**Technische Universität München**

TUM School of Life Sciences

# The Dynamic Personality of the Amyloid Precursor Protein's Transmembrane Domain

Alexander Götz, M.Sc.

Vollständiger Abdruck der von der TUM School of Life Sciences der Technischen Universität München zur Erlangung eines

**Doktors der Naturwissenschaften (Dr. rer. nat.)**

genehmigten Dissertation.

| | |
|---|---|
| Vorsitz: | Prof. Dr. Dmitrij Frischmann |
| Prüfer der Dissertation: | 1. Prof. Dr. Dieter Langosch |
| | 2. Prof. Dr. Martin Zacharias |

Die Dissertation wurde am 30.03.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Life Sciences am 08.08.2022 angenommen.

This work is dedicated to my beloved ones who had to do without their daddy and husband many times during the last years. I love you!

## Acknowledgments

The present work would never been possible without constant support by several persons which should be mentioned in the following.

First, I want to thank my doctoral advisor Prof. Dr. Dieter Langosch for his supervision during this work and the many stimulating discussions on the dynamic personalities of various TMDs. I also want to thank Prof. Dr. Martin Zacharias for examining this thesis as well as the very helpful discussions in advance and Prof. Dr. Dmitrij Frishman for being the chair of the examining committee.

Next, I want to thank Dr. Christina Scharnagl for being the best supervisor and mentor I could ever have. Without her patience, help and expertise this thesis and related publications would never have been possible. It has been a great time working together with you and I will never forget the many discussions about protein dynamics, membrane proteins and the larger field of statistical mechanics and biophysics during our long coffee sessions. In this sense, I also want to thank my former colleagues at the Chair of Biopolymer Chemistry for the stimulating discussions, help and various other activities outside of science. In particular, I want to thank Philipp Högel for sharing his mass spectrometry data with me and our work on several publications. Likewise, I want to thank my FOR2290 collaborators for sharing their data and insights with me as well as the encouraging discussions during our regular meetings. Especially I must thank Mara Silber, Claudia Muhle-Goll, Fritz Kamp and Alexander Vogel. Finally, I want to thank all my students that did their own thesis together with me.

I am deeply grateful for all the help my beloved wife Barbara Götz provided me with sorting my partially confusing thoughts and helping me in my constant battles with Microsoft Word. I know that the present work took a lot of our time together and it was sometimes a challenge with me, however, you were always encouraging me to finish this chapter of my life, keeping my back free from any disturbances. I know that often I could not be the husband you deserved and will try to compensate for all the deprivations of the last years. The same accounts for my little son, who missed a lot of hours with his dad over the last two years, hours that sometimes cannot be compensated anymore. In this sense I also want to thank my parents who have been supporting me with everything I needed from the first day of my life till the present day and always encouraged me to successfully finish this chapter of my life, thank you! Also I need to thank my family-in-law for taking care of me not to get completely engrossed by science and work and always remember me that there is more in life.

Finally, I want to thank all those people who supported me in the last year and always encouraged me to finish my thesis, including my former colleagues at the Leibniz Supercomputing Centre as well as my current colleagues at Machine Learning Reply. Special thanks goes to the Leibniz

Supercomputing Centre for providing me with computing time and support at any time during this thesis.

# Table of Content

## Abstract

The dynamic properties of substrates are a matter of increasing importance for substrate selection and processing by enzymes in general, and intramembrane cleaving proteases in particular. Intramembrane cleaving proteases hydrolyze their substrates in the hydrophobic environment of the cellular membrane. The amyloid precursor protein is one of the substrates whose dynamic properties apparently impact its catalytic cleavage by the intramembrane cleaving proteases γ-secretase at the stage of initial endoproteolysis and/or subsequent carboxypeptidase-like trimming steps, generating Aß-peptides of variable length. Due to the location of the cleavage sites in the substrate's transmembrane domain, a series of studies investigated the transmembrane domain's conformation as well as some of its major dynamic characteristics that seem to shape its dynamic personality. However, the mechanism by which familial Alzheimer's disease mutations of the amyloid precursor protein's transmembrane domain affect its ε-endoproteolysis is only poorly understood. While all of them alter the disease-related Aß42/Aß40 ratio towards higher Aß42 levels, their majority shifts also the preference of initial ε-cleavage from position ε49 to ε48. Others like the I45T mutation, do not alter ε-cleavage site preference, but reduce ε-efficiency drastically. Previous models tried to explain such observations by alterations in the extent and direction of the transmembrane domain's characteristic bending motion, steered by its central di-glycine hinge. However, structures of NOTCH1, another substrate of γ-secretase, detected neither a bend nor an exceptionally flexible transmembrane domain. This raised several questions about (1) the impact of the amyloid precursor protein's di-glycine motif on its cleavage by γ-secretase, (2) the general requirement of a bending motif in other γ-secretase substrates, and (3) other dynamic characteristics that impact cleavage specificity. As the transmembrane domain of the amyloid precursor protein is well characterized by functional assays and structural analysis, the present work focuses on this substrate and tries to determine (1) its entire dynamic personality beyond the dominant bending motion, (2) the impact of familial Alzheimer's disease mutations on those, and (3) the relation to γ-secretase cleavage. To get detailed insights into global as well as local dynamics, atomistic molecular dynamic simulations were used, many of them carefully validated against experimental results. The amyloid precursor proteins wild-type transmembrane domain was compared to seven familial Alzheimer's disease mutants and two artificial mutants of the di-glycine hinge. Consistently, familial Alzheimer's disease mutations in the cleavage-domain enhance hydrogen-bond fluctuations upstream of the ε-sites but maintain strong helicity at the cleavage sites. Functional mode analysis confirmed that this hydrogen-bond shifting is correlated with conformational fluctuations like local bending upstream of the ε-sites but downstream to the central di-glycine motif. These motions are likely to be correlated with substrate binding and positioning of the cleavage domain in the active site of γ-secretase, as supported by perturbation-response scanning. Thus, investigated familial Alzheimer's disease mutations impact ε-endoproteolysis by a modulation of the cleavage site's presentation, thus shifting ε-cleavage site selection and finally leading to early onset, fast progressing forms of Alzheimer's disease.

## Zusammenfassung

Die dynamischen Eigenschaften von Substraten sind von zunehmender Bedeutung für die Auswahl und Prozessierung selbiger durch Enzyme im Allgemeinen und Intramembranproteasen im Besonderen. Intramembranproteasen hydrolysieren ihre Substrate in der hydrophoben Umgebung der Zellmembran. Das Amyloid-Vorläuferprotein ist eines der Substrate, dessen dynamische Eigenschaften offenbar seine katalytische Spaltung durch die Intramembranproteasen γ-Sekretase im Stadium der initialen Endoproteolyse und/oder nachfolgenden schrittweisen Spaltung beeinflussen. Letztere ist ähnlich der einer Caboxypeptidase ist und führt zur Entstehung von Aß-Peptiden variabler Länge. Aufgrund der Lage der Spaltstellen in der Transmembrandomäne des Substrats untersuchte eine Reihe von Studien die Struktur der Transmembrandomäne sowie einige ihrer wichtigsten dynamischen Eigenschaften, die ihre dynamische Persönlichkeit zu prägen scheinen. Der Mechanismus, durch den familiäre Alzheimer-Mutationen der Transmembrandomäne des Amyloid-Vorläuferproteins dessen ε-Endoproteolyse beeinflussen, ist jedoch nur unzureichend verstanden. Während alle das für die Alzheimer Krankheit relevante Aß42/Aß40-Verhältnis in Richtung höherer Aß42-Konzentrationen verändern, verschiebt ihre Mehrheit auch die Präferenz der initialen ε-Spaltung von Position ε49 auf ε48. Andere Mutationen, wie beispielsweise die I45T-Mutation, verändern die Präferenz der ε-Spaltstelle nicht, reduzieren aber die ε-Effizienz drastisch. Frühere Modelle versuchten, diese Beobachtungen durch Veränderungen des Ausmaßes und der Richtung der charakteristischen Biegebewegung der Transmembrandomäne zu erklären, welche durch ein zentrales di-Glyzin-Scharnier gesteuert wird. Strukturen von NOTCH1, einem weiteren Substrat der γ-Sekretase, zeigten jedoch weder eine gebogene noch eine außergewöhnlich flexible Transmembrandomäne. Dies warf mehrere Fragen auf über (1) den Einfluss des di-Glyzin-Motivs des Amyloid-Vorläuferproteins auf seine Spaltung durch die γ-Sekretase, (2) die generelle Notwendigkeit einer tendenziell eher gebogenen Struktur in anderen γ-Sekretase-Substraten und (3) andere dynamische Merkmale, die die Spaltungsspezifität beeinflussen. Da die Transmembrandomäne des Amyloid-Vorläuferproteins durch funktionelle Assays und Strukturanalysen gut charakterisiert ist, konzentriert sich die vorliegende Arbeit auf dieses Substrat und versucht, (1) seine gesamte dynamische Persönlichkeit über die dominante Biegebewegung hinaus, (2) die Auswirkungen von familiären Alzheimer-Mutationen auf diese und (3) die Beziehung zur γ-Sekretase-Spaltung zu bestimmen. Um detaillierte Einblicke sowohl in die globale als auch in die lokale Dynamik zu erhalten, wurden atomistische Molekulardynamiksimulationen durchgeführt, von denen einige sorgfältig gegen experimentelle Ergebnisse validiert wurden. Die Transmembrandomäne des Amyloid-Vorläuferproteine Wildtyps wurde mit sieben familiären Alzheimer-Mutanten und zwei künstlichen Mutanten des Di-Glycin-Scharniers verglichen. Übereinstimmend verstärken die familiären Alzheimer-Mutationen in der Spaltdomäne die Fluktuationen von Wasserstoffbrückenbindungen oberhalb der ε-Spaltstellen, erhalten aber die ausgeprägte Helikalität an den Spaltstellen. Die Analyse der funktionalen Bewegungen bestätigte,

dass diese Verschiebung der Wasserstoffbrücken mit strukturellen Veränderungen wie lokalen Biegungen oberhalb der ε-Sites, aber unterhalb des zentralen di-Glyzin-Motiv korreliert ist. Die Analyse der dynamischen Antwort auf Störungen durch Wechselwirkungen beim Binden legt nahe, dass diese Bewegungen wahrscheinlich mit der Substratbindung und der Positionierung der Spaltdomäne im aktiven Zentrum der γ-Sekretase verknüpft sind. Somit beeinflussen die untersuchten familiären Alzheimer-Mutationen die ε-Endoproteolyse durch eine veränderte Präsentation der Spaltstelle, wodurch sich die Auswahl der ε-Spaltstelle verschiebt, was letztendlich zu früh einsetzenden, schnell fortschreitenden Formen der Alzheimer-Krankheit führt.

## Abbreviations

| | |
|---|---|
| **Aβ** | Amyloid β Peptide |
| **Aβ-ratio** | Ratio between the concentration of Aβ42 and Aβ40 |
| **AICD** | APP Intracellular Domain |
| **AD** | Alzheimer's Disease |
| **ADAM** | A Disintegrin and Metalloprotease |
| **APH1** | Anterior Pharynx Defective Protein 1 |
| **APLP1** | Amyloid Precursor Like Protein 1 |
| **APP** | Amyloid Precursor Protein |
| **BACE** | β-Site APP Cleaving Enzyme |
| **BCMA** | B-Cell Maturation Antigen |
| **CG** | Coarse-Grained or Coarse Grain |
| **cryoEM** | Cryogenic Electron Microscopy |
| **CTF** | C-Terminal Fragment of Presenilin, Covering Transmembrane Domains 7-9 |
| **C83** | 83 Amino Acid Long C-Terminal Fragment of APP |
| **C99** | 99 Amino Acid Long C-Terminal Fragment of APP After Cleavage by BACE |
| **DAFT** | Docking Assay for Transmembrane Components |
| **DHX** | Deuterium Hydrogen Exchange |
| **DPC** | Dodecylphosphocholine |
| **DSSP** | Define Secondary Structure of Proteins Algorithm |
| **ECD** | Extra Cellular Domain |
| **ESI** | Electrospray Ionization |
| **ETD** | Electron Transfer Dissociation |
| **FAD** | Familial Alzheimer's Disease |
| **FMA** | Functional mode analysis |
| **$f_{open}$** | Fraction of Times a Backbone Hydrogen Bonds Was Found to be Open |
| **GCS** | Gauss Centre for Supercomputing |
| **G38L** | Glycine to Leucine Mutation of the APP TMD at $G_{38}$ (C99 Numbering) |
| **G38P** | Glycine to Proline Mutation of the APP TMD at $G_{38}$ (C99 Numbering) |
| **H-Bond** | Hydrogen Bond |
| **HDX** | Hydrogen Deuterium Exchange |
| **HPC** | High Performance Computing |
| **I-CLiP** | Intramembrane Cleaving Protease |
| **ITGB1** | Integrin β1 |
| **I45T** | Isoleucine to Threonine Mutation of the APP TMD at $I_{45}$ (C99 Numbering) |
| **JMD** | Juxta-Membrane Domain |
| **$k_{ch}$** | Chemical Rate for Deuterium to Hydrogen Exchange |

| | |
|---|---|
| **$k_{DHX}$** | Deuterium Hydrogen Exchange Rate Constant |
| **$k_{HDX}$** | Hydrogen Deuterium Exchange Rate Constant |
| **LMPG** | Lyso-myristoyl Phosphatidylglycerol |
| **LRZ** | Leibniz Supercomputing Centre |
| **MD** | Molecular Dynamics |
| **MS** | Mass Spectrometry |
| **MSF** | Mean-Squared Fluctuations |
| **MSM** | Markov State Model |
| **NCT** | Nicastrin |
| **NMR** | Nuclear Magnetic Resonance |
| **NOTCH1** | Neurogenic Locus Notch Homolog Protein 1 |
| **NPRA** | Atrial Natriuretic Peptide Receptor 1 |
| **NTF** | N-Terminal Fragment of Presenilin, Covering Transmembrane Domains 1-6 |
| **PEN2** | Presenilin Enhancer Protein 2 |
| **POPC** | 1-Palmitoyl-2-oleoylphosphatidylcholine |
| **PRS** | Pertubation Response Scanning |
| **PSN** | Presenilin |
| **RIP** | Regulated Intramembrane Proteolysis |
| **RMSD** | Root-Mean Squared Deviation |
| **SPP** | Signal Peptide Peptidase |
| **SPPL** | Signal Peptide Peptidase Like |
| **SREBP-1** | Sterol Regulatory Element-Binding Protein 1 |
| **TFE** | 2,2,2-trifluoroethanol (TFE) |
| **TMD** | Transmembrane Domain |
| **TM-N** | N-terminal Domain of the Amyloid Precursor Protein's Transmembrane Domain |
| **TM-C** | C-terminal Domain of the Amyloid Precursor Protein's Transmembrane Domain |
| **TOF** | Time of Flight (TOF) |
| **vdW** | Van der Waals |
| **WT** | Wild-type APP TMD |

Amino acids are referred to by their 1-letter code.

## List of Figures

# 1 Introduction

Despite their discovery over 20 years ago, intramembrane cleaving proteases (I-CLiPs) and the process of regulated intramembrane proteolysis (RIP) are still only poorly understood today. I-CLiPs occur in all kingdoms of life with different structural and mechanistic features.[1] However, their role in cellular processes is still unknown and subject to an intense debate. γ-Secretase is by far the most prominent I-CLiP due to its involvement in the generation of Amyloid β (Aβ) peptides, which seem to be involved in Alzheimer's disease (AD).[2–4] Aβ-peptides of varying lengths result from the consecutive cleavage of the transmembrane domain (TMD) of the Amyloid Precursor Protein (APP) by γ-secretase in the cell membrane.[2,5] Besides APP, γ-secretase cleaves ~150 other substrates which all share a common topology as type I integral transmembrane proteins.[6–8] Although this substrate repertoire covers several proteins essential for life, such as neurogenic locus notch homolog protein 1 (NOTCH1), there is no common motif that allows to distinguish known substrates from the remaining ~ 1500 type I integral transmembrane proteins found in the human proteome,[9] which may act as further substrates.

## 1.1 Intramembrane Cleaving Proteases

The first report about proteolysis in a cellular membrane dates back to 1990[10] and for a long time I-CLiPs have been considered as rare exceptions.[11,12] Since that time, a series of I-CLiPs was identified,[13] which can be categorized in four major classes according to their type of active site residues, which catalyse hydrolysis:[14] (1) metalloproteases (e.g. site-2 protease (S2P) harbouring zinc ions),[12,15] (2) aspartate proteases (e.g. presenilin, the active subunit of γ-secretase, signal peptide peptidase (SPP), and signal peptide peptidase-like (SPPL) which have a consensus GxGD motif in their active site),[16–20] (3) rhomboids (serine proteases with a serine-histidine catalytic dyad),[10,13] and (4) glutamate proteases (e.g. Rce1).[21] As summarized by Urban (2016), different I-CliPs can be found in the cellular membranes of different cellular compartments, what might provide a hint for their specific functionality.[22]

Like their soluble counterparts, I-CLiPs catalyse the hydrolysis of peptide bonds. Although proteolytic cleavage requires the availability of water in close distance to the active site, the process of regulated intramembrane proteolysis, as catalysed by I-CLiPs, takes place in the hydrophobic environment of cell membranes.[22] This hydrophobic environment raised the question for the availability and transport of water molecules, necessary for hydrolytic cleavage. While location of the active sites close to the membrane boundary may provide an explanation, the determination of structural features of several I-CLiPs revealed the active sites to be located deep in the membrane plane.[19,20,23,24] Cysteine labeling studies of γ-secretase indicated that water molecules can be located in its interior,[25,26] while other studies found ion channel-like activity for two other I-CLiPs[27] and recently for the APH-1B subunit of γ-secretase, too.[28,29] While ion channel-like functionality may

just be a potential side effect of required water permeability, the main functionality of I-CLiPs is the proteolytic cleavage of other transmembrane proteins. As such, they are involved in many biochemical processes and therefore of fundamental physiological relevance in health and disease and their malfunction leads to a series of major civilization disease, ranging from AD and Parkinson's disease to immune deficiencies and type-2 diabetes.[14] As they control protein degradation in the cell membrane, they take on a central role in the regulation of cellular processes through activation of membrane-tethered transcription factors and transcriptional activators or maturation of bacterial translocation channels. Therefore, I-CLiPs have become prominent targets in drug research as summarized by Verhelst (2017).[14]

Although the structure is solved only for a few I-CLiPs,[20,21,23,30] they all consist of multiple helical TMDs, spanning the cellular membrane and connected by flexible outer membrane loops. The highly dynamic loop regions make the I-CLiPs conformationally flexible, as summarized for γ-secretase and its presenilin subunit by Aguayo-Ortiz et al. (2018)[31] and Zoltowska et al. (2018).[32] However, most of these studies have been carried out in rather homogeneous model membranes. In contrast to such model membranes, cellular membranes consist of a vast number of different components and their composition significantly differs between different cell compartments.[33] The severe effects of lipid composition on activity and functionality of various I-CLiPs was recently reviewed by Paschkowsky et al. (2018).[33]

A rather unique feature of I-CLiPs is their dependence on other membrane bound proteases, which initially cleave off the ectodomain of most I-CLiP substrates. Such initial cleavage is termed as shedding and is done by a series of so called sheddases like proteases of the A disintegrin and metalloprotease (ADAM family) and β-site APP cleaving enzyme (BACE) families. The details of shedding and the outcome of malfunction of these enzymes and their relevance for RIP have recently been reviewed by Lichtenthaler, Lemberg and Fluhrer (2018).[34] However, shedding does not seem to be a requirement for I-CLiP substrates in general. Transmembrane proteins with a rather short ectodomains can be cleaved without previous shedding as shown for B-cell maturation antigen (BCMA) and γ-secretase by Laurent et al. (2015).[35] The same study revealed that extensions of BCMA's ectodomain lead to inhibition of γ-secretase cleavage. While this would indicate a steric blockage, preventing the longer substrates from reaching the active site of the enzyme, this is questioned by a very recent study, which indicated that the amyloid precursor like protein 1 (APLP1) is directly cleaved by γ-secretase, despite its long ECD.[36] However, this appeared only to happen to a very minor extent compared to its usual cleavage.[36]

## 1.2 The Aspartyl Intramembrane Cleaving Protease γ-Secretase

γ-Secretase is by far the most prominent I-CLiP, which is related to its involvement in AD and its large repertoire of known substrates,[6,8,37] suggesting a functionality as the proteasome of the

membrane.[38] However, until today proteasome (like) functionality could not be confirmed due to a significant lack of validated substrates and, in particular, non-substrates (see **section 1.3**). γ-Secretase's structure consists of 20 TMDs, arranged in a 230 kDa complex of four individual subunits in a 1:1:1:1 stoichiometry. The TMDs form a horseshoe like conformation covered by a bulky, lid-like extracellular structure.[20] The four subunits are presenilin (PSN) with isoforms 1 and 2 (PSN1 and PSN2), the presenilin enhancer protein 2 (PEN-2), the anterior pharynx defective protein 1 (APH-1a or b) and nicastrin (NCT).[39] PSN1 is the major isoform of γ-secretase.

PSN consists of 9 TMDs, connected by multiple flexible loops. However, native PSN is inactive and requires autoproteolytic cleavage by PEN-2 between its TMD6 and TMD7.[40,41] This autoproteolysis happens after the assembly of all of its four components in the endoplasmic reticulum[41,42] and leads to a N-terminal fragment (NTF, PSN TMDs 1-6) and a C-terminal fragment (CTF, with PSN TMDs 7-9), which form the active PSN. Both catalytic aspartates are located close to the NTF-CTF interface with $Asp_{257}$ at TMD6 and $Asp_{385}$ at TMD7.[41] Both residues are buried in the interior of γ-secretase and, therefore, excluded from the membrane-exposed surface of the enzyme. The CTF-bound active site aspartate on TMD7 is part of a GxGD motif which is highly conserved between currently known intramembrane aspartyl proteases. Mutations of the motif's glycine residues to any other residue than alanine showed to significantly disrupt the function of γ-secretase.[43] Mutations of the native leucine at the x-position to hydrophobic as well as hydrophilic residues are generally well tolerated in terms of APP processing.[44] Besides mutations of the GxGD motif, many other PSN mutations are known and most of them are related to familial forms of AD (FAD).[11,45] However, the location and type of mutations significantly vary and the mechanism which finally leads to FAD is mostly unknown.[46] While some of them significantly disrupt the processivity of γ-secretase, others do not seem to trigger a deficiency in the carboxypeptidase function of γ-secretase.[46]

The subunits APH-1 and PEN-2 have important functions for assembly[39,47] and stability[39] of the γ-secretase complex. Like for PSN, multiple isoforms of APH-1 exist, with APH-1a being the major one in the γ-secretase complex.[48] A lack of APH-1 leads to a dramatic reduction of the NCT, PSN1 and PEN-2 levels and therefore, a significant drop of γ-secretase complexes.[48] In addition, APH-1 seems to be involved in substrate recruitment.[49] Like APH-1, the PEN2 subunit is involved in stabilization of the complex[40,47] and plays a decisive role for APP substrate recruitment.[50] In addition, it is involved in the endoproteolytic activation of the PSN subunit.[47,51]

NCT is a type I glycoprotein and contains a type 1 transmembrane domain and a large extracellular domain. NCT's TMD interacts with PSN1 and APH-1 linking NCT to the γ-secretase complex, which gains additional stability from NCT.[52] The glycosylated extracellular domain (ECD) of NCT seems to be involved in substrate recognition and, possibly, substrate placement in γ-secretase's active site.[20,50,53,54] NCT was suggested to act as a steric gatekeeper for γ-secretase, blocking access

for substrates with too large extracellular domains.[55,56] The detailed mechanism remains elusive, since other reports show that absence of NCT does not affect γ-secretase activity.[52,57,58]

The composition of the surrounding lipid bilayer has a major impact on γ-secretase activity and cleavage.[59] Several studies reported γ-secretase to be located in lipid rafts.[33,39] As lipid rafts are special areas in the membrane, which are characterized by an increased thickness of the membrane bilayer, this agrees well with studies performed by Winkler et al. (2012) which report the activity of γ-secretase to decrease for shorter lipid poly-acyl chains, while longer ones seem to increase γ-secretase activity.[60]

## 1.3 Known Substrates and Non-Substrates of γ-Secretase

As previously mentioned, γ-secretase cleaves a wide spectrum of ~150 substrates,[6–8] while only a very limited number of non-substrates has been described in the literature.[37] A common feature of all γ-secretase substrates is their type-I topology, which is also shared by known non-substrates.[6–8] The large variety of substrates as well as their related functionalities led to the idea of γ-secretase being "the proteasome of the membrane",[38] implying a general role in protein turnover (see **section 1.2**). However, 150 substrates are just a minor subset of the ~1500 type I integral transmembrane proteins found in the human proteome,[9] which all may act as further γ-secretase substrates. This raised the question for a shared motif being present in substrates but not in non-substrates, which would allow to classify the remaining ~1500 TMDs into groups of substrates and non-substrates. Due to the diverse functional implications of TMDs[14] such classification is of urgent importance in the development of γ-secretase modulators, targeting AD. While bioinformatics approaches did not evince a consensus motif in their TMD, a cluster of positively charged residues at the C-terminal end of the TMD seemed to be a shared feature of γ-secretase substrates.[61] However, charged residues at both termini are a common feature of TMDs, as they stabilize the TMD in the membrane bilayer. In addition, the occurrence of positively charged residues at the intracellular membrane border is not unexpected as type-I TMDs follow the well-known positive inside rule, which makes positively charged residues at this position a mandatory requirement.[62]

Only the atrial natriuretic peptide receptor 1 (NPRA) and integrin beta 1 (ITGB1) could be annotated as non-substrates. However, NPRA can be converted in a substrate by shortening its intracellular domain,[37] and ITGB1 fails to be shed by BACE.[37] However, as mentioned in **section 1.1**, shedding may not be required in all cases. As a final consequence, even more than a decade after the discovery of γ-secretase and many of its substrates, it is still rather unclear what qualifies a substrate as such. A recent development in the discussion of characteristic features shared by γ-secretase substrates is the concept of conformational dynamics of the TMDs as a discriminating factor for substrate recognition.[63,64] It is hypothesized that the mutual adaption of substrate and enzyme during binding and processing is enabled by the bending flexibility of the substrate's TMD.[63] This model was

deduced from experimental and simulation studies of the APP TMD, which found a flexible di-glycine hinge in the centre of the TMD which coordinates helix bending.[65,66] Due to their large-scale nature, these bending motion dominate the dynamic phenotype of the APP TMD. Initially, this lead to the assumption that bending movements, coordinated by a central hinge might be a common feature of γ-secretase substrates but should not be present in non-substrates.[63] However, this hypothesis was recently challenged by NMR investigations of the γ-secretase substrates NOTCH1, which could neither detect a hinge nor an exceptionally flexible TMD.[67] The later one was recently challenged by Stelzer and Langosch (2019) revealing a flexible site potentially acting as a hinge within the Notch1 TMD.[68] Although, large scale bending motions as well as central hinges might not be of primary importance, dynamics may still provide a rationale for a shared motif. Functional modes might be related to lower-amplitude, more localized motions that are hidden in the conformational ensemble of the unbound state, but might be activated upon binding to the enzyme.[69]

## 1.4 The Amyloid Precursor Protein

Human APP is part of a protein family consisting of APP, APLP1 and APLP2.[70] They form a group of membrane proteins, which are highly conserved across several species excluding prokaryotes, plants, and yeasts.[70] While their biological function is still not fully uncovered, knockout experiments showed a lack of APP to lead to development deficiencies in different organisms.[70] In the same way double knockouts of APP and either APLP1 or APLP2 led to early death after birth.[70] In particular, APP seems to modulate interactions with intracellular signaling systems, implicated in the growth of axonal and dendritic processes and in the support of a variety of functions involved in synaptic maintenance. In its native state, APP consists of 770 amino acids, which are grouped in three major domains. With a total of ~700 amino acids, APP's ECD is by far the largest domain. The shorter intra-cellular domain (ICD) consists of ~45 amino acids and seems to interact directly with the membrane bilayer.[65] Both domains are connected by 24 residues, which form a helical domain that spans the lipid bilayer of cellular membranes. Although APP has attracted significant interest due to its relation to AD, only little is known about the structural properties of the ECD and the ICD. However, structures of several fragments were resolved, allowing to predict a potential overall structure of APP.[71] In contrast to its ECD and ICD, a series of rather high resolution conformations exists for the APP TMD, which harbours the cleavage sites.[65,72] While all of them agree in terms of a helically shaped TMD, they significantly differ in terms of their structure in the membrane. While Barret et. al. (2012)[65] found a kink in the TM-helix of APP located at a di-glycine motif ($G_{708}$ and $G_{709}$), this was not the case in the study of Nadezhdin et al. (2011).[72] This discrepancy was later traced back to the different sizes of the micelles used in the experiments (see 1.4.3). Besides those di-glycine motifs, two additional glycine residues ($G_{700}$ and $G_{704}$) are located in the N-terminal part of the APP TMD (TM-N). These two glycine residues seem to be involved in cholesterol binding[65] as well as formation an APP homo-dimer.[66,73–76]

### 1.4.1 Cleavage of the Amyloid Precursor Protein by γ-Secretase

In contrast to most of the other substrates of γ-secretase, APP is well characterized in terms of cleavage sites, cleavage efficiency, cleavage pathways and cleavage kinetics. After shedding off the ECD by BACE a 99 amino acids long membrane bound C-terminal fragment is released, referred to as C99. This C99 fragment is processed by γ-secretase. Therefore, C99 numbering will be used in the following (e.g., $G_{708}$ becomes $G_{37}$). Although C99 can assemble to homodimers,[66,73–76] cross-linking experiments could reject the hypothesis, that C99 is cleaved in its dimeric form.[73] Therefore, the following sections only covers the properties of the monomeric C99 and its TMD.

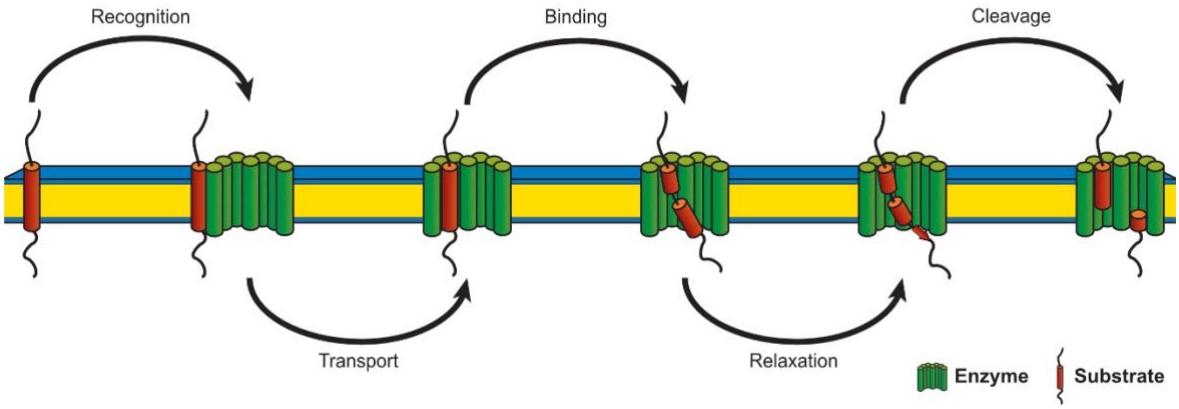


**Figure 1: Model for Catalytic Processing of the C99 TMD by γ-Secretase.**
The cleavage process can be divided into six consecutive steps: (1) substrate recognition at an exosite, (2) substrate transport to the final docking site, (3) formation of the enzyme-substrate complex, (4) conformational reorganization of the enzyme-substrate complex, and (5) unwinding of the cleavage domain into the active site and, (6) hydrolysis and release of cleavage products. This model has been proposed in Götz & Scharnagl (2018)[79] and is based on Langosch et al. (2015)[63] and photo-crosslinking experiments by Fukimori & Steiner.[50] Figure reprinted and adapted from "Dissecting conformational changes in APP's transmembrane domain linked to ε-efficiency in familial Alzheimer's disease" by Götz, A. and Scharnagl, C. 2018, PLoS One, 13(7), e0200077. Licensed under CC BY 4.0, https://doi.org/10.1371/journal.pone.0200077.g002.

The cleavage process can be separated in multiple steps, which cover (1) binding of the substrate, (2) its positioning and (3) chemical cleavage and release of the cleavage product, schematically shown in **Figure 1**. After initial endoproteolytic cleavage releases the intracellular domain of (AICD), subsequent carboxypeptidase-like trimming of the N-terminal fragments releases Aβ-peptides of varying lengths.[2,5,77] Kamp et al. (2015) showed that cleavage of the APP TMD is an unusually slow process with turnover rates in the hours range, compared to soluble proteases which exhibit turnover rates in the seconds to minutes range.[78] A similarly slow processing was observed also for NOTCH1.[55] The reasons for this slow turnover are unknown and may be related to translocation between several binding intermediates during the substrate-enzyme association process. Recent experimental studies suggested

that the initial encounter can occur at various sites in the NCT, PEN-2 and PSN1 subunits,[50,80,81] which was confirmed by docking simulations in Götz et. al. (2019b)[82]. According to cross-linking experiments by Fukumori and Steiner (2016),[50] the substrate makes first contact with PEN-2, and is then transported on the outside of γ-secretase, towards an exosite located in the NTF of PSN1, before its cleavage domain makes contact to interaction sites in the PSN1 CTF close to the active site (but still on the outside of γ-secretase). To the current date it is unclear how long the translocation process may take. According to structural analysis, a certain distance must be overcome between the potential exosites and the active site of γ-secretase. Successful translocation might require a certain flexibility of C99's helical domain to allow its C-terminal domain (TM-C), which harbors the cleavage sites, to get into contact with γ-secretase's active site. Such "swing-in" has been suggested several years ago[83] and seemed to be in good agreement with recent studies on the flexibility of the APP TMD.[79,82,84,85]

After positioning of the cleavage domain close to in the catalytic cavity of γ-secretase's, its unwinding into the active site is a key requirement to allow access of the active aspartates to the scissile bonds as already described for soluble proteases.[86] Recently, experimental studies using RAMAN and NMR spectroscopy as well as cryogenic electron microscopy (cryoEM) were able to confirm that substrates of I-CLiPs are unfolded around their scissile bonds when bound to the enzyme's active site.[87–90] Similar results were obtained from MD simulations, showing that binding of the substrate induces its unfolding at the active site.[91] Cleavage of the TMD happens in a sequential manner, with initial cleavage taking place at the ε-cleavage sites in the TM-C of the C99 TMD (schematically shown in **Figure 2**). Initial cleavage of C99 releases the APP-intracellular domain (AICD). The length of the AICD as well as the remaining N-terminal fragment, depends on the selection of the initial ε-cleavage site. Two such ε-cleavage sites exist, which are either located between $T_{48}$ and $L_{49}$ (ε48) or $L_{49}$ and $V_{50}$ (ε49). If initial cleavage takes place at ε49, a 50-residue long AICD (AICD50) is released and γ-secretase follows the Aβ40 production line, which includes cleavage at ε49-ζ46-γ43-γ40 and finally leads to the non-toxic Aβ40 fragment. If cleavage happens at $\varepsilon_{48}$, AICD49 is released and further cleavage takes place at ζ45 and γ42, which also referred to as the Aβ42 production line.[2] Aβ42 is also referred to as the toxic Aβ species due to its high potential to form aggregates and its occurrence in neurotoxic plaques, which are related to AD.[92–94] The ratio between Aβ42 and Aβ40 (Aβ-ratio) is a central indicator for AD, which is correlated to an increased amount of Aβ42 relative to Aβ40.[6,45,95–97]

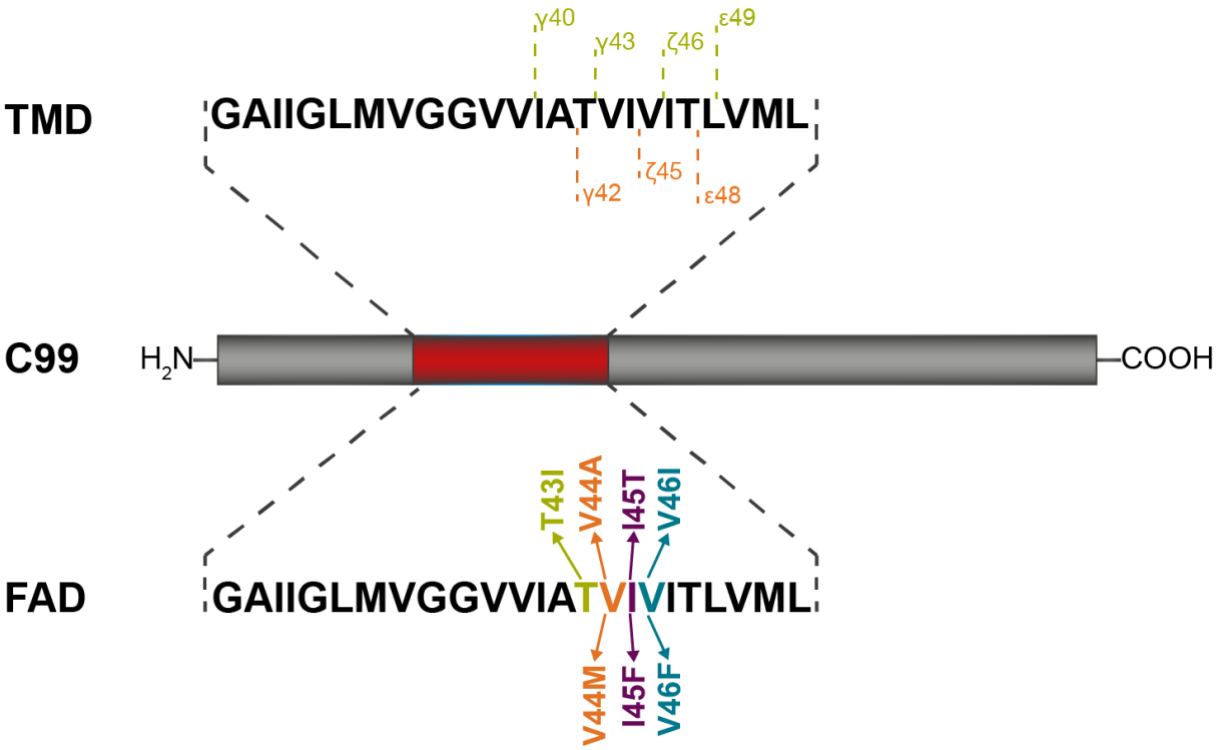


**Figure 2: Schematic Representation of the Amyloid Precursor Protein's C99 Fragment.**
The upper panel shows the sequence of the transmembrane domain of C99 (residues $G_{29}$ to $L_{52}$) together with the two main production lines which lead to formation of Aβ40 and Aβ42 peptides, respectively. The lower panel shows the location of FAD mutations investigated in Götz & Scharnagl (2018)[79] and Götz et. al (2019a).[85] Figure reprinted from "Dissecting conformational changes in APP's transmembrane domain linked to ε-efficiency in familial Alzheimer's disease" by Götz, A. and Scharnagl, C. 2018, PLoS One, 13(7), e0200077. Licensed under CC BY 4.0, https://doi.org/10.1371/journal.pone.0200077.g001.

### 1.4.2 Familial Forms of the Amyloid Precursor Protein Transmembrane Domain

A series of APP FAD mutations exist, which are mainly located in the cleavage domain (TM-C) of the C99 TMD.[97] Reduced ε-cleavage efficiencies are generally observed for FAD mutations and express themselves through reduced AICD and/or total Aβ levels.[45,96,98] In general, all of them shift the Aβ production towards higher Aβ42/Aβ40 ratios, which are correlated with severely progressing forms of AD.[6,45,95,96,99–101] The observed increase in Aβ42/Aβ40 ratios are commonly linked to an altered ε-cleavage site preference, with higher cleavage propensities for the ε48 site observed in FAD mutants. However, the final extent of observed changes behaves rather differently for individual FAD mutations.[102–105] For example, the V46F FAD mutant drastically shifts ε-cleavage to ε49 (increased AICD50) and enters the Aβ42 product line.[95] Nevertheless, it generates substantial amounts of Aβ40 indicating that it also used the alternative ε48-ζ46-γ43-γ40 cleavage steps.[95,106,107] Other exceptions are the I45T and I45F mutations as they both exhibit an increased Aβ42/Aβ40 ratio but only negligible changes in its ε-cleavage site propensity.[98] This contradicts the previously described product line preference (see. 1.3.1). In case of the I45T mutation, the observed alterations in Aβ42/Aβ40 ratios, without altered ε-cleavage site propensity, were related to pathway switching that happens after the initial ε-cleavage.[98] Similar decoupling between the initial ε-cleavage and entry

into specific product lines was also reported for artificial mutations.[95,98,102] However, ε-cleavage is not unaffected as dramatically reduced ε-cleavage efficiencies can be observed.[45,96,98]

### 1.4.3 Dynamic Properties of the Amyloid Precursor Protein's Transmembrane Domain

Since the publication of the first experimentally determined conformations of the C99 TMD,[65,72] its dynamic properties attracted significant interest.[66,72,82,84,85,108–111] In contrast to the strongly kinked conformation reported by Barret et al. (2012),[65] the conformations reported by Nadezhdin et al. (2011),[72] found the C99 TMD to have a mostly straight helical shape. Most MD simulation studies of the C99 TMD could not detect a strongly kinked conformation,[66,72,84,108,112,113] however, they found the di-glycine motif to be a very flexible region in the centre of the TMD.[66,75,113] Lemmin et al. (2014) were able to reveal, that the discrepancy was a consequence of the curvature and related thickness of the micelles that were used in the two experiments to mimic the native lipid environment.[108,112] However, this discrepancy indicated also that the C99 TMD owns the ability to adapt to its surrounding environment by changing its shape. This was of major interest as the proposed "swing-in" model of substrate positioning would require such a structural rearrangement to bring the cleavage sites in contact with the active cavity of γ-secretase.[63,83]

Such a structural flexibility is enabled through the previously mentioned di-glycine hinge.[65,108,112] The glycine residues $G_{37}$ and $G_{38}$ induce a flexible region in the helix centre which allows the otherwise straight helix to change its conformation.[66,108] Several studies showed that this di-glycine motif acts as a flexible hinge, which coordinates bending and twisting motions of the neighbouring, quasi-rigid helical segments of the C99 TMD.[79,82,84,85,108,112] The extent of bending depends on the surrounding environment,[108,112] while the direction can additionally be altered by Thr to Val mutations in the TM-C.[84,114] Considering the observed changes of Aβ42/Aβ40 ratios and ε-cleavage site preference[115] it was argued that perturbations of the bending motion might be related to the observed shifted ε-cleavage site preference in FAD mutants.[63,64] In addition, it was proposed, that the presence of hinges and the motions they coordinate might be the distinguishing factor between substrates and non-substrates of the γ-secretase.[63,64]

To act as a hinge, the flexible di-glycine motif needs to be flanked by two less flexible (quasi-rigid) domains. In case of the C99 TMD, these are the TM-N and TM-C domains. The TM-N domain is the shorter one and forms a small helix which is significantly less stable than the TM-C domain.[66] It contains several glycine residues, making this part of the TMD less hydrophobic.[73,116] In contrast to the rather flexible TM-N helix, the TM-C domain is the most rigid part of the helix, with a stability maximum at the ε-cleavage sites.[66,72,84,108,110] This is rather unexpected as hydrolysis is assumed to require local unwinding of the cleavage domain in order to make the scissile bonds accessible to the active aspartate residues of γ-secretase (see **Section 1.4.1**). Bolduc et. al. [98] presented a simple pocket model in which unwinding of the helix at the ε-cleavage sites is coupled with the entry of TM-C into

the active site of γ-secretase and binding to the S1´-S2´-S3´pockets of γ-secretase. This pocket model together with the requirement of cleavage site unwinding led to the alternative idea that differences of helix flexibility at the ε-cleavage sites in the enzyme-bound state provide a rationale for observed alterations of ε-cleavage site preference in FAD mutants.

## 1.5 Resulting Objectives for the Present Study

The described peculiarities of I-CLiPs and γ-secretase, shift the focus of interest away from sequence towards more collective substrate properties like structure and dynamics and the sequence motifs that control them. Unfortunately, both, structure and dynamics of membrane proteins are difficult to access experimentally. However, *in-silico* techniques like molecular dynamics (MD) simulations, in combination with the available capacity in high-performance computing (HPC), provide a valuable tool to access these properties in a (relatively) fast and reliable way. They gain even more importance if they can be carefully validated by experimental results. Therefore, it will be the aim of the present work to provide detailed insights into the dynamic properties of the APP TMD by MD simulations, which were carefully validated by amide hydrogen-exchange experiments using mass spectroscopy (MS) and nuclear magnetic resonance (NMR) in an organic solvent environment consisting of 80% 2,2,2-trifluoroethanol (TFE) and 20% water (v/v).[117–119] Although the experiments were carried out by collaboration partners, they are an integral part of the overall study, as they will proof the MD simulations to provide a suitable model to investigate the dynamic properties of TMDs. The used isotropic solution has shown to be a good mimetic for the water-filled active site of γ-secretase.[66,79,84,120–122] Usage of such a mimetic was necessary due to a lack of detailed information on the binding of TMDs to γ-secretase at the begin of the study. Only very recently this information became available for NOTCH1 and APP.[87,88] A subset of mutants as well as other putative substrates will be studied also in a model membrane consisting of a 1-Palmitoyl-2-oleoylphosphatidylcholine (POPC) bilayer. In addition to the biophysical measurements, the MD simulations will be compared to the outcome of biochemical cleavage assays. This is intended to provide new mechanistic insights into the cleavage of APP by γ-secretase, the role of a previously described di-glycine hinge flexibility as well as other structural elements,[65,79,85] together determining the dynamic personality of the APP TMD.[117] Although APP's prominent di-glycine hinge plays an important role in its cleavage by γ-secretase, it will be shown that hinges and the resulting possibility of bent conformations are not a general requirement for substrates as discussed before.[63,64,79,85] This is supported by supplementary results for 30 other known or putative substrates of γ-secretase included in the appendix.[123]

# 2 Materials and Methods

The following section summarizes methods which have been used by the author of this thesis in the publications presented in **Chapter 3**. All simulations were performed on the (now decommissioned) SuperMUC high performance computing (HPC) system and used computing resource kindly provided by the Leibniz Supercomputing Centre (LRZ) via projects pr42ri, pr27wa and the Gauss Centre for Supercomputing (GCS) via project pr48ko. Methods that have only been used in some of the presented publications, have been explicitly mentioned in the corresponding section. If a publication used a different implementation of an analysis algorithm or used other parameters than the previous ones, this is also stated in the corresponding paragraph. All presented algorithms are custom implementations if not mentioned otherwise. Implementations were done in the Python programming language version 3 and used the NumPy[124] and SciPy[125] numeric library and related linear algebra algorithms as well as the Mdtraj library[126] for reading and writing trajectory files as well as distance computations. Plots and figures have been created using Matplotlib,[127] while schemes were created by Adobe Illustrator CS6 and Affinity Designer. Visualizations of conformations were done by VMD 1.9.3.[128]

### 2.1.1 Investigated Peptides and Sequences

Two different kinds of peptides were investigated in this thesis. In Högel et al. (2018),[118] the MD simulations modelled a 23 residues long model peptide (termed LV16) that was used to investigate the impact of glycine on its dynamic properties. The helix itself consisted of eight consecutive leucine-valine (LV) repeats. To be compatible with experiments, a KKWK tag was added at the N-terminus and a KKK tag at the C-terminus. By sequentially exchanging leucine to glycine along the TMD backbone, a series of mutants was created, referred to as LXG, with X representing the glycine position in the LV-repeat. For detailed sequences please see Table 1 in Högel et al. (2018).[118]

In Götz & Scharnagl (2018) the native residues $K_{28}$–$K_{55}$ of C99 were investigated.[79] In order to get in agreement with previous publications,[66,84,120] an additional KKW tag was added at the N-terminus (e.g. for wild-type C99: KKW$^{28}$K-GAIIGLMVGGVVIATVIVITLVML-KK$^{55}$K). These short peptides, mainly consisting of the C99 TMD, were shown to be good substrates for γ-secretase.[109,129] Seven FAD mutants (T43I, V44A, V44M, I45T, I45F, V46I and V46F) were compared to WT (see **Figure 2**). These mutations have been selected from literature because they were previously characterized with respect to cleavage efficiency, product-line preference and Aβ-ratios by several groups.[45,95,96,98,109,130]

Instead of the non-native KKW tag as used in Götz & Scharnagl (2018),[79] Götz et al. (2019a/b)[82,119] used the native residues $S_{26}$-$K_{55}$ of C99 to investigate the impact of the I45T FAD mutation (Götz et al. 2019a)[85] and two artificial mutations of glycine at position 38 (C38P, G38L) in the C99 TMD

(Götz et al. 2019b).[82] Instead of charged termini, blocked termini were used to achieve more native-like conditions. For this purpose, the N-terminus was acetylated, the C-terminus was amidated using patches ACE and CT2 from the CHARMM force-field, respectively.

### 2.1.2 Generation of Initial TMD Conformations for MD Simulations

As no structures were available for the investigated peptides from www.rscb.org[131], initial start configurations were generated by two different approaches. Högel et al. (2018)[118] and Götz & Scharnagl (2018)[79] modelled the TMDs as ideal α-helices using general backbone dihedral angles for TMDs as determined by NMR.[132] Geometries of side chains were modelled with the parameters from the CHARMM22[133] force field. Götz et al. (2019a/b)[82,85] as well as simulations presented in the appendix applied a different approach which uses stochastic sampling of side chain dihedral angle combinations from databases in combination with simulated annealing. Like in Högel et al. (2018)[118] and Götz & Scharnagl (2018),[79] an ideal α-helix in terms of backbone dihedral angles is assembled first. Side chain orientations are then adjusted according to their weight within the Dunbrack backbone-dependent rotamer database.[134] The helix is finally placed with the TMD's centre of mass (according to the notation in www.uniprot.org[135]) in the middle of in an implicit membrane slap, which is mimicked by the GBSW model with parameters as described in **Table 1**.

**Table 1:** Settings for the GBSW Model used in the Structure Sampling Protocol

| Parameter | Setting | Description |
|---|---|---|
| SW | 0.3 Å | Half smoothing length |
| SGAMMA | 0.03 kcal $mol^{-1}$ $Å^{-2}$ | Nonpolar surface tension coefficients |
| DGP | 1.5 Å | Grid spacing for lookup table |
| TMEMB | 40.0 Å | Thickness of low-dielectric membrane slab |
| MSW | 2.5 Å | Half membrane switching length |
| NANG | 50 | Number of angular integration points |

The helix axis as computed by the CHARMM molecular dynamics package[133] was aligned with the z-axis. The final system was minimized for 100 steps using CHARMM's steepest descent algorithm, which is followed by minimization with the adopted basis Newton-Raphson algorithm (ABNR).

Subsequently, the system was subjected to a simulated annealing protocol, which consists of 25 temperature steps. Initially, the system is heated up from 0 K to 800 K in 400 ps, followed by a 10 ps long holding step at 800 K. After that, the system is cooled down to 300 K with a stepwise profile proposed by Kannan & Zacharias (2009),[136] which uses 23 steps of 100 ps length (800 K, 755 K, 710 K, 670 K, 640 K, 615 K, 590 K, 565 K, 540 K, 520 K, 500 K, 480 K, 460 K, 440 K, 420 K, 400 K, 385 K, 370 K, 355 K, 340 K, 325 K, 310 K, 300 K). A Berendsen thermostat was used to

keep temperature fluctuations within ±5 K. Integration was performed in steps of 2 ps, using the Leapfrog algorithm and SHAKE. Non-bonded neighbour lists were computed up to 20 Å. Van der Waals (vdW) interactions were switched off between 15.9 Å and 16 Å. The TMD's backbone dihedral angle fluctuations were constrained by 4.9 (Φ) and 2.2 kcal/mol (Ψ), respectively. The whole sampling procedure was repeated ~1500 times for each peptide to obtain a heterogeneous set of conformations. To keep computing times short, the protocol has been ported to the SuperMUC HPC system. A clustering approach was used to select a series of representatives from the 1500 structures. For this purpose, all conformations were clustered by their pairwise Cα -RMSD of the annotated TMD. Initially, the pairwise distance matrix $d_{ij}$ is computed between all sampled conformations of a peptide. If only a single representative has been required, hierarchical clustering as implemented in SciPy[125] has been used. The number of clusters was determined by the elbow in the cost-function plot and representatives (named Centroid in the following) for each cluster were computed from the pairwise distances between structures in the cluster $d_{ij}$ as shown in **Equation 1**.

$$Centroid = argmax_i \sum_j e^{-d_{ij}/\sigma(d_{ij})} \quad (1)$$

If multiple starting conformations are required (e.g. for ensemble runs) the conformations were clustered by affinity propagation clustering[137] as implemented in Scikit-learn[138]. Because affinity propagation works with similarities $s_{ij}$ between data points $d_{ij}$ was transformed according to $s_{ij} = -d_{ij}$. The initial preference was set to the highest similarity $\max\{s_{ij}\}$ and the damping factor λ to 0.98. Centroids were used as determined by the algorithm.

## 2.2 Molecular Dynamics Simulation

### 2.2.1 Simulations in an Isotropic Solvent Mixture

#### 2.2.1.1 Single-Run Simulations

Simulations in Högel et al. (2018)[118] and Götz & Scharnagl (2018)[79] were performed as described in Pester et al. (2012)[66] In brief, ideal α-helices, were placed in a pre-equilibrated, rectangular solvent box (6*6*10.4 $nm^3$) that contained 80% TFE and 20% water (v/v). Initially, the system was equilibrated over a total of 1.2 ns, using multiple short steps with varying constrains. Production runs were performed in an NPT ensemble (T = 293 K, p = 0.1 MPa) using NAMD2.9[139] and the CHARMM22[133] force field with CMAP corrections.[140] The total simulation time was ~200 ns. To guarantee a sufficiently equilibrated system, only the last 150 ns of each trajectory were subjected to further analysis.

#### 2.2.1.2 Ensemble Simulations

For each investigated peptide 78 simulations were performed, each with a total length of 200 ns length. For this purpose, the cluster centroids as previously determined by affinity propagation clustering[137] (se 2.2) were used as start conformations. The same parameters as described in **Section 2.2.1** were used for each simulation. All equilibration and production runs were performed using NAMD 2.11[139] and the CHARMM36 force field.[141] To keep the number of runs to be submitted to the SuperMUC HPC cluster manageable, a job-farming approach as developed by LRZ was used, which allows to control an unlimited number of jobs with varying job sizes, runtimes and dependencies.[142] Using a graph-like approach, the dependencies between several jobs can easily be managed, which allowed to join hundreds of jobs together to single large job, being much more suitable for the queuing system. As a result, a total aggregated simulation time of 15.6 µs was collected for each peptide. The last 150 ns of each simulation were subjected to analysis, leading to an aggregated analysis over 11.7 µs for each peptide. To reduce the amount of data for analysis, trajectories were reduced after sampling and stored on a local network attached storage (NAS) filer hosted by LRZ, while the full trajectories were stored to attached tape libraries. Reduction included removal of solvent and ion atoms from the trajectory and storing frames every 10 ps instead of every ps as in the raw trajectory files.

### 2.2.2 Simulations in a POPC Bilayer

Simulations performed in POPC bilayers used the centroid conformation of the highest populated cluster as obtained by hierarchical clustering (see **Section 2.1.2**) as start conformation. The structure was placed in a POPC bilayer, consisting of 128 POPC lipids, using procedures as provided by CHARMM-GUI.[143] Simulations of 2.5 µs length (T=303.15 K, p = 0.1 MPa) were performed, using NAMD 2.12[144] in combination with the CHARMM36 force field.[141] Similar to simulations in **Section 2.2.2** only the coordinates of the peptide and phosphate atoms are stored on the local NAS system, using a stride of 10 ps between frames. Only the last 1.5 µs of each trajectory were subjected to analysis.

## 2.3 Investigation of Parameters Describing Local Helix Flexibility

To obtain a detailed dynamic profile of the investigated TMDs, a series of parameters was calculated which describe the sequence-dependent backbone dynamics. Besides being easily accessible for interpretation, some parameters can be compared to experimental investigations by either electrospray ionization (ESI) time of flight (TOF) mass spectrometry (MS) (deuterium hydrogen exchange rate constant $k_{DHX}$) or NMR (hydrogen deuterium exchange rate constant $k_{HDX}$, helicity) experiments.

### 2.3.1 Helicity of the TMD

As the TMD annotation in the UniProt database often does not match with experimental or *in silico* determinations of the TM-helix,[145] the content of helical structure of each residue has been determined. For this purpose, the Define Secondary Structure of Proteins (DSSP) algorithm[146] as implemented in the Mdtraj library[126] was used in its simplified version. The helicity of each residue was calculated as the percentage of helical structure reported during the simulation. Residues, which showed > 95% helicity were considered as part of the TM-helix.

### 2.3.2 Intrahelical Hydrogen Bonds Occupancy

To investigate H-bond stabilities, the fraction of closed hydrogen bonds was computed for α H-bonds (O(i) to NH(i+4)) and $3_{10}$ H-bonds (O(i) to NH(i+3)) by a combined distance and angle cut-off criterion. An amide H-bond was considered as closed if the $O \cdots H$ distance was < 0.26 nm and the $O \cdots H - N$ angle was within the range of 180° ± 60°.

### 2.3.3 Calculation of Deuterium Hydrogen Exchange Kinetics from MD Simulations

In order to validate the MD simulations against experimental data, deuterium-hydrogen exchange (DHX) kinetics were reconstituted from the MD simulations as described in Pester et al. (2012).[66] In brief, the exchange rate $k_{DHX(i)}$ of any backbone amide at position $i$ depends on the stability of the emanating H-bond, the chemical exchange rate $k_{ch}$.as well as the concentration of available exchange catalyst $[OH^-]$. The fraction of an H-bond in its open state ($f_{open}$) characterizes its stability. $k_{ch}$ depends on the pH, and the chemical nature of the nearest-neighbour residues.[20]

$$k_{DHX(i)} = \mathrm{f}_{\mathrm{open(i)}} k_{ch(i)} [OH^-] \tag{2}$$

Opening of the amide H-bond is important in two different ways as it (1) allows the amide to exchange, and (2) provides access of the exchange catalyst to the amide deuteron. Therefore, opening of the H-bond is monitored in the MD simulations by a distance cut-off criterion. An H-bond is considered to be in the closed state if the distance between the amide hydrogen at position $i$ and the carbonyl oxygen at $i + 4$ (α H-bond) or $i + 3$ ($3_{10}$ H-bond) is below the cut-off value. In Högel et al. (2018)[118] as well as Götz et al. (2019a/b)[82,85] distances $\leq 0.35\ nm$ characterized a closed H-bond, which is 0.05 nm larger than in Pester et al. (2012).[66] The concentration of exchange catalyst $[OH^-]$ was computed as described in Pester et al. (2012).[66] As the water dissociation constant in TFE is not known and to account for TFE effects on the chemical exchange rate $k_{ch}$ a correction factor $f$ as described in Pester et. al. (2012)[66] is introduced which is determined by a least-square optimization procedure. Therefore, the reconstituted overall exchange kinetics $D(t)$, which is the sum of

individual exchange kinetics of each residue, is optimized to fit the experimental data by least squares optimization of **Equation 3**.

$$D(t) = \sum_{i=1}^{n_{res}} \left( a * e^{-k_{DHX(i)} * f * (t+t_0)} + c \right) \quad (3)$$

Högel et al. (2018)[118] used the original fitting protocol as described in Pester et al. (2012)[66], setting the amplitude ($a$) to 1.0, the baseline ($c$) and the time correction factor ($t_0$) to 0.0. This was changed with Götz et al. (2019a) which introduced $t_0$ as well as modified parameters for $a$ and $c$.[85] The time correction $t_0$ was introduced since the measured data always included minor time shift, which made the first time point at $t = 0$ useless for fitting. To account for 5% of non-deuterated peptide in the experiment, the amplitude (a) was set to 0.95 and a baseline (c) of 0.05 was used. The quality of the MD-derived prediction of exchange kinetics was assessed by the normalized mean-squared deviation ($\chi^2$) of the averaged D(t) values with respect to the experimental averages.

### 2.3.4 Flexibility Profiles from Backbone Mean-Squared Fluctuations

To characterize flexibility of backbone Cα atoms, their mean-squared fluctuations (MSF) were calculated block wise (see **Section 2.7** for details on block averaging). For each block, an average structure was determined iteratively.[147] All conformations of the block were aligned to the average structure by their root-mean squared deviations (RMSD). Mean square fluctuations of the block were computed by using the main diagonal of the block's covariance matrix. To compare MSF between different peptides and environments, the MSF were normalized by their mean $\mu_{MSF}$ and standard deviation $\sigma_{MSF}$ as shown in **Equation 3**, excluding outliers that are detected by the median absolute deviation, using an $M_{score}$ larger than 3.5.

$$MSF_{norm}(i) = \frac{MSF(i) - \mu_{MSF}}{\sigma_{MSF}} \quad (3)$$

## 2.4 Investigation of Helix Geometry and Structural Dynamics

While dynamics might be altered by mutations and different solvent environments this is often not the case for structural properties like helix shape. Therefore, several structural parameters have been investigated which are of specific interest for TM-helices.

### 2.4.1 Side Chain Packing Score

Densely packed side chains are important stabilizing elements of TM-helices and provide steric shielding of the amide proton, limiting solvent access.[148] In order to access the packing of side chains around backbone amide protons and carbonyl oxygen atoms, we computed the packing score $S_i$ as proposed by Grossfield et al. (2006).[149] $S_i$ is defined as the sum of the inverse of the pairwise distance $d_{ij}$ between the according backbone atom $i$ and all other atoms $j$ raised to the sixth power. As this gives higher weight to small distances, backbone atoms belonging to the same residue $i$ as the atom of interest were excluded from the computation to avoid bias caused by the length of the side chain. In some cases (see **Section 3.1**), usage of a normalized packing score $S_{i,norm}$ instead of absolute values is of advantage to avoid numerical instabilities. Normalization of $S_i$ is achieved by dividing $S_i$ by the sum of $S_i$ over all residues $j$.

### 2.4.2 Determination of Local Helix Axis and Rise Per Residue

The rise between two consecutive residues (RPR, pitch) is a valuable descriptor of deviations from an ideal helix. RPR values for ideal α-helices are 1.52 Å and 1.96 Å for $3_{10}$ helices, respectively. For the calculation a differential geometric algorithm was applied, which has been describe previously by Guo et al. (2013).[150] The algorithm uses the cartesian coordinates of Cα atoms as anchor points for cubic splines to generate a coarse-grained representation of the helix. From these analytically determined splines, an orthogonal coordinate system (Frenet-Serret frame) can be defined analytically for each Cα atom. This vector frame allows to determine several geometric properties of the helix like its RPR, torsion and spoke angles. The original algorithm is implemented in the HAXIS software package.[150] However, this package was not optimized for analysis of trajectories. As the analysis of ensemble simulations required the investigation of > 1 million conformations per peptide, all algorithms had to be optimized and parallelized. To overcome the limitations of the original implementation and increase its flexibility, the algorithm was re-implemented as a Python3 package, using the NumPy and SciPy libraries for linear algebra operations, replacing the inefficient Python3 loops.

A central part in the algorithm is the determination of local helix axis points $a_i$ for each residue (for details see Guo et al. (2013)).[150] The vector $A_i$ between $a_i$ and a consecutive axis point $a_{i+1}$ describes the direction of the local helix axis. Unfortunately, this is rather prone to noise from small distortions. However, helix axis points $a_i$ showed to be rather suitable to determine the axis of a helical segment with $n \geq 3$ residues in a rather stable and fast way compared to other algorithms. Therefore, the geometric mean of $n$ axis points $a_i$ is calculated as origin of the helix axis. In the next step, singular value decomposition (SVD) is performed to obtain a set of three vectors that describe the principal

axis and related eigenvalues. The vector with the largest eigenvalue represents the axis of the helical segment which can be used for further computations as shown in **Sections 2.4.3** and **2.4.4**.

Furthermore, $a_i$ represents the origin of a circle which crosses the Cα atom and local axis vector $A_i$ as its normal. Hence, it can be used to compute the RPR, which is equal to the distance $d_{i,i+1}$ between consecutive axis points $a_i$ and $a_{i+1}$.

### 2.4.3 Extent and Direction of Global Helix Bending

Bending of its TM-helix has been assumed to be a major feature of APP and essential for cleavage by γ-secretase.[63–65,84] As bending of the APP TMD is thought to control the position of the cleavage domain relative to a putative binding domain in TM-C, two segments were defined. The TM-N segment covered residue $Ala_{30}$-$Leu_{34}$ in Götz & Scharnagl (2018)[79] and $Ile_{31}$-$Met_{35}$ in Götz et al. (2019a/b)[82,85] while the TM-C segment covered $Ile_{47}$-$Met_{51}$ in all papers.

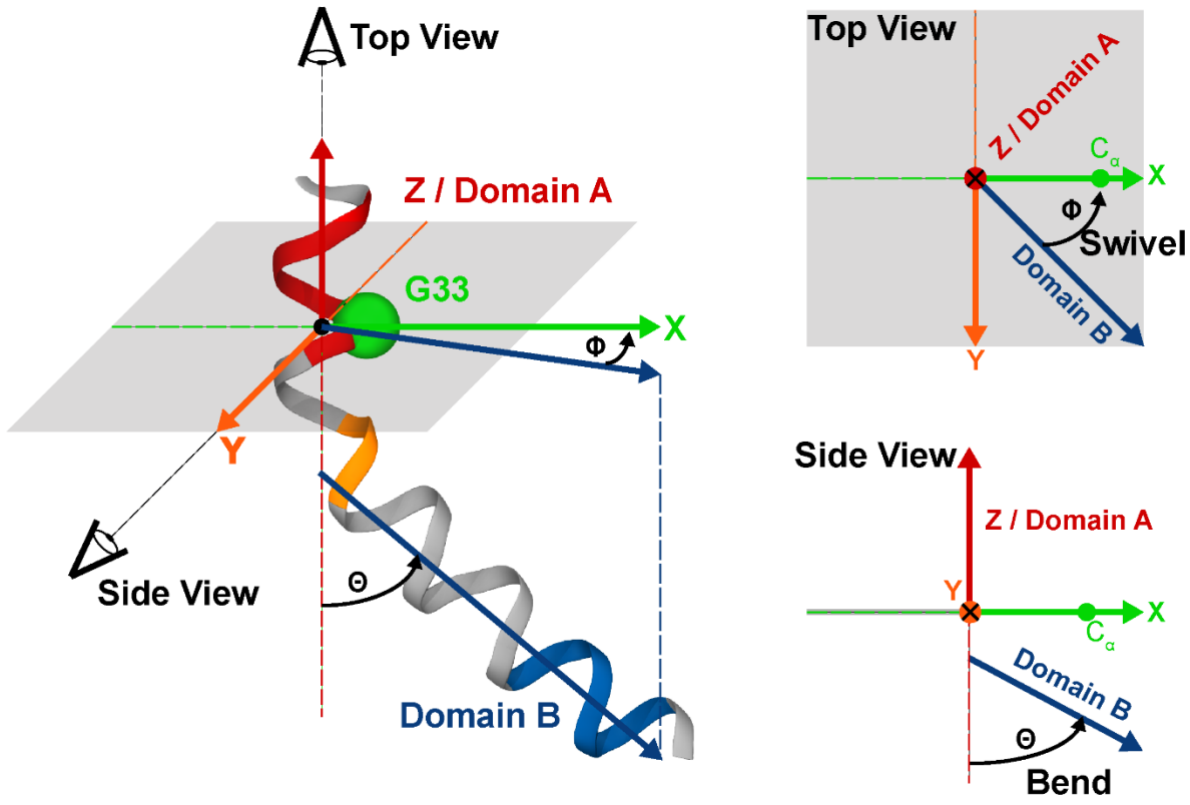


**Figure 3: Geometric Definition of Bending (Θ) and Swivel Angle (Φ).**
These angles describe the orientation of a helical segment located in TM-N (domain A, residues $A_{30}$-$L_{34}$) with respect to a helical segment in TM-C (domain B, residues $I_{47}$-$M_{51}$). The bending (Θ) angle is defined as the angle between the axis of both helical segments. The swivel angle (Φ) is defined as the angle between the vector orthogonal to the TM-N helix axis, crossing the Cα atom of $G_{33}$, and the projection of the helix axis of TM-C onto the plane perpendicular to the helix axis of TM-N. Figure reprinted from "Dissecting conformational changes in APP's transmembrane domain linked to ε-efficiency in familial Alzheimer's disease" by Götz, A. and Scharnagl, C. 2018, PLoS One, 13(7), e0200077. Licensed under CC BY 4.0, https://doi.org/10.1371/journal.pone.0200077.s002.

In Götz & Scharnagl (2018)[79] nearly the same, rather time consuming algorithm was used as described in Scharnagl et al. (2014).[84] Ensemble simulations in Götz et al. (2019a/b) required an adapted version in order to reduce the computing time drastically.[82,85] For both segments, the helix axis vectors of the TM-N domain $\vec{h}_{TM-N}$ (domain A in Figure 3) and the TM-C domain $\vec{h}_{TM-C}$ (domain B in Figure 3) were computed as described in **Section 2.4.2**. The bending angle Θ is calculated from the inner product of both helix axis vectors. The direction of bending as defined by the swivel angle Φ required the additional definition of a reference vector. For this purpose, the Cα atom of $G_{33}$ was chosen as a reference point to compute the reference vector $\vec{r}_{G33}$, which is perpendicular to $\vec{h}_{TM-N}$ and crosses the reference atom. $\vec{h}_{TM-C}$ is then projected in the plane perpendicular to $\vec{h}_{TM-N}$ which results in a vector $\vec{h}^{*}_{TM-C}$. The swivel angle Φ is then computed as the inner product between $\vec{r}_{G33}$ and $\vec{h}^{*}_{TM-C}$ (see Top View in Figure 3). The sign of Φ is given by the inner product of $\vec{h}_{TM-N}$ and the cross product $\vec{r}_{G33} \times \vec{h}^{*}_{TM-C}$.

### 2.4.4 Extent and Direction of Local Helix Bending

In order to access local bending of the TM-helix in Högel et al. (2018),[118] a new approach has been implemented, which is rather similar to the one used for global helix bending. Instead of using domains in TM-N and TM-C, local helix bending at the residue of interest $i$ is defined by domain A ranging from residue $i-1$ to $i-4$ and domain B ranging from $i+1$ to $i+4$. Bending and swivel angles are defined as shown in **Figure 3**.

## 2.5 Interactions Between the TMD and its Environment

### 2.5.1 Solvent Coordination Around the TMD

As water can act as donor of H-bonds to carbonyl oxygens as well as acceptor for H-bonds emanating from backbone amides the occupancy of these H-bonds has been calculated by the same criteria as used for intrahelical H-bonds (see **Section 2.3.2**).

### 2.5.2 TMD Orientation in a Lipid Bilayer

In Götz et al. (2019a/b),[82,85] the orientation of the TMDs (residues $G_{29}$-$L_{52}$ of C99) relative to the surrounding membrane bilayer was accessed by their tilt ($\tau$) and azimuthal ($\rho$) rotation angles. Computation of $\tau$ was done by calculating the inner product between the TM-helix axis vector, as determined by the method described in **Section 2.4.2** and the membrane normal vector (z-axis). The corresponding azimuthal rotation angle $\rho$ was calculated as described in Strandberg et al. (2009)[151], using the Cα atom of $G_{33}$ as a reference instead of the first residue of the helix.

Insertion depths were calculated from the z coordinate of Cα atoms in a coordinate system with the geometric centre of the membrane phosphate atoms as origin.

## 2.6 Collective Helix Dynamics

This section summarizes the analysis methods used to investigate the collective helix motion. The helical part of the TMD was determined from the residue's helicity (see **Section 2.3.1)**.

### 2.6.1 Identification and Classification of Hinges in the TMD

To identify and classify hinge regions in the TM-helix, the Dyndom program was used.[152] Conformations taken every 50 ps (Högel et al. (2018)[118] and Götz & Scharnagl (2018)[79]) or every 100 ps (Götz et al. (2019a/b)[82,85] and **Appendix**) were subjected to analysis, using a sliding window of five residues in order to identify quasi-rigid body domains that consisted of at least four residues. Motions of the TMD were compared to a reference structure obtained from the conformation with the lowest RMSD to an iteratively determined average structure.[147] In order to identify the type of hinge motion, classification was done according to the orientation of the screw axis relative to the helix axis. Screw axes which are mainly perpendicular to the helix axis ($\%_{closure} > 50\%$) are classified as bending motions, while twisting motions are classified by a screw axis which is mainly parallel to the helix axis ($\%_{closure} \leq 50\%$).

### 2.6.2 Identification of Functional Modes by Partial Least-Squares

To identify motions that were correlated with alterations between wild-type APP and its FAD mutants, the functional mode analysis (FMA) tool using partial least-squares (PLS) was used, kindly provided by Bert de Groot.[153] Fluctuations of heavy backbone atoms of residues $G_{29}$ to $L_{52}$ were used as input features, time-series of intrahelical H-bond occupancies (α or $3_{10}$ closed) define the functional order parameters. In Götz & Scharnagl (2018)[79] H-bond occupancies of two regions were individually summed up. The first region (I) contained H-bonds spanning residues $G_{33}$-$V_{42}$, while the second region (II) spans residues $V_{40}$-$I_{47}$. In Götz et al. (2019a)[85] only region II was used for the analysis as no differences could be observed for H-bonds in region I. In both publications, the first half of the (aggregated) trajectory was used for model training, while the second half was used for cross-validation of the PLS-FMA model. To determine the required number of PLS components, convergence of the Pearson correlation coefficient between training data and model ($R_m$) was monitored as a function of the number of components. Structural changes causing substantial variation in the order parameters were characterized by the ensemble-weighted, maximally correlated motions (ewMCM). For visualization, trajectories along the ewMCM vectors interpolating from low to high value of occupancies were used. Characterization of the ewMCM was done by Dyndom (see

**Section 2.6.1**), subjecting the conformations with the highest and lowest extent along the ewMCM to analysis. Similarities of ewMCM vectors were quantified by their inner product.

## 2.7 Statistical Analysis

Mean values, standard errors of the mean (SEM), and standard deviations (SD) were determined from block-averaging over 30 ns in all publication. This block size was > $2\tau$ in all investigated blocks, with $\tau$ representing the first zero passage time of the block timeseries' autocorrelation function. Gaussian error propagation was used for cases with two or more variables. In Götz et. al. (2019a/b)[82,85] bias-corrected and accelerated bootstrap resampling was used to compute mean values and 95% confidence intervals.[154] Instead of Gaussian error propagation Monte-Carlo sampling was used. For bootstrapping as well as for Monte-Carlo sampling, $10^4$ samples were generated to achieve sufficient convergence of the sampled distributions.

# 3 Results

The following chapter summarizes the publications which this thesis is based on and provide information about the contributions of the author of this thesis. The publications are presented in chronological order as they were published in international, peer-reviewed journals. The original articles can be found after the according summary. The supplementary information of each publication are available on the journal websites.

1. Högel, P., Götz, A., Kuhne, F., Ebert, M., Stelzer, W., Rand, K. D., Scharnagl, C., & Langosch, D. (2018). Glycine Perturbs Local and Global Conformational Flexibility of a Transmembrane Helix. *Biochemistry*, 57, 8, 1326-1337, https://doi.org/10.1021/acs.biochem.7b01197
2. Götz, A., & Scharnagl, C. (2018). Dissecting conformational changes in APP's transmembrane domain linked to ε-efficiency in familial Alzheimer's disease. *PLOS ONE*, 13(7), e0200077. https://doi.org/10.1371/journal.pone.0200077
3. Götz, A., Högel, P., Silber, M., Chaitoglou, I., Luy, B., Muhle-Goll, C., Scharnagl, C., & Langosch, D. (2019). Increased H-Bond Stability Relates to Altered ε-Cleavage Efficiency and Aβ Levels in the I45T Familial Alzheimer's Disease Mutant of APP. *Scientific Reports*, 9, 5321 (2019), https://doi.org/10.1038/s41598-019-41766-1
4. Götz, A., Mylonas, N., Hoegel, P., Silber, M., Heinel, H., Menig, S., Vogel, A., Feyrer, H., Huster, D., Luy, B., Langosch, D., Scharnagl, C., Muhle-Goll, C., Kamp, F., & Steiner, H. (2019). Modulating hinge flexibility in the APP transmembrane domain alters γ-secretase cleavage, *Biophysical Journal*, 116, 11, 2103-2120, https://doi.org/10.1016/j.bpj.2019.04.030

Besides those publications, a further publication is summarized in the following, which does contain MD simulation results and analysis of the author, who did not contribute to manuscript writing.

- Hitzenberger, M., Götz, A., Menig, S., Brunschweiger, B., Zacharias, M., & Scharnagl, C. (2020). The dynamics of γ-secretase and its substrates. *Seminars in Cell & Developmental Biology*, S108495211830274X. https://doi.org/10.1016/j.semcdb.2020.04.008

This publication was added as it extends the analysis of TMD dynamics to other substrates and non-substrates. The detailed results are included in the **Appendix** and are subject to the discussion.

## 3.1 Högel et. al. (2018), Glycine Perturbs Local and Global Conformational Flexibility of a Transmembrane Helix, *Biochemistry*

Helix flexibility is assumed to be a key requirement for efficient cleavage of C99 by γ-secretase. C99's central di-glycine hinge raised the interest for allowing the APP TMD to bend anisotropically over this hinge.[65,66,84] Since small conformational transitions of individual TMDs in membrane proteins are often functionally relevant,[155–160] flexible regions are of significant importance as they allow such transitions at a lower cost of energy. Especially kinks and flexible hinges are frequently observed in functionally relevant domains of membrane proteins. While proline's role as a helix breaker is well known, the impact of glycine on helix geometry and/or flexibility is less obvious.[161–165] In order to study glycine's role on TM-helix dynamics systematically, its impact on a low-complexity leucine-valine model TMD (LV16) was used. Substitution of individual leucine residues to glycine along LV16, allowed to study its impact on helix geometry and dynamics dependent on the location of the perturbation. These models were termed LXG peptides, with X representing the position of the leucine-to-glycine substitution in LV16. A total of eight mutations of LV16 were compared to LV16 wild-type as well as two previously established models of higher (LV16-G8P9) and lower (L16) flexibility.[148] By a joint experimental and theoretical approach, combining DHX and electron transfer dissociation (ETD) experiments with all-atom MD simulations, we investigated the dynamics of the LV16 model TMD and its leucine to glycine substitutes. Excellent agreement between MD simulations and the experiments was achieved for overall DHX kinetics of LV16 and the seven LXG peptides. Increased exchange kinetics near the mutation sites indicated increased flexibility at and around the glycine site, potentially being related to weakened H-bonds. The simulations confirmed lower occupancies of backbone H-bonds downstream of the mutation site. The atomistic resolution of the simulations revealed a severe packing defect at the glycine site due to its lacking sidechain. This structural characteristic of glycine allowed the helix structure to change without breaking its overall helical conformation by local shifting from α- to $3_{10}$-helical H-bonds, consistent with an increased rise-per-residue. In addition, the hydrophilic nature of glycine as well as the reduced packing density allowed for increased hydration at the glycine site. The enhanced flexibility enables larger bending over the insufficiently packed glycine and leads to a concave shape at the substitution site. Due to the long-range impact, the flexible hinge in the centre of the TMD coordinating global helix bending is shifted in the direction of the glycine substitution.

**Contributions by Alexander Götz:** Performed MD simulations, developed algorithms, analysed and visualized the simulation data. Together with P. Högel and D. Langosch he drafted and wrote the manuscript.

## 3.2 Götz & Scharnagl (2018). Dissecting Conformational Changes in APP's Transmembrane Domain Linked to ε-Efficiency in Familial Alzheimer's Disease, *PLOS ONE*

Cleavage of APP's TMD by γ-secretase is a central hallmark in the pathogenesis of AD.[166] While most AD cases occur sporadically, a series of single site mutations in the APP TMD is related to familial forms of AD (FAD). These mutations affect ε-endoproteolysis of the APP TMD, leading to altered ratios of Aβ cleavage products that finally cause AD. Reduced ε-efficiency is generally observed for FAD mutations in the APP TMD.[45,95] In addition, some of these FAD mutations shift the preferential initial ε-cleavage site from Aβ49 towards Aβ48 while some others do not. With cleavage taking place in the TMD of a 99-residue long C99 fragment of APP, changes within the TM-helix's structure and dynamics by mutations seem to be related to its faulty cleavage. At the beginning of the study, it was unknown how C99 is bound to γ-secretase and how the TMD is located in its active site. To overcome these limitations, this study used all-atom MD simulations of the wild-type C99-TMD and seven FAD mutants in its cleavage domain (TM-C). An isotropic solvent with a low amount of water was used to mimic the environment in the water-filled active site cleft of γ-secretase. The used solvent consisted of 80% TFE in a mixture with 20% (v/v) water and was already successfully applied in previous investigations.[66,84,120] Such a mixture allowed to mimic hydrophobic interactions in the interior of globular proteins without known specific interactions.[122,167,168] By dissecting the recorded dynamics in a series of local and global features, significant but heterogeneous differences between wild-type and the FAD mutations were observed in the simulations. While previous publications focused on the impact of mutations on C99's di-glycine hinge,[63,84] the current work did not show any significant difference between the wild-type TMD and its FAD mutants at this site. Instead, significantly decreased H-bond stabilities upstream of the ε-cleavage site, but downstream the di-glycine hinge, were consistently observed in all FAD mutations. Various parameters indicated slight deviations of the TM-C toward a more $3_{10}$ like structure. Compensation of weak α-helical H-bonds by $3_{10}$-helical H-bonds was in perfect agreement with the previous studies.[118,169] In order to examine how interactions between the C99 TMD and γ-secretase affect its dynamics, a dynamic perturbation-response approach was applied. Using this analytical tool, it was observed that FAD mutations did not alter the large-scale bending motions of the APP TMD, described in previous publications.[65,84,112] Instead, the mutations impact otherwise hidden lower-amplitude motions in the cleavage domain (TM-C) which are utilized in the bound state. Those motions provided a mechanistic model for the proposed coupling between binding and ε-cleavage that significantly differs between the FAD mutants and wild-type APP.

**Contributions by Alexander Götz:** Performed MD simulations, analysed and visualized the simulation data, drafted and wrote the manuscript together with C. Scharnagl.

## 3.3 Götz et. al. (2019a). Increased H-Bond Stability Relates to Altered ε-Cleavage Efficiency and Aβ Levels in the I45T Familial Alzheimer's Disease Mutant of APP, *Scientific Reports*

Building on the screening of a series of FAD mutations in C99's cleavage domain for their structural and dynamic impact by Götz & Scharnagl (2018),[79] the present study focused on the I45T mutation. While most FAD mutations shift the initial ε-cleavage site from the ε49 cleavage site towards ε48, the I45T and I45F mutations behave differently.[45] They seem to shift γ-secretase product line preference decoupled from the initial ε-cleavage site, indicating switching between product lines during the cleavage process.[98] However, they both exhibit increased Aβ42/Aβ40 ratios[98], while for the I45T mutation, dramatically reduced ε-cleavage efficiency was observed[45]. For the I45F mutation, Bolduc et. al.[98] provided an explanation by a pocket model, consisting of a small-large-small motif at the S1-S2-S3 sites, which is able to explain the observed behaviour of many other known FAD mutations in the TM-C of APP. However, it cannot explain the I45T mutation, which would perfectly fulfil all requirements of the model, but seems to change cleavage preference anyhow.[45,98] A possible explanation is a product line switching at the γ- or ζ-cleavage sites instead of the ε-sites,[98] which is in agreement with previously observed pathway switching.[102] To complement the investigations in Götz & Scharnagl (2018),[79] the current study combined amide hydrogen exchange experiments by NMR and ESI-TOF MS in a TFE/$H_2O$ solvent mixture with multi μs long ensemble MD simulations in this solvent as well as in a POPC bilayer. Good agreement was achieved between the experiments and simulations concerning the overall dynamic profile of the C99 TM-helix, revealing a rather flexible TM-N helix coupled to a very rigid TM-C helix by a flexible di-glycine hinge. Both, experiments and simulations, showed the I45T mutation to lead to increased stability of amide hydrogen bonds at the ζ- and γ-cleavage sites. This was also observed in a POPC bilayer and is related to an additional H-bond between the T45 side chain and the TMD backbone. This H-bond alters the dynamics within the cleavage domain and may provide a rationale for observed γ-secretase product-line switching in the I45T mutation.[98] The increased H-bond stability inhibits an upward movement of the ε-cleavage sites, which is observed in the APP WT but only to a rather minor extent in the I45T mutant. Therefore, presentation of the ε-sites to the active site of γ-secretase is restricted due to increased local stability. In agreement with the previous work on FAD mutations, such a mechanism provides a rationale for reduced ε-cleavage efficiency of the I45T mutant. However, this mechanism remains rather speculative, due to the lack of the native environment of the active site of γ-secretase.

**Contributions by Alexander Götz:** Performed MD simulations, analysed the simulation data, made the figures, drafted and wrote the manuscript together with the other authors.

## 3.4 Götz et. al (2019b). Modulating Hinge Flexibility in the APP Transmembrane Domain Alters γ-Secretase Cleavage, *Biophysical Journal*

Götz & Scharnagl (2018)[79] and Götz et. al. (2019a)[85] dealt with FAD mutations in the TM-C of the APP TMD, which showed negligible small impact on the central di-glycine hinge of APP, but affected motions which were located in the TM-C domain of the APP TMD. However, the functional relevance of the APP hinge for γ-secretase cleavage was still not fully determined.[63,64] After the experimental description by Barrett et. al.[65], the di-glycine hinge of APP had attracted interest as it allows the substrate TMD to undergo structural changes.[66,84,112] This helix bending was suggested to provide the necessary flexibility[66,108,112] for a putative "swing-in" of APP's TM-C into the active site of γ-secretase.[63,83] Such a "swing-in" is frequently discussed, as current initial docking sites seem to be too far away from the active site[50,170] to be reached without structural rearrangement of the substrate.[19] The existence of only one known mutation in APP's di-glycine hinge, which seems to have no impact on AD,[171] strongly questioned the importance of the di-glycine hinge. Hence a broad range of different experimental and theoretical approaches was combined, ranging from cellular cleavage assays to coarse-grained MD simulations. Two artificial mutants of the APP TMD have been designed which substituted $G_{38}$ to either leucine or proline. The initial concept behind both mutations was to decrease (leucine) or increase (proline) helix flexibility, which was predicted to lead to either increased or decreased cleavage efficiency, respectively, as well as altered Aβ42/Aβ40 ratios. Following the approach as presented in Götz et. al. (2019a)**,**[85] multi µs long, all-atom ensemble MD simulations were validated against amide hydrogen exchange experiments by NMR and ESI-TOF MS in a TFE/$H_2O$ solvent mixture, as well as solid-state NMR experiments in a POPC bilayer. Simulation results were in excellent agreement with the experiments. In contrast to the initial assumption, both mutations reduced ε-cleavage efficiency and significantly altered γ-secretase's cleavage specificity. Using coarse-grained MD simulations, altered binding propensities of the mutants' APP TMD to the γ-secretase complex were excluded. This suggests changes in structure and/or dynamics as being responsible for observed changes in γ-secretase cleavage. As in our previous study, H-bond stability of the initial ε-cleavage sites was not affected.[79,84,85] Instead, changed flexibility at the di-glycine sites changed location of the hinge and impacts extent and direction of the global bending motion. A major change in the direction of the TM-C may lead to a non-productive presentation of the cleavage domain to the active site of γ-secretase, thus reducing the likelihood of a successful formation of the enzyme substrate complex.

**Contributions by Alexander Götz:** Performed all-atom MD simulations, analysed the simulation data, made the related figures, drafted and wrote the manuscript together with all other authors.

## 3.5 Comparing the TMD Dynamics of APP to Those of Other γ-Secretase Substrates and non-Substrates. *Hitzenberger et al. (2020) and Appendix*

Using APP's TMD as a paradigm for general substrate TMD dynamics, it was suggested that the large-scale bending motion might be a common flexibility motif of other γ-secretase substrates, too.[63,64] However, NMR investigations of two other substrates, NOTCH1 (5KZO[67]) and the insulin receptors (2MFR[172]), did neither report a bend nor a very flexible TM helix. To investigate the hypothesis of a common flexibility motif only shared by substrates, 26 known substrates were compared to three putative non-substrates (NPRA, ITGB1, DAB12) with respect to local parameters and global bending and twisting motions (see **Figures A2-A9** in the **Appendix**). Candidates were selected from a pool of known substrates and a few non-substrates.[6,37] Selection was either made according to similarity to APP (APLP1, APLP2), already existing or intended experimental investigations (e.g. NOTCH1, BCMA), occurrence of poly-glycine motifs (e.g. Ecadherin, ErbB4), or other characteristic sequence motifs like poly-leucine (e.g. CSF1R). A full list is provided in **Table A1** in the Appendix. Concerning the selected non-substrates, one has to note that NPRA can be converted to a substrate by shortening its intracellular domain,[37] and ITGB1 fails to be shed by BACE.[37] Similar to Götz et. al (2019a/b) MD simulations were carried out in a POPC bilayer as well as in a TFE/water mixture. Comparison of TMD dynamics of APP, NOTCH1 and ITGB1 was recently published by Hitzenberger et al (2020), revealing a set of features:[1,2]

(1) All TMDs form stable transmembrane helices, stabilized by intrahelical α-helical H-bonds. However, switching between α- and $3_{10}$ H-bonds can be observed regularly, with enhanced flexibility in regions with increased contents of $3_{10}$ H-bonds.

(2) Helix stability at known γ-secretase ε-cleavage sites is rather high even in TFE/water without any tendency for unfolding. At the intracellular membrane interface, approximately one turn downstream to the ε-cleavage sites, loss of α-helix stability is compensated by a $3_{10}$ H-bond, favouring the formation of a ß-turn which is even more pronounced in TFE/water. In the apparent non-substrate ITGB1, a stretch of C-terminal leucine residues protects the TM helix from unfolding.[123]

(3) As TM helices generally behave like elastic rods, helix bending and twisting motions are common features of all TM helices, but never as dominant as seen in the case of APP's di-glycine hinge. This emphasized the relevance of sequence for helix flexibility, highlighting the APP TMD as a special case among γ-secretase substrates.

While these observations need to be placed in the context with experimental observations, a common pattern is that the dynamic personalities of substrates are as diverse as their sequences. This might provide a hint towards a proteasome-like functionality of γ-secretase.

**Contributions by Alexander Götz:** Performed all-atom MD simulations, analysed the simulation data and made the related figures.

# 4 Discussion

Although a continually-increasing number of studies opened new insights into the γ-secretase cleavage mechanism, several essential steps like substrate selection, regulation of cleavage efficiencies and pathway preference, are still outstanding problems.[84,95,109,114,115] The dynamic properties of substrates had previously been proposed to be key to their recognition and processing by an enzyme (see Götz & Scharnagl (2018)[79] and references cited therein). Due to the involvement of the cleavage products in AD, γ-secretase cleavage within APP's TMD and the impact of FAD mutations were studied exhaustively in previous work,[6,45,66,95–97,99–101,109,110,116,173,174] some of it focusing on the homo-dimer of the APP TMD.[66,109,110,116,174] As recently uncovered, γ-secretase is not able to cleavage the APP TMD in its homo-dimeric form.[73,175] Therefore, dynamic properties were investigated only for the cleavage competent monomeric APP TMD in this work. Little is known about the recognition, binding and positioning steps of the substrate that must take place before substrate hydrolysis. While C99 recruitment was investigated by photo-crosslinking experiments[50] and coarse-grained docking simulations[123] the penultimate contact site, steering the substrate into the active site of the enzyme, is still elusive. The results of these experiments and simulations suggested multiple exosites for different substrates.[50,82] However, potential exosites seem to be located at a certain distance from the catalytic aspartate residues of γ-secretase. As a consequence, a major structural rearrangement of γ-secretase or a "swing-in" of the substrate's cleavage domain seem to be mandatory in order to get the scissile bond in contact with γ-secretase's catalytic aspartates.[83] While recent cryoEM studies were able to identify the finally positioned substrate in the active site cavity of γ-secretase,[87,88] they were not able to capture the way from the outer surface of the enzyme into its active site cleft. MD simulation, biasing the process of substrate entry into the active site cavity of γ-secretase, revealed a variety of transient conformational switches during this step.[91]

## 4.1 The di-Glycine Motif Determines APP's Obvious Dynamic Personality

How can the dynamic personality of the APP TMD be properly characterized? This question arises when considering the various parameters investigated in the course of the present work. Hence, they have been grouped into two major classes, (1) local parameters, describing structural or dynamic changes with at individual amino-acids (e.g. opening and closing of a single intrahelical H-bond), and (2) more collective global parameters describing the movement of larger segments relative to each other (e.g. hinge bending).[152] However, one has to keep in mind that, despite being local, the residue-resolved parameters are not independent of each of them. For example, the network of backbone H-bonds along the whole helix must cooperate to enable the large-scale motions. From the localized parameters, a dynamic fingerprint can be observed for the APP TMD. In general, the TMD consists of a single α-helix.[72,112] TM-N and TM-C domains are characterized by different flexibility profiles.[65,66,75,84,112] The TM-C domain, harbouring the γ-secretase cleavage sites, is a straight, rather

rigid α-helix. In contrast, the TM-N domain showed to be significantly less stable in the simulations, as well as in the validating experiments.[79,82,85] However, while experiments and simulations agreed on a less stable TM-N, the extent of flexibility showed significant differences.[82,85,118] The origins of these differences were outlined in Götz et al. (2019a)[85] and were related to either (1) difficulties in capturing weak H-bonds in the experiments due to their very rapidly exchanging amides, (2) difference in TM-N structure between the $C99_{26-55}$ peptides used in the simulations and the KKW-tagged $C99_{28-55}$ peptides which had to be used in the ETD measurements, or (3) incomplete sampling of unfolded or slightly disordered regions in the simulations. However, (3) does not seem to be a rationale as NMR spectroscopy suggested slower amide exchange within TM-N closer to the simulation results.[82,85] In addition, previous simulations and measurements in micelles, lipid bilayers as well as TFE/water mixtures showed similar results as the current simulations.[72,75,112]

Surprisingly, all simulations and experiments in this work,[79,82,85] as well as in previous works,[66,72,75,84,112,114] are in contrast to a recent study, which investigated D/H fractionation factors Φ derived from experimental $k_{HDX}/k_{DHX}$ ratios, in the C99 TMD in LMPG micelles.[111] Cao et al. (2017) reported $\Phi < 1$ in TM-N and $\Phi > 1$ for some residues in TM-C. As $\Phi > 1$ indicates weak H-bonds while lower Φ values report higher H-bond stabilities, this led to the conclusion that the C99 TMD has strong H-bonds in TM-N and rather weak H-bonds for some residues ($T_{43}$, $V_{44}$, and $T_{48}$) in TM-C.[111] Through a straight forward analysis of the underlying chemical and physical mechanisms of the exchange reaction, we were able to argue that Φ values of Cao et al. (2017)[111] potentially do not reflect the stability of an intrahelical amide-to-carbonyl H-bond but those of an amide-to-solvent H-bond (for further details please see Götz et. al. (2019a)).[119] Obviously, the distribution of Φ values as reported by Cao et al. (2017) rather closely follows the hydration levels observed in TFE/$H_2O$ mixtures[66,79] and POPC membranes.[66] Hydration of TM-C showed to be significantly lower than for TM-N,[66,79] which parallels observed weak amide H-bonds by Cao et al. (2017) for TM-C.[111] However, according to the simulations in the present work and supporting NMR and ETD experiments,[79,82,85] the TM-C domain of APP is the most stable part of the TM-helix, reaching its highest stability at the ε-cleavage sites.

Normally, one would expect the ε-cleavage sites to be less stable than other parts of the helix, as cleavage requires access of the catalytic aspartates to the helix backbone. However, most studies of the APP TMD, except a few ones including Cao et al. (2017)[109–111] found the TM-C of the APP TMD to be exceptionally stable, with the highest stability observed at or close to the ε-cleavage sites.[66,72,79,82,85,95] A stable helix near scissile bonds was also found for NOTCH1[67] and the sterol regulatory element-binding protein 1 (SREBP-1) substrate of the intramembrane site-2 protease.[176] Taking into account the latest cryoEM studies of bound C83 and NOTCH100, which found an unfolded TM-C domain around the scissile bonds together with a β-strand downstream the cleavage sites,[87,88] this is a rather unexpected result. Comparing the stability of APP's TM-C to observations from simulations of other γ-secretase substrates (see **Appendix Figures A5-7**), it becomes obvious

that stability of their TM-C domain is a common feature, also shared by apparent non-substrates. However, this indicates that unwinding and accompanying structural rearrangement of TM-C must be triggered by specific interactions with γ-secretase, a result that has also been observed in simulations, using artificially biasing potentials, as well as recent spectroscopic investigations.[89,91,123]

While the TM-N and TM-C domain mainly behave like stable TM helices, the dynamic personality of the whole APP TMD is dominated by its central double-glycine motif ($G_{37}G_{38}$) which interconnects both sub-domains. This highly flexible element, can easily be detected in nearly all parameters as it stands out significantly.[79,82,84] From previous investigations it is known that such a double-glycine motif can act as a flexible hinge.[65] Hinges are defined as flexible regions that permit the rotation of flanking quasi-rigid segments around a screw axis passing through the flexible region. Accordingly, a helix that preserves its H-bonding structure (like for TM-N and TM-C of the APP TMD), provides mechanical hinges at the flexible sites like APP's di-glycine motife.[177] As previously mentioned, APP's di-glycine hinge attracted certain interest as it obviously dominates the flexibility profile of the APP TMD.[84] Since glycine does not feature any side chain groups, it induces a significant packing defect in the helical structure of the TMD, that allows the helix to bend anisotropically over the hinge.[85,118] As TM-helices are significantly stabilized by their (mostly) tightly packed side chains along the backbone,[148] a lack of such interactions consequently leads to an increase in helix dynamics as observed for TM-helices of APP and perfectly reproduced within LV16 model peptides.[79,85,118] Tight packing of side chains does not only stabilize the helix through non-covalent interactions, it further protects the backbone H-bonds from being destabilized by backbone-to-solvent interactions as experimentally observed by NMR.[111] In terms of the APP TMD, the tightly packed TM-C showed nearly no hydration, while the poorly packed TM-N domain showed higher hydration levels, with the highest hydration levels being observed at the di-glycine motif in model membranes,[66,111] as well as in TFE/$H_2O$ mixtures.[66,79,84] As a consequence, the substitution of a single amino acid by glycine already induces a significant increase in flexibility of the TM-helix.[118] The site-specific impact of glycine was investigated in low-complexity LV16 model helix, consisting of eight leucine/valine repeats. Substitution of the bulky leucine by glycine was related to increased dynamics at and around the substitution site. As TM helices generally behave like elastic rods, they naturally feature a hinge in their centre (see supplementary content of Götz & Scharnagl (2018)).[79] By increasing the flexibility upstream or downstream of the centre of the helix in LV16 model peptides, we were able to shift the hinge location along the TM helix while only slightly changing the type of hinge motion.[118]

## 4.2 Motions Coordinated by the di-Glycine Hinge are Modulated by the Environment

During catalytic processing, C99 (see **Figure 1**) is translocated from the membrane environment to the interior of γ-secretase harbouring the water-filled active site cleft. Therefore, studying the response of dynamic properties to the changed environment is of major importance. Unfortunately, a study of the enzyme-bound substrate was not possible in the course of this work. Only recently did cryo-EM determined structures for the bound APP fragment C83 became available,[88] and the first simulation studies of the enzyme-substrate complex were being published after work leading to the present thesis was finished.[91,178] The present study used an implicit approach instead of an explicit description of the native environment of cell membrane, enzyme and functionally relevant water in the active site cavity of γ-secretase.[25,26] An isotropic solvent mixture, consisting of 80% TFE and 20% water (v/v), was used to mimic such an environment. While this non-native environment certainly introduces a bias, it allowed the direct comparison between experimental and simulation results, which showed excellent agreement in most cases.[82,85,113,118] A low amount of water in the hydrophobic, helix-stabilizing TFE matrix[167] mimics the presence of functionally important water molecules in the enzyme's active cavity.[25,26,179] With TFE's dielectric constant being intermediate ($\varepsilon = 8.55$) between that of a dry ($\varepsilon \sim 4$) and that of a solvated ($\varepsilon \sim 12$) protein interior, it reproduced the electrostatic properties in the catalytic cleft of an intramembrane protease. Hence, in combination with 20% water, this solvent mixture presented a rational approach to the active site cleft of γ-secretase.[25,26,79,122,179] Structural and dynamic properties in the TFE/water mixture were compared to those in a POPC model membrane. POPC has been successfully used in experiments and simulations as a reasonable model of a synaptic membrane.[66,108,112,180]

Low helix stability at the APPs di-glycine motif can be observed in all our simulations of the APP TMD, regardless of the environment. However, the extent of stability reduction and hinge motions as well as their type, significantly vary between the investigated POPC membrane environment and the isotropic solvent mixture. However, the environment had a pronounced impact on the site-resolved flexibility profiles while preserving their general shape. The TFE/water environment amplified effects related to low-stability regions, and thus the profiles encoded more entropy as compared to the more flattened profiles in POPC. Similar observations were made for other TMDs that were simulated in TFE/$H_2O$ and POPC for 2μs (**Appendix Figure A4**). Also, global dynamics changed between TFE/$H_2O$ and POPC as large scale motions are restricted by the tightly packed membrane bilayer.[82,85] In particular, the hinge location was more distributed, and the preferred motion was a twisting of TM-C vs. TM-N coordinated by the same hinge as the large-scale bending.

Investigations of the helix orientation in membrane did not reveal any significant difference between the APP WT TMD and the I45T FAD as well G38P and G38L mutants, which were investigated in Götz et al (2019a) and (2019b).[82,85] However, it should be mentioned, that those investigations were

carried out in the POPC model membrane that neither represents the native conditions in a cellular environment nor the membrane properties around the enzyme, which might show membrane thinning or thickening as well as different electrostatic properties. Hence, differences to native properties in the human body are likely to occur, therefore, requiring investigation of these properties in a more native lipid mixture.

## 4.3 The Dynamic Personality of APP's TMD is Unique Among γ-Secretase Substrates.

One of the initial working hypotheses of this thesis was that a dominant hinge motif, controlling bending motions, might discriminate between substrates and non-substrates of γ-secretase.[84] This is supported by the large distance between the putative exosites of γ-secretase and its active aspartate residues,[50,82] as well as the proposed "swing-in" mechanism for substrate posititioning.[112] As such, the ability to undergo structural rearrangement through large scale bending motions is expected to be apparent in the dynamic identities of all γ-secretase substrates, but not in non-substrates. To clarify this question, a series of known and putative γ-secretase substrates and non-substrates (for a full list see **Appendix Table A1**) has been studied in the same way as the APP TMD. Although these results have only been published for NOTCH1 and the apparent non-substrate integrin ß1 (ITGB1) at the time of writing this thesis, they are of interest for this discussion.[123] Therefore, **Appendix Figures A2-8** were added which show the analysis of helicity, H-bond occupancies, MSF, and hinge propensities. By concentrating on H-bond occupancies as valuable reporters for helix flexibility,[79,119] it can be recognized that most of the other TMDs exhibit less dramatic decrease in H-bond occupancies than observed around APP's di-glycine motif, even if they contain several glycine residues in a row (e.g., APLP1, harbouring three consecutive glycine residues). Similar observations were made for the occurrence and location of hinge sites along the TMD (**Appendix Figure A8**) as well as MSF profiles (**Appendix Figure A4**). What can be concluded from these results? Even if they are not analysed in full depth to the present day, lacking analysis of functional motions, the results in the **Appendix** question the original hypothesis that a hinge being as pronounced as in the APP TMD qualifies as a discriminating factor between substrates and non-substrates. This is supported by NMR studies of two other substrates, NOTCH1 (5KZO[67]) and the insulin receptor (2MFR[172]), and the apparent non-substrate ITGB1,[181] which all feature a straight TM helix, similar to those found in the presented simulations.[67,172] This also opens the question about the relevance of APPs di-glycine hinge flexibility as the only requirement for cleavage by γ-secretase. It should be mentioned, that the recently found β-strand downstream the cleavage sites of NOTCH1 and APP in the active cavity of γ-secretase could not be detected in any of the simulations.[87,88] Therefore, it is not possible to make any statement from the present simulations about β-strand formation, especially as it is very likely that observed changes require key interactions between enzyme and substrate that

can only be mimicked by an according model.[123] As mentioned previously, the later one was not available during data generation for this work. However, it would be of major interest if the investigated substrates show the same structural rearrangement within the vicinity of γ-secretase's active aspartates. In addition, it would also be of interested if consecutive cleavage products show the same β-strand formation downstream to the according cleavage sites.

## 4.4 γ-Secretase Cleavage of FADs is Not Determined by Di-Glycine Hinge Dynamics

The initial hypothesis of APP's di-glycine hinge being an integral part for γ-secretase cleavage,[63] was derived from the results of Barret et al (2012),[65] Scharnagl et. al (2014)[84] and Oesterreich et. al (2015).[115] While the first study suggested a kink located at the two glycine residues $Gly_{37}$ and $Gly_{38}$, the second one found the hydrophilic threonine residues in TMD of APP to be essential for the stabilization of the di-glycine hinge, while the third one showed that artificial mutations of these threonine residues (namely $Thr_{43}$ and $Thr_{48}$) to the hydrophobic amino acid valine, led to significantly altered Aβ40 and Aβ42 levels. As shown in Götz & Scharnagl (2018), most FAD mutations that are found natively in the TM-C of the APP TMD[79] and that are known to alter γ-secretase cleavage efficiency and preference without affecting hinge motions.[45] Comparing seven native FAD mutations, no consistent impact on the dynamics of the di-glycine hinge was observed (see **Figure 4**), although variations in the extent and direction of bending can be observed in some of the investigated FADs as recently confirmed by NMR and previously by ETD.[113,114] Even introduction of a further helix stabilizing threonine residues in the FAD mutant I45T only negligibly affected dynamics at the di-glycine hinge.[79,85,113] In case that di-glycine hinge dynamics would be a major discriminating factor, consistent deviations from the WT in the investigated parameters need to be observed, however, as visible in **Figure 4** only some FADs show major differences in their distributions when compared to WT (e.g. T43I and V46I), while others show only very substile changes. Similar observations can be made for more localised parameters as shown in Götz & Scharnagl (2018) and Götz et. al (2019a).[79,85] Insufficient sampling could be excluded to a certain degree for WT and the I45T mutation as even their several μs long ensemble simulations did not reveal any significant differences to the results from the shorter 200ns simulations in Götz & Scharnagl (2018).[79,85] Here it must be mentioned that the chosen ensemble approach, using varying conformations as starting points allows to quickly sample conformations space and partially compensate some of the know limitations of current force fields, still tending towards overstabilization of conformations.[141,182] However, such limitations of current force fields become more of an issue for simulations in the regime of several tens of μs to ms or simulations of intrinsically disordered proteins. Ensemble approaches as used in the present work can partially compensate this as the initial starting points of the ensemble provide the possibility to reach lower populated states much easier as simulations do not only start from highest stability/low energy states. Instead, ensembles used to seed the simulations can include higher energy states, only hardly sampled

in common MD simulations but reachable by enhanced sampling approaches (e.g., replica exchange). Starting from higher energy states increases the probability to reach other minima in the free energy landscape describing a (bio-)molecules conformational variability. Considering the I45T mutations, similar results were recently observed by NMR.[113] This also supports the significance of simulations for other FAD mutations in this publication and led to the conclusion that the dynamics of TM-helices mainly occurs on the ns timescale. This allows to capture a sufficiently large ensemble of conformations with rather short simulation times in the lower µs range. This result also supports the validity of the just 2µs long simulations of various γ-secretase substrates and non-substrates shown in the supplementary materials.

In contrast to results for FAD mutations, artificial mutations at the $G_{38}$ site to either leucine (G38L) or proline (G38P) led to drastically reduced γ-secretase cleavage efficiencies while altering dynamics at the di-glycine hinge.[82] As both location reside close to well-known TMD-TMD interaction interfaces in APP (i.e., $G_{29}XXXG_{33}$, $G_{33}XXXG_{37}$, and $G_{38}XXXA_{42}$)[66,73,108] altered contact preferences with γ-secretase might be a rationale for impaired cleavage of the $G_{38}$ mutants. By using the well-established docking assay for transmembrane components (DAFT),[80,183] which used a coarse-grained description of POPC lipids, water, the $C99_{26–55}$ TMD and γ-secretase, alterations of the initial contact sites of γ-secretase and the $C99_{26–55}$ TMD by these mutations could be excluded.[82] MD simulations as well as solution NMR and ETD experiments consistently agreed on altered helix stability downstream to the $G_{38}$ mutation site. As expected, the G38L mutation showed higher stability at this site. Because of restricted helix flexibility, the hinge is shifted downstream and the direction of bending changes counterclockwise. However, the effects on hinge dynamics were rather small for G38L compared to the dramatic decrease in γ-secretase cleavage efficiencies. The G38P mutation stands in strong contrast to G38L, as it exhibited rather extreme bending angles (>60°) as well as a significant, counterclockwise shift of the bending direction as indicated by swivel angle (Φ) shifted by more than 40°. Such counterclockwise shift in orientation of the scissile bonds in TM-C, relative to a putative binding site in TM-N may be associated with their misdirected presentation to the active site of γ-secretase in case of a “swing-in” event.[82,112]

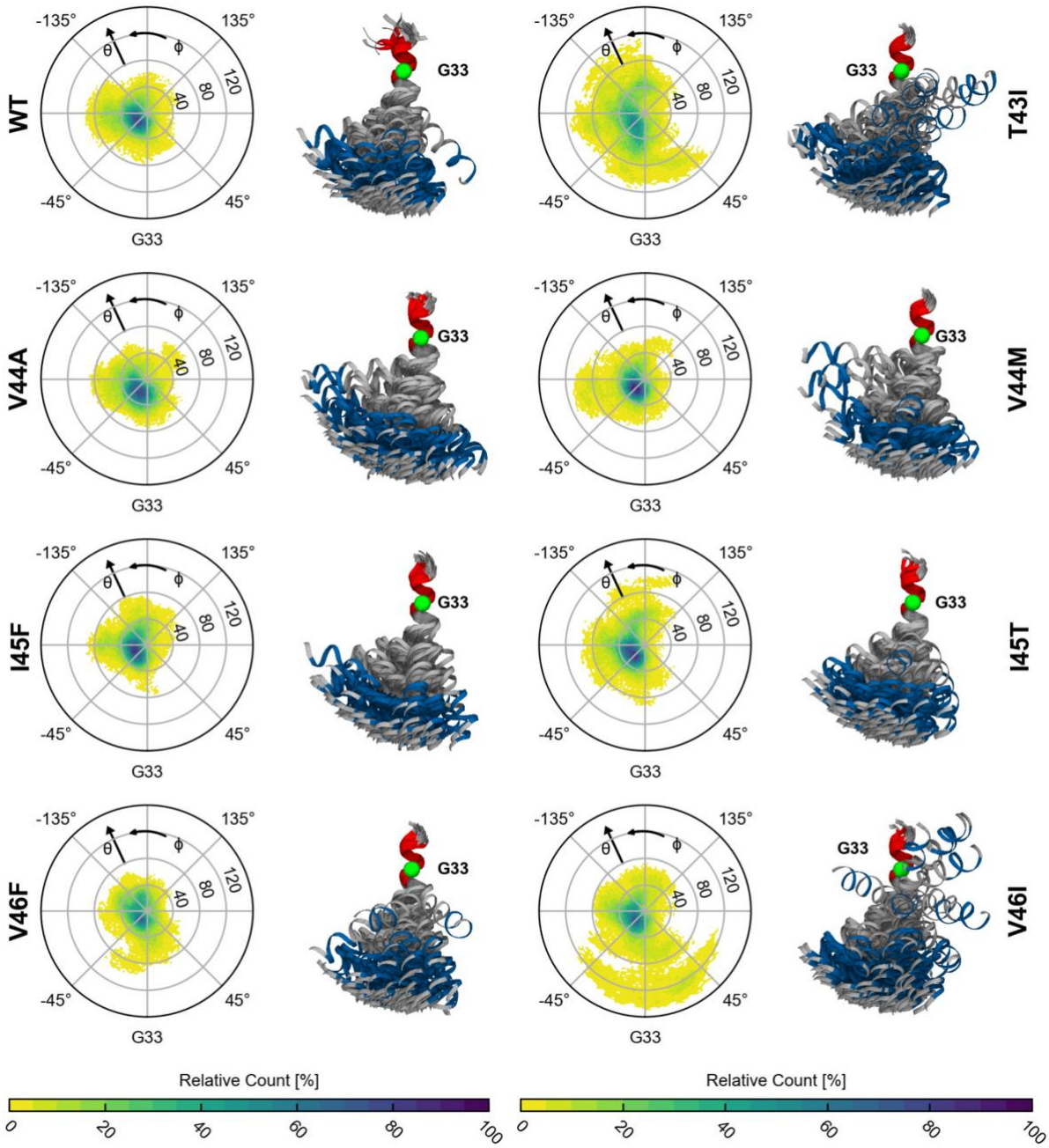


**Figure 4: Orientation of the Cleavage Domain in the APP TMD WT and FAD Mutants.**
Probability density distributions quantify extent and direction of TM helix bending in a TFE/water mixture (80/20% v/v) by bending (Θ) and swivel (Φ) angles as defined in Figure 3. Anisotropic bending is reflected by a swivel angle range restricted between 0° and -135°. Visualizations are obtained by overlaying 150 frames with 1 ns spacing. Color coding as in Figure 3. Figure reprinted from "Dissecting conformational changes in APP's transmembrane domain linked to ε-efficiency in familial Alzheimer's disease" by Götz, A. and Scharnagl, C. 2018, PLoS One, 13(7), e0200077. Licensed under CC BY 4.0, https://doi.org/10.1371/journal.pone.0200077.g006.

As a consequence, the scissile bonds rarely get in close vicinity of the catalytic aspartates in the active cavity of γ-secretase's PSN subunit and, therefore, the enzyme-substrate complex rarely reaches a cleavage competent state.[82] However, this model only accounts for the artificial G38L and G38P mutations and likely for the artificial T43V and T48V mutations.[84] In this context, it must be

noticed that only a single native mutation in the di-glycine hinge of APP, substituting $G_{38}$ by serine, has been described yet. However, this G38S mutation is not associated with FAD but with Parkinson disease.[171] This raises the question of the relevance of results for mutations of $G_{38}$ with regard to FAD mutations and further γ-secretase substrates. While results for artificial mutations in the APP TMD might not be of highest relevance for general γ-secretase substrate processing, they are a significant indicator that "swing-in" of the TM-C domain into the active site of γ-secretase is an essential step for the formation of a productive enzyme substrate complex between γ-secretase and the APP TMD.[63,83] As mentioned in the previous section, this model is challenged by the straight helical TMDs of the γ-secretase substrate NOTCH1 and the insulin receptor, as revealed by NMR.[67,172] A recent paper by Stelzer and Langosch (2019) detected a flexible region in the NOTCH1 TMD solvated in TFE/water (80/20% v/v).[68] This stands in steep contrast to the results of the previous NMR studies of NOTCH1 in micelles.[67] Whether the detected flexible region has hinge functionality, i.e. it is able to correlate motions of flanking helical segments, remains an open question. The simulation results for the NOTCH1 TMD in the same solvent as used in the experiments confirm lower H-bond occupancies in the flexible region but could not detect hinge propensities beyond noise.[123]

Taken together, results for FAD mutants seem to challenge the involvement of the bending motion controlled by the di-glycine in γ-secretase product line preference. This is further supported by missing large scale TMD bending motions in further substrates, as revealed by NMR studies of NOTCH1 and the insulin receptor, [67,172] as well as in our comparative simulation study (see **Figure A8**). However, this implies that the obvious large-scale dynamics of the APP TMD might not be the only determinant for γ-secretase cleavage, raising the question of a lower-amplitude, "hidden" property in the dynamic personality of the APP TMD possibly also present in other substrates and activated upon binding to the enzyme

## 4.5 FAD Mutants Provide a Hint Towards Functional Relevant Dynamics

The previous sections enforced the rejection of models for γ-secretase cleavage of the APP TMD, which are mainly based on properties of the di-glycine hinge. In addition, they also questioned the concept of a general substrate-distinguishing role of large-scale hinge bending in γ-secretase cleavage. However, the question of key elements that control γ-secretase cleavage of the APP TMD still remained open. This required to shift the focus away from the (obvious) di-glycine hinge dynamics, towards more substile characteristics, which were not that easily accessible from standard analysis procedures and the experimental approaches. Comparison of the H-bond dynamics of FAD mutations, as done in Götz & Scharnagl (2018),[79] did not indicate any significant impact at the di-glycine hinge site for FADs except for the T43I one. Instead, they showed consistent alterations in the APP's TM-C domain close to the γ-cleavage site in the APP TMD.[79,85] Solid-state NMR found a

mixture of helical and non-helical conformations in the vicinity of this γ-cleavage site,[174,184] which is consistent with the simulations in Götz & Scharnagl (2018)[79] and Götz et. al (2019a).[85] While differences between FAD mutants were found to be rather subtle in the vicinity of γ-cleavage sites, they had to be put in the context of the correlated H-bond network that stabilizes the TM-helix. The question is which of the lower-amplitude backbone motions is correlated with the H-bond flexibility perturbed by the mutations. However, due to a potential non-linear relationship between changes in the H-bond network and modes of motion of lower amplitude,[153,185] it is difficult to detect such correlations. A first hint provided the analysis of motions controlled by a pair of hinges, which indicated significant changes in their propensity and location, especially in the TM-C domain, when compared to WT.[79,82,85] By using a partial-least-squares method to account for non-linearities[153] and correlating the relative content of affected H-bonds in the investigated FAD mutants (residues $V_{44}$ to $I_{47}$) with modes of motions, it was possible to show that these H-bonds were mainly linked to motions, controlled by a hinge in the TM-C domain and less to motions controlled by the central di-glycine hinge, as detailed in Figures 5 as well as in 8c and d of Götz & Scharnagl (2018).[79,85] Although the original analysis in Götz & Scharnagl (2018)[79] is based on rather short simulation times of ~200 ns, the observed differences in the H-bond network could be perfectly reproduced in the multi μs long ensemble simulations for WT and the I45T mutation.[85] While the large-amplitude motions are the functionally relevant ones in many bio-molecular functions,[155,156,160] NMR techniques also revealed the functional importance of sparsely populated conformations (also referred to as invisible or hidden) for binding and recognition of soluble proteins.[69] Therefore it can be concluded that such lower populated conformations may also exist in the conformational ensemble of the APP TMD and provide insights into substrate positioning (see **Section 4.4**) as well as explain experimentally observed difference in substrate cleavage.

A potential hypothesis might be that such diffuse variations of lower amplitude motions lead to altered cleavage of FAD mutants. Observed perturbations might interfere with steering the cleavage sites into the active site cavity of γ-secretase before transition into a more extended, cleavage competent conformation,[90] therefore leading to altered cleavage efficiencies and pathway preferences. This model is based on the observation that lower amplitude motions might be selectively utilized for optimization and relaxation steps after substrate-enzyme binding, where obvious large-scale bending motions are obstructed by a tightly packed environment.[156,186–188] At the time of the investigation, no structural model of the γ-secretase enzyme-substrate complex was available except a conformation complexed with a co-purified α-helical peptide.[19] In this case, the N-terminal segment of the helix is tightly embraced by the PSN1-NTF, which provides various hydrophobic and polar contacts (for details see Figure. 9A in Götz & Scharnagl (2018)[79]). The functional relevance of this distinct environment was supported by the high amount of PSN1 disease mutations located in this section.[97] Recent computational investigations supported this assumption further.[91] The proposed model for binding of the TM-N can be translated into a simple analytical

model that allows to study the response of the TMD's motions to binding interactions with the enzyme.

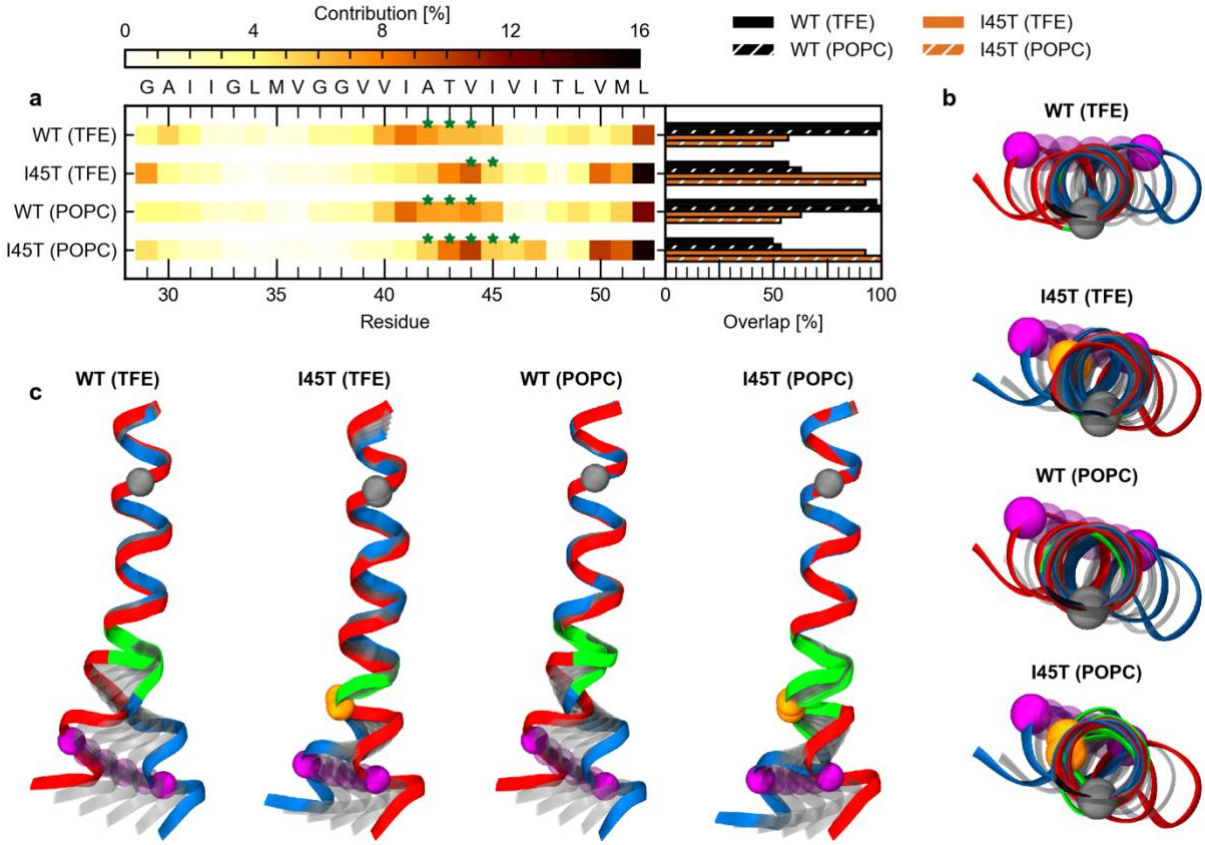


**Figure 5: Functional Mode Analysis of I45T and WT.**
(a) Contribution of each residue to the motion maximally correlated (ewMCM) with occupancy variations of intrahelical amide H-bonds spanning residues V44 - I47 are shown in the left panel. Green stars indicate residues that act as hinges. The right panel quantifies the similarity of the motions in the different sequences and environments by the overlap (inner product) between the corresponding ewMCM vectors. (b,c) Top (b) and front (c) view of motions along the ewMCM vectors interpolated between the maximum (red) and minimum (blue) displacement from the mean structure.. Structures were overlaid onto residues $I_{31}$ – $M_{35}$. Grey spheres represent the Cα atom of $G_{33}$, orange spheres the Cα atom of $T_{45}$ and purple spheres the Cα atom of $L_{49}$. Residues classified as hinge residues (see a) are highlighted in green. Figure reprinted from "Increased H-Bond Stability Relates to Altered ε-Cleavage Efficiency and Aβ Levels in the I45T Familial Alzheimer's Disease Mutant of APP" by Götz, A. Högel, P., Silber, M. et al. 2019, Sci Rep, 9, 5321 e0200077. Licensed under CC BY 4.0, https://www.nature.com/articles/s41598-019-41766-1/figures/4.

Among the FAD mutants, the I45T mutation plays a special role, as it does not exhibit significantly altered ε-site preferences[98] but strongly reduced ε-cleavage efficiency and increased Aβ42/Aβ40 ratio.[45,96,98] This was explained by a switching of cleavage pathways after initial ε-cleavage.[98] The simulations in Götz & Scharnagl (2018)[79] and Götz et. al (2019)[85] found an additional H-bond from the T45 side chain that binds to the $T_{43}$ main chain and stabilizes the helix around the $\zeta$- and $\gamma$-cleavage sites. This back-bonding might inhibit unfolding within the catalytic cleft of presenilin and thus limit its access to the ζ46 scissile bond. As a result, γ-secretase may switch from the ε49-ζ46-γ43-γ40 pathway to the ε48-ζ45-γ42 pathway, explaining the almost complete lack of Aβ43.[45] This hypothesis is supported by experimental results, which showed increased stability at the cleavage site to decrease cleavage efficiency in γ-secretase[130] and even inhibit cleavage in rhomboids.[89]

## 4.6 How Small Dynamic Variations Can Drastically Impact γ-Secretase Cleavage

The results from FMA as well as PRS shifted the focus towards lower amplitude motions in TM-C as major difference between the APP WT and the investigated FAD mutants. From the observed changes a model of how γ-secretase cleavage is affected by changes of substrate TMD dynamics could be derived. It is based on advanced models for enzyme catalysis (e.g. models that combine induced fit and conformational selection) that provide evidence that the intrinsic conformational dynamics of substrates and enzymes plays a key role for recognition and catalytic steps.[156–159] Considering APP TMD dynamics and its perturbations by FAD mutations in the TMD, the overall process leading to a cleavage competent state can roughly be separated in two consecutive steps. Both taking place at the level of substrate transfer from exosites to the active site as well as at the level of substrate fitting into the active site, previous to observed local unfolding and β-strand formation.[87,88] Both steps may impose different requirements on the TMD dynamics and multiple conformational selection steps may play a decisive role on whether cleavage takes place or not.[156] Thus, the importance of rigidity and flexibility and its distribution along the helix backbone can significantly exceed that of local flexibility at the ε-cleavage sites.

During the first step, taking place after substrate encounter at an exosite (see **Figure 1**), large-scale shape fluctuations dominate substrate selection by steering the substrate towards a conformation that allows entry into the enzyme's active site. Hinges provide the necessary bending and twisting flexibility for orienting the reaction partners properly. In the case of C99, this is mainly enabled by the flexible di-glycine hinge. While most FAD mutations only show minor impact on di-glycine hinge dynamics,[113] artificial mutants like G38L and G38P as well as the previously investigated T43I mutation, showed significant impact on the orientation of the ε-cleavage sites and therefore seem to provide the rationale for the observed reduced cleavability.[82,84,113] This is particularly obvious for G38L and G38P as both exhibit counterclockwise shifts in the orientation of the their ε-cleavage sites. However, this does not account for most of the investigated FAD mutants, and nearly all other investigated substrates do not exhibit hinge motions comparable to the C99 TMD (see **Appendix Figure A8** as well as Hitzenberger et. al (2020)[123]).

After binding and positioning, more localized, small amplitude motions become of relevance. However, their analysis requires advanced techniques, since their low amplitude nature lets them vanish within the MD simulation's "background noise". The simulations in Götz & Scharnagl (2018)[79] revealed that the residues $T_{43}V_{44}I_{45}$ upstream of the ε-sites provide additional hinge flexibility. In the rather densely packed environment of the bound state[87,88,91] large-scale bending motions, as coordinated by the $G_{37}G_{38}$ hinge,[79,85] get obstructed. As a direct consequence, those more localized motions, coordinated by the hinge in TM-C are likely to help to optimize and stabilize the enzyme-bound intermediate states (see **Figure 6**). In case of any perturbation of the dynamic

properties of the substrate, the fine-tuned interplay between substrate and enzyme, being necessary to sample productive intermediate states, gets disrupted. Therefore, the propensity for miss-positioning of the ε-sites cleavage sites in the active cavity of γ-secretase gets increased, something that has recently be shown in simulations of the V46I mutant bound to γ-secretase.[199] In the best case, the substrate is simply not cleaved by γ-secretase, leading to a significantly reduced cleavage efficiency. However, faulty positioning can also stabilize any other intermediate state, which relates to cleavage at ε48 instead of ε49, leading to selection of the Aβ42 production line, increasing the amount of Aβ42 released from the cleavage reaction.

As both steps involve dynamic processes, the resulting distribution of conformations translates directly into a diversity of ε-site orientations relative to the catalytic aspartate residues in the catalytic cavity of γ-secretase. As the chemical reaction is thought to be a rare, yet rapid, event that occurs only after sufficient conformational sampling of the enzyme-substrate complex, the process of generating a configuration that is conducive to the chemical reaction seems to be the limiting step in the reaction.[179] According to studies on the kinetics of this reaction, sampling of conformations seems to be an slow process which takes a rather long time as indicated by very long reaction times.[78] While the study can qualitatively be connected with the observed dynamic personality of the C99 TMD, it is not possible to derive whether the first or second step is the limiting one due to the simplistic nature of the Michaelis Menten kinetic model used in the study, which tries to describe the complex kinetic process consisting of multiple steps in just two parameters, lacking necessary resolution. In order to produce a reliable evidence on the rate limiting step from the perspective of a kinetic model, significantly more advanced models as well as high resolution NMR investigations and trapping of intermediates are required.[179]

This is further surrendered by the observation for C83 as well as NOTCH1 to form a β-strand downstream the cleavage sites, when bound in the active cavity of γ-secretase.[87,88] From the current data it is not possible to make any valid statement if β-strand formation is related to observed changes in lower amplitude motions and how this does relate to the kinetics of the cleavage reaction. Instead, it is very likely that β-strand formation is first induced when the cleavage sites get near the catalytic aspartates of γ-secretase due to key interactions between the enzyme and its substrate. While the ability to form a β-strand like conformation might be a requirement for cleavage by γ-secretase and FAD mutations might affect the kinetic of such formation this cannot be concluded from the present data as they investigate the overall backbone dynamics. So, it needs to be pointed out that the present work deals with dynamic processes, being related to positioning of the substrate within the active site of γ-secretase prior to β-strand formation and chemical cleavage.

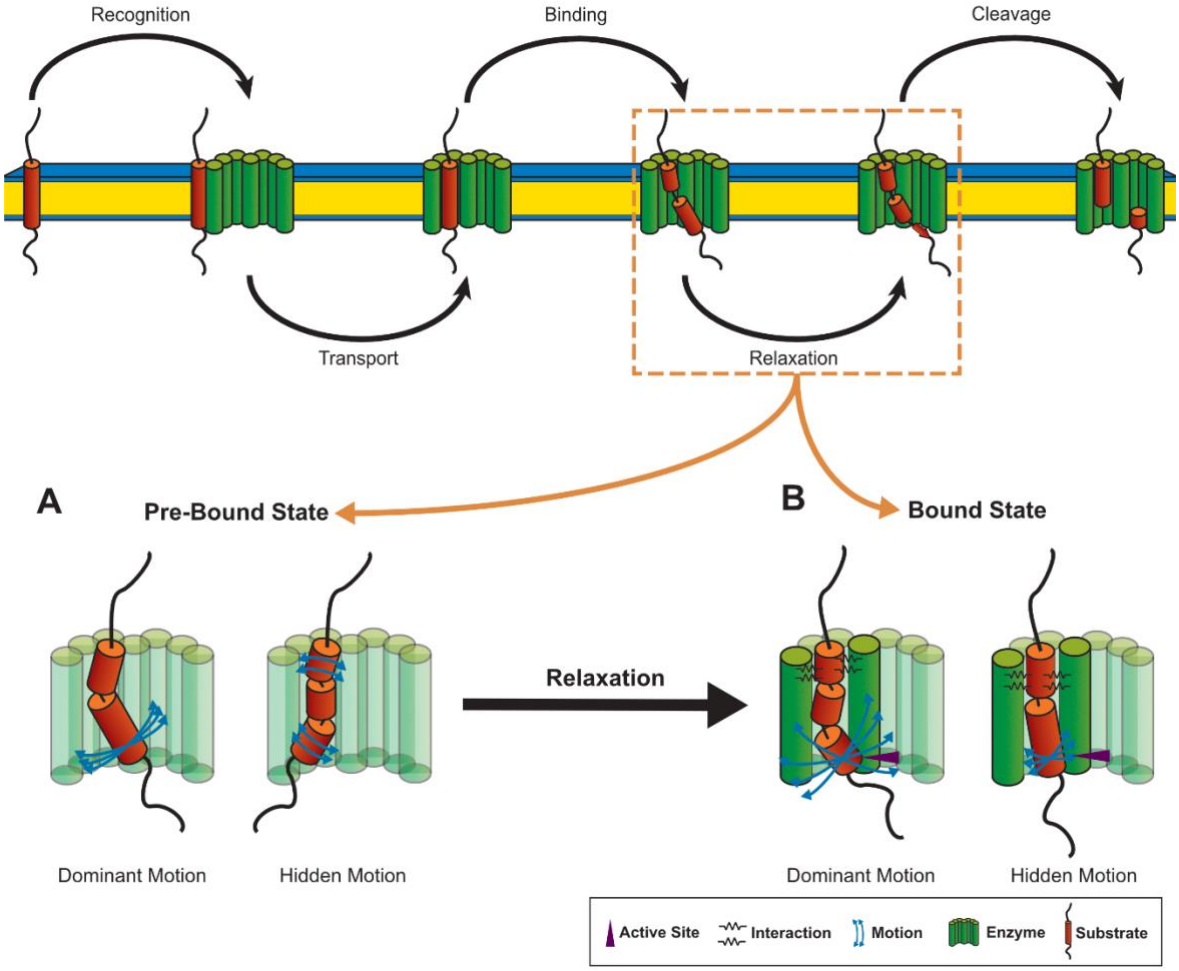


**Figure 6: Reorganization of Conformational Dynamics of the APP TMD as Induced by Binding to γ-secretase.** (A) Large-scale bending coordinated by the di-glycine hinge prevails over low-amplitude motions in the pre-bound state. Hinges as detected in **Figure 5C** and **D** of Götz & Scharnagl (2018)[79] are symbolized by broken connections in the helix (red cylinders). Blue arrows symbolize extent and direction of the bending motion. (B) Binding interactions (represented by black springs) selectively enhance higher order bending at a hinge localized upstream to the ε-sites. Modifying extent of bending and/or shifting hinge location by FAD mutations in TM-C affect the way in which the ε-cleavage sites interact with the enzyme's active site (purple arrow). Figure reprinted and adapted from "Dissecting conformational changes in APP's transmembrane domain linked to ε-efficiency in familial Alzheimer's disease" by Götz, A. and Scharnagl, C. 2018, PLoS One, 13(7), e0200077. Licensed under CC BY 4.0, https://doi.org/10.1371/journal.pone.0200077.g012.

## 4.7 Lessons Learned from APP - What Qualifies a TMD as γ-Secretase Substrate?

While this work mainly focuses on the APP TMD, it gave also new insights on the dynamic properties of several other γ-secretase substrates. While analysis of most of these substrates is at an early stage (only the comparison between APP, NOTCH1 and the apparent non-substrate integrin ß1 was recently published in Hitzenberger et. al (2020)[123]) and in-depth analysis of dynamic properties like functional mode analysis or perturbation-response scanning are missing, the results presented in the **Appendix** indicate both similarities and differences.

(1) For substrates with known cleavage sites, the often-assumed instabilities at the initial cleavage sites cannot be detected, even not in water-containing solvent environment. This makes it likely

that a specific interaction is required to trigger the necessary unfolding upon binding to the active site, as discussed in the previous chapters and indicated by spectroscopic investigations, *in silico* modelling[91] and the recently published structure of APP and NOTCH1 fragments bound to γ-secretase.[88–90]

(2) Second, a common hinge motif in γ-secretase substrate TMDs is missing.[63,84] Furthermore, the APP TMD seems to be an outlier due to its hinge-promoting di-glycine hinge motif, that might be of relevance for initial substrate positioning according to the so called "swing-in" model.[63,64,83] The TM helices behave mainly like elastic rods with low-amplitude bending around the centre, comparable to the behaviour found for the NOTCH1 TMD.[123] This raises the question, whether all substrates encounter the enzyme at the same exosite and are translocated along the same pathway to the active site.[123] Docking at different exosites with varying distances to the active site might place different requirements on the flexibilities of the TMDs. Such a model is supported by recent *in silico* docking simulations revealing different exosite preferences of NOTCH1 and APP.[123]

(3) For almost all of the substrates we detected a pronounced tendency to shift a few H-bonds located close to the membrane interface and approximately one turn downstream to the initial cleavage site (if known) from the α to the $3_{10}$ type, forming a single ß-turn. In the water-containing TFE solvent, local $3_{10}$ H-bond content is greatly enhanced, what might indicate the onset of ß-sheet formation once the substrate TMD contacts the hydrophilic active site of the enzyme. Increased $3_{10}$ helix propensity was also detected as precursor of helix unwinding in resonance Raman spectroscopic investigations.[90] In the apparent non-substrate ITGB1, a stretch of leucine residues protects the C-terminus and stabilizes the α-H-bond from unfolding.[123]

From a very simplified point of view, the consequence of these observations might be that all TMDs which are not part of a "larger" complex (as stable dimers are not cleaved[73]) and provide a sufficiently short ECD are potential substrates to γ-secretase. A potential non-substrate might either (i) not populate a cleavage-competent substrate monomer, (ii) fail to interact with and translocate to the catalytic PSN subunit of γ-secretase, or (iii) resist C-terminal unwinding of their TM helix when exposed to the enzyme's active site.[123]

What would be the consequence of the absence of a motif common for substrates? The most promising and simplest answer to this question would be that γ-secretase does not have a substrate specific role. Instead, the idea of γ-secretase as proteasome of the membrane would be a rationale for the moment.[38] As such γ-secretase would act as the final degrading step in the membrane bilayer after trimming of the substrate was done by other proteins like BACE. However, this is a rather harsh simplification and does not cover observations that different substrates are cleaved with different efficiency. Especially different efficiencies are likely to be related to different dynamic properties of the substrates as they can (1) only less often reach the active site, (2) reaching of a cleavage competent state is much more difficult due to their dynamic properties and (3) they do not unfold in

the proximity of the catalytic aspartates of γ-secretase, therefore, not being cleavable. Recently (3) came into focus due to cryoEM observations showing β-strands to be present for C83 as well as NOTCH1 when located in the active cavity of γ-secretase. However, as pointed out in the previous section of the work, the current simulations are not able to make any assertion about unfolding and β-strand formation due to a lack of the native conditions within the simulations.

# 5 Conclusion

Starting with the intention to proof the hypothesis that the central di-glycine hinge steers cleavage of the APP TMD by γ-secretase and its dynamic properties act as a feature shared by other γ-secretase substrates, the results presented in this work changed this. A comparative analysis of seven FAD mutants revealed that the obvious dynamics of the APP TMD, controlled by its di-glycine hinge, is not the only determinant for successful cleavage by γ-secretase. This partially contradicted previous investigations of artificial T43V and T48V mutants, as well as present results for artificial mutations of the $G_{38}$ residue, namely G38L and G38P. However, especially G38P as well as the T43V mutation (and to some extend the T43I as well) induced significant structural distortions, which might prevent the APP TMD from properly reaching the active cavity. Comparison of the APP TMD dynamics with that of other γ-secretase substrate TMDs revealed that neither a central hinge, nor large scale bending motions are mandatory for cleavage, a finding that is also supported by the few available NMR investigations of other γ-secretase substrates (e.g., NOTCH1 and the insulin receptor). This marks APP as a special case among the γ-secretase substrates in terms of its dynamics. However, the demand for a common large-scale TMD dynamics would only be valid if all substrates are translocated to the active site along the same pathway. If different substrates encounter γ-secretase at different exosites[123] the requirements on TMD dynamics might vary. While large-scale conformational fluctuations may support the cleavage process at the level of substrate transfer from the exosite to the active site, lower amplitude, more localized motion may allow fitting of the substrate's cleavage domain into the active site. This is concluded from a comparison of WT and FAD mutant TMDs. Not the large-scale bending, but more subtle dynamic properties in the TM-C domain of the APP TMD showed to be linked to γ-secretase cleavage of FADs. Such dynamics is controlled by a domain located close to the γ-cleavage sites in the APP TMD, which has already attracted some interest in previous NMR studies due to its tendency towards a $3_{10}$-helical fold. The simulations showed that transitions between α- and $3_{10}$-helical H-bonds occur at higher propensities, a behaviour characteristic for dynamic hinges as shown by LV16 model peptides.[118] The hinges detected in this region, control motions of much lower amplitude as the obvious di-glycine hinge, hence the motions seem to be "hidden" below the dominant large-scale bending. Due the low amplitude nature of these motions, advanced analysis methods like functional mode analysis and perturbation-response scanning must be used to detect the relationship between perturbed H-bond fluctuations and functional relevant mode. The *in-silico* analysis of a series of different binding scenarios clearly pointed to enhancement of motions controlled by hinges in the TM-C domain of the APP WT and its FAD mutants. In summary, the large-scale bending dynamics controlled by APP's di-glycine hinge might be selected to allow coarse positioning of the cleavage domain in an orientation that allows entry into the active site of the enzyme. However, final cleavage is determined by lower-amplitude, more localized motions that help to optimize and stabilize the enzyme-bound intermediate states. While the second one should be general for other γ-secretase substrates, the first

one is specific to the APP TMD, therefore large-scale dynamics does not allow to discriminate substrates of γ-secretase from non-substrates.

# 6 Outlook

The present work tried to uncover the dynamic personality of the APP TMD from a comparison of mutation-induced flexibility changes with the response of endoproteolytic cleavage specificity. Unfortunately, the overall picture of how dynamics relates to cleavage of the TMD by γ-secretase is still incomplete. Recent structural investigations of enzyme-substrate complexes with bound fragments of C99[88,90] and NOTCH1[87] revealed unfolded TM-helices around the initial cleavage sites. These results seem to contradict this and previous work's findings of the initial cleavage site being located in a stable α-helical fold.[65,66,72,84] This conundrum was resolved by the results from spectroscopic[89,90] and in silico[91] investigations revealing that unwinding is induced upon entry of the substrate's cleavage domain into the hydrophilic active-site cleft of presenilin.

The present work's results allowed to construct a rough model of the role played by TMD flexibility in the step of positioning and correct presentation of the cleavage domain to the enzyme. Based on the finding that the TMD flexibility profiles of many substrates largely differ from that of the APP TMD, the "swing-in" model for substrate entry was rejected as a common feature of γ-secretase substrates. This also raised the question whether substrate encounter and following positioning steps proceed along the same pathways for all substrates. Unfortunately, up to the present time, substrate encounter and downstream translocation steps are characterized experimentally only for C99.[50]

Furthermore, the present simulation studies were biased using an unbound TMD including only a few juxtramembrane residues as well as by the artificial solvent environment mimicking the conditions in the active site cleft. Several observations hint towards a major role of the domains flanking the TMD in γ-secretase substrate recognition as: (i) the length of the ECD after shedding impacts cleavage specificity of APP,[200–202] (ii) mutations in APP's ECD negatively modulate γ-secretase cleavage activity,[203,204] and (iii) domain swapping of ECDs between substrates can inhibit cleavage of their TMD.[205,206]

Future work will have to investigate a significantly more complete model of C99 and other substrates in the native environment. Techniques to assemble full-length models of membrane proteins are available.[71,180,207] The conformational space sampled by a large number of substrate TMDs will provide a suitable starting point for homology modelling. Such models, in combination with high resolution models of γ-secretase will allow new insights into the role of dynamics for substrate recognition by γ-secretase.

# 7 References


1. Wolfe, M. S. Intramembrane-cleaving Proteases. *J. Biol. Chem.* **284**, 13969–13973 (2009).

2. Lichtenthaler, S. F., Haass, C. & Steiner, H. Regulated intramembrane proteolysis - lessons from amyloid precursor protein processing: Regulated intramembrane proteolysis and APP processing. *J. Neurochem.* **117**, 779–796 (2011).

3. De Strooper, B., Iwatsubo, T. & Wolfe, M. S. Presenilins and -Secretase: Structure, Function, and Role in Alzheimer Disease. *Cold Spring Harb. Perspect. Med.* **2**, a006304–a006304 (2012).

4. Selkoe, D. J. & Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* **8**, 595–608 (2016).

5. Haass, C., Kaether, C., Thinakaran, G. & Sisodia, S. Trafficking and Proteolytic Processing of APP. *Cold Spring Harb. Perspect. Med.* **2**, a006270–a006270 (2012).

6. Haapasalo, A. & Kovacs, D. M. The Many Substrates of Presenilin/γ-Secretase. *J. Alzheimers Dis.* **25**, 3–28 (2011).

7. Jurisch-Yaksi, N., Sannerud, R. & Annaert, W. A fast growing spectrum of biological functions of γ-secretase in development and disease. *Biochim. Biophys. Acta BBA - Biomembr.* **1828**, 2815–2827 (2013).

8. Güner, G. & Lichtenthaler, S. F. The substrate repertoire of γ-secretase/presenilin. *Semin. Cell Dev. Biol.* **105**, 27–42 (2020).

9. Almén, M., Nordström, K. J., Fredriksson, R. & Schiöth, H. B. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biol.* **7**, 50 (2009).

10. Lemberg, M. K. Sampling the membrane: function of rhomboid-family proteins. *Trends Cell Biol.* **23**, 210–217 (2013).

11. De Strooper, B. *et al.* Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. *Nature* **391**, 387–390 (1998).

12. Rawson, R. B. *et al.* Complementation Cloning of S2P, a Gene Encoding a Putative Metalloprotease Required for Intramembrane Cleavage of SREBPs. *Mol. Cell* **1**, 47–57 (1997).

13. Strisovsky, K. Why cells need intramembrane proteases - a mechanistic perspective. *FEBS J.* **283**, 1837–1845 (2016).

14. Verhelst, S. H. L. Intramembrane proteases as drug targets. *FEBS J.* **284**, 1489–1502 (2017).

15. Ye, J., Dave, U. P., Grishin, N. V., Goldstein, J. L. & Brown, M. S. Asparagine-proline

sequence within membrane-spanning segment of SREBP triggers intramembrane cleavage by Site-2 protease. *Proc. Natl. Acad. Sci.* **97**, 5123–5128 (2000).

16. Steiner, H. *et al.* Glycine 384 is required for presenilin-1 function and is conserved in bacterial polytopic aspartyl proteases. *Nat. Cell Biol.* **2**, 848–851 (2000).

17. Jurisch-Yaksi, N., Sannerud, R. & Annaert, W. A fast growing spectrum of biological functions of γ-secretase in development and disease. *Biochim. Biophys. Acta BBA - Biomembr.* **1828**, 2815–2827 (2013).

18. Wolfe, M. S. Structure, mechanism and inhibition of γ-secretase and presenilin-like proteases. *Biol. Chem.* **391**, (2010).

19. Bai, X., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. Sampling the conformational space of the catalytic subunit of human γ-secretase. *eLife* **4**, (2015).

20. Bai, X. *et al.* An atomic structure of human γ-secretase. *Nature* **525**, 212–217 (2015).

21. Manolaridis, I. *et al.* Mechanism of farnesylated CAAX protein processing by the intramembrane protease Rce1. *Nature* **504**, 301–305 (2013).

22. Urban, S. SnapShot: Cartography of Intramembrane Proteolysis. *Cell* **167**, 1898-1898.e1 (2016).

23. Li, X. *et al.* Structure of a presenilin family intramembrane aspartate protease. *Nature* **493**, 56–61 (2012).

24. Strisovsky, K. Structural and mechanistic principles of intramembrane proteolysis - lessons from rhomboids. *FEBS J.* **280**, 1579–1603 (2013).

25. Tolia, A., Chávez-Gutiérrez, L. & De Strooper, B. Contribution of Presenilin Transmembrane Domains 6 and 7 to a Water-containing Cavity in the γ-Secretase Complex. *J. Biol. Chem.* **281**, 27633–27642 (2006).

26. Sato, C., Morohashi, Y., Tomita, T. & Iwatsubo, T. Structure of the Catalytic Pore of -Secretase Probed by the Accessibility of Substituted Cysteines. *J. Neurosci.* **26**, 12081–12088 (2006).

27. Kuo, I. Y., Hu, J., Ha, Y. & Ehrlich, B. E. Presenilin-like GxGD Membrane Proteases Have Dual Roles as Proteolytic Enzymes and Ion Channels. *J. Biol. Chem.* **290**, 6419–6427 (2015).

28. Aguayo-Ortiz, R. & Dominguez, L. APH-1A Component of γ-Secretase Forms an Internal Water and Ion-Containing Cavity. *ACS Chem. Neurosci.* **10**, 2931–2938 (2019).

29. Dehury, B. & Kepp, K. P. Membrane dynamics of γ-secretase with the anterior pharynx-defective 1B subunit. *J. Cell. Biochem.* **122**, 69–85 (2021).

30. Wang, Y., Zhang, Y. & Ha, Y. Crystal structure of a rhomboid family intramembrane protease. *Nature* **444**, 179–180 (2006).

31. Aguayo-Ortiz, R. & Dominguez, L. Simulating the γ-secretase enzyme: Recent advances and future directions. *Biochimie* **147**, 130–135 (2018).

32. Zoltowska, K. M. & Berezovska, O. Dynamic Nature of presenilin1/γ-Secretase: Implication for Alzheimer's Disease Pathogenesis. *Mol. Neurobiol.* **55**, 2275–2284 (2018).

33. Paschkowsky, S., Oestereich, F. & Munter, L. M. Embedded in the Membrane: How Lipids Confer Activity and Specificity to Intramembrane Proteases. *J. Membr. Biol.* **251**, 369–378 (2018).

34. Lichtenthaler, S. F., Lemberg, M. K. & Fluhrer, R. Proteolytic ectodomain shedding of membrane proteins in mammals—hardware, concepts, and recent developments. *EMBO J.* e99456 (2018) doi:10.15252/embj.201899456.

35. Laurent, S. A. *et al.* γ-secretase directly sheds the survival receptor BCMA from plasma cells. *Nat. Commun.* **6**, (2015).

36. Schauenburg, L. *et al.* APLP1 is endoproteolytically cleaved by γ-secretase without previous ectodomain shedding. *Sci. Rep.* **8**, (2018).

37. Hemming, M. L., Elias, J. E., Gygi, S. P. & Selkoe, D. J. Proteomic Profiling of γ-Secretase Substrates and Mapping of Substrate Requirements. *PLoS Biol.* **6**, e257 (2008).

38. Kopan, R. & Ilagan, Ma. X. G. γ-Secretase: proteasome of the membrane? *Nat. Rev. Mol. Cell Biol.* **5**, 499–504 (2004).

39. Yang, G., Zhou, R. & Shi, Y. Cryo-EM structures of human γ-secretase. *Curr. Opin. Struct. Biol.* **46**, 55–64 (2017).

40. Steiner, H. *et al.* PEN-2 Is an Integral Component of the γ-Secretase Complex Required for Coordinated Expression of Presenilin and Nicastrin. *J. Biol. Chem.* **277**, 39062–39065 (2002).

41. Wolfe, M. S. *et al.* Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and g-secretase activity. **398**, 5 (1999).

42. Fukumori, A., Fluhrer, R., Steiner, H. & Haass, C. Three-Amino Acid Spacing of Presenilin Endoproteolysis Suggests a General Stepwise Cleavage of -Secretase-Mediated Intramembrane Proteolysis. *J. Neurosci.* **30**, 7853–7862 (2010).

43. Pérez-Revuelta, B. I. *et al.* Requirement for small side chain residues within the GxGD-motif of presenilin for γ-secretase substrate cleavage. *J. Neurochem.* **112**, 940–950 (2010).

44. Kretner, B. *et al.* Important functional role of residue x of the presenilin GxGD protease active site motif for APP substrate cleavage specificity and substrate selectivity of γ-secretase. *J.*

*Neurochem.* **125**, 144–156 (2013).

45. Chávez-Gutiérrez, L. *et al.* The mechanism of γ-Secretase dysfunction in familial Alzheimer disease: Mechanisms of Alzheimer disease-causing mutations. *EMBO J.* **31**, 2261–2274 (2012).

46. Wolfe, M. S. Dysfunctional γ-Secretase in Familial Alzheimer's Disease. *Neurochem. Res.* (2018) doi:10.1007/s11064-018-2511-1.

47. Tolia, A. & De Strooper, B. Structure and function of γ-secretase. *Semin. Cell Dev. Biol.* **20**, 211–218 (2009).

48. Ma, G. APH-1a Is the Principal Mammalian APH-1 Isoform Present in -Secretase Complexes during Embryonic Development. *J. Neurosci.* **25**, 192–198 (2005).

49. Chen, A. C., Guo, L. Y., Ostaszewski, B. L., Selkoe, D. J. & LaVoie, M. J. Aph-1 Associates Directly with Full-length and C-terminal Fragments of γ-Secretase Substrates. *J. Biol. Chem.* **285**, 11378–11391 (2010).

50. Fukumori, A. & Steiner, H. Substrate recruitment of γ-secretase and mechanism of clinical presenilin mutations revealed by photoaffinity mapping. *EMBO J.* **35**, 1628–1643 (2016).

51. Luo, W. *et al.* PEN-2 and APH-1 Coordinately Regulate Proteolytic Processing of Presenilin 1. *J. Biol. Chem.* **278**, 7850–7854 (2003).

52. Zhao, G., Liu, Z., Ilagan, Ma. X. G. & Kopan, R. γ-Secretase Composed of PS1/Pen2/Aph1a Can Cleave Notch and Amyloid Precursor Protein in the Absence of Nicastrin. *J. Neurosci.* **30**, 1648–1656 (2010).

53. Audagnotto, M., Kengo Lorkowski, A. & Dal Peraro, M. Recruitment of the amyloid precursor protein by γ-secretase at the synaptic plasma membrane. *Biochem. Biophys. Res. Commun.* **498**, 334–341 (2018).

54. Lee, J. Y., Feng, Z., Xie, X.-Q. & Bahar, I. Allosteric Modulation of Intact γ-Secretase Structural Dynamics. *Biophys. J.* **113**, 2634–2649 (2017).

55. Bolduc, D. M., Montagna, D. R., Gu, Y., Selkoe, D. J. & Wolfe, M. S. Nicastrin functions to sterically hinder γ-secretase–substrate interactions driven by substrate transmembrane domain. *Proc. Natl. Acad. Sci.* **113**, E509–E518 (2016).

56. De Strooper, B. Nicastrin: Gatekeeper of the γ-Secretase Complex. *Cell* **122**, 318–320 (2005).

57. Chávez-Gutiérrez, L. *et al.* $Glu^{332}$ in the Nicastrin Ectodomain Is Essential for γ-Secretase Complex Maturation but Not for Its Activity. *J. Biol. Chem.* **283**, 20096–20105 (2008).

58. Futai, E., Yagishita, S. & Ishiura, S. Nicastrin Is Dispensable for γ-Secretase Protease

Activity in the Presence of Specific Presenilin Mutations. *J. Biol. Chem.* **284**, 13013–13022 (2009).

59. Aguayo-Ortiz, R., Straub, J. E. & Dominguez, L. Influence of membrane lipid composition on the structure and activity of γ-secretase. *Phys. Chem. Chem. Phys.* **20**, 27294–27304 (2018).

60. Winkler, E. *et al.* Generation of Alzheimer Disease-associated Amyloid $\beta_{42/43}$ Peptide by γ-Secretase Can Be Inhibited Directly by Modulation of Membrane Thickness. *J. Biol. Chem.* **287**, 21326–21334 (2012).

61. Yan, Y., Xu, T.-H., Melcher, K. & Xu, H. E. Defining the minimum substrate and charge recognition model of gamma-secretase. *Acta Pharmacol. Sin.* **38**, 1412–1424 (2017).

62. von Heijne, G. Membrane protein structure prediction. *J. Mol. Biol.* **225**, 487–494 (1992).

63. Langosch, D., Scharnagl, C., Steiner, H. & Lemberg, M. K. Understanding intramembrane proteolysis: from protein dynamics to reaction kinetics. *Trends Biochem. Sci.* **40**, 318–327 (2015).

64. Langosch, D. & Steiner, H. Substrate processing in intramembrane proteolysis by γ-secretase – the role of protein dynamics. *Biol. Chem.* **398**, (2017).

65. Barrett, P. J. *et al.* The Amyloid Precursor Protein Has a Flexible Transmembrane Domain and Binds Cholesterol. *Science* **336**, 1168–1171 (2012).

66. Pester, O. *et al.* The Backbone Dynamics of the Amyloid Precursor Protein Transmembrane Helix Provides a Rationale for the Sequential Cleavage Mechanism of γ-Secretase. *J. Am. Chem. Soc.* **135**, 1317–1329 (2013).

67. Deatherage, C. L. *et al.* Structural and biochemical differences between the Notch and the amyloid precursor protein transmembrane domains. *Sci. Adv.* **3**, e1602794 (2017).

68. Stelzer, W. & Langosch, D. Conformationally Flexible Sites within the Transmembrane Helices of Amyloid Precursor Protein and Notch1 Receptor. *Biochemistry* **58**, 3065–3068 (2019).

69. Sekhar, A. & Kay, L. E. NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proc. Natl. Acad. Sci.* **110**, 12867–12874 (2013).

70. van der Kant, R. & Goldstein, L. S. B. Cellular Functions of the Amyloid Precursor Protein from Development to Dementia. *Dev. Cell* **32**, 502–515 (2015).

71. Coninck, D. de, Schmidt, T. H., Schloetel, J.-G. & Lang, T. Packing Density of the Amyloid Precursor Protein in the Cell Membrane. *Biophys. J.* **114**, 1128–1141 (2018).

72. Nadezhdin, K. D., Bocharova, O. V., Bocharov, E. V. & Arseniev, A. S. Structural and dynamic study of the transmembrane domain of the amyloid precursor protein. *Acta Naturae* **3**, 69–76 (2011).

73. Winkler, E., Julius, A., Steiner, H. & Langosch, D. Homodimerization Protects the Amyloid Precursor Protein C99 Fragment from Cleavage by γ-Secretase. *Biochemistry* **54**, 6149–6152 (2015).

74. Beel, A. J. *et al.* Structural Studies of the Transmembrane C-Terminal Domain of the Amyloid Precursor Protein (APP): Does APP Function as a Cholesterol Sensor? † ‡. *Biochemistry* **47**, 9428–9446 (2008).

75. Miyashita, N., Straub, J. E. & Thirumalai, D. Structures of β-Amyloid Peptide 1−40, 1−42, and 1−55—the 672−726 Fragment of APP—in a Membrane Environment with Implications for Interactions with γ-Secretase. *J. Am. Chem. Soc.* **131**, 17843–17852 (2009).

76. Munter, L.-M. *et al.* GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of Aβ42. *EMBO J.* **26**, 1702–1712 (2007).

77. Kaether, C., Haass, C. & Steiner, H. Assembly, Trafficking and Function of γ-Secretase. *Neurodegener. Dis.* **3**, 275–283 (2006).

78. Kamp, F. *et al.* Intramembrane Proteolysis of β-Amyloid Precursor Protein by γ-Secretase Is an Unusually Slow Process. *Biophys. J.* **108**, 1229–1237 (2015).

79. Götz, A. & Scharnagl, C. Dissecting conformational changes in APP's transmembrane domain linked to ε-efficiency in familial Alzheimer's disease. *PLOS ONE* **13**, e0200077 (2018).

80. Audagnotto, M., Kengo Lorkowski, A. & Dal Peraro, M. Recruitment of the amyloid precursor protein by γ-secretase at the synaptic plasma membrane. *Biochem. Biophys. Res. Commun.* **498**, 334–341 (2018).

81. Li, S., Zhang, W. & Han, W. Initial Substrate Binding of γ-Secretase: The Role of Substrate Flexibility. *ACS Chem. Neurosci.* **8**, 1279–1290 (2017).

82. Götz, A. *et al.* Modulating hinge flexibility in the APP transmembrane domain alters γ-secretase cleavage. *Biophys. J.* S0006349519303741 (2019) doi:10.1016/j.bpj.2019.04.030.

83. Tian, G. *et al.* Linear Non-competitive Inhibition of Solubilized Human γ-Secretase by Pepstatin A Methylester, L685458, Sulfonamides, and Benzodiazepines. *J. Biol. Chem.* **277**, 31499–31505 (2002).

84. Scharnagl, C. *et al.* Side-Chain to Main-Chain Hydrogen Bonding Controls the Intrinsic Backbone Dynamics of the Amyloid Precursor Protein Transmembrane Helix. *Biophys. J.* **106**, 1318–1326 (2014).

85. Götz, A. *et al.* Increased H-Bond Stability Relates to Altered ε-Cleavage Efficiency and Aβ Levels in the I45T Familial Alzheimer's Disease Mutant of APP. *Sci. Rep.* **9**, 5321 (2019).

86. Robertson, A. L. *et al.* Protein unfolding is essential for cleavage within the α-helix of a

model protein substrate by the serine protease, thrombin. *Biochimie* **122**, 227–234 (2016).

87. Yang, G. *et al.* Structural basis of Notch recognition by human γ-secretase. *Nature* (2018) doi:10.1038/s41586-018-0813-8.

88. Zhou, R. *et al.* Recognition of the amyloid precursor protein by human γ-secretase. *Science* **363**, eaaw0930 (2019).

89. Brown, M. C. *et al.* Unwinding of the Substrate Transmembrane Helix in Intramembrane Proteolysis. *Biophys. J.* **114**, 1579–1589 (2018).

90. Clemente, N., Abdine, A., Ubarretxena-Belandia, I. & Wang, C. Coupled Transmembrane Substrate Docking and Helical Unwinding in Intramembrane Proteolysis of Amyloid Precursor Protein. *Sci. Rep.* **8**, (2018).

91. Hitzenberger, M. & Zacharias, M. Structural Modeling of γ-Secretase $A\beta_{n}$ Complex Formation and Substrate Processing. *ACS Chem. Neurosci.* **10**, 1826–1840 (2019).

92. Saito, T. *et al.* Potent amyloidogenicity and pathogenicity of Aβ43. *Nat. Neurosci.* **14**, 1023–1032 (2011).

93. Sandebring, A., Welander, H., Winblad, B., Graff, C. & Tjernberg, L. O. The Pathogenic Aβ43 Is Enriched in Familial and Sporadic Alzheimer Disease. *PLoS ONE* **8**, e55847 (2013).

94. Castro, M. A., Hadziselimovic, A. & Sanders, C. R. The vexing complexity of the amyloidogenic pathway. *Protein Sci.* pro.3606 (2019) doi:10.1002/pro.3606.

95. Dimitrov, M. *et al.* Alzheimer's disease mutations in APP but not γ-secretase modulators affect epsilon-cleavage-dependent AICD production. *Nat. Commun.* **4**, (2013).

96. Xu, T.-H. *et al.* Alzheimer's disease-associated mutations increase amyloid precursor protein resistance to γ-secretase cleavage and the Aβ42/Aβ40 ratio. *Cell Discov.* **2**, (2016).

97. Alzforum Mutation Database. https://www.alzforum.org/.

98. Bolduc, D. M., Montagna, D. R., Seghers, M. C., Wolfe, M. S. & Selkoe, D. J. The amyloid-beta forming tripeptide cleavage mechanism of γ-secretase. *eLife* **5**, (2016).

99. Kakuda, N. *et al.* Equimolar Production of Amyloid β-Protein and Amyloid Precursor Protein Intracellular Domain from β-Carboxyl-terminal Fragment by γ-Secretase. *J. Biol. Chem.* **281**, 14776–14786 (2006).

100. Weggen, S. & Beher, D. Molecular consequences of amyloid precursor protein and presenilin mutations causing autosomal-dominant Alzheimer's disease. *Alzheimers Res. Ther.* **4**, 9 (2012).

101. Page, R. M. *et al.* β-Amyloid Precursor Protein Mutants Respond to γ-Secretase Modulators. *J. Biol. Chem.* **285**, 17798–17810 (2010).

102. Olsson, F. *et al.* Characterization of Intermediate Steps in Amyloid Beta (Aβ) Production under Near-native Conditions. *J. Biol. Chem.* **289**, 1540–1550 (2014).

103. Sato, T. *et al.* Potential Link between Amyloid β-Protein 42 and C-terminal Fragment γ 49–99 of β-Amyloid Precursor Protein. *J. Biol. Chem.* **278**, 24294–24301 (2003).

104. Qi-Takahara, Y. Longer Forms of Amyloid Protein: Implications for the Mechanism of Intramembrane Cleavage by -Secretase. *J. Neurosci.* **25**, 436–445 (2005).

105. Takami, M. *et al.* -Secretase: Successive Tripeptide and Tetrapeptide Release from the Transmembrane Domain of -Carboxyl Terminal Fragment. *J. Neurosci.* **29**, 13042–13052 (2009).

106. Matsumura, N. *et al.* γ-Secretase Associated with Lipid Rafts: MULTIPLE INTERACTIVE PATHWAYS IN THE STEPWISE PROCESSING OF β-CARBOXYL-TERMINAL FRAGMENT. *J. Biol. Chem.* **289**, 5109–5121 (2014).

107. Szaruga, M. *et al.* Alzheimer's-Causing Mutations Shift Aβ Length by Destabilizing γ-Secretase-Aβn Interactions. *Cell* **170**, 443-456.e14 (2017).

108. Dominguez, L., Meredith, S. C., Straub, J. E. & Thirumalai, D. Transmembrane Fragment Structures of Amyloid Precursor Protein Depend on Membrane Surface Curvature. *J. Am. Chem. Soc.* **136**, 854–857 (2014).

109. Chen, W. *et al.* Familial Alzheimer's mutations within APPTM increase Aβ42 production by enhancing accessibility of ε-cleavage site. *Nat. Commun.* **5**, (2014).

110. Sato, T. *et al.* A helix-to-coil transition at the -cut site in the transmembrane dimer of the amyloid precursor protein is required for proteolysis. *Proc. Natl. Acad. Sci.* **106**, 1421–1426 (2009).

111. Cao, Z., Hutchison, J. M., Sanders, C. R. & Bowie, J. U. Backbone Hydrogen Bond Strengths Can Vary Widely in Transmembrane Helices. *J. Am. Chem. Soc.* **139**, 10742–10749 (2017).

112. Lemmin, T., Dimitrov, M., Fraering, P. C. & Dal Peraro, M. Perturbations of the Straight Transmembrane α-Helical Structure of the Amyloid Precursor Protein Affect Its Processing by γ-Secretase. *J. Biol. Chem.* **289**, 6763–6774 (2014).

113. Silber, M., Hitzenberger, M., Zacharias, M. & Muhle-Goll, C. Altered Hinge Conformations in APP Transmembrane Helix Mutants May Affect Enzyme–Substrate Interactions of γ-Secretase. *ACS Chem. Neurosci.* **11**, 4426–4433 (2020).

114. Stelzer, W., Scharnagl, C., Leurs, U., Rand, K. D. & Langosch, D. The Impact of the 'Austrian' Mutation of the Amyloid Precursor Protein Transmembrane Helix is Communicated to

the Hinge Region. *ChemistrySelect* **1**, 4408–4412 (2016).

115. Oestereich, F. *et al.* Impact of Amyloid Precursor Protein Hydrophilic Transmembrane Residues on Amyloid-Beta Generation. *Biochemistry* **54**, 2777–2784 (2015).

116. Perrin, F. *et al.* Dimeric Transmembrane Orientations of APP/C99 Regulate γ-Secretase Processing Line Impacting Signaling and Oligomerization. *iScience* **23**, 101887 (2020).

117. Goetz, A. *et al.* Stabilization / destabilization of the APP transmembrane domain by mutations in the di-glycine hinge alter helical structure and dynamics, and impair cleavage by γ-secretase. (2018) doi:10.1101/375006.

118. Högel, P. *et al.* Glycine Perturbs Local and Global Conformational Flexibility of a Transmembrane Helix. *Biochemistry* (2018) doi:10.1021/acs.biochem.7b01197.

119. Götz, A. *et al.* Increased H-Bond Stability Relates to Altered ε-Cleavage Efficiency and Aβ Levels in the I45T Familial Alzheimer's Disease Mutant of APP. *Sci. Rep.* **9**, 5321 (2019).

120. Pester, O., Götz, A., Multhaup, G., Scharnagl, C. & Langosch, D. The Cleavage Domain of the Amyloid Precursor Protein Transmembrane Helix Does Not Exhibit Above-Average Backbone Dynamics. *ChemBioChem* **14**, 1943–1948 (2013).

121. Schutz, C. N. & Warshel, A. What Are the Dielectric 'Constants' of Proteins and How To Validate Electrostatic Models? *Proteins Struct. Funct. Genet.* **44**, 400–417 (2001).

122. Buck, M. Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* **31**, 297–355 (1998).

123. Hitzenberger, M. *et al.* The dynamics of γ-secretase and its substrates. *Semin. Cell Dev. Biol.* S108495211830274X (2020) doi:10.1016/j.semcdb.2020.04.008.

124. Oliphant, T. E. Python for Scientific Computing. *Comput. Sci. Eng.* **9**, 10–20 (2007).

125. SciPy 1.0 Contributors *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

126. McGibbon, R. T. *et al.* MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **109**, 1528–1532 (2015).

127. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

128. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

129. Yin, Y. I. *et al.* γ-Secretase Substrate Concentration Modulates the Aβ42/Aβ40 Ratio: *IMPLICATIONS FOR ALZHEIMER DISEASE*. *J. Biol. Chem.* **282**, 23639–23644 (2007).

130. Fernandez, M. A. *et al.* Transmembrane Substrate Determinants for γ-Secretase Processing of APP CTFβ. *Biochemistry* **55**, 5675–5688 (2016).

131. Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

132. Page, R. C., Kim, S. & Cross, T. A. Transmembrane Helix Uniformity Examined by Spectral Mapping of Torsion Angles. *Structure* **16**, 787–797 (2008).

133. Brooks, B. R. *et al.* CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).

134. Shapovalov, M. V. & Dunbrack, R. L. A Smoothed Backbone-Dependent Rotamer Library for Proteins Derived from Adaptive Kernel Density Estimates and Regressions. *Structure* **19**, 844–858 (2011).

135. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).

136. Kannan, S. & Zacharias, M. Simulated annealing coupled replica exchange molecular dynamics—An efficient conformational sampling method. *J. Struct. Biol.* **166**, 288–294 (2009).

137. Frey, B. J. & Dueck, D. Clustering by Passing Messages Between Data Points. *Science* **315**, 972–976 (2007).

138. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Mach. Learn. PYTHON* 6.

139. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).

140. Best, R. B. *et al.* Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain $\chi_1$ and $\chi_2$ Dihedral Angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).

141. Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).

142. Jamitzky, Ferdinand. *redisexec*. (Leibniz Supercomputing Centre).

143. Lee, J. *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

144. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–802 (2005).

145. Bugge, K., Lindorff-Larsen, K. & Kragelund, B. B. Understanding single-pass transmembrane receptor signaling from a structural viewpoint-what are we missing? *FEBS J.* **283**, 4424–4451 (2016).

146. Kabsch, W. & Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).

147. Grossfield, A., Feller, S. E. & Pitman, M. C. Convergence of molecular dynamics simulations of membrane proteins. *Proteins Struct. Funct. Bioinforma.* **67**, 31–40 (2007).

148. Quint, S. *et al.* Residue-Specific Side-Chain Packing Determines the Backbone Dynamics of Transmembrane Model Helices. *Biophys. J.* **99**, 2541–2549 (2010).

149. Grossfield, A., Feller, S. E. & Pitman, M. C. A role for direct interactions in the modulation of rhodopsin by -3 polyunsaturated lipids. *Proc. Natl. Acad. Sci.* **103**, 4888–4893 (2006).

150. Guo, Z., Kraka, E. & Cremer, D. Description of local and global shape properties of protein helices. *J. Mol. Model.* **19**, 2901–2911 (2013).

151. Strandberg, E., Esteban-Martín, S., Salgado, J. & Ulrich, A. S. Orientation and Dynamics of Peptides in Membranes Calculated from 2H-NMR Data. *Biophys. J.* **96**, 3223–3232 (2009).

152. Hayward, S. & Lee, R. A. Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graph. Model.* **21**, 181–183 (2002).

153. Krivobokova, T., Briones, R., Hub, J. S., Munk, A. & de Groot, B. L. Partial Least-Squares Functional Mode Analysis: Application to the Membrane Proteins AQP1, Aqy1, and CLC-ec1. *Biophys. J.* **103**, 786–796 (2012).

154. DiCiccio, T. J. *et al.* Better Bootstrap Confidence Intervals. *Stat. Sci.* **11**, 189–228 (1996).

155. Haliloglu, T. & Bahar, I. Adaptability of protein structures to enable functional interactions and evolutionary implications. *Curr. Opin. Struct. Biol.* **35**, 17–23 (2015).

156. Ma, B. & Nussinov, R. Enzyme dynamics point to stepwise conformational selection in catalysis. *Curr. Opin. Chem. Biol.* **14**, 652–659 (2010).

157. Henzler-Wildman, K. A. *et al.* A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913–916 (2007).

158. Nashine, V. C., Hammes-Schiffer, S. & Benkovic, S. J. Coupled motions in enzyme catalysis. *Curr. Opin. Chem. Biol.* **14**, 644–651 (2010).

159. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).

160. Wei, G., Xi, W., Nussinov, R. & Ma, B. Protein Ensembles: How Does Nature Harness Thermodynamic Fluctuations for Life? The Diverse Functional Roles of Conformational Ensembles in the Cell. *Chem. Rev.* **116**, 6516–6551 (2016).

161. Hall, S. E., Roberts, K. & Vaidehi, N. Position of helical kinks in membrane protein crystal structures and the accuracy of computational prediction. *J. Mol. Graph. Model.* **27**, 944–950 (2009).

162. Langelaan, D. N., Wieczorek, M., Blouin, C. & Rainey, J. K. Improved Helix and Kink Characterization in Membrane Proteins Allows Evaluation of Kink Sequence Predictors. *J. Chem. Inf. Model.* **50**, 2213–2220 (2010).

163. Meruelo, A. D., Samish, I. & Bowie, J. U. TMKink: A method to predict transmembrane helix kinks. *Protein Sci.* **20**, 1256–1264 (2011).

164. Wilman, H. R., Shi, J. & Deane, C. M. Helix kinks are equally prevalent in soluble and membrane proteins. *Proteins Struct. Funct. Bioinforma.* **82**, 1960–1970 (2014).

165. Mai, T.-L. & Chen, C.-M. Computational prediction of kink properties of helices in membrane proteins. *J. Comput. Aided Mol. Des.* **28**, 99–109 (2014).

166. Selkoe, D. J. Light at the End of the Amyloid Tunnel: Published as part of the *Biochemistry* series "Biochemistry to Bedside". *Biochemistry* **57**, 5921–5922 (2018).

167. Roccatano, D., Colombo, G., Fioroni, M. & Mark, A. E. Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. *Proc. Natl. Acad. Sci.* **99**, 12179–12184 (2002).

168. Gente, G. Water–Trifluoroethanol Mixtures: Some Physicochemical Properties. 14.

169. Cao, Z. & Bowie, J. U. Shifting hydrogen bonds may produce flexible transmembrane helices. *Proc. Natl. Acad. Sci.* **109**, 8121–8126 (2012).

170. Li, S., Zhang, W. & Han, W. Initial Substrate Binding of γ-Secretase: The Role of Substrate Flexibility. *ACS Chem. Neurosci.* **8**, 1279–1290 (2017).

171. Schulte, E. C. *et al.* Rare variants in β-Amyloid precursor protein (APP) and Parkinson's disease. *Eur. J. Hum. Genet.* **23**, 1328–1333 (2015).

172. Li, Q., Wong, Y. L. & Kang, C. Solution structure of the transmembrane domain of the insulin receptor in detergent micelles. *Biochim. Biophys. Acta BBA - Biomembr.* **1838**, 1313–1321 (2014).

173. Richter, L. *et al.* Amyloid beta 42 peptide (A 42)-lowering compounds directly bind to A and interfere with amyloid precursor protein (APP) transmembrane dimerization. *Proc. Natl. Acad. Sci.* **107**, 14597–14602 (2010).

174. Itkin, A. *et al.* Structural Characterization of the Amyloid Precursor Protein Transmembrane Domain and Its γ-Cleavage Site. *ACS Omega* **2**, 6525–6534 (2017).

175. Jung, J. I. *et al.* Independent Relationship between Amyloid Precursor Protein (APP) Dimerization and γ-Secretase Processivity. *PLoS ONE* **9**, e111553 (2014).

176. Linser, R. *et al.* The membrane anchor of the transcriptional activator SREBP is characterized by intrinsic conformational flexibility. *Proc. Natl. Acad. Sci.* **112**, 12390–12395 (2015).

177. Hayward, S. Structural principles governing domain motions in proteins. 11.

178. Dehury, B., Tang, N., Mehra, R., Blundell, T. L. & Kepp, K. P. Side-by-side comparison of Notch- and C83 binding to γ-secretase in a complete membrane model at physiological temperature. *RSC Adv.* **10**, 31215–31232 (2020).

179. Warshel, A. & Bora, R. P. Perspective: Defining and quantifying the role of dynamics in enzyme catalysis. *J. Chem. Phys.* **144**, 180901 (2016).

180. Pantelopulos, G. A., Straub, J. E., Thirumalai, D. & Sugita, Y. Structure of APP-C99 1–99 and implications for role of extra-membrane domains in function and oligomerization. *Biochim. Biophys. Acta BBA - Biomembr.* (2018) doi:10.1016/j.bbamem.2018.04.002.

181. Lu, Z. *et al.* Implications of the differing roles of the β1 and β3 transmembrane and cytoplasmic domains for integrin function. *eLife* **5**, e18633 (2016).

182. Huang, J. *et al.* CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).

183. Wassenaar, T. A. *et al.* High-Throughput Simulations of Dimer and Trimer Assembly of Membrane Proteins. The DAFT Approach. *J. Chem. Theory Comput.* **11**, 2278–2291 (2015).

184. Lu, J.-X., Yau, W.-M. & Tycko, R. Evidence from Solid-State NMR for Nonhelical Conformations in the Transmembrane Domain of the Amyloid Precursor Protein. *Biophys. J.* **100**, 711–719 (2011).

185. Hub, J. S. & de Groot, B. L. Detection of Functional Modes in Protein Dynamics. *PLoS Comput. Biol.* **5**, e1000480 (2009).

186. Marcos, E., Crehuet, R. & Bahar, I. Changes in Dynamics upon Oligomerization Regulate Substrate Binding and Allostery in Amino Acid Kinase Family Members. *PLoS Comput. Biol.* **7**, e1002201 (2011).

187. Lezon, T. R. & Bahar, I. Using Entropy Maximization to Understand the Determinants of Structural Dynamics beyond Native Contact Topology. *PLoS Comput. Biol.* **6**, e1000816 (2010).

188. Zen, A., Micheletti, C., Keskin, O. & Nussinov, R. Comparing interfacial dynamics in protein-protein complexes: an elastic network approach. *BMC Struct. Biol.* **10**, 26 (2010).

189. Guarnera, E. & Berezovsky, I. N. Structure-Based Statistical Mechanical Model Accounts for the Causality and Energetics of Allosteric Communication. *PLOS Comput. Biol.* **12**, e1004678 (2016).

190. Ettayapuram Ramaprasad, A. S., Uddin, S., Casas-Finet, J. & Jacobs, D. J. Decomposing Dynamical Couplings in Mutated scFv Antibody Fragments into Stabilizing and Destabilizing Effects. *J. Am. Chem. Soc.* **139**, 17508–17517 (2017).

191. Zheng, W. & Brooks, B. R. Probing the Local Dynamics of Nucleotide-Binding Pocket Coupled to the Global Dynamics: Myosin versus Kinesin. *Biophys. J.* **89**, 167–178 (2005).

192. Zheng, W., Liao, J.-C., Brooks, B. R. & Doniach, S. Toward the mechanism of dynamical couplings and translocation in hepatitis C virus NS3 helicase using elastic network model. *Proteins Struct. Funct. Bioinforma.* **67**, 886–896 (2007).

193. Ming, D. & Wall, M. E. Interactions in Native Binding Sites Cause a Large Change in Protein Dynamics. *J. Mol. Biol.* **358**, 213–223 (2006).

194. Erman, B. The Gaussian Network Model: Precise Predictions of Residue Fluctuations and Application to Binding Problems. *Biophys. J.* **91**, 3589–3599 (2006).

195. Atilgan, A. R., Aykut, A. O. & Atilgan, C. Subtle *p* H differences trigger single residue motions for moderating conformations of calmodulin. *J. Chem. Phys.* **135**, 155102 (2011).

196. Ikeguchi, M., Ueno, J., Sato, M. & Kidera, A. Protein Structural Change Upon Ligand Binding: Linear Response Theory. *Phys. Rev. Lett.* **94**, 078102 (2005).

197. Fuchigami, S., Fujisaki, H., Matsunaga, Y. & Kidera, A. Protein Functional Motions: Basic Concepts and Computational Methodologies. in *Advances in Chemical Physics* (eds. Komatsuzaki, T., Berry, R. S. & Leitner, D. M.) 35–82 (John Wiley & Sons, Inc., 2011). doi:10.1002/9781118087817.ch2.

198. Echave, J. & Fernández, F. M. A perturbative view of protein structural variation. *Proteins Struct. Funct. Bioinforma.* **78**, 173–180 (2010).

199. Dehury, B., Somavarapu, A. K. & Kepp, K. P. A computer-simulated mechanism of familial Alzheimer's disease: Mutations enhance thermal dynamics and favor looser substrate-binding to γ-secretase. *J. Struct. Biol.* **212**, 107648 (2020).

200. Petit, D. *et al.* Extracellular interface between APP and Nicastrin regulates Aβ length and response to γ-secretase modulators. *EMBO J.* **38**, (2019).

201. Funamoto, S. *et al.* Substrate ectodomain is critical for substrate preference and inhibition of γ-secretase. *Nat. Commun.* **4**, 2529 (2013).

202. Siegel, G. *et al.* The Alzheimer's Disease γ-Secretase Generates Higher 42:40 Ratios for β-Amyloid Than for p3 Peptides. *Cell Rep.* **19**, 1967–1976 (2017).

203. Tang, T.-C. *et al.* Conformational Changes Induced by the A21G Flemish Mutation in the Amyloid Precursor Protein Lead to Increased Aβ Production. *Structure* **22**, 387–396 (2014).

204. Tian, Y., Bassit, B., Chau, D. & Li, Y.-M. An APP inhibitory domain containing the Flemish mutation residue modulates γ-secretase activity for Aβ production. *Nat. Struct. Mol. Biol.* **17**, 151–158 (2010).

205. van Tetering, G. & Vooijs, M. Proteolytic Cleavage of Notch: 'HIT and RUN'. *Curr. Mol. Med.* **11**, 255–269 (2011).

206. Wilhelmsen, K. & van der Geer, P. Phorbol 12-Myristate 13-Acetate-Induced Release of the Colony-Stimulating Factor 1 Receptor Cytoplasmic Domain into the Cytosol Involves Two Separate Cleavage Events. *Mol. Cell. Biol.* **24**, 454–464 (2004).

207. Leman, J. K. & Bonneau, R. A Novel Domain Assembly Routine for Creating Full-Length Models of Membrane Proteins from Known Domain Structures. *Biochemistry* (2017) doi:10.1021/acs.biochem.7b00995.

# Appendix

This supplementary material contains data for 30 known/putative (non-)substrates of γ-secretase, according to Hemming et. al 2008[37], Haapasalo and Kovacs 2011[6] and Güner and Lichtenthaler (2020).[8] While a sub-set was recently published (see Hitzenberger et.al. 2020[123]), the vast majority is not. The main results are collected in chapter 3.5 in the thesis. Python source codes and jupyter notebooks to run the analysis and create the figures can be found at https://github.com/ag1989/gsec-substrates (publicly available with publication of this thesis.)

## A1. Investigated Proteins, Peptide Design and MD simulations

Candidates were selected from a pool of known (non-)substrates as described in Hemming et. al 2008[37] and Haapasalo and Kovacs 2011[6]. Selection was either made according to similarity to APP (APLP1, APLP2), already existing experimental investigations (Notch1, Notch2), occurrence of poly-glycine motifs like in the APP TMD (e.g., Ecadherine or Erbb4), other characteristic motifs like poly-leucine, or according to experimental plans by collaboration partners. To allow comparisons, only the TMD subsection of each candidate was investigated. Therefore, slices of 35 amino acids that cover the TMD as well as a certain amount of N- and C-terminal residues were designed. Juxtamembrane domains were selected according to charge distribution between the N- and C-terminus to allow proper placement in the membrane bilayer. All sequences, their corresponding UniProt IDs, as well as the TMD domain as annotated in UniProt are collected in Table A1. Initial start configurations were generated as described in section 2.2 of the thesis for single conformations. Single run simulations in TFE/water (80%/20% v/v) and POPC were conducted as described in (2.3.1 and 2.3.2) in the main thesis without changes. The simulation time for each run was 2 µs, the last 1 µs were subjected to analysis. Due to the charged glutamate residue in the centre to the DAB12 TMD, the peptide unfolded in TFE/water and this run was not analysed further. The analysis tools are used as describe in (2.4 – 2.7) of the main text. The only exception concerns the definition of intrahelical H-bonds, with are carboxyl-based in the subsequent result, i.e., α H-bonds extend from carboxyl O(i) to amide hydrogen HN (i+4), while $3_{10}$ H-bonds spans O(i) to HN(i+3).

## A 2. Determination of Helicity and the Membrane Spanning Helix (TM Helix)

The TM Helix is defined as the helix that spans the membrane bilayer. Therefore, it is determined by the helicity of individual residues along the investigated peptide. A residue's helicity is computed as the amount of times the residues is detected to be in helical conformation by the simplified DSSP algorithm as implemented in MDtraj[126], during a certain time window. The TM Helix is then defined as the first and last residues along the investigated peptide that shows a helicity > 95% during the investigated simulation time.

### A.3 Determination of a Residue's z-Position in the Membrane Bilayer

z-Positions were determined relative to the membrane bilayers centre. Therefore, lipid phosphate atoms as well as Cα atoms of the according peptides were extracted from the simulations. A reference coordinate system is defined with its origin at the center of the lipid phosphate atoms. After translating the atom coordinates to the center of the coordinate system, the normal of the plane described by the phosphate atoms is computed and the system is rotated to align the normal with the z-axis of the coordinate system. z-positions are then calculated as the z-shift relative to origin of the coordinate system.

**Table A1: Proteins Investigated by Molecular Dynamics Simulations:** The table shows the name of the investigated proteins, their shortcut used within the results, their corresponding UniProt ID and the sequence used for the simulations. Residues highlighted in green represent the TMD as annotated within the UniProt database

| Protein | Shortcut | UniProt ID | Sequence |
|---|---|---|---|
| Amyloid-Like Protein 1 | APLP1 | P51693 | TGVSREAVSGLLIMGAGGGSLIVLSMLLLRRKKPY |
| Amyloid-Like Protein 1 | APLP2 | Q06481 | DFSLSSSALIGLLVIAVAIATVIVISLVMLRKRQY |
| Amyloid-β Precursor Protein | APP | P05067 | AEDVGSNKGAIIGLMVGGVVIATVIVITLVMLKKK |
| B-Cell Maturation Protein | BCMA | Q02223 | VKGTNAILWTCLGLSLIISLAVFVLMFLLRKINSE |
| CD44 Antigen | CD44 | P16070 | RTPQIPEWLIILASLLALALILAVCIAVNSRRRCG |
| Alcadein-α | CLSTN1 | Q94985 | FAVVPSTATVVIVVCVSFLVFMIILGVFRIRAAHR |
| Alcadein-γ | CLSTN2 | Q9H4D0 | QHSSVVPSIATVVIIISVCMLVFVVAMGVYRVRIA |
| Macrophage Colony-Stimulating Factor 1 Receptor | CSF1R | P07333 | DEFLFTPVVVACMSIMALLLLLLLLLLYKYKQKPK |
| Tyro Protein Kinase-Binding Protein | DAP12 | Q43914 | SCSTVSPGVLAGIVMGDLVLTVLIALAVYFLGRLV |
| Nectrin Receptor DCC | DCC | P43146 | QKNSNLLVIIVVTVGVITVLVVVIVAVICTRRSSA |
| Dystroglycan | DG | Q14118 | SEDDVYLHTVIPAVVVAAILLIAGIIAMICYRKKR |
| Delta & Notch-like EGF-Related Receptor | DNER | Q8NFT8 | MPRHSLYIIIGALCVAFILMLIILIVGICRISRIE |
| Desmoglein-2 | DSG2 | Q14126 | DSYVGLGPAAIALMILAFLLLLLVPLLLLMCHCGK |
| Cadherin-1 | ECADHERIN | P12830 | AGLQIPAILGILGGILALLILILLLLLFLRRRAVV |
| Receptor Tyrosine-Protein Kinase | ERBB4 | Q15303 | QHARTPLIAAGVIGGLFILVIVGLTFAVYVRRKSI |

| | | | |
|---|---|---|---|
| Glycophorin A | GPA | P02724 | HFSEPEITLIIFGVMAGVIGTILLISYGIRRLIKK |
| HLA Class 1 Histocompatible Antigen α | HLA-A | P01891 | PTIPIVGIIAGLVLFGAVITGAVVAAVMWRRKSSD |
| Integrin β 1 | ITGB1 | P05556 | CPTGPDIIPIVAGVVAGIVLIGLALLLIWKLLMII |
| Protein Jagged 2 | JAGGED2 | Q9Y219 | VTGGSSTGLLVPVLCGAFSVLWLACVVLCVWWTRK |
| Low-Density Lipoprotein Receptor | LDLR | P01130 | KPSSVRALSIVLPIVLLVFLCLGVFLLWKNWRLKN |
| Cadherin-2 | NCADHERIN | P19022 | AGLGTGAIIAILLCIIILLILVLMFVVWMKRRDKE |
| Neurogenic Locus Notch Homolog Protein 1 | NOTCH1 | P46531 | PAQLHFMYVAAAAFVLLFFVGCGVLLSRKRRRQHG |
| Neurogenic Locus Notch Homolog Protein 2 | NOTCH2 | Q04721 | LTPERTQLLYLLAVAVVIILFIILLGVIMAKRKRK |
| Atrial natriuretic peptide receptor A | NPRA | P16066 | DHLSTLEVLALVGSLSLLGILIVSFFIYRKMQLEK |
| Atrial natriuretic peptide receptor 3 | NPRC | P17342 | GGLEESAVTGIVVGALLGAGLLMAFYFFRKKYRIT |
| Plexin domain-containing protein 2 | PLXDC2 | Q6UX71 | KGGTLHAGLIIGILILVLIVATAILVTVYMYHHPT |
| Syndecan-1 | SDC1 | P18827 | DRKEVLGGVIAGGLVGLIFAVCLVGFMLYRMKKKD |
| Syndecan-2 | SDC2 | P34741 | LFKRTEVLAAVIAGGVIGFLFAIFLILLLVYRMRK |
| Tumor necrosis factor receptor superfamily member 12A | TNR12 | Q9NP84 | APFRLLWPILGGALSLTFVLGLLSGFLVWRRCRRR |
| Vasorin | VASN | Q6EMK4 | QAREGNLPLLIAPALAAVLLAALAAVGAAYCVRRG |

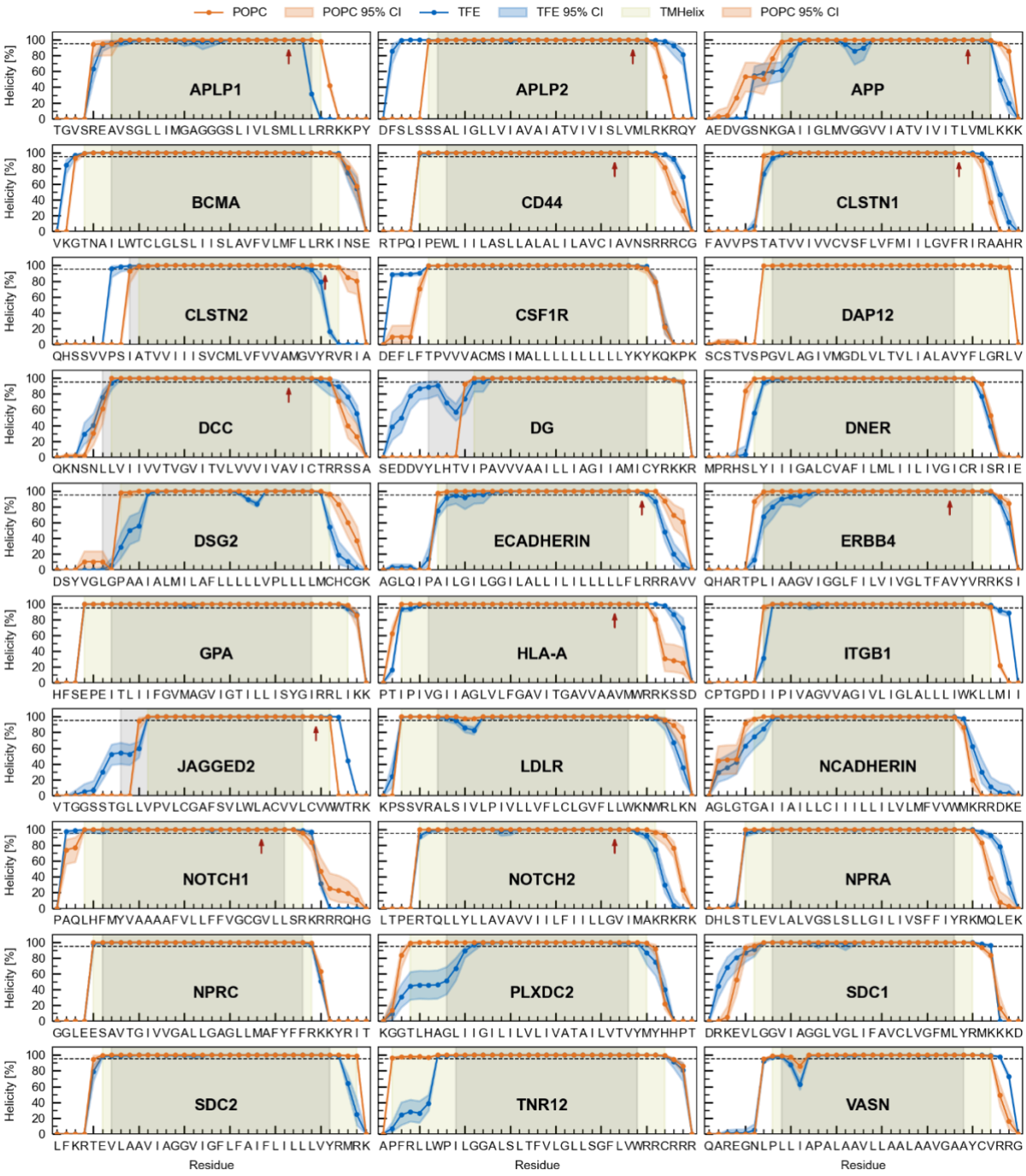


**Figure A2: Helicity of Residues Along the Investigated Sequence:** Helicity was calculated by the simplified DSSP algorithm as implemented in MDtraj.[126] For further details please refer to section 2.3. Blue lines show profiles for simulations in TFE/water, orange lines for simulations in POPC. Coloured areas indicate the 95 % confidence interval as computed by bootstrap resampling. Grey areas show the TM-Helix as determined from MD simulations. Red arrows indicate known endoproteolytic cleavage sites.

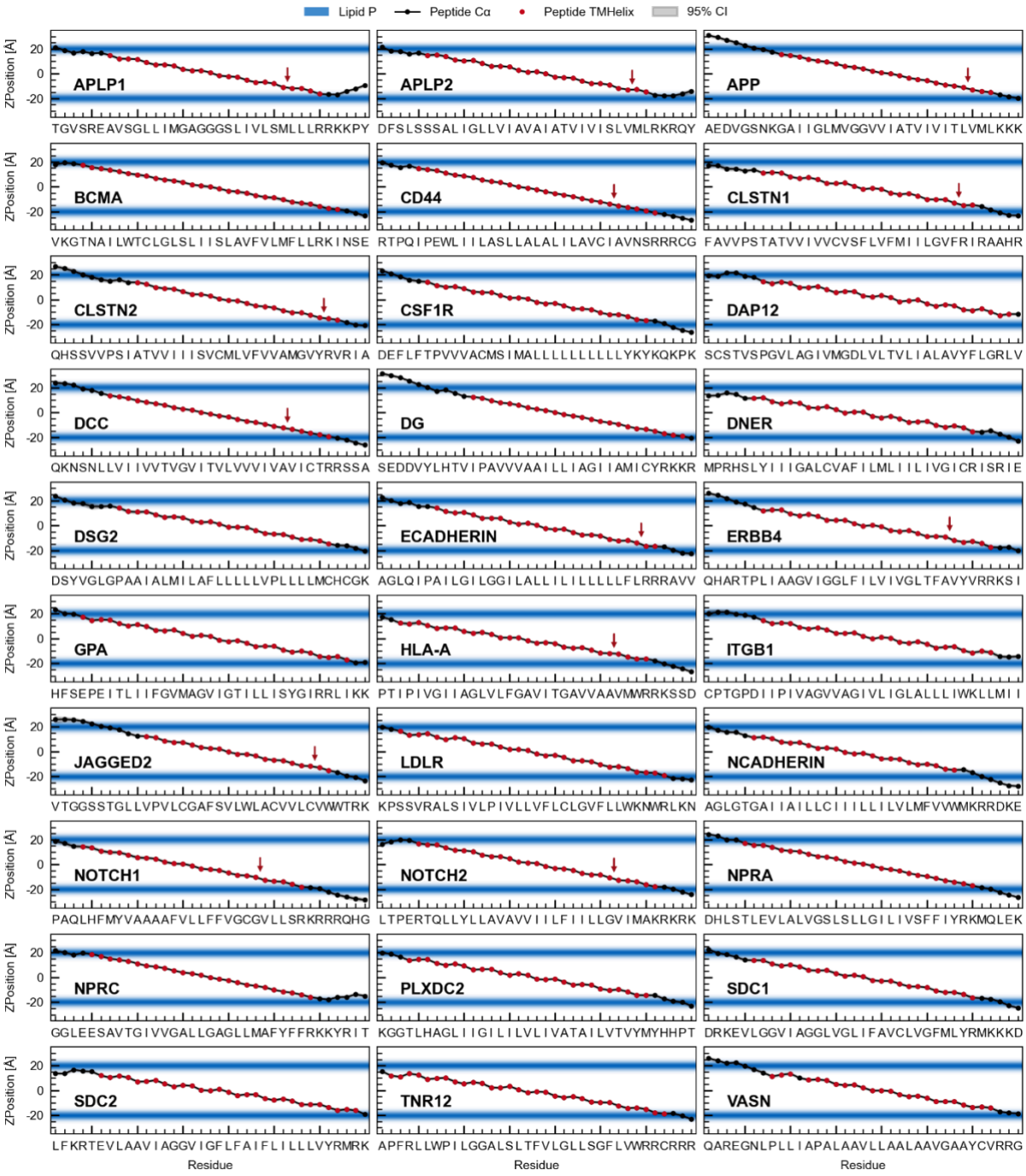


**Figure A3: z-Location of Residues Relative to the Membrane Bilayer Centre:** Blue colour gradients show the distribution of z-positions of lipid phosphate atoms relative to the membrane centre, grey lines indicate z-positions of residues and red dots mark residues that are part of the TM-helix. Red arrows indicate known endoproteolytic cleavage sites.

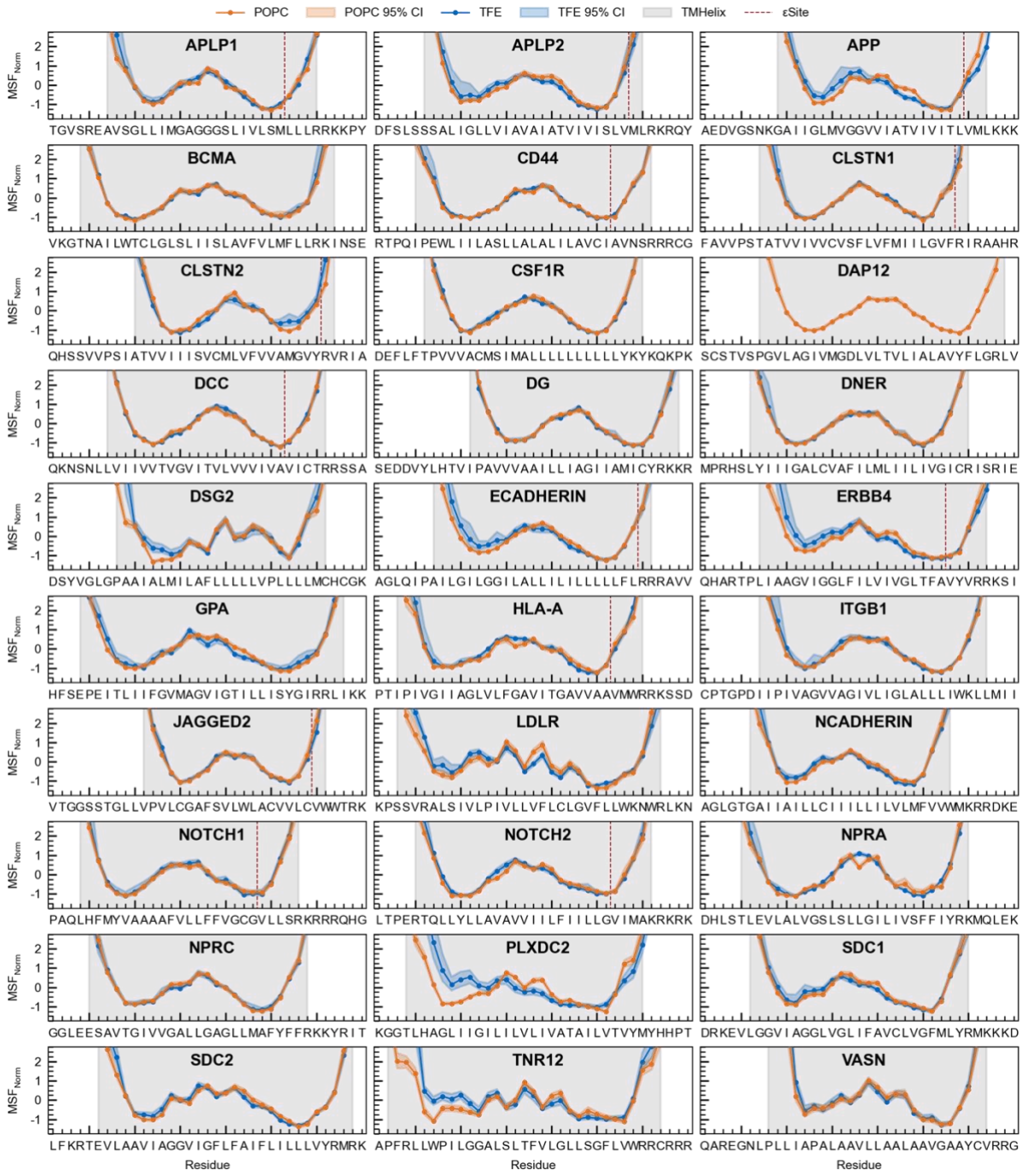


**Figure A4: Normalized Mean Squared Fluctuations:** Blue lines show profiles for simulations in TFE/water, orange lines for simulations in POPC. Coloured areas indicate the 95 % confidence interval as computed by bootstrap resampling. Grey areas show the TM-Helix as determined by MD simulations. Red dashed lines indicate known endoproteolytic cleavage sites.

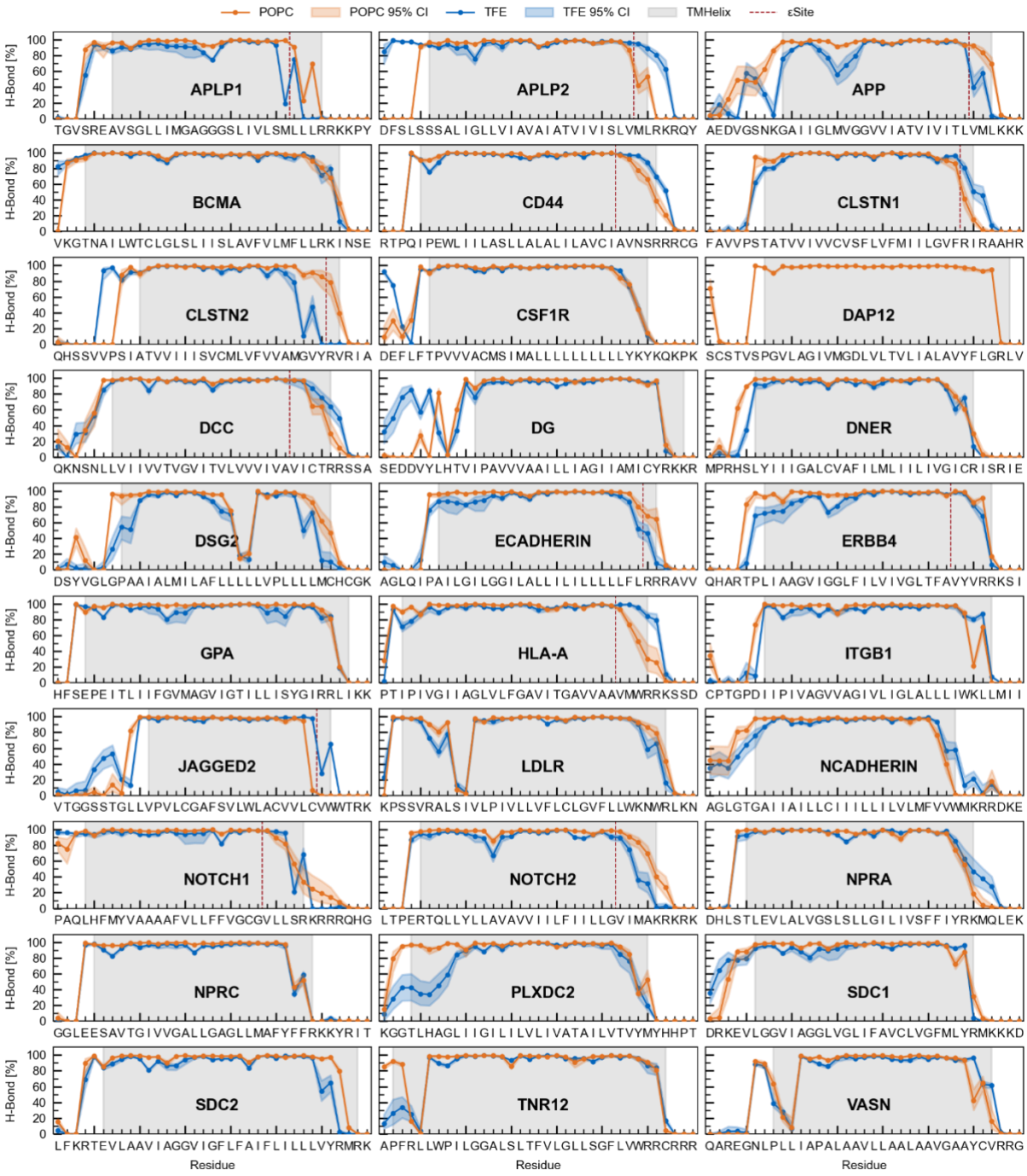


**Figure A5: H-Bond Occupancies:** An H-Bond is considered to be closed if either its α or $3_{10}$ H-Bond is closed. Blue lines show profiles for simulations in TFE/water, orange lines for simulations in POPC. Coloured areas indicate the 95 % confidence interval as computed by bootstrap resampling. Grey areas show the TM-Helix as determined by MD simulations. For individual results for α and $3_{10}$ H-Bond occupancies, please see **Figures A6** and **A7.** Red dashed lines indicate known endoproteolytic cleavage sites.

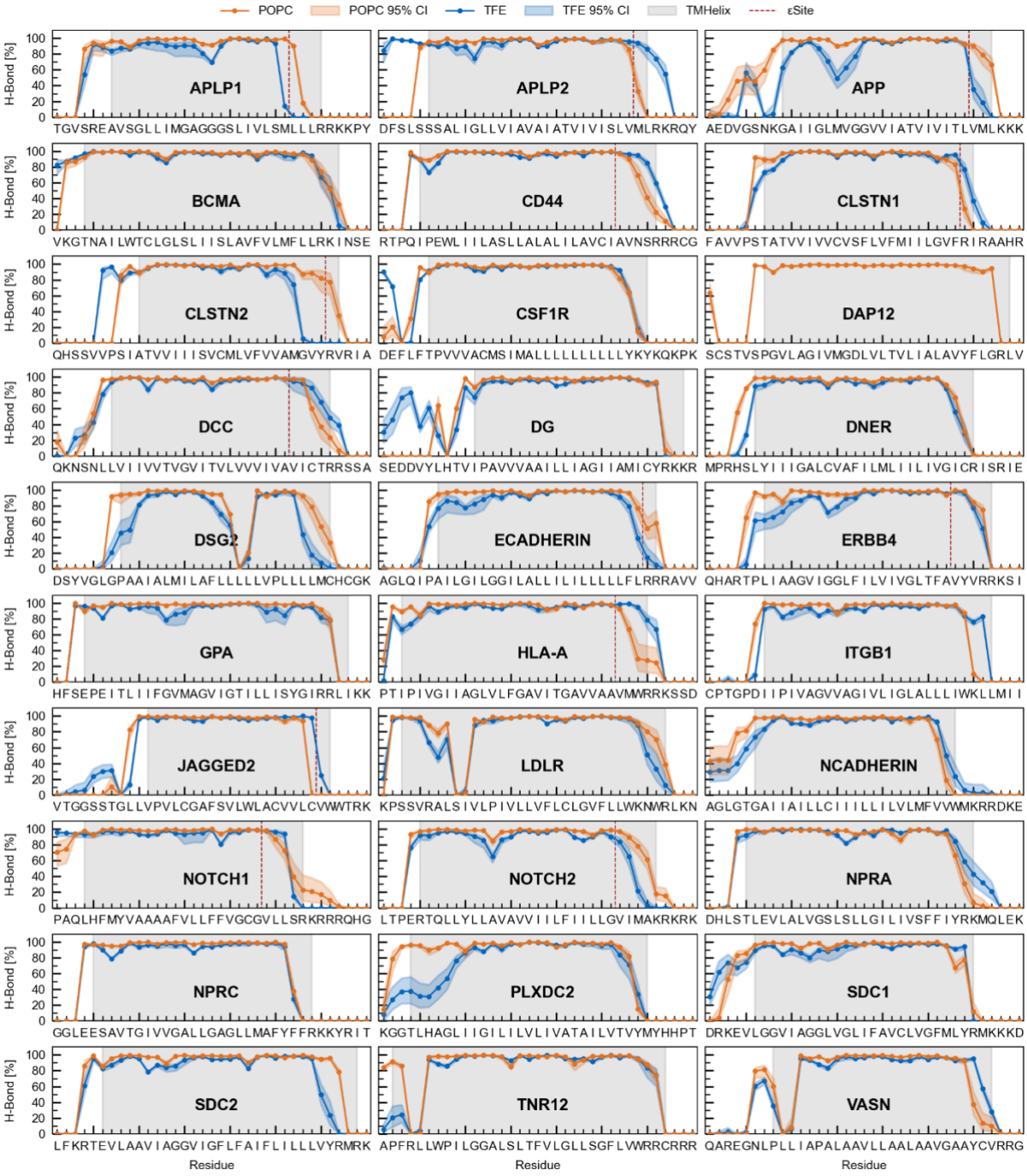


**Figure A6: αH-Bond Occupancies:** Blue lines show profiles for simulations in TFE/water, orange lines for simulations in POPC. Coloured areas indicate the 95 % confidence interval as computed by bootstrap resampling. Grey areas show the TM-Helix as determined by MD simulations. Red dashed lines indicate known endoproteolytic cleavage sites.

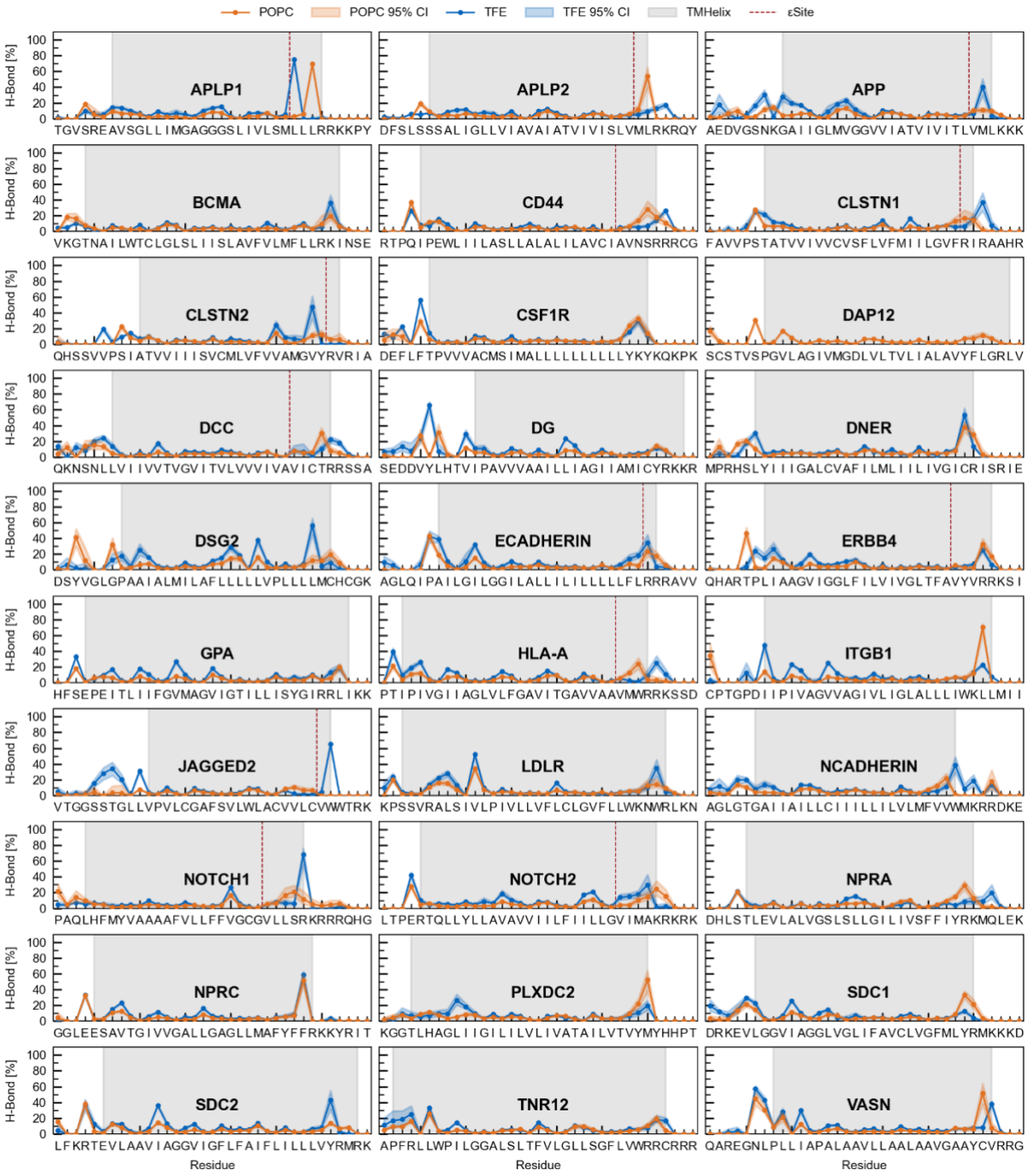


**Figure A7: $3_{10}$ H-Bond Occupancies:** Blue lines show profiles for simulations in TFE/water, orange lines for simulations in POPC. Coloured areas indicate the 95 % confidence interval as computed by bootstrap resampling. Grey areas show the TM-Helix as determined by MD Simulations. Red dashed lines indicate known endoproteolytic cleavage sites.

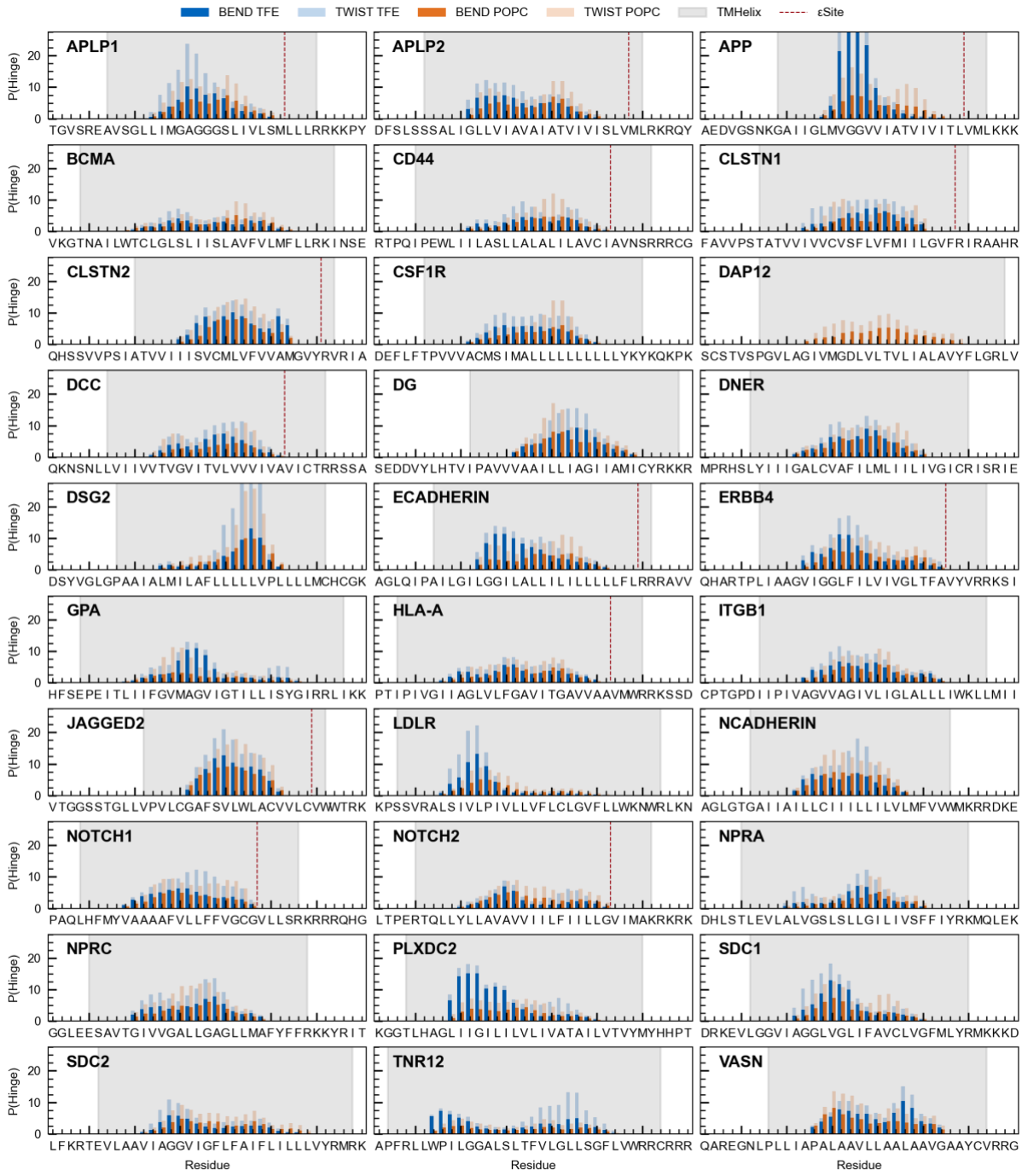


**Figure A8: Hinge Propensities:** Grey areas show the TM-Helix as determined by MD Simulations. Bars show propensity of individual residues to be part of a hinge controlling bending or twisting motions. Red dashed lines indicate known endoproteolytic cleavage sites. Hinge bending propensities around the di-glycine motif in APP's TMD have been cut at 27.5% (maximal values reached in TFE/water are ~38%).

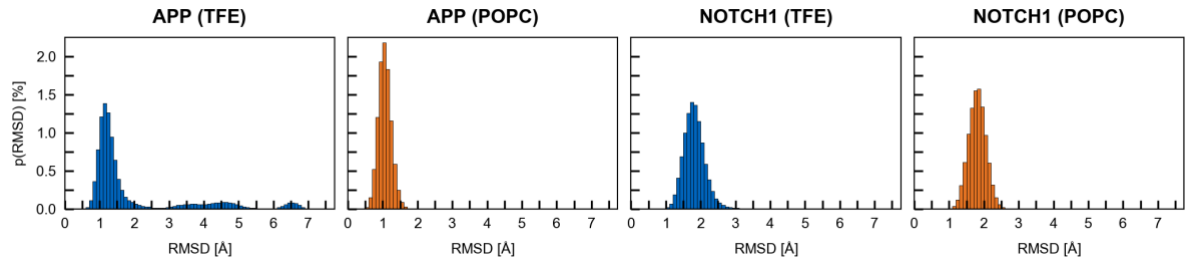


**Figure A9: RMSD to NMR structures:** Distribution of root-mean squared deviations between the simulated ensemble of conformations and conformations as determined by NMR spectroscopy. The RMSD has been computed for backbone heavy atoms of all conformations included in the corresponding PDB files of APP (2LLM[72]) and NOTCH1 (5KZO[67]). For other investigated TMDs no monomeric TMD conformations have been published in the RCSB database.

# List of Publications

### Publications in International, Peer Reviewed Journals and Pre-Prints

Spitz, C., Schlosser, C., Guschtschin-Schmidt, N., Stelzer, W., Menig, S., **Götz, A.,** Haug-Kröper, M., Scharnagl, C., Langosch, D., Muhle-Goll, C., & Fluhrer, R. (2020). Non-canonical Shedding of TNFα by SPPL2a Is Determined by the Conformational Flexibility of Its Transmembrane Helix. *IScience*, *23*(12), 101775. https://doi.org/10.1016/j.isci.2020.101775

Hitzenberger, M., **Götz, A.,** Menig, S., Brunschweiger, B., Zacharias, M., & Scharnagl, C. (2020). The dynamics of γ-secretase and its substrates. *Seminars in Cell & Developmental Biology*, S108495211830274X. https://doi.org/10.1016/j.semcdb.2020.04.008

**Goetz, A**., Mylonas, N., Hoegel, P., Silber, M., Heinel, H., Menig, S., Vogel, A., Feyrer, H., Huster, D., Luy, B., Langosch, D., Scharnagl, C., Muhle-Goll, C., Kamp, F., & Steiner, H. (2019). Modulating Hinge Flexibility in the APP Transmembrane Domain Alters γ-Secretase Cleavage. *Biophysical Journal*, 116(11) https://doi.org/10.1016/j.bpj.2019.04.030

**Götz, A.,** Högel, P., Silber, M., Chaitoglou, I., Luy, B., Muhle-Goll, C., Scharnagl, C., & Langosch, D. (2018). Increased H-Bond Stability Relates to Altered ε-Cleavage Efficiency and Aβ Levels in the I45T Familial Alzheimer's Disease Mutant of APP. *Scientific Reports 9 (5321)*, https://doi.org/10.1038/s41598-019-41766-1

**Götz, A.**, & Scharnagl, C. (2018). Dissecting conformational changes in APP's transmembrane domain linked to ε-efficiency in familial Alzheimer's disease. *PLOS ONE*, *13*(7), e0200077. https://doi.org/10.1371/journal.pone.0200077

Högel, P., **Götz, A.**, Kuhne, F., Ebert, M., Stelzer, W., Rand, K. D., Scharnagl, C., & Langosch, D. (2018). Glycine Perturbs Local and Global Conformational Flexibility of a Transmembrane Helix. *Biochemistry*. https://doi.org/10.1021/acs.biochem.7b01197

Leeb, E., **Götz, A.**, Letzel, T., Cheison, S. C., & Kulozik, U. (2015). Influence of denaturation and aggregation of β-lactoglobulin on its tryptic hydrolysis and the release of functional peptides. *Food Chemistry*, *187*, 545–554. https://doi.org/10.1016/j.foodchem.2015.04.034

Scharnagl, C., Pester, O., Hornburg, P., Hornburg, D., **Götz, A**., & Langosch, D. (2014). Side-Chain to Main-Chain Hydrogen Bonding Controls the Intrinsic Backbone Dynamics of the Amyloid Precursor Protein Transmembrane Helix. *Biophysical Journal*, *106*(6), 1318–1326. https://doi.org/10.1016/j.bpj.2014.02.013

Pester, O., **Götz, A**., Multhaup, G., Scharnagl, C., & Langosch, D. (2013). The Cleavage Domain of the Amyloid Precursor Protein Transmembrane Helix Does Not Exhibit Above-Average Backbone Dynamics. *ChemBioChem*, *14*(15), 1943–1948. https://doi.org/10.1002/cbic.201300322

Wani, A. A., Sogi, D. S., Singh, P., & **Götz, A.** (2013). Impacts of Refining and Antioxidants on the Physico-Chemical Characteristics and Oxidative Stability of Watermelon Seed Oil. *Journal of the American Oil Chemists' Society*, *90*(9), 1423–1430. https://doi.org/10.1007/s11746-013-2277-1

Wani, A. A., Singh, P., Shah, M. A., Wani, I. A., **Götz, A.**, Schott, M., & Zacherl, C. (2013). Physico-chemical, thermal and rheological properties of starches isolated from newly released rice cultivars grown in Indian temperate climates. *LWT - Food Science and Technology*, *53*(1), 176–183. https://doi.org/10.1016/j.lwt.2013.02.020

**Book contribution**

**Götz, A**., Wani, A. A., Langowski, H.-C., & Wunderlich, J. (2014). Food Technologies: Aseptic Packaging. In *Encyclopedia of Food Safety* (pp. 124–134). Elsevier. https://doi.org/10.1016/B978-0-12-378612-8.00274-2

**Conference contributions**

Ritz, R., Götz, A., and Hachinger, S. (2020). Methods for Federating Data Repositories Across Domains and Sites, *International FAIR Convergence Symposium 2020,* Online, **Session Organiser**

**Götz, A.,** Munke, J., Hayek, M., Nguyen, H., Weber, T., Hachinger, S., and Weismüller, J. (2020). A Lightweight, Microservice-Based Research Data Management Architecture for Large Scale Environmental Datasets, *EGU General Assembly 2020,* Online, **Poster**

Munke, J., **Götz, A.**, Heller, H., Hachinger, S., Laux, D., Goussev, O., Handschuh, J., Wüst, S., Bittner, M., Mair, R., Wittmann, B., Rehm, T., Beck, I., & Neumann, M. (2020). Connecting data streams with On-Demand Services in the Alpine Environmental Data Analysis Centre, *EGU General Assembly 2020,* Online, **Poster**

Bittner, M., Laux, D., Goussev, O., Wüst, S., Handschuih, J., **Götz, A.,** Heller, H., Munke, J., Mair, R., Wittmann, B., Beck, I., Neumann, M., & Rehm, T. (2020). The Alpine Environmental Data Analysis Centre (www.alpendac.eu) – A Component of the Virtual Alpine Observatory (VAO) (www.vao.bayern.de), EGU General Assembly 2020, *Online*, **Poster**

**Götz, A.,** Weber, T., & Hachinger S. (2019). Let The Data Sing - A Scalable Architecture to Make Data Silos FAIR. *Research Data Alliance 14th Plenary*, Helsinki, **Poster**

**Götz, A.**, Menig, S., & Scharnagl C. (2018). What Does the Dynamic Heterogeneity of Its Substrates Tell us About γ-Secretase‘s Role in the Cell Membrane? *SuperMUC Status and Results Workshop 2018,* Garching, **Talk**

**Götz, A.**, & Scharnagl, C. (2018). What Does the Dynamic Heterogeneity of Its Substrates Tell us About γ-Secretase‘s Role in the Cell Membrane? *Workshop on Computer Simulation and Theory of Macromolecules, Hünfeld, **Poster + Talk***

**Götz, A.**, & Scharnagl, C. (2017). The Dynamic Architectures of γ-Secretase Substrates: Discriminating Factor or Just Another Hint Towards a Membrane FOR2290 Proteasome? *Understanding Intramembrane Proteolysis From Substrates to Enzymes, Regensburg, **Poster***

**Götz**, **A.**, Heinel, H., Högel, P., Vogel, A., Langosch, D., Huster, D., & Scharnagl, C. (2017). Teaming up Molecular Dynamics Simulations with Mass-Spectrometry and ssNMR to Reveal the

Dynamic Architecture of the Amyloid Precursor Protein's Transmembrane Domain, *Conformational Ensembles from Experimental Data and Computer Simulations, Berlin,* ***Poster***
**Götz, A.**, & Scharnagl, C. (2017), Hierarchical TMD dynamics provides a rationale for presentation of the APP ε-sites to γ-secretase, *19th IUPAB congress and 11th EBSA congress, Edinburgh,* ***Poster***

**Götz, A.**, & Scharnagl, C. (2017). Hunting for the Link Between Familial Azheimer's Disease Mutations of the Amyloid Precursor Protein and Initial Cleavage by γ-Secretase, *Workshop on Computer Simulation and Theory of Macromolecules, Hünfeld,* ***Poster***

**Götz, A.**, & Scharnagl, C. (2016), Global Substrate Dynamics Provides a Rationale for Perturbed γ-Secretase Cleavage Patterns, *Annual Meeting of the German Biophysical Society, Erlangen,* ***Poster***

**Götz, A.**, & Scharnagl, C. (2016). Multiscale Modeling of Transmembrane Domain Dynamics: From Normal Modes to Markov State Models, *Workshop on Computer Simulation and Theory of Macromolecules, Hünfeld,* ***Poster + Talk***

**Götz, A.**, & Scharnagl, C. (2016), Dynamics of Gamma-Secretase Substrates on Multiple Scales: From Normal Modes to Markov State Models, *SuperMUC Status and Results Workshop 2016, Garching,* ***Talk***

**Götz, A.**, Scharnagl, C., & Langosch, D. (2015). Investigating Dynamic Motifs in Amyloid Precursor Protein Mutants as Impact Factors for Alzheimer's Disease, *Workshop on Computer Simulation and Theory of Macromolecules, Hünfeld,* ***Poster + Talk***

**Götz, A.**, & Scharnagl, C. (2014), How Long is Long Enough? Using SuperMUC to Break Time Related Frontiers of Membrane Protein's Molecular Dynamics Simulations, *SuperMUC Status and Results Workshop 2014, Garching,* ***Talk***

**Götz, A.**, Uhrmann, H., Scharnagl, C., & Langosch, D. (2014). Investigating Dynamic Motifs in Amyloid Precursor Protein Mutants as Impact Factors for Alzheimer's Disease, *3rd Annual CCP-BioSim Conference: Frontiers of Biomolecular Simulations, Edinbugh,* ***Poster***

Scharnagl, C., Mörch, M., **Götz, A.**, & Langosch, D. (2014). How Dynamic Transmembrane Helices Can Influence Peptide/Lipid Interactions: A Molecular Dynamics Study, *3rd Annual CCP-BioSim Conference: Frontiers of Biomolecular Simulations, Edinbugh,* ***Poster***