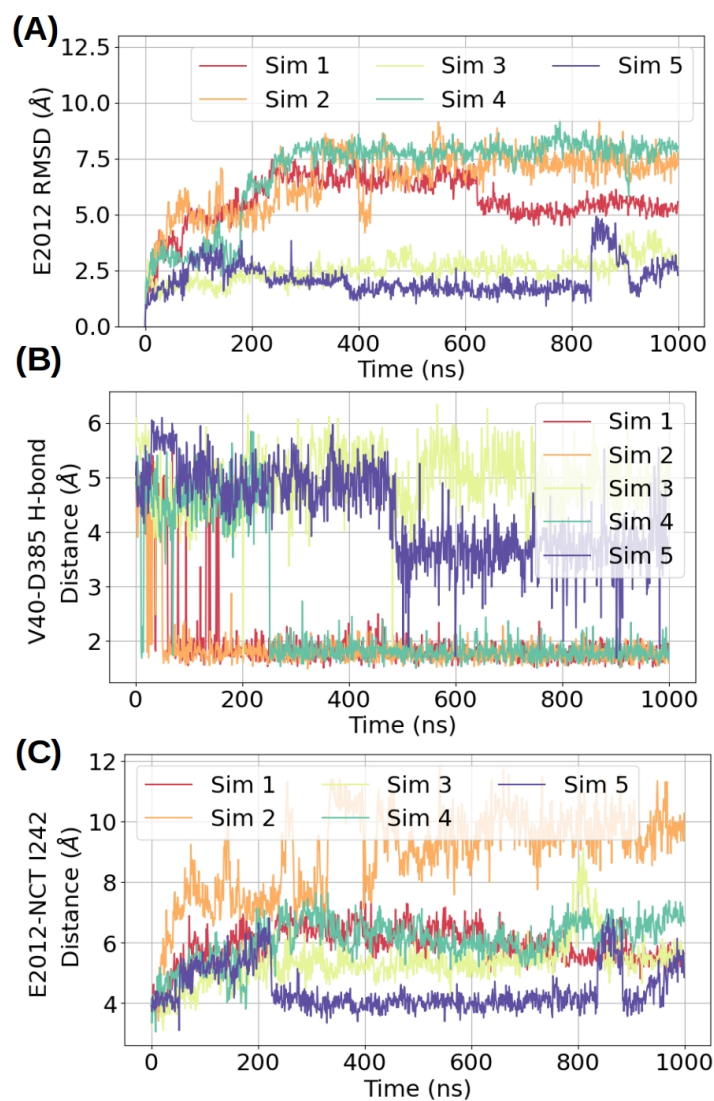


Figure D.14.: Dynamics and stability in the E2012-A $\beta$ 43-bound  $\gamma$ -secretase simulations. (A) Time evolution of the catalytic hydrogen bond between A $\beta$ 43 V40 and the protonated D385 of PS1. (B) Time evolution of RMSD in five simulations against the starting structure generated by molecular docking. (C) Time evolution of the distance between the hydrogen bond formed between E2012 and PS1 Y106.

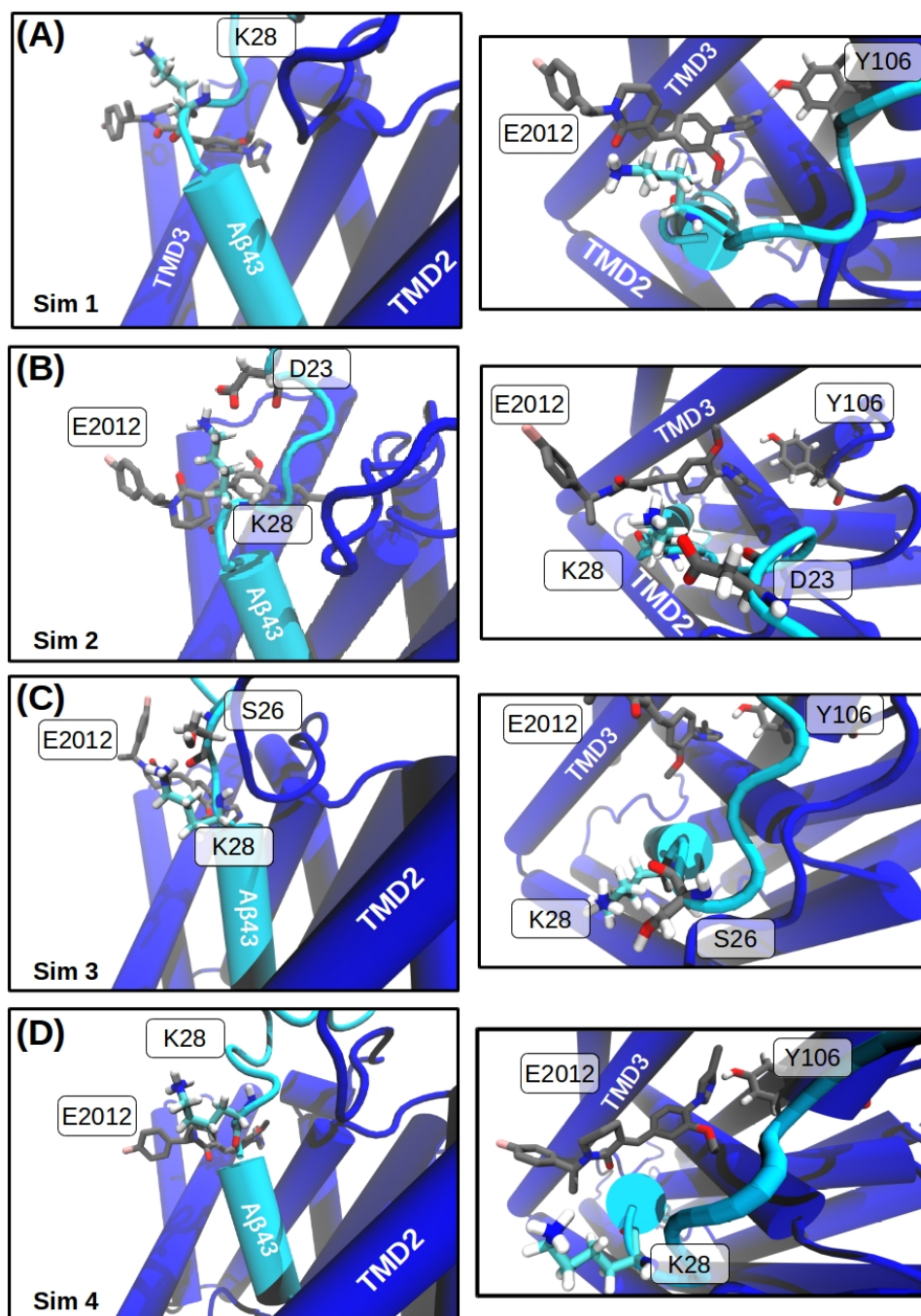


Figure D.15.: Snapshots of the binding poses in revealed in the last frame of the first four runs of the E2012-A $\beta$ 43-bound  $\gamma$ -secretase simulations. (A) Binding pose in run1. (B) Binding pose in run2. (C) Binding pose in run3. (D) Binding pose in run4. The substrate is depicted in light blue. D23 and S26 of A $\beta$ 43, Y106 of PS1, and E2012 are shown with the gray carbons. K28 of A $\beta$ 43 is shown in with the cyan carbons.

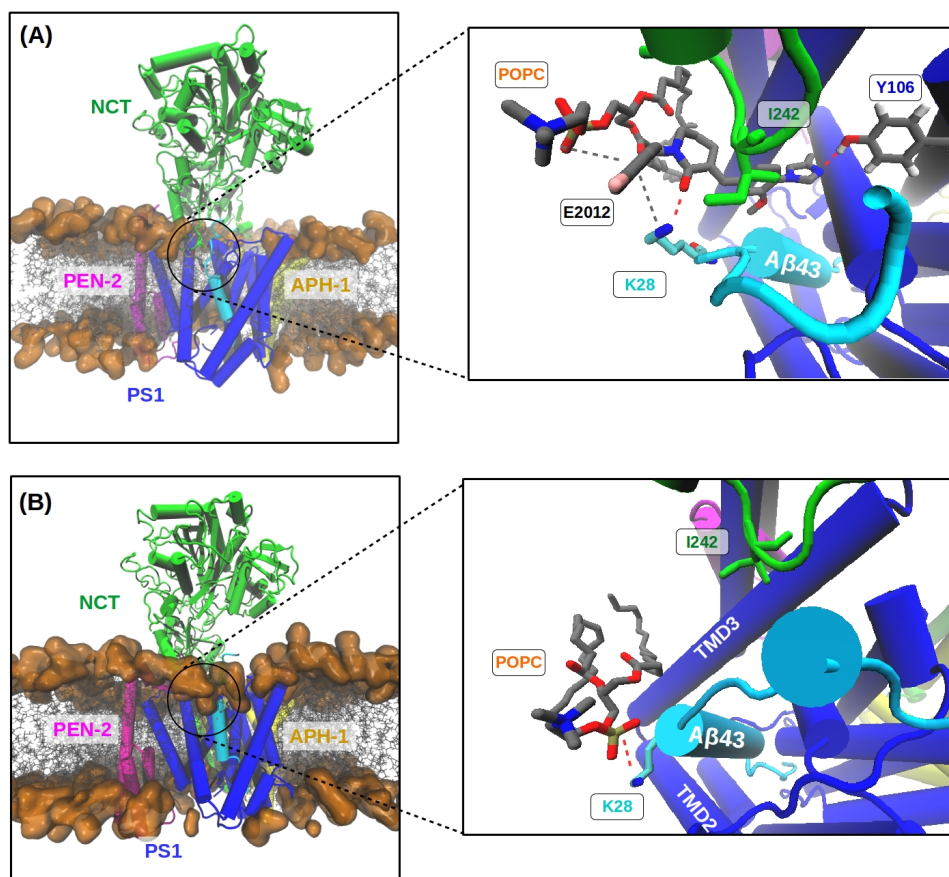


Figure D.16.: Influence of the binding of E2012 in the interaction between lipid phosphate and K28 of A $\beta$ 43. (A) The E2012-A $\beta$ 43- $\gamma$ -secretase ternary complex in the membrane bilayer with the zoom-in view in the right panel. E2012, stabilized by a hydrogen bond with PS1 Y106, sat between K28 and POPC and formed a hydrogen bond with K28 and cation- $\pi$  interactions with POPC and K28. (B) The A $\beta$ 43-bound  $\gamma$ -secretase with the zoom-in view in the right panel. K28 of  $\beta$ 43 formed ionic interaction with the lipid phosphate group. Hydrogen bonds and ionic interactions are shown in red dashed lines and the cation- $\pi$  interactions are shown in the gray dashed lines. I242 of NCT is shown in the green licorice representation. Carbons of POPC, PS1 Y106, and E2012 are shown in gray and the carbons of K28 are shown in cyan. All hydrogen atoms are hidden.

## D.1. SUPPLEMENTARY FIGURES

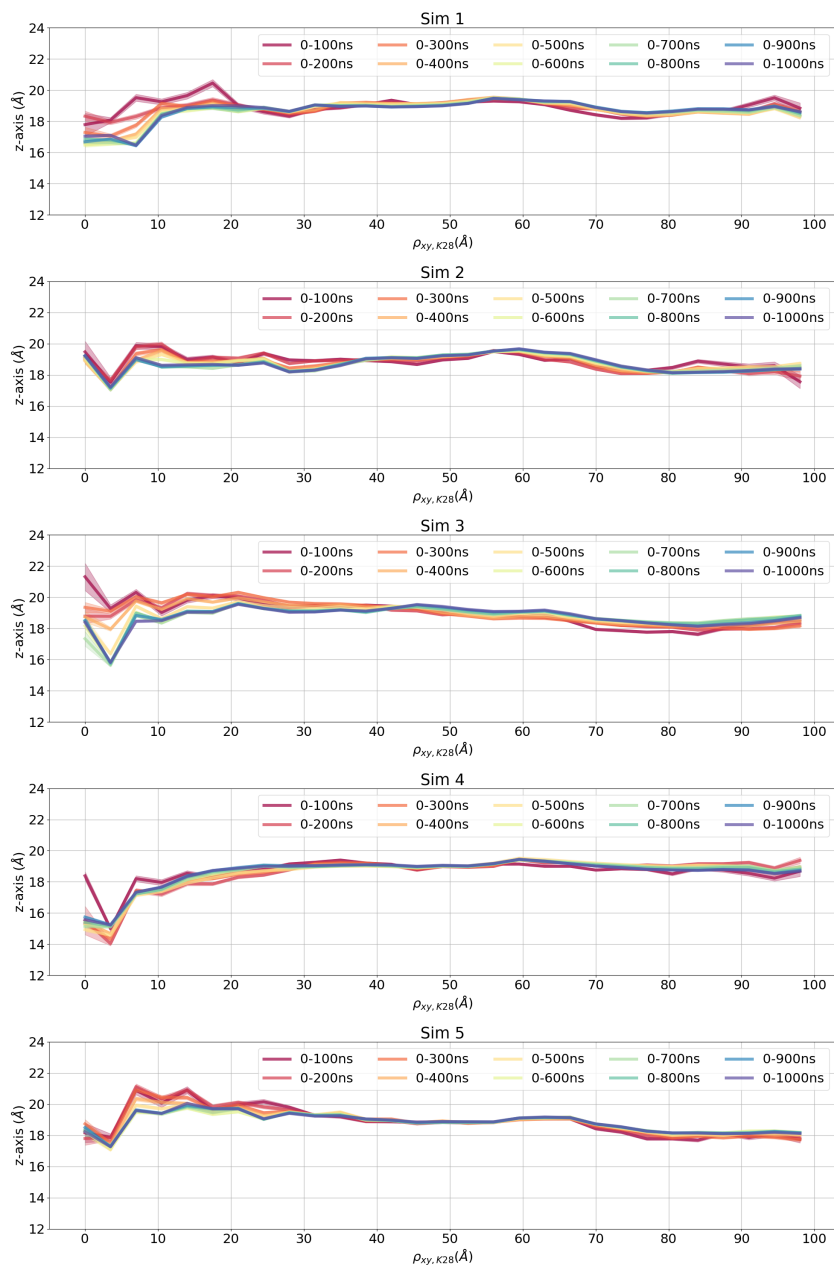


Figure D.17.: Convergence check of the membrane curvature in each E2012-A $\beta$ 43-bound  $\gamma$ -secretase simulation with an time increment of 100ns.

## E. Supplementary Information for Chapter 9: What makes a good protein-protein interaction stabilizer: Analysis and Application of the Dual-Binding Mechanism

### E.1. Supplementary Figures

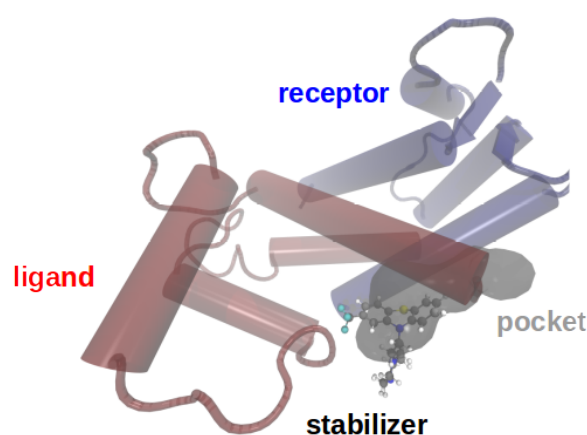


Figure E.1.: An illustration of the best-scored (pocket score) pocket detected from the C2 RL complex (PDBID: 3cga) using Fpocket with a ligand coverage fraction of 60%.



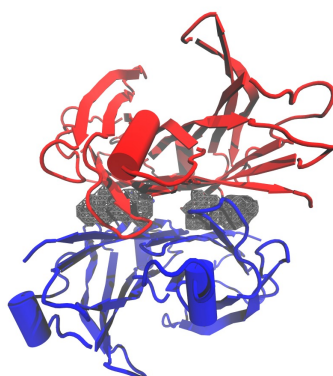


Figure E.2.: An illustration of the probes generated with 2Å around the ligand coordinates with TTR complex (PDBID 3tct).

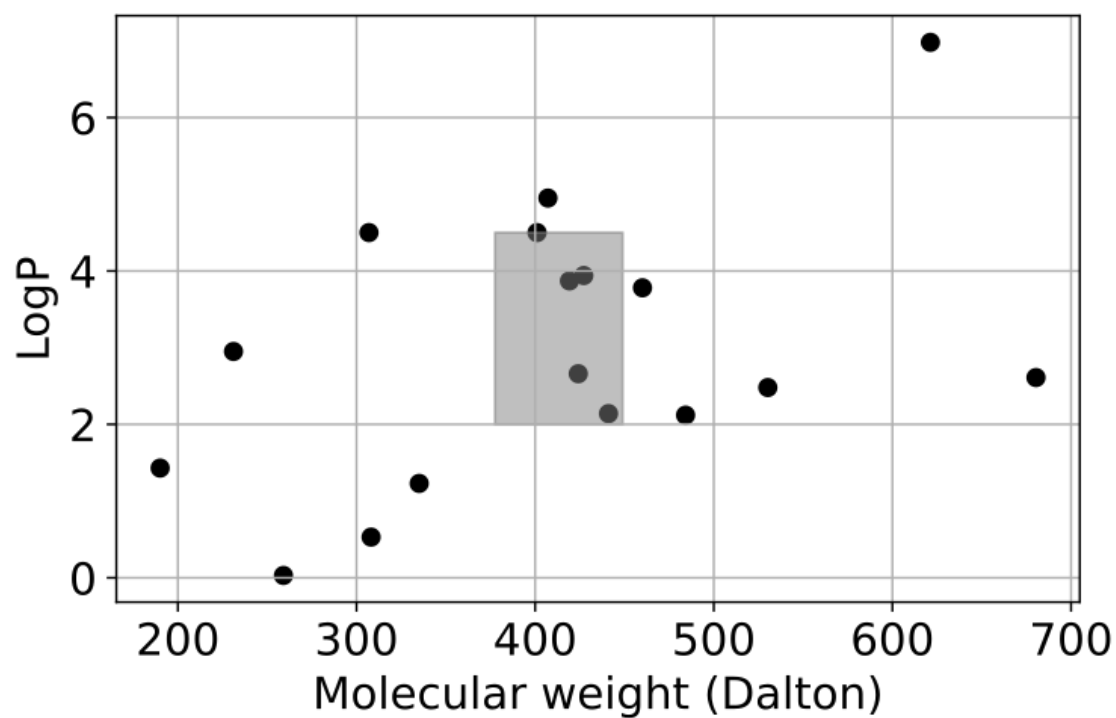


Figure E.3.: Population of the molecular weight and LogP of the stabilizers present in this study. The grey area indicates the range of molecular weight and LogP used to for virtual screening.

## E.1. SUPPLEMENTARY FIGURES

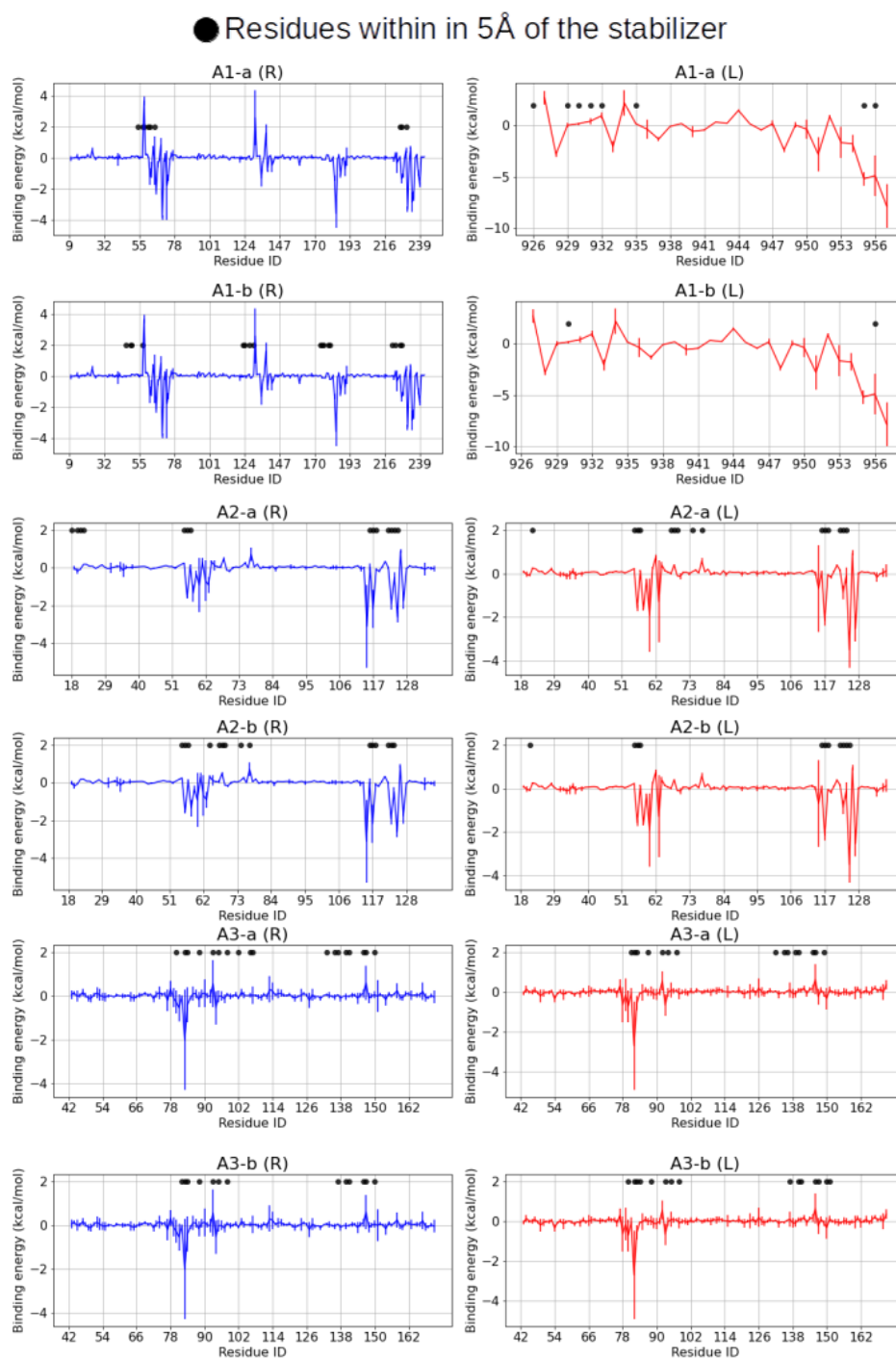


Figure E.4.: Residue-wise RL binding energy decomposition in set A complexes. The black dots represent the residues within 5Å of the stabilizer.

## E.1. SUPPLEMENTARY FIGURES

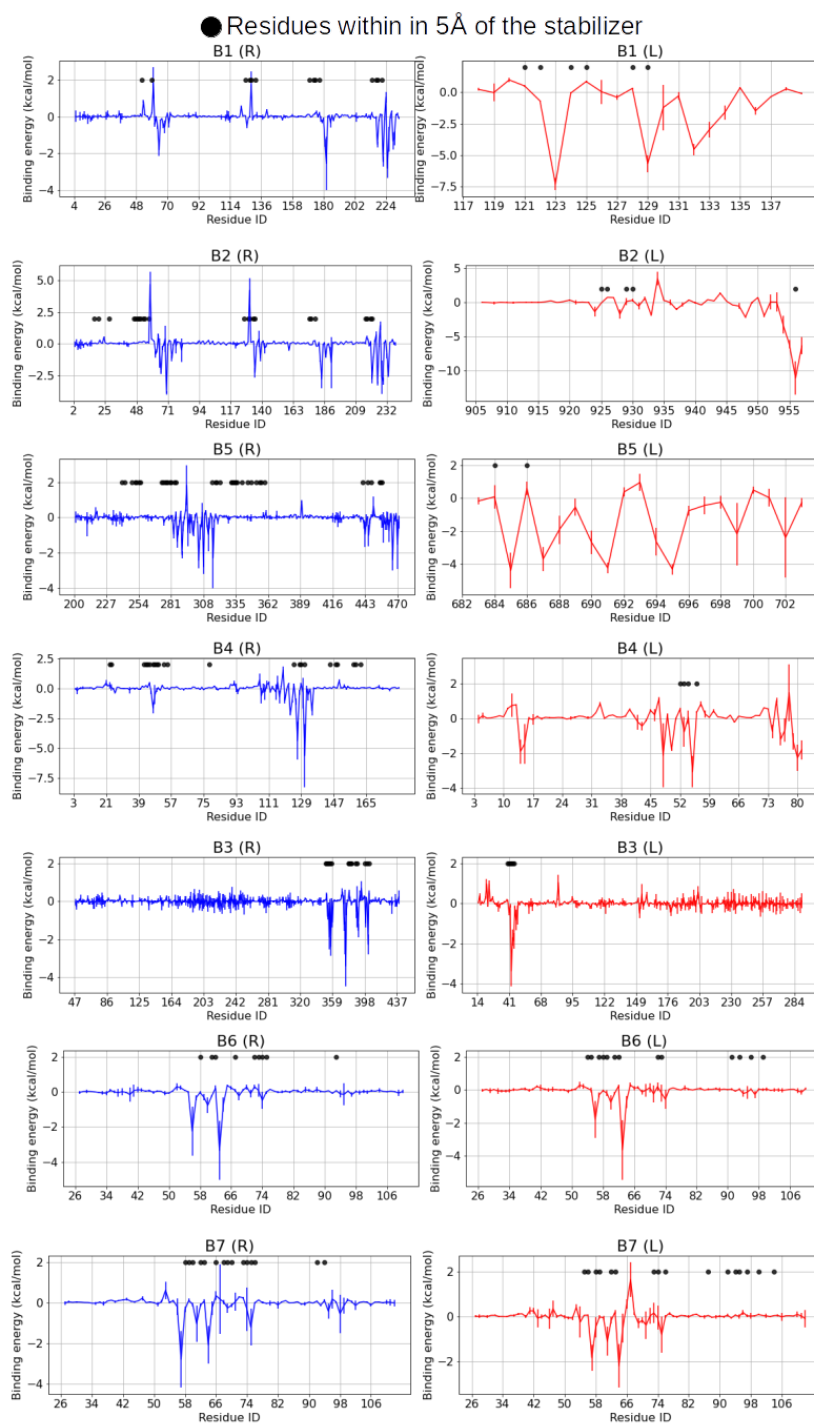


Figure E.5.: Residue-wise RL binding energy decomposition in set B complexes. The black dots represent the residues within 5Å of the stabilizer.

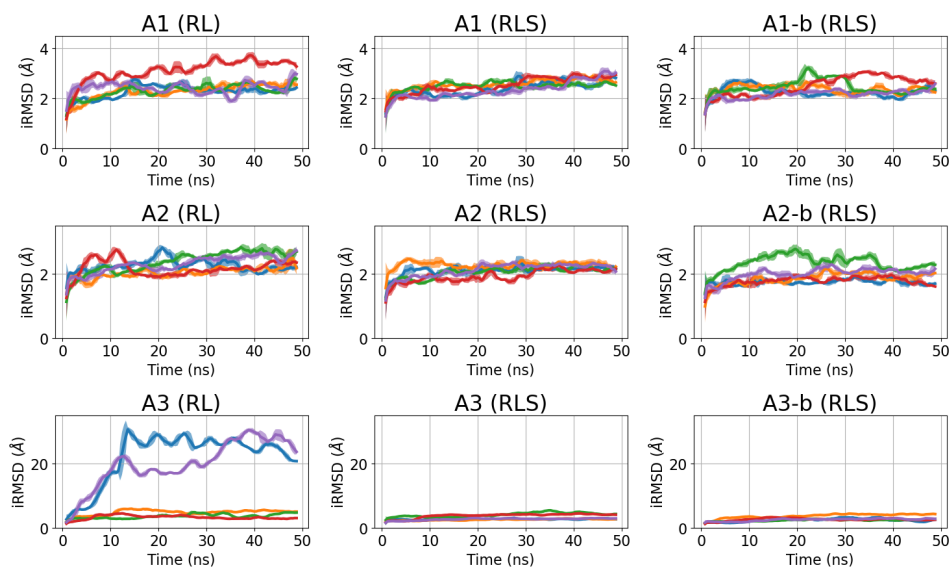


Figure E.6.: Interface RMSD (iRMSD) of the complexes in set A. Simulations without the stabilizers are shown in the left panel and simulations with the stabilizers are shown in the right panels. Five different colors represent five individual simulations. Data are shown as rolling averages over every 1 ns with standard deviation indicated by the shaded area.

Ligand	$\Delta\Delta G_{RS}$ (kcal/mol)	$\Delta\Delta G_{LS}$ (kcal/mol)	$ \Delta\Delta G_{RS} - \Delta\Delta G_{LS} $ (kcal/mol)	$\Delta\Delta G_{(RL)S}$ (kcal/mol)
1	<b>-11.11 ± 1.34</b>	-26.75 ± 2.2	15.65 ± 3.42	-39.19 ± 0.73
2	<b>-24.4 ± 1.69</b>	<b>-19.03 ± 2.95</b>	5.9 ± 3.57	-50.71 ± 3.13
3	<b>-15.2 ± 4.04</b>	-33.21 ± 1.68	18.01 ± 5.5	-58.41 ± 4.53
4	<b>-17.77 ± 1.92</b>	-23.67 ± 2.85	5.89 ± 3.27	-43.5 ± 3.52
5	<b>-13.25 ± 3.04</b>	-24.68 ± 1.93	11.44 ± 3.36	-41.48 ± 5.14
6	<b>-15.78 ± 3.91</b>	-22.8 ± 3.17	8.39 ± 4.96	-45.41 ± 4.35
7	<b>-18.77 ± 1.17</b>	-24.76 ± 4.28	6.0 ± 3.87	-43.9 ± 7.04
8	-27.49 ± 2.97	<b>-13.28 ± 3.22</b>	14.22 ± 5.19	-45.91 ± 4.11
9	<b>-20.41 ± 2.88</b>	-20.8 ± 3.07	4.8 ± 3.38	-45.99 ± 2.46
10	<b>-11.48 ± 6.29</b>	-26.48 ± 5.41	17.13 ± 8.21	-36.31 ± 2.42

Table E.1.: Binding free energy of the 10 ligands binding to the interface PDBID 1AK4. Structures of the ligands are shown in Figure E.15. The binding affinity between the compound and the more weakly bound protein is emphasized with bold font. The 3 compounds with the lowest  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$  are expected to be the most promising stabilizers and are highlighted in gray.

## E.1. SUPPLEMENTARY FIGURES

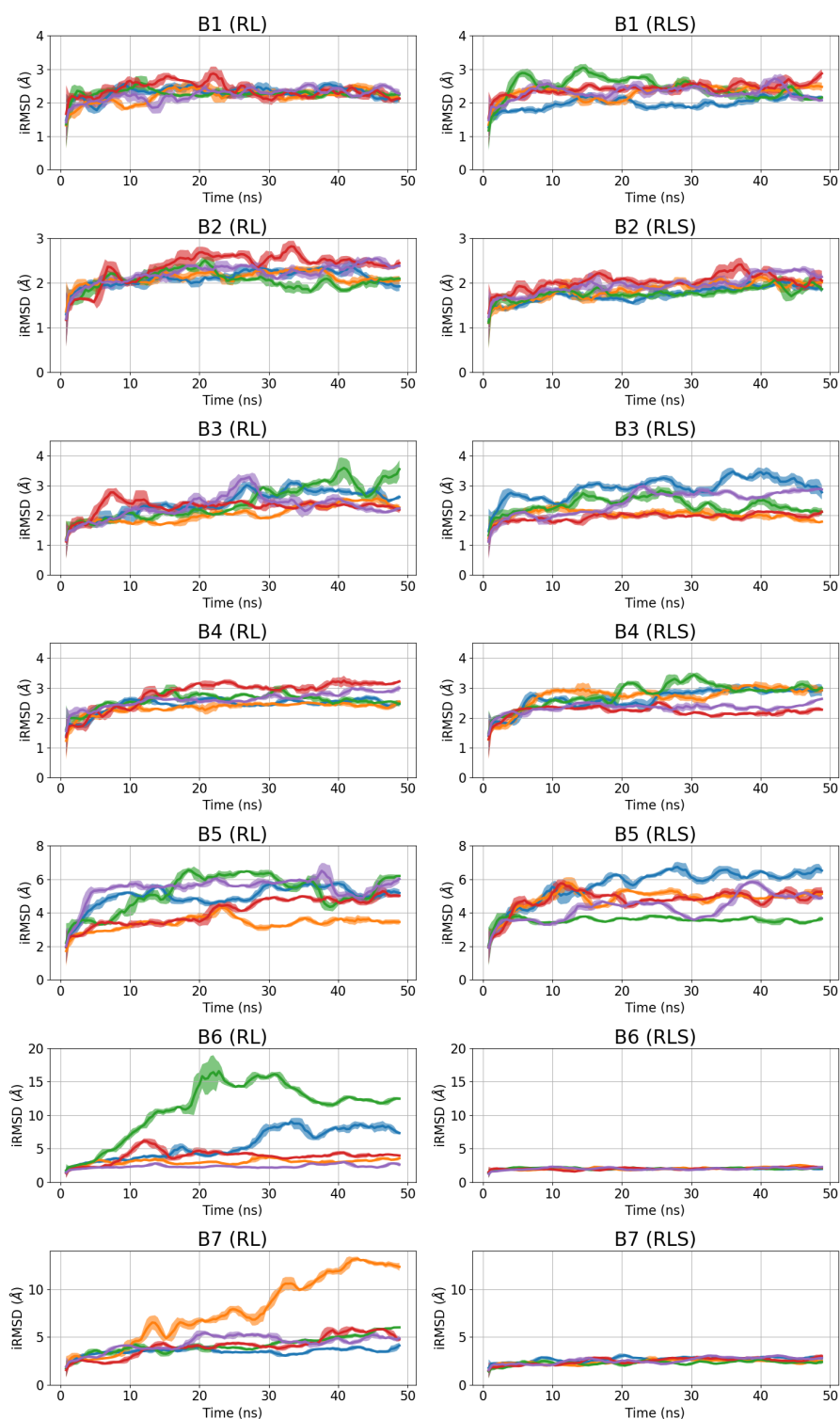


Figure E.7.: Interface RMSD (iRMSD) of the complexes in set B. Simulations without the stabilizers are shown in the left panel and simulations with the stabilizers are shown in the right panel. Five different colors represent five individual simulations. Data are shown as rolling averages over every 1 ns with standard deviation indicated by the shaded area.

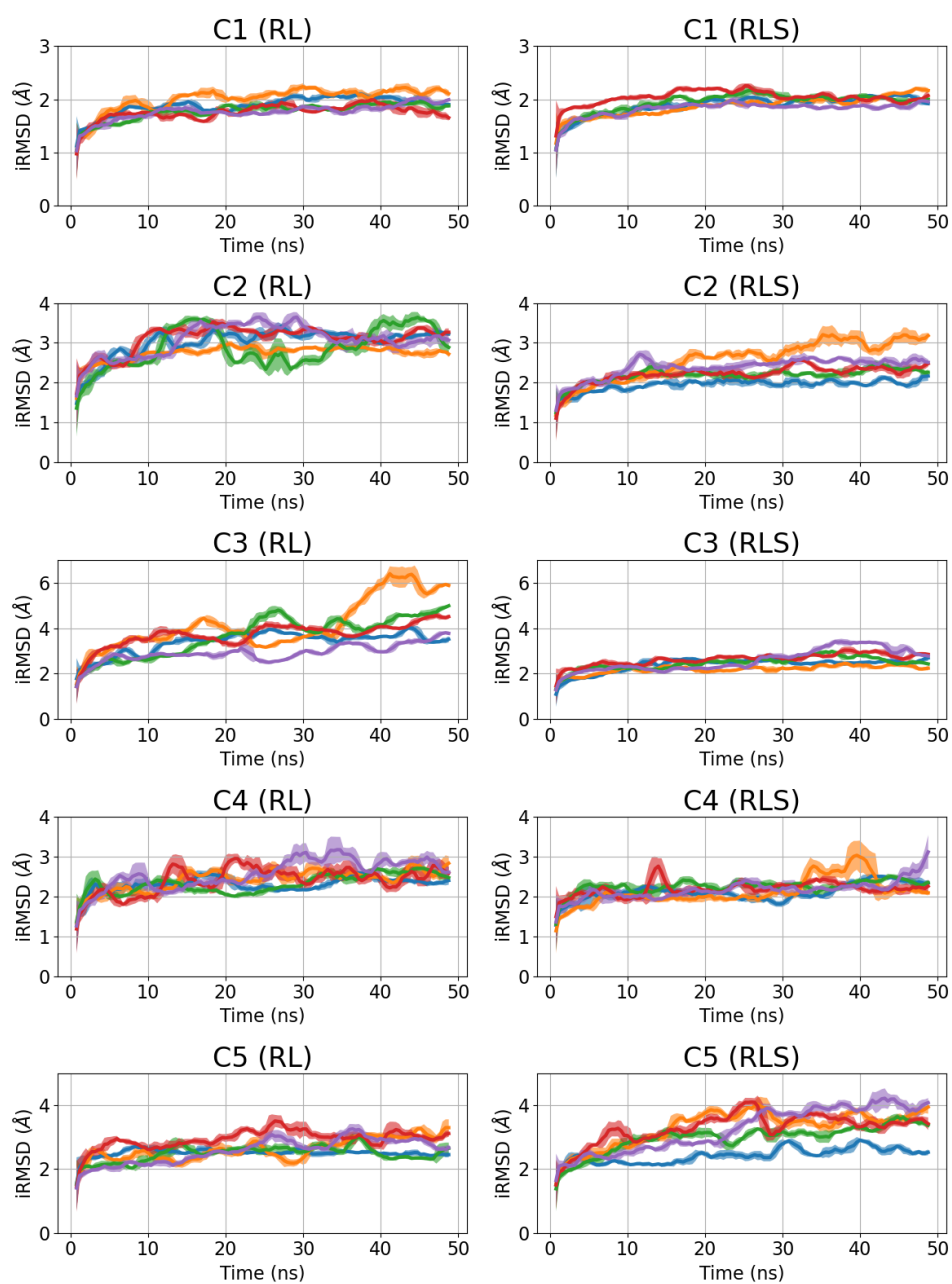


Figure E.8.: Interface RMSD (iRMSD) of the complexes in set C. Simulations without the stabilizers are shown in the left panel and simulations with the stabilizers are shown in the right panel. Five different colors represent five individual simulations. Data are shown as rolling averages over every 1 ns with standard deviation indicated by the shaded area.

## E.1. SUPPLEMENTARY FIGURES

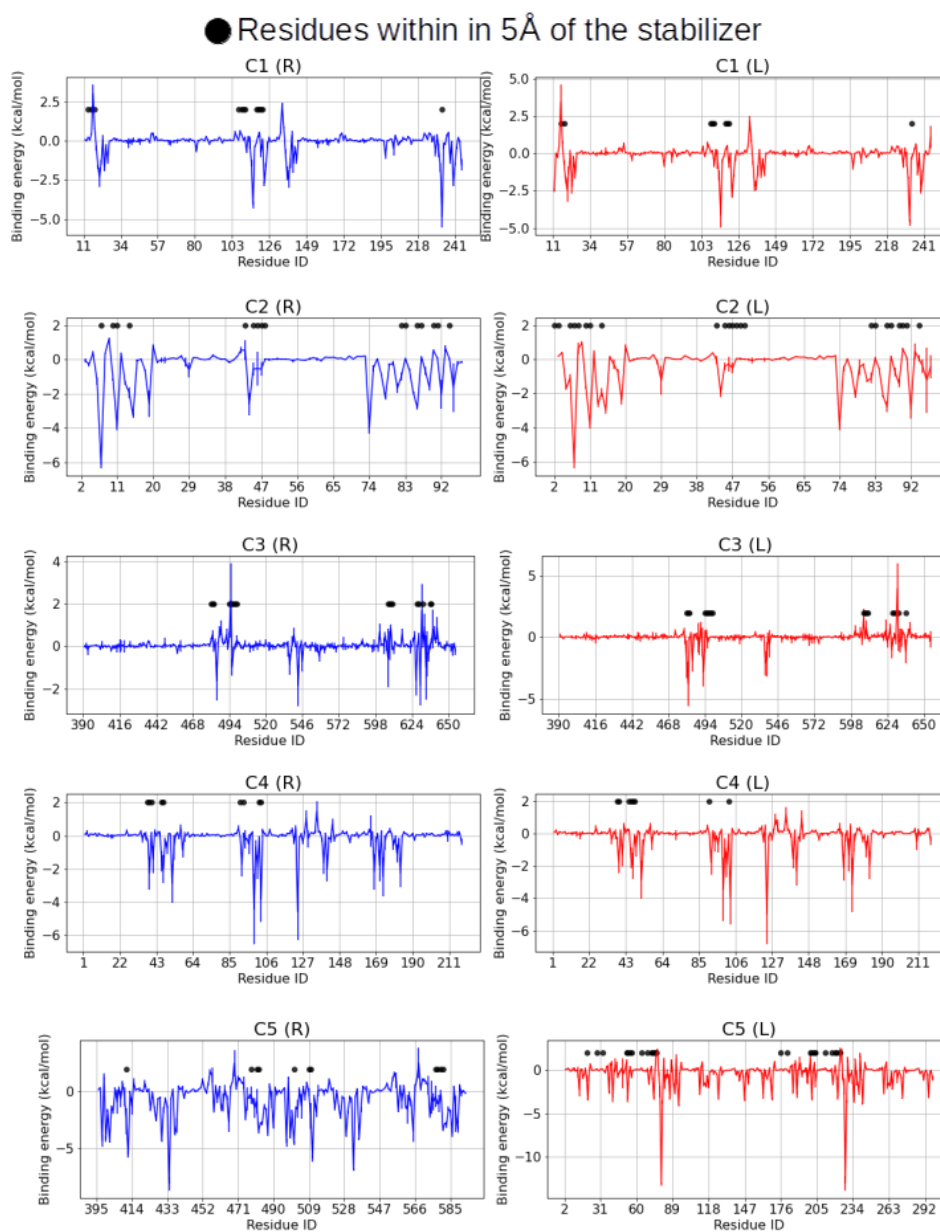


Figure E.9.: Residue-wise RL binding energy decomposition in set C complexes. The black dots represent the residues within 5Å of the stabilizer.

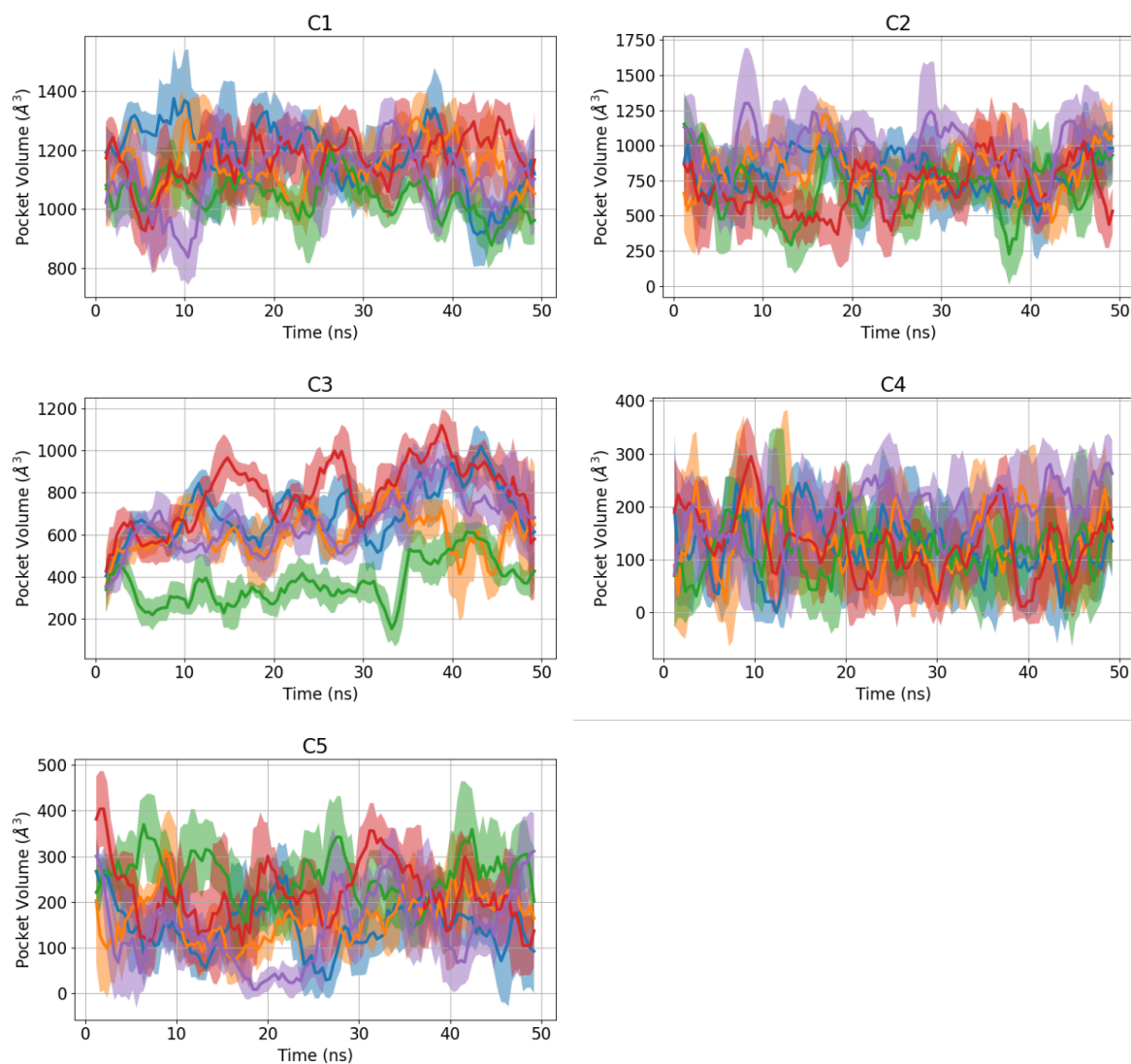


Figure E.10.: Time-evolution of the stabilizer-binding pocket volume in set C complexes without stabilizers. Data from five individual simulations are color-coded the same way as in Figure E.9. Data are shown as rolling averages over every 1 ns with standard deviation indicated by the shaded area.



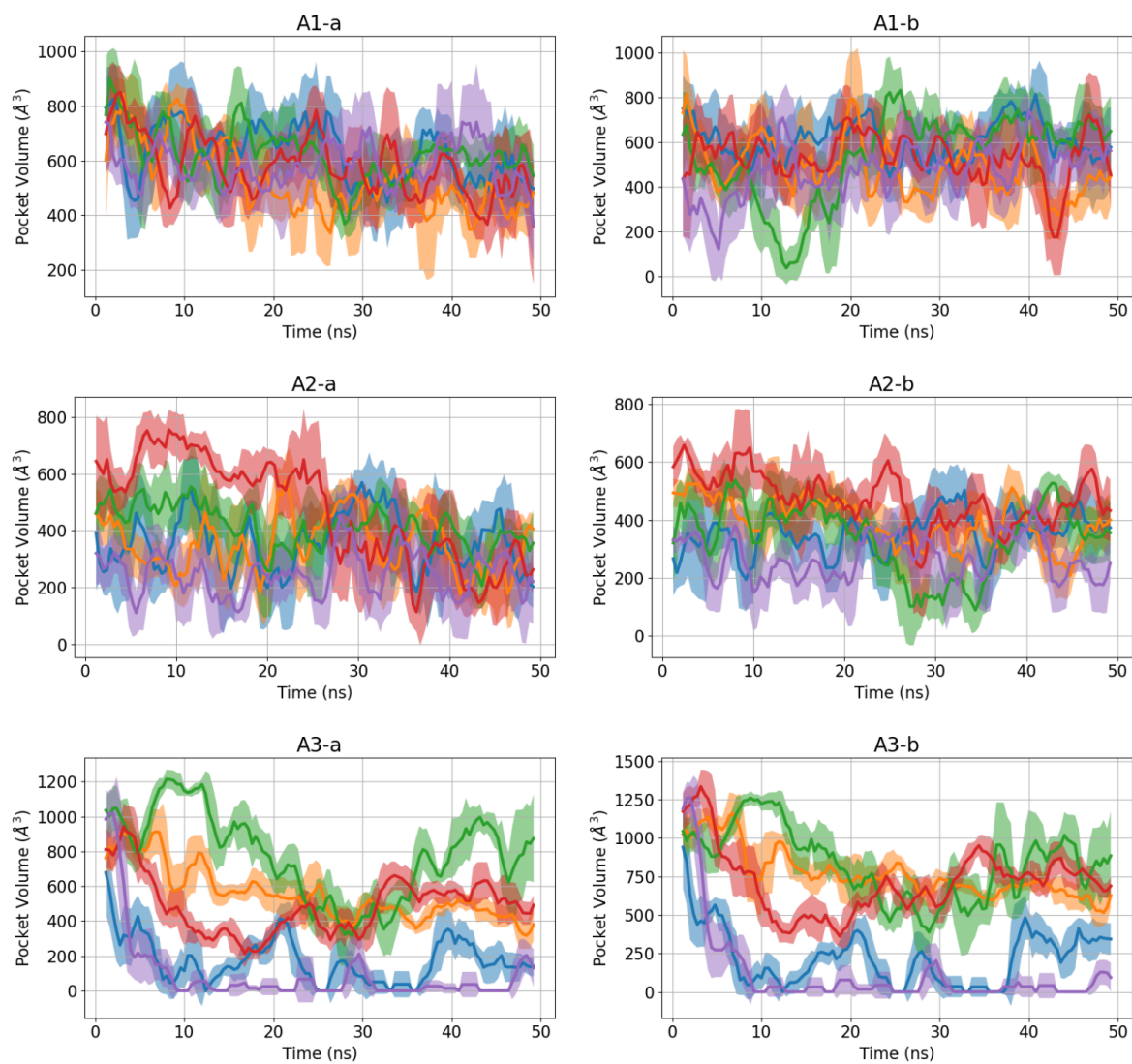


Figure E.11.: Time-evolution of the stabilizer-binding pocket volume in set A complexes without stabilizers. Data from five individual simulations are color-coded the same way as in Figure E.6. Data are shown as rolling averages over every 1 ns with standard deviation indicated by the shaded area.

E.1. SUPPLEMENTARY FIGURES

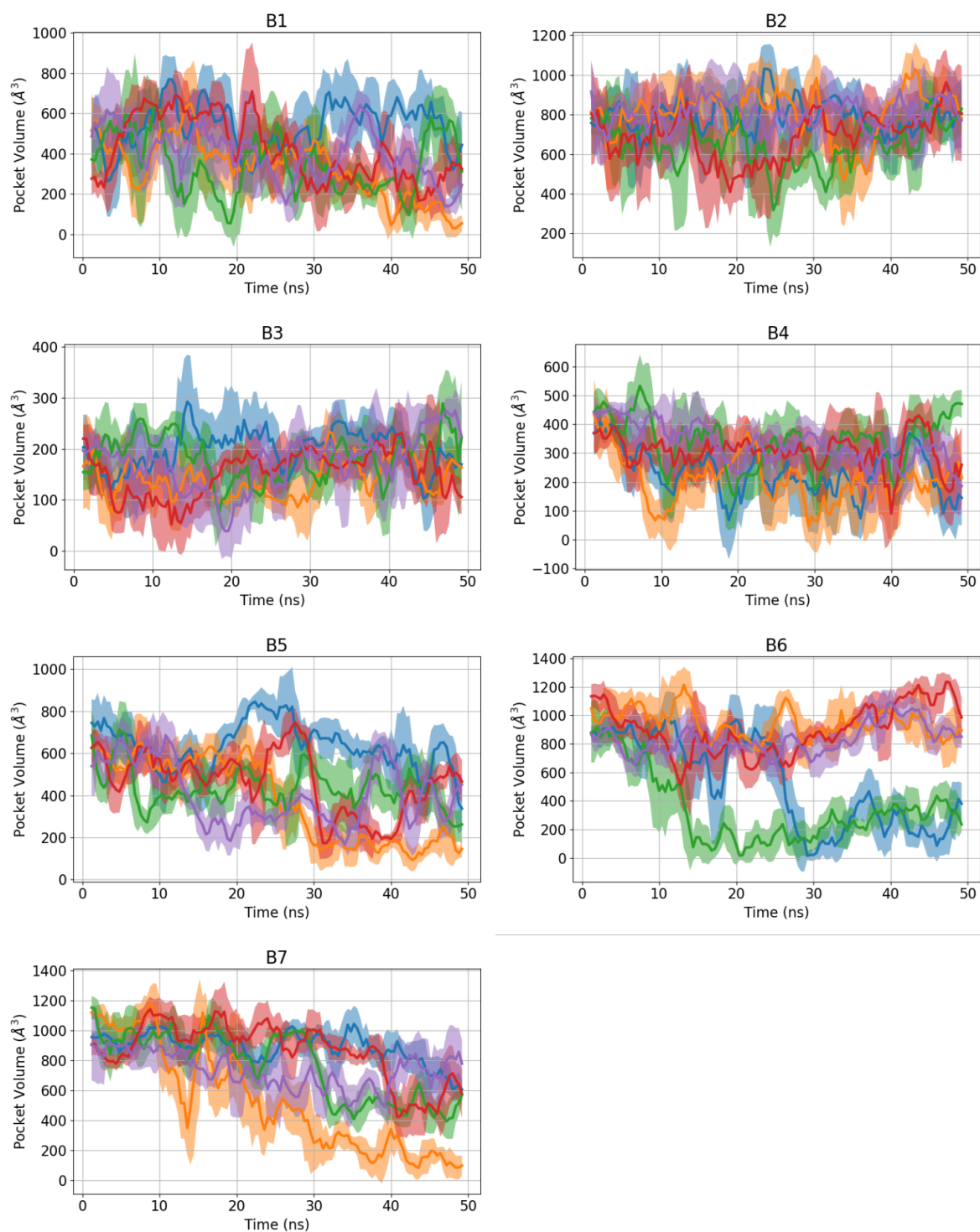


Figure E.12.: Time-evolution of the stabilizer-binding pocket volume in set B complexes without stabilizers. Data from five individual simulations are color-coded the same way as in Supplementary Figure E.7. Data are shown as rolling averages over every 1 ns with standard deviation indicated by the shaded area.

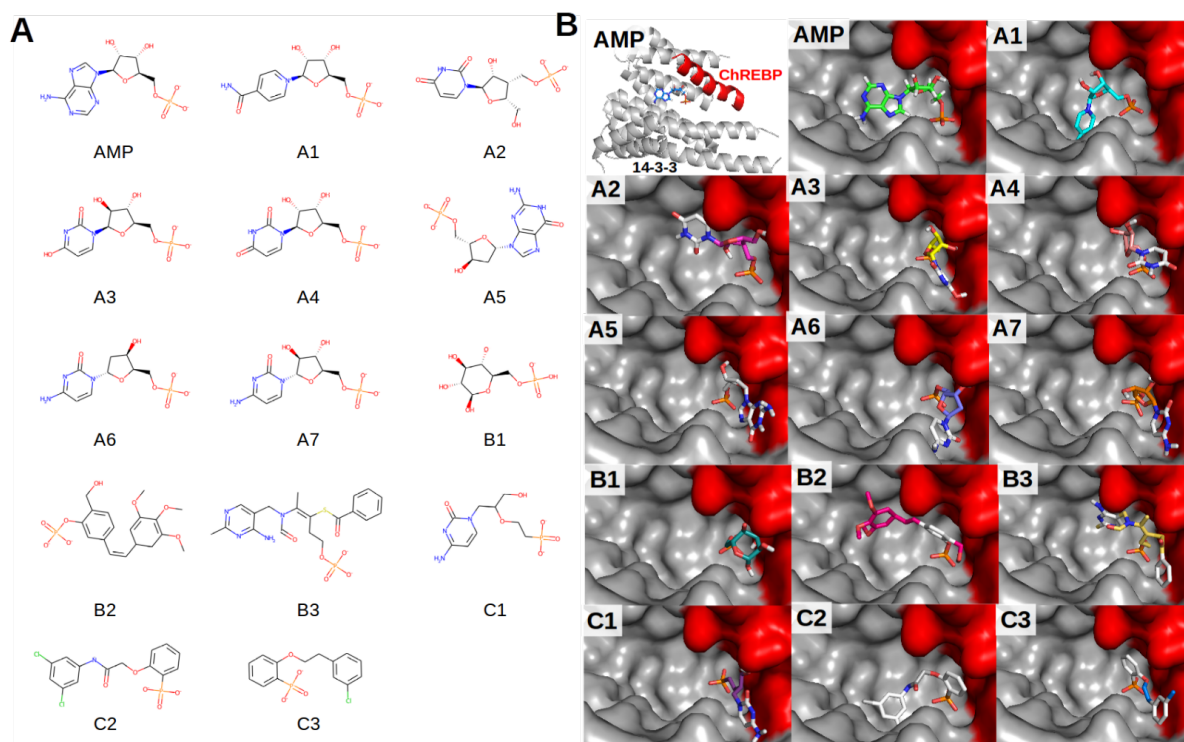


Figure E.13.: MD simulations with preselected compounds binding to 14-3-3/ChREBP complex. (A) Chemical structures of the 13 compounds investigated. (B) Docking pose of AMP and the 13 compounds.

Ligand	$\Delta\Delta G_{RS}$ (kcal/mol)	$\Delta\Delta G_{LS}$ (kcal/mol)	$ \Delta\Delta G_{RS} - \Delta\Delta G_{LS} $ (kcal/mol)	$\Delta\Delta G_{(RL)S}$ (kcal/mol)
1	<b>-14.37 ± 1.93</b>	-21.05 ± 2.67	6.68 ± 2.83	-44.04 ± 5.05
2	<b>-15.23 ± 2.31</b>	-20.31 ± 2.15	5.08 ± 3.06	-39.07 ± 3.37
3	-20.17 ± 3.41	<b>-18.7 ± 7.02</b>	6.11 ± 4.62	-45.67 ± 8.23
4	-20.59 ± 3.11	<b>-16.75 ± 1.88</b>	4.28 ± 3.1	-45.99 ± 3.92
5	<b>-19.63 ± 2.91</b>	-21.36 ± 3.92	4.66 ± 3.32	-46.68 ± 2.41
6	<b>-22.95 ± 3.68</b>	-21.33 ± 3.39	5.7 ± 2.98	-55.41 ± 3.49
7	<b>-20.45 ± 1.8</b>	-26.34 ± 5.44	5.9 ± 3.88	-57.03 ± 7.17
8	-22.84 ± 2.23	<b>-19.06 ± 2.57</b>	4.23 ± 3.43	-47.07 ± 2.46
9	-22.91 ± 2.73	<b>-20.97 ± 2.96</b>	4.05 ± 3.36	-53.1 ± 3.57
10	-26.74 ± 6.37	<b>-10.94 ± 2.53</b>	15.8 ± 8.64	-45.17 ± 5.68

Table E.2.: Binding free energy of the 10 ligands binding to the interface of PDBID 1Y64. The structures of the ligands are shown in Figure E.16. The binding affinity between the compound and the more weakly bound protein is emphasized with bold font. The 3 compounds with the lowest  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$  are expected to be the most promising stabilizers and are highlighted in gray.

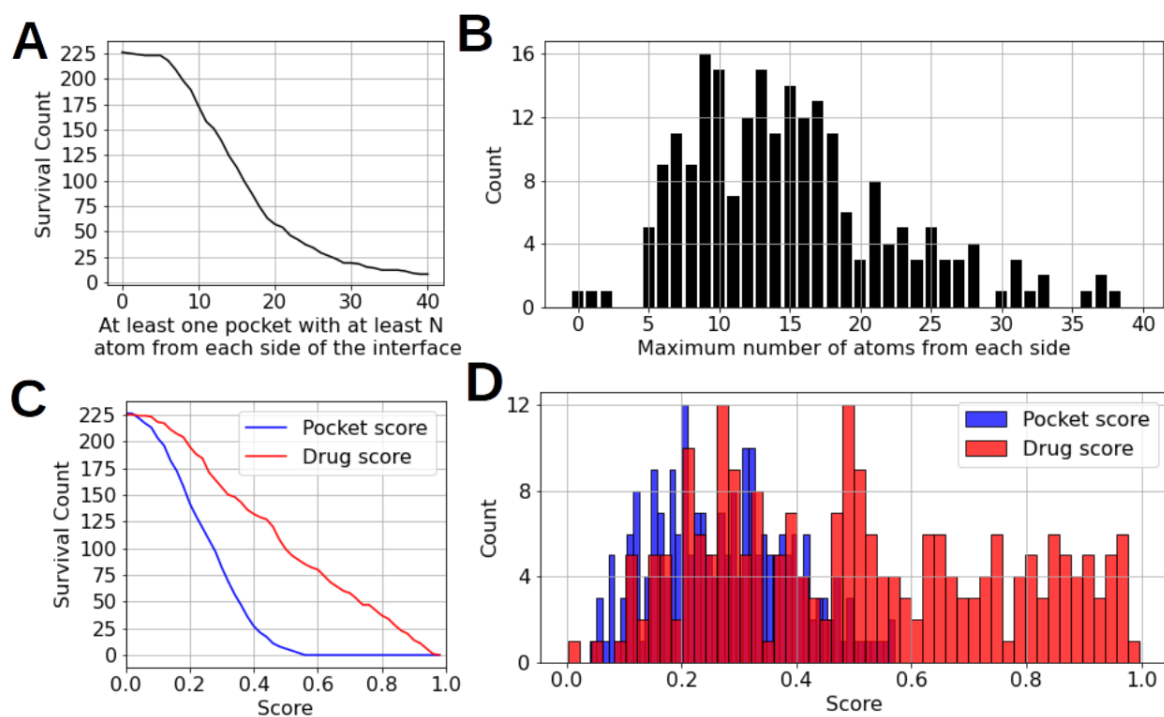


Figure E.14.: Pocket detection with Fpocket on the 226 PPI complexes from protein binding data set 5.0. (A) Survival curve of number of complexes with at least one pocket in contact with at least N atoms from each protein. (B) Histogram of the number of complexes with interface pocket maximally contacting N atom from each protein. (C) Survival curve of the number of complexes with maximum pocket score (blue) or drug score (red). (D) Histogram of the number of complexes with maximum pocket score (blue) or drug score (red). (C) and (D) consider only pockets contacting at least 5 atoms from each side of the protein-protein complex.

E.1. SUPPLEMENTARY FIGURES

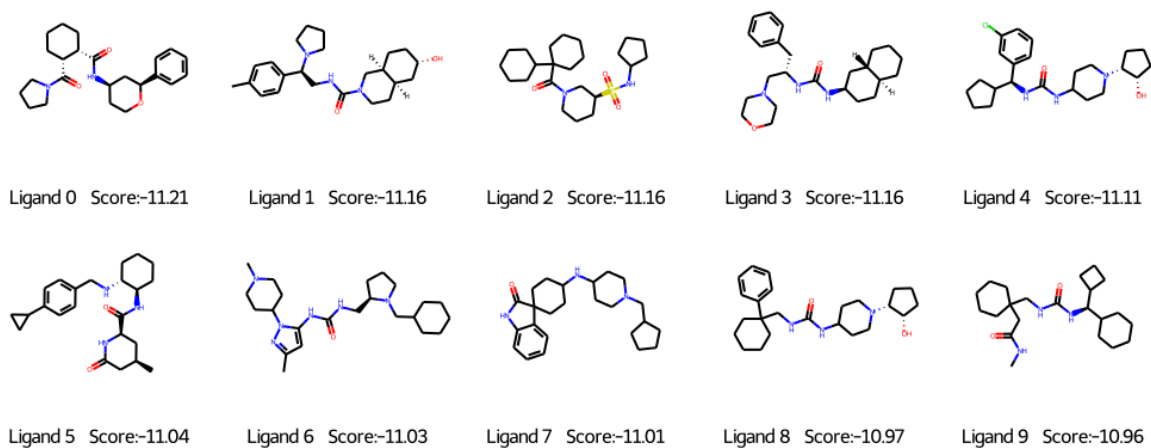


Figure E.15.: The best-docked 10 ligands to PDBID 1AK4 from virtual screening. The docking score is shown in a unit of kcal/mol.

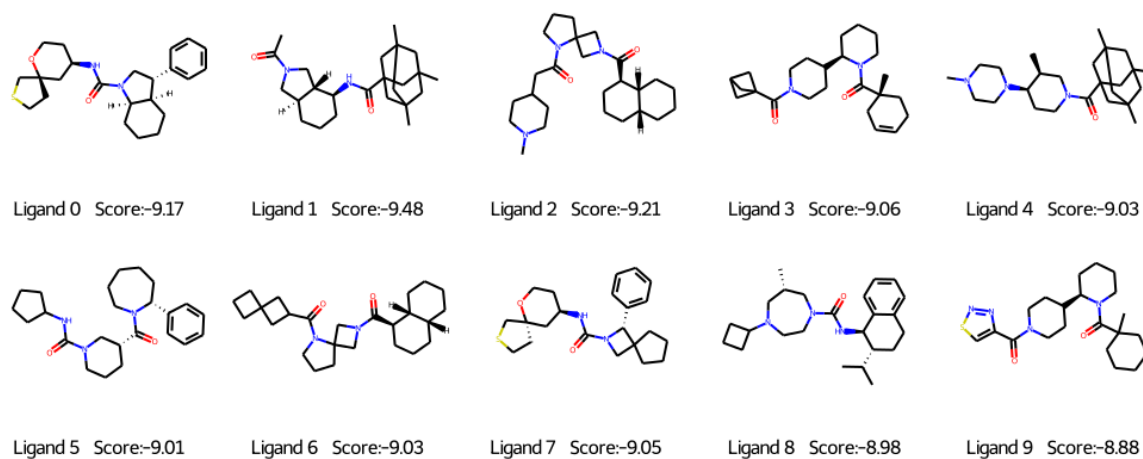


Figure E.16.: The best-docked 10 ligands to PDBID 1Y64 from virtual screening. The docking score is shown in a unit of kcal/mol.

E.1. SUPPLEMENTARY FIGURES

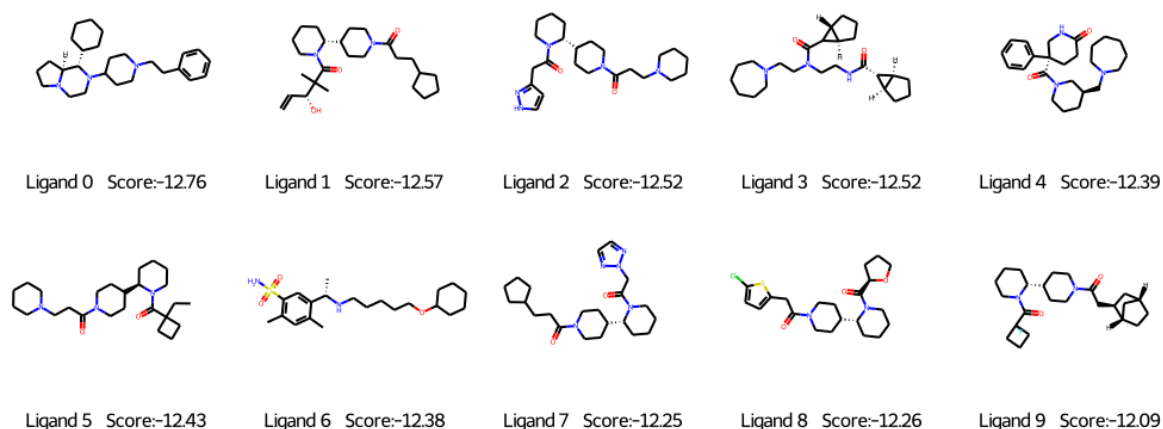


Figure E.17.: The best-docked 10 ligands to PDBID 2GAF from virtual screening. The docking score is shown in a unit of kcal/mol.

Ligand	$\Delta\Delta G_{RS}$ (kcal/mol)	$\Delta\Delta G_{LS}$ (kcal/mol)	$ \Delta\Delta G_{RS} - \Delta\Delta G_{LS} $ (kcal/mol)	$\Delta\Delta G_{(RL)S}$ (kcal/mol)
1	$-42.45 \pm 5.45$	<b><math>-13.7 \pm 6.18</math></b>	$28.75 \pm 7.68$	$-67.99 \pm 11.8$
2	$-26.7 \pm 4.5$	<b><math>-17.63 \pm 1.86</math></b>	$9.07 \pm 5.33$	$-51.99 \pm 2.79$
3	<b><math>-20.28 \pm 4.76</math></b>	$-27.79 \pm 4.31$	$8.69 \pm 6.96$	$-50.1 \pm 1.49$
4	$-30.76 \pm 3.7$	<b><math>-19.89 \pm 2.26</math></b>	$10.89 \pm 5.6$	$-56.84 \pm 3.32$
5	$-29.24 \pm 3.9$	<b><math>-24.56 \pm 3.42</math></b>	$7.54 \pm 2.65$	$-60.16 \pm 6.78$
6	$-31.96 \pm 3.02$	<b><math>-22.51 \pm 2.6</math></b>	$9.45 \pm 5.29$	$-60.84 \pm 4.49$
7	$-31.59 \pm 2.57$	<b><math>-22.07 \pm 3.36</math></b>	$9.52 \pm 5.4$	$-63.33 \pm 4.81$
8	$-30.9 \pm 5.24$	<b><math>-20.59 \pm 3.35</math></b>	$10.31 \pm 7.51$	$-56.93 \pm 4.51$
9	$-27.68 \pm 4.06$	<b><math>-21.54 \pm 1.82</math></b>	$6.79 \pm 4.44$	$-55.31 \pm 4.56$
10	$-27.11 \pm 4.78$	<b><math>-21.79 \pm 3.82</math></b>	$7.78 \pm 1.91$	$-57.1 \pm 6.32$

Table E.3.: Binding free energy of the 10 ligands binding to the interface of PDBID 2GAF. The structures of the ligands are shown in Figure E.17. The binding affinity between the compound and the more weakly bound protein is emphasized with bold font. The 3 compounds with the lowest  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$  are expected to be the most promising stabilizers and are highlighted in gray.

E.1. SUPPLEMENTARY FIGURES

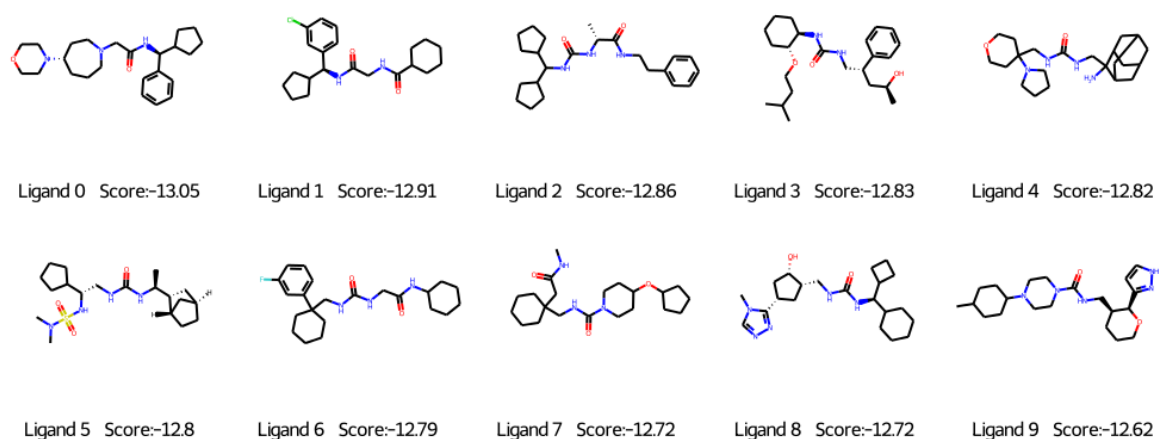


Figure E.18.: The best-docked 10 ligands to PDBID 3H11 from virtual screening. The docking score is shown in a unit of kcal/mol.

Ligand	$\Delta\Delta G_{RS}$ (kcal/mol)	$\Delta\Delta G_{LS}$ (kcal/mol)	$ \Delta\Delta G_{RS} - \Delta\Delta G_{LS} $ (kcal/mol)	$\Delta\Delta G_{(RL)S}$ (kcal/mol)
1	<b>-20.95 ± 4.48</b>	-26.44 ± 2.24	5.49 ± 6.51	-50.62 ± 2.8
2	<b>-22.39 ± 1.68</b>	-25.49 ± 4.41	5.09 ± 3.27	-54.93 ± 4.55
3	<b>-14.98 ± 2.07</b>	-27.6 ± 2.4	12.62 ± 3.3	-48.89 ± 2.68
4	<b>-24.25 ± 2.9</b>	-29.86 ± 2.29	5.93 ± 3.34	-66.44 ± 5.65
5	<b>-16.45 ± 2.66</b>	-16.94 ± 3.69	4.22 ± 4.31	-42.73 ± 2.03
6	<b>-40.25 ± 13.28</b>	<b>-27.84 ± 2.7</b>	14.49 ± 11.44	-92.08 ± 14.15
7	<b>-18.57 ± 4.29</b>	-27.75 ± 0.94	9.17 ± 4.51	-55.27 ± 5.81
8	-22.57 ± 3.7	<b>-21.23 ± 1.49</b>	2.49 ± 1.72	-43.76 ± 4.42
9	<b>-22.88 ± 6.16</b>	-23.25 ± 4.85	10.24 ± 2.72	-54.42 ± 3.44
10	-20.72 ± 0.82	<b>-20.32 ± 3.81</b>	3.44 ± 2.46	-45.76 ± 4.39

Table E.4.: Binding free energy of the 10 ligands binding to the interface of PDBID 3H11. The structures of the ligands are shown in Figure E.18. The binding affinity between the compound and the more weakly bound protein is emphasized with bold font. The 3 compounds with the lowest  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$  are expected to be the most promising stabilizers and are highlighted in gray.

E.1. SUPPLEMENTARY FIGURES

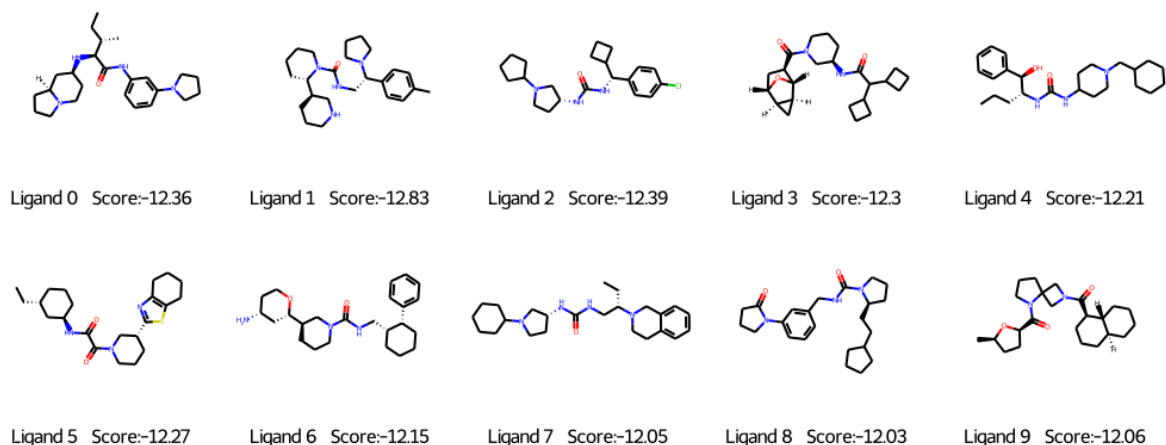


Figure E.19.: The best-docked 10 ligands to PDBID 4FZA from virtual screening. The docking score is shown in a unit of kcal/mol.

Ligand	$\Delta\Delta G_{RS}$ (kcal/mol)	$\Delta\Delta G_{LS}$ (kcal/mol)	$ \Delta\Delta G_{RS} - \Delta\Delta G_{LS} $ (kcal/mol)	$\Delta\Delta G_{(RL)S}$ (kcal/mol)
1	<b>-17.0 ± 3.98</b>	-28.95 ± 2.56	11.95 ± 5.2	-60.3 ± 6.6
2	<b>-23.23 ± 9.39</b>	-26.16 ± 2.28	8.18 ± 4.26	-53.9 ± 13.63
3	<b>-20.91 ± 1.48</b>	-23.36 ± 1.78	2.81 ± 1.65	-47.8 ± 2.73
4	<b>-16.59 ± 3.09</b>	-18.98 ± 3.79	2.41 ± 1.66	-40.9 ± 7.28
5	<b>-24.63 ± 1.53</b>	<b>-24.61 ± 0.96</b>	1.73 ± 1.1	-55.47 ± 1.2
6	-23.49 ± 4.4	<b>-22.02 ± 3.36</b>	5.49 ± 3.74	-55.59 ± 4.39
7	<b>-16.44 ± 3.16</b>	-30.32 ± 2.4	13.88 ± 3.24	-42.95 ± 7.69
8	<b>-25.84 ± 4.15</b>	-28.66 ± 2.8	3.63 ± 3.46	-60.94 ± 6.97
9	<b>-25.47 ± 2.02</b>	-26.72 ± 2.6	1.51 ± 1.31	-63.62 ± 6.45
10	-23.62 ± 1.98	<b>-21.83 ± 1.66</b>	2.41 ± 1.58	-57.46 ± 3.16

Table E.5.: Binding free energy of the 10 ligands binding to the interface of PDBID 4FZA. The structures of the ligands are shown in Figure E.19. The binding affinity between the compound and the more weakly bound protein is emphasized with bold font. The 3 compounds with the lowest  $\max\{\Delta\Delta G_{RS}, \Delta\Delta G_{LS}\}$  are expected to be the most promising stabilizers and are highlighted in gray.



## List of Publications

1. How Mutations Perturb  $\gamma$ -Secretase Active Site Studied by Free Energy Simulations. **SY Chen**, M Zacharias, ACS Chemical Neuroscience 11 (20), 3321-3332 (2020)
2. An Internal Docking Site Stabilizes Substrate Binding to  $\gamma$ -Secretase: Analysis by Molecular Dynamics Simulations. **SY Chen**, M Zacharias, Biophysical Journal 121 (12), 2330-2344. (2022)
3. Active Site Geometry Stabilization of a Presenilin Homolog by the Lipid Bilayer Promotes Intramembrane Proteolysis. LP Feilen, **SY Chen**, A Fukumori, R Federle, M Zacharias, H Steiner, Elife 11, e76090 (2022)
4. Functional Role of the Enzyme-Substrate Hybrid  $\beta$ -Strand in  $\gamma$ -Secretase Proteolysis. **SY Chen**, LP Feilen, H Steiner, M Zacharias, Communications Biology 6 (1), 670 (2023)
5. What Makes a Good Protein-Protein Interaction Stabilizer: Analysis and Application of the Dual-Binding Mechanism. **SY Chen**, M Zacharias, ACS Central Science 9 (5), 969-979 (2023)
6. A Chemical Proteomic Strategy Reveals Inhibitors of Lipoate Salvage in Bacteria and Parasites. JN Dienemann, **SY Chen**, M Hitzenberger, M Sievert, M Groll, A König, S Prigge, S Hacker, M Zacharias, S Sieber, Angewandte Chemie International Edition, e202304533. (2023)
7. Molecular Dynamics Simulations for Predicting the Stability of Ligand-Receptor Interactions. **SY Chen**, M Zacharias (submitted)
8. APP substrate ectodomain defines A $\beta$  length by restraining  $\gamma$ -secretase processivity and facilitating product release. Matthias Koch, Thomas Enzlein, **SY Chen**, Dieter Petit, Sam Lismont, Martin Zacharias, Carsten Hopf, L. Chàvez-Gutiérrez. (submitted)
9. Modulation mechanism of  $\gamma$ -secretase modulator E2012. **SY Chen**, M. Koch, L. Chàvez-Gutiérrez, M. Zacharias. (in preparation)