



UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS QUÍMICAS

CARACTERIZACIÓN DINAMICA DE SITIOS MUTACIONALES EN LA
SUBUNIDAD CATALÍTICA DE γ -SECRETASA

TESIS

PARA OPTAR POR EL GRADO DE

MAESTRO EN CIENCIAS

PRESENTA

FÍSICA CECILIA CHÁVEZ GARCÍA

DRA. LAURA DOMÍNGUEZ DUEÑAS
FACULTAD DE QUÍMICA

Ciudad Universitaria, Cd.Mx., Junio 2017



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

JURADO ASIGNADO

Presidente	Dr. Ramón Garduño Juárez
Vocal	Dr. Marcelo Enrique Galván Espinosa
Vocal	Dr. Rogelio Rodríguez Sotres
Vocal	Dr. Juan Raúl Álvarez Idaboy
Secretario	Dra. Karina Martínez Mayorga

LUGAR DONDE SE DESARROLLÓ EL TEMA

El presente trabajo de tesis se realizó en la Facultad de Química, Departamento de Físicoquímica, cubículo sin número del edificio F, Universidad Nacional Autónoma de México, bajo la dirección de la Dra. Laura Domínguez Dueñas.

Los resultados de esta tesis dieron lugar a la siguiente publicación:

- Cecilia Chávez-García, Rodrigo Aguayo-Ortiz, Laura Dominguez. Quantifying correlations between mutational sites in the catalytic subunit of γ -secretase. *En proceso*.

Se colaboró en la siguiente publicación relacionada con el tema de tesis:

- Rodrigo Aguayo-Ortiz, Cecilia Chávez-García, John E. Straub, Laura Dominguez. Characterizing the structural ensemble of γ -secretase using a multiscale molecular dynamics approach. *Chemical Science*, **2017**.

Parte de este proyecto se presentó en el siguiente congreso:

- “Quantifying correlations between mutational sites in the catalytic subunit of γ -secretase”. Cartel. Tercer simposio: Tendencias actuales en la búsqueda y desarrollo de fármacos, celebrado en la Facultad de Química, Ciudad Universitaria, del 14 al de 16 junio de **2017**.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca otorgada para la realización de mis estudios de maestría (No. 663225/583333) y por la Beca Mixta de Movilidad en el Extranjero que me permitió realizar una estancia académica de tres meses en la Universidad de Harvard. Asimismo, agradezco al Posgrado en Ciencias Químicas por el apoyo económico brindado para la impresión de mi tesis.

Agradezco a la Dirección General de Cómputo y de Tecnologías de Información (LANCAD-UNAM-DGTIC-306) por el soporte recibido para el uso de la supercomputadora HP Cluster Platform 3000SL (Miztli).

Agradezco también al Programa de Apoyo a la Investigación y el Posgrado (PAIP IARA204716) y al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT IA204716).

ÍNDICE GENERAL

Abreviaturas y acrónimos.....	1
RESUMEN	2
ABSTRACT	3
1. ANTECEDENTES	4
1.1 Enfermedad de Alzheimer.....	4
1.1.1 Síntomas y patogenia	4
1.1.2 Hipótesis amiloidogénica	5
1.2 γ -secretasa.....	5
1.2.1 Características	5
1.2.2 Procesamiento de APP por γ -secretasa.....	7
1.3 Mutaciones.....	8
1.3.1 Mutaciones en APP.....	8
1.3.2 Mutaciones en PS1	8
1.4 Herramientas computacionales	9
1.4.1 Dinámica molecular.....	9
1.4.2 Modelo de grano grueso	9
1.5 Definiciones.....	10
1.5.1 POPC	10
1.5.2 Correlación	11
2. Planteamiento del problema.....	12
3. Hipótesis.....	12
4. Objetivos.....	13
4.1 Objetivo general	13
4.2 Objetivos particulares	13
5. Métodos	14
5.1 Metodología general	14
5.2 Preparación de los modelos	15
5.3 Dinámicas moleculares atomísticas	15
5.4 Cálculo de la información mutua	16
5.5 Dinámicas moleculares de grano grueso.....	16
5.6 Correlación cruzada de las simulaciones de grano grueso.....	18
6. Resultados	19

6.1 Correlación entre residuos de PS1	19
6.2 Perturbaciones en PS1 introducidas por mutaciones sencillas.....	23
6.3 Alteraciones en la estabilidad de PS1.....	27
7. Conclusiones.....	34
8. Perspectivas.....	34
Bibliografía.....	35

Abreviaturas y acrónimos

A β	β -amiloide
APH-1A	Faringe anterior defectuosa 1A
APP	Proteína precursora amiloide
DM	Dinámica Molecular
ECD	Dominio extracelular
FAD	Enfermedad de Alzheimer familiar
NCT	Nicastrina
NFT	Ovillos neurofibrilares
PDB	Protein Data Bank
PEN-2	Potenciador de presenilina 2
PS1	Presenilina 1
RMSD	Desviación cuadrática media
RMSF	Fluctuación cuadrática media
STD	Desviación estándar
TM	Hélice transmembranal

RESUMEN

La presenilina 1 (PS1) es la subunidad catalítica del complejo γ -secretasa, la cual está involucrada en el procesamiento de la proteína precursora amiloide (APP) para la generación del péptido β -amiloide (A β). Las mutaciones sencillas en PS1 alteran el patrón de corte del sustrato y llevan a la formación de péptidos A β aberrantes. A la fecha, más de doscientas mutaciones en esta subunidad han sido asociadas al desarrollo temprano de la enfermedad de Alzheimer.

Varios estudios han demostrado que las mutaciones pueden alterar la actividad enzimática a través de redes alostéricas. Por consiguiente, en el presente trabajo de tesis se estudiaron los movimientos correlacionados de los aminoácidos de la proteína nativa y de 13 modelos con mutaciones en PS1, utilizando una aproximación de dinámica molecular multi-escala.

Se utilizaron dos modelos de γ -secretasa (PDBs: 5FN2 y 5FN3) para realizar simulaciones de dinámica molecular atomística de 1 μ s cada una. Se calculó la correlación dinámica entre todos los aminoácidos de PS1 durante los últimos 500 ns de cada trayectoria. Los aminoácidos con mayor frecuencia de aparición en pares de correlación con distancias grandes y correlaciones altas, presentes en ambos modelos y con mutaciones reportadas, fueron seleccionados para construir los 13 modelos con mutaciones.

Los 13 modelos mutados y la proteína nativa fueron simulados con dinámica molecular de grano grueso durante 1 μ s con 20 réplicas cada uno. La correlación entre los aminoácidos de PS1 fue calculada de nueva cuenta, y los residuos con un gran número de correlaciones altas fueron filtrados. Los residuos seleccionados fueron comparados entre los modelos mutados y la proteína nativa.

En este trabajo se identificaron 13 mutaciones relacionadas con la enfermedad de Alzheimer en aminoácidos con correlaciones altas y distancias grandes entre pares de aminoácidos de dos modelos de γ -secretasa. Doce de estas mutaciones fueron clasificadas de acuerdo a sus redes de correlación con aminoácidos en las interfaces (1) PS1-APH1-A y (2) PS1-PEN2, (3) el sitio catalítico, (4) el sitio de entrada del sustrato, y (5) el sitio de reconocimiento del sustrato. Asimismo, se identificaron los cambios provocados por las mutaciones en la estabilidad de PS1 y en la inclinación de sus hélices.

ABSTRACT

Presenilin 1 (PS1) is the catalytic subunit of the γ -secretase complex involved in the cleavage of the amyloid precursor protein (APP) for the generation of the amyloid- β peptide (A β). Single point mutations in PS1 alter the cleavage pattern of the substrate and lead to the formation of aberrant A β peptides. To date, more than two hundred mutations in this subunit have been associated to the development of early onset Alzheimer's disease.

Several studies have demonstrated that mutations may alter the enzyme catalytic activity through allosteric networks. Therefore, in this work we studied the amino acid correlated motions in the wild type and 13 mutant models of PS1 employing a multi-scale molecular dynamics simulation approach.

We employed two γ -secretase models (PDB IDs: 5FN2 and 5FN3) to perform all-atom molecular dynamics simulations of 1 μ s each. The correlation between the amino acid dynamics of the catalytic subunit was computed for the last 500 ns of each trajectory. The amino acids with largest frequency of appearance in correlation pairs with strong correlation and long distance, present in both models and with reported mutations, were selected to build the 13 mutant models.

The 13 mutant models plus the wild type were simulated with coarse grained molecular dynamics during 1 μ s and with 20 replicas each. The correlation between PS1 amino acids was computed again and amino acids with the largest frequency of significant correlations were filtered. The selected residues were compared between the mutated models and the wild type.

In this work, 13 mutations related to Alzheimer's disease were identified in amino acids with high correlations and long distances between the pairs in two γ -secretase models. Twelve of these mutations were classified according to their correlation networks with amino acids in the interfaces (1) PS1-APH1-A and (2) PS1-PEN2, (3) the catalytic site, (4) the substrate entry site, and (5) the substrate recognition site. Furthermore, changes in PS1 stability and helix tilts induced by mutations were identified.

1. ANTECEDENTES

1.1 Enfermedad de Alzheimer

1.1.1 Síntomas y patogenia

De acuerdo con la Organización Mundial de la Salud alrededor de 47.5 millones de personas padecen de demencia. La causa de demencia más común es la enfermedad de Alzheimer, a la cual se le adjudican entre un 60% y un 70% de dichos casos. En México más de 350,000 personas son afectadas por la enfermedad de Alzheimer y al año mueren por ella 2,030 pacientes.¹

Los síntomas iniciales de la enfermedad de Alzheimer son la pérdida de memoria a corto plazo y confusión. Posteriormente surgen cambios importantes de comportamiento y de personalidad, así como dificultad para resolver problemas y problemas de lenguaje. En la última etapa de la enfermedad, se pierde el control corporal y se requiere atención constante. Eventualmente el paciente sucumbe a un estado de coma y muere, pero en la mayoría de los casos mueren de neumonía, deshidratación o desnutrición.

Patológicamente, ésta enfermedad está definida por una extensa pérdida neuronal y por la aparición de dos tipos de agregados de proteínas: placas seniles y ovillos neurofibrilares (NFT). Los NFT están compuestos por depósitos intraneuronales de la proteína *tau* hiperfosforilada y las placas seniles en depósitos extracelulares de diferentes isoformas del péptido β -amiloide ($A\beta$).² Aunque las placas seniles y los NFTs son patognomónicos, no son los únicos cambios funcionales y estructurales en un cerebro con Alzheimer. Otras alteraciones incluyen respuestas inflamatorias, estrés oxidativo y pérdida del tejido cerebral, lo que provoca un agrandamiento de los ventrículos. El cerebro de un paciente con Alzheimer puede pesar una tercera parte menos que el de una persona de la misma edad sin la enfermedad (fig. 1).³



Figura 1. Imágenes comparativas de un cerebro sano y uno con la enfermedad de Alzheimer. (A) Cerebro sano, (B) cerebro con Alzheimer y (C) superposición de ambos cerebros. (Adaptada de ALZFORUM).⁴

1.1.2 Hipótesis amiloidogénica

La hipótesis amiloidogénica postula que la acumulación de diferentes isoformas del péptido A β en el cerebro es el evento central en la patología de la enfermedad de Alzheimer. A pesar de que esta idea ha sido controversial, la evidencia que proviene de laboratorios y clínicas en todo el mundo apoyan el concepto de que la dyshomeostasis de A β es un factor temprano y a menudo desencadenante en la enfermedad de Alzheimer.⁵

El mecanismo de generación de A β involucra el procesamiento proteolítico de la proteína precursora amiloide (APP), el cual se lleva a cabo en dos pasos. En el primer paso, un corte de APP por la proteína β -secretasa libera 99 aminoácidos del fragmento C-terminal de APP (C99).^{6,7} Este fragmento, permanece anclado a la membrana en donde es procesado, en un segundo paso, mediante una serie de cortes proteolíticos secuenciales por la enzima γ -secretasa.³ Dado que γ -secretasa no es muy regioselectiva, genera las diferentes isoformas de A β .

La acumulación de agregados amiloidogénicos de los fragmentos A β inicia cuando γ -secretasa genera una mayor proporción de las isoformas A β_{42} y A β_{43} ; éstas tienen una alta propensión a agregarse mientras que A β_{40} podría incluso ser anti-amiloidogénico.⁸ La isoforma de A β más común es A β_{40} . Los dos aminoácidos que hacen la diferencia entre las isoformas A β_{40} y A β_{42} son una Alanina e Isoleucina, ambos hidrofóbicos. Estos dos residuos provocan un cambio en la población de láminas β entre ensamblajes estructurales de A β_{40} y A β_{42} , aumentando la tasa de formación de fibras de ésta última.⁹

1.2 γ -secretasa

1.2.1 Características

La enzima γ -secretasa es una aspartil proteasa incrustada en la membrana celular que corta un gran número de proteínas sustrato en sus regiones transmembranales. Es una proteasa poco común, su sitio activo está embebido en la membrana celular a pesar de que requiere la presencia de al menos una molécula de agua para llevar a cabo su actividad catalítica.¹⁰ A la fecha han sido descritos más de 60 sustratos de esta proteasa, sin embargo, ninguno de ellos comparte algún motivo estructural reconocible.^{11,12}

La enzima γ -secretasa consiste en cuatro componentes: presenilina 1 (PS1), faringe anterior defectuosa 1A (APH-1A), potenciador de presenilina 2 (PEN-2) y nicastrina (NCT) (fig. 2A).^{13,14}

PS1 es el componente catalítico, fue inicialmente identificado como el blanco de mutaciones causantes de la enfermedad de Alzheimer Familiar (FAD, por sus siglas en inglés).^{15,16} La enfermedad de Alzheimer familiar y la de inicio tardío (LOAD, por sus siglas en inglés) son esencialmente la misma enfermedad pero el proceso que las genera es distinto. FAD es la consecuencia del malfuncionamiento de genes mutados mientras que LOAD proviene del acumulación gradual de fallas relacionadas con la edad.

PS1 tiene nueve hélices transmembranales (TM), sus dos residuos catalíticos (Asp257 y Asp385) se encuentran en las TMs 6 y 7 respectivamente (fig. 2B).¹⁷ Se ha propuesto que PEN-2, con sus dos TMs, facilita la maduración de PS y aumenta la actividad de la γ -secretasa.¹⁸ APH-1 es una proteína con siete TMs que estabiliza el complejo. La nicastrina es una glicoproteína transmembranal de tipo I con un gran dominio extracelular (ECD, por sus siglas en inglés) y una única hélice TM en su extremo C-terminal.

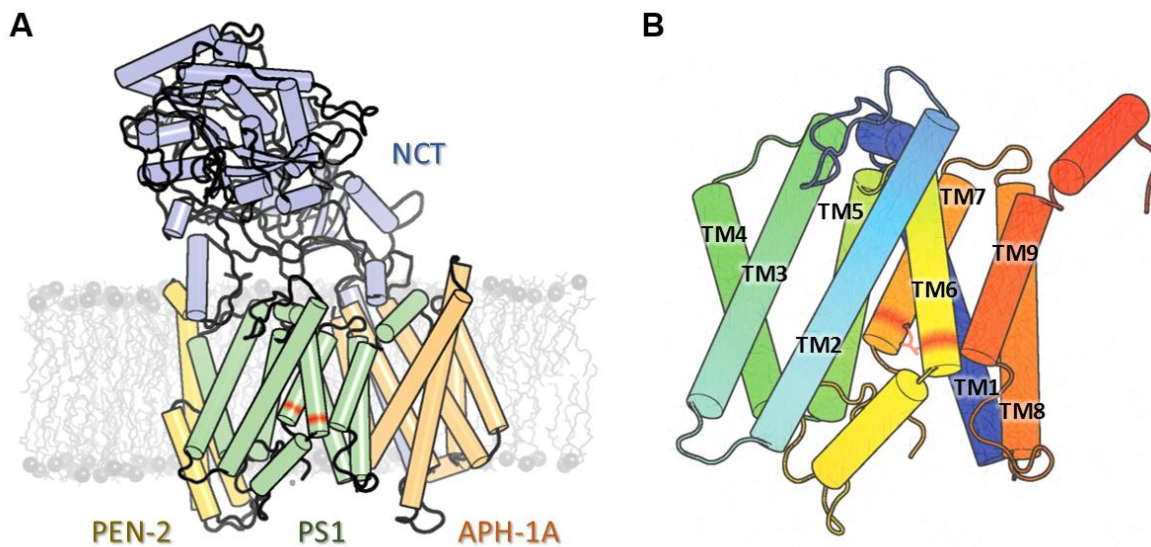


Figura 2. (A) Estructura de γ -secretasa coloreada por componentes: NCT, PEN-2, PS1 y APH1-A. (B) Representación de las nueve hélices TMs de PS1. La posición de los residuos catalíticos se presentan marcados en rojo.

Uno de los mayores desafíos en el desarrollo de inhibidores y moduladores de γ -secretasa ha sido que hasta mediados de 2015, no se contaba con una estructura de este blanco terapéutico. Recientemente, seis estructuras de crio-microscopía electrónica (3.4 – 4.4 Å) fueron depositadas en el Protein Data Bank (PDBs: 4UIS, 5A63, 5FN2, 5FN3, 5FN4 y 5FN5). Las estructuras 5FN2 y 5FN3 tienen los esqueletos más completos de la proteína, las

cadena lateral de las 20 TMs de γ -secretasa asignadas y el ECD de NCT. Al resto de las estructuras les hace falta TM2 de PS1, lo cual refleja su alta movilidad y posible implicación en el mecanismo de unión del sustrato.^{19,20}

1.2.2 Procesamiento de APP por γ -secretasa

Un requisito que deben cumplir los sustratos de γ -secretasa para poder ser procesada, es haber sufrido un corte cercano a la membrana, liberando un gran fragmento extracelular. El ectodominio residual es típicamente menor a 50 residuos, lo cual aumenta dramáticamente la capacidad de γ -secretasa de reconocer al sustrato.¹² APP es procesado por γ -secretasa únicamente después de haber perdido su ectodominio bajo la acción de β -secretasa.

Al igual que a varios de sus sustratos, γ -secretasa corta a APP en múltiples sitios. El sitio de corte ϵ se encuentra entre los residuos 48/49 (de acuerdo con la numeración de A β), el ζ entre los residuos 45/46 y el sitio de corte γ entre los residuos 42/43 y 38/40 (fig. 3). Estos cortes dan lugar a pequeños fragmentos de tres aminoácidos, con periodicidad similar a la de una hélice α .^{21,22}

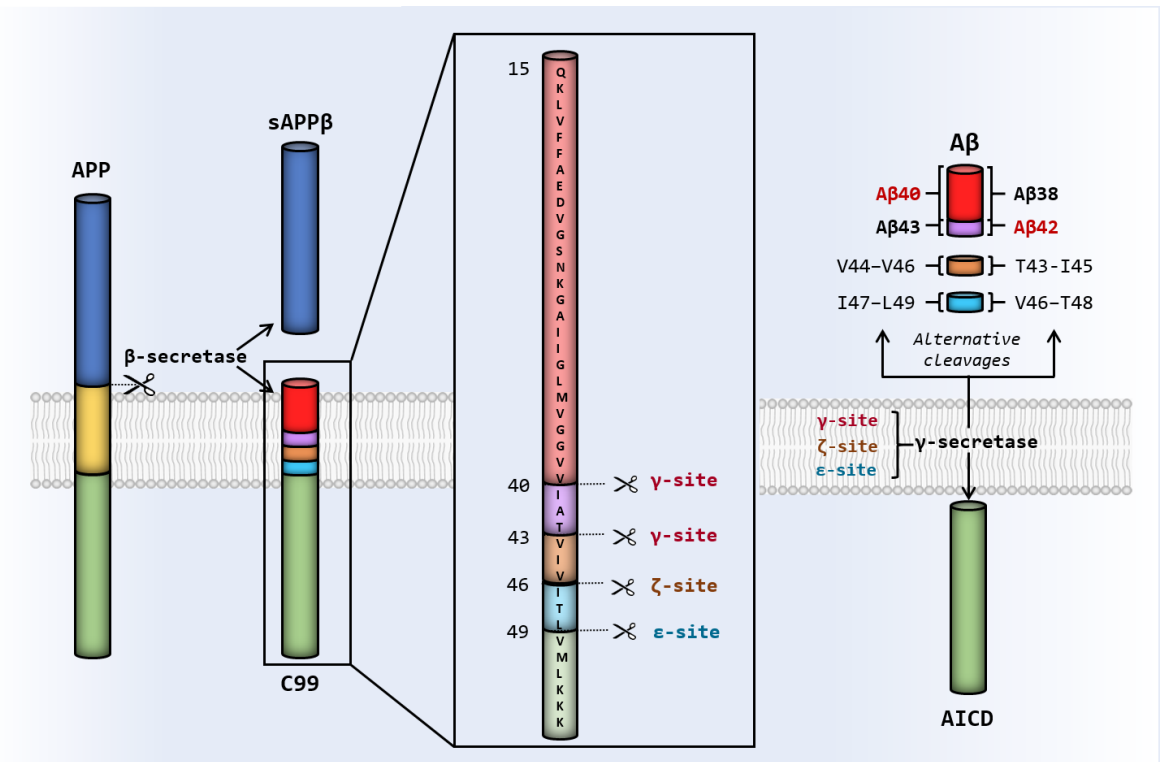


Figura 3. Pérdida de ectodominio y procesamiento secuencial de la Proteína Precursora Amiloide (APP). Series de corte alternativos de los sitios de corte ϵ , ζ y γ conducen a la formación de diferentes productos A β . (Adaptada de Aguayo-Ortiz, R. & Dominguez, L.)²⁰

1.3 Mutaciones

1.3.1 Mutaciones en APP

Existen por lo menos 46 mutaciones en APP que resultan en la enfermedad de Alzheimer autosómica dominante de inicio precoz.²³ Estas mutaciones se localizan principalmente cerca de los sitios de corte de β - y γ -secretasa.

En pacientes con síndrome de Down, la sobreexpresión de APP resulta en depósitos de $A\beta$ en el cerebro a edades cercanas a los 30 años. Posteriormente, desarrollan ovillos neurofibrilares que correlacionan con el inicio de deterioro mental común en estos individuos.²⁴

1.3.2 Mutaciones en PS1

La mayoría de las mutaciones en PSEN1 son mutaciones sencillas, que resultan en la sustitución de un solo aminoácido en PS1. Algunas son más complejas e incluyen ablaciones, inserciones y splicing; sin embargo, las consecuencias bioquímicas de estas mutaciones en el ensamblado de γ -secretasa son limitadas.²⁵ A la fecha, se han identificado más de 230 mutaciones en casi un tercio de los 467 aminoácidos de este componente que inducen FAD (fig. 4).⁴ Todas estas mutaciones llevan a un fenotipo similar: aumento en la proporción de $A\beta_{42}$ respecto a la de $A\beta_{40}$, aumento en el depósito de placa amiloide en el cerebro y un desarrollo en edades tempranas de la enfermedad. A la fecha no se ha logrado explicar cómo es que mutaciones distribuidas a lo largo de toda la proteína y distantes al sitio activo causan un aumento similar de la función tóxica en esta enzima.²⁵

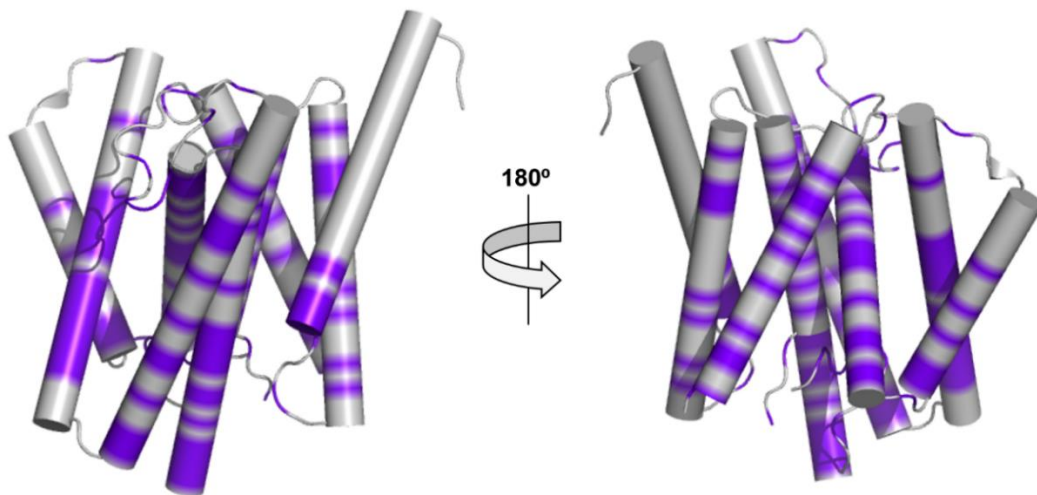


Figura 4. Representación tridimensional de PS1. Los aminoácidos con mutaciones reportadas como causantes de FAD se muestran en morado.

1.4 Herramientas computacionales

1.4.1 Dinámica molecular

La dinámica molecular es un tipo de simulación computacional que permite analizar la evolución en el tiempo de un sistema. Consiste en generar las trayectorias de las N partículas que lo componen a partir de resolver de forma numérica las ecuaciones de movimiento de cada partícula, sin tomar en cuenta efectos cuánticos o relativistas. La dinámica molecular es una técnica determinística: dado un conjunto de posiciones y velocidades iniciales, su subsecuente evolución en el tiempo está (en principio) determinado.²⁶

La dinámica molecular de sistemas en equilibrio por lo general corresponde al ensamble microcanónico de la mecánica estadística (NVE). Los sistemas que se pueden estudiar actualmente pueden contener alrededor de 10^4 - 10^5 átomos y se pueden simular durante periodos de tiempo del orden de nanosegundos (10^{-9} s) a microsegundos (10^{-6} s).²⁷

Cada simulación requiere definir la función potencial o campo de fuerza. Éste es la fórmula matemática y el conjunto de parámetros empleados para describir la energía de una proteína en función de sus coordenadas atómicas.

1.4.2 Modelo de grano grueso

Un gran problema es la brecha que existe entre las escalas de tiempo de los métodos computacionales y los experimentales. Una manera de resolver este problema es utilizar modelos de grano grueso, los cuales representan al sistema con un número reducido de grados de libertad (comparando con una descripción atomística). Grupos de átomos son reemplazados por pseudo-átomos, construyendo un modelo más simple de la estructura molecular pero reteniendo las características críticas de las moléculas que determinan la física de la escala de tiempo bajo estudio. Una simulación de este tipo necesita menos recursos de cómputo y avanza más rápido, incrementando en órdenes de magnitud el tiempo de simulación.²⁸

El campo de fuerza de grano grueso Martini fue construido para hacer simulaciones de dinámica molecular de sistemas biomoleculares. Este modelo usa un mapeo 4 a 1, es decir, en promedio cuatro átomos pesados y los hidrógenos asociados a éstos son representados por un solo centro de interacción (fig. 20). Este mapeo fue elegido como punto óptimo entre eficiencia computacional y representabilidad química.²⁹

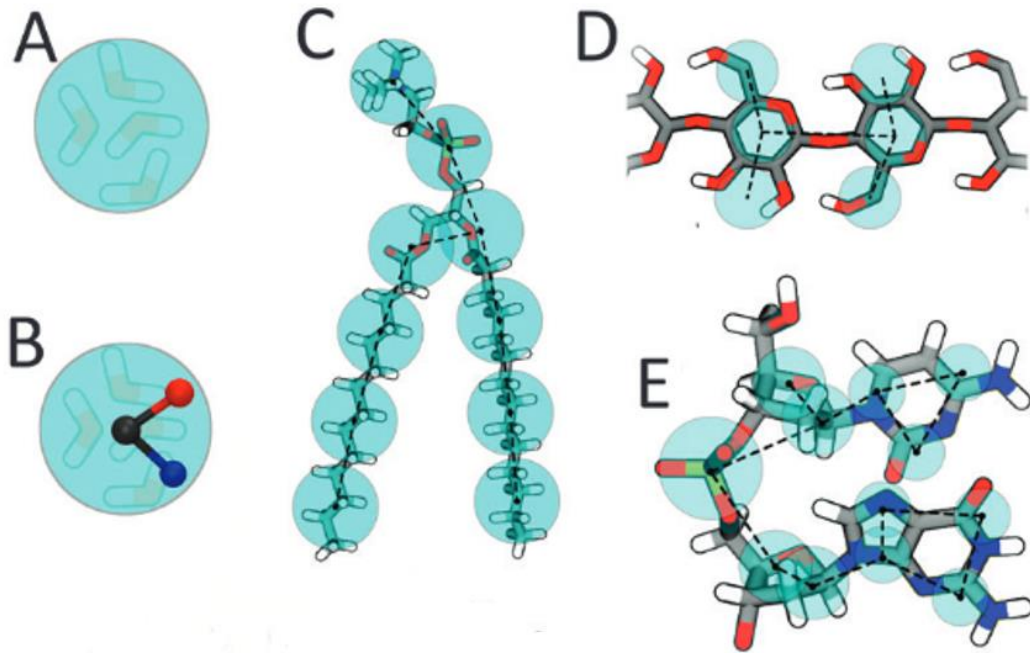


Figura 20. Ejemplos de mapeo con el campo de fuerza Martini. A) Partícula de agua estándar que representa cuatro moléculas de agua, B) partícula de agua polarizable con cargas embebidas, C) lípido DMPC, D) fragmento de polisacárido y E) péptido. Las cuentas azules representan las cuentas de grano grueso de Martini.

1.5 Definiciones

1.5.1 POPC

POPC es considerado uno de los lípidos modelo para experimentos biofísicos. Es un lípido con cabeza de fosfatidilcolina y cola de 1-palmitoil-2-oleoil (fig. 21). La fosfatidilcolina es el tipo de lípido más abundante en la membrana celular de los eucariontes.³⁰ El POPC tiene una sola insaturación, lo que aumenta el área lateral por lípido (65.4 \AA^2) en la bicapa. Las membranas de POPC tienen un grosor (cabeza polar a cabeza polar) de 37.3 \AA a 303 K .³¹ Existen datos experimentales de γ -secretasa en POPC.³²

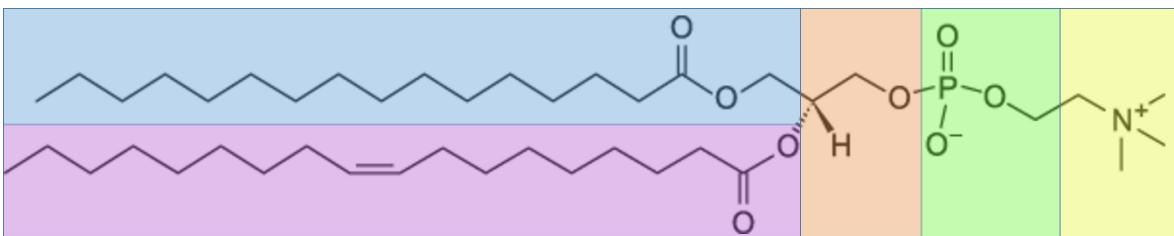


Figura 21. Estructura de una molécula de POPC. Se muestra en amarillo la colina, en verde el fosfato, en rojo el glicerol, en azul el palmitoil y en morado el oleoil.

1.5.2 Correlación

La correlación es una relación estadística que involucra dependencia y que indica qué tan lineal es la relación entre dos variables. Puede tomar valores entre -1 y 1 (fig. 22). El signo positivo indica que las variables correlacionadas aumentan al mismo tiempo y el signo negativo indica que cuando una de las variables aumenta la otra disminuye.

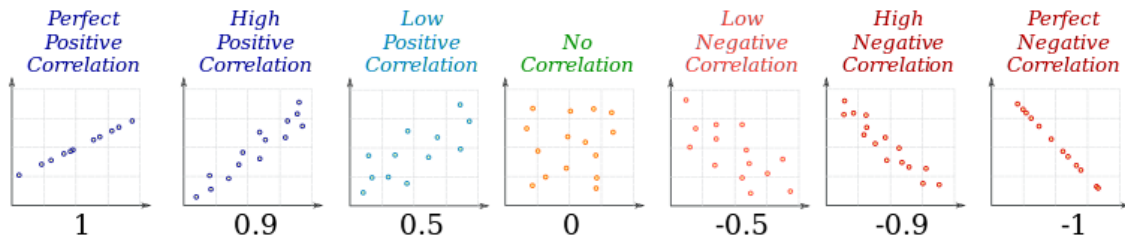


Figura 22. Ejemplos de distintos valores de correlación entre dos variables. (Tomada de mathsisfun)³³

2. Planteamiento del problema

Presenilin-1 es el componente catalítico de la enzima γ -secretasa. Fue inicialmente identificado como el blanco de mutaciones relacionadas con la enfermedad de Alzheimer. A la fecha se han reportado más de 230 mutaciones en PS1, distribuidas en casi una tercera parte de sus aminoácidos. Sin embargo, no se ha podido explicar cómo es que mutaciones distribuidas a lo largo de esta proteína y lejos de los residuos catalíticos pueden modificar el patrón de corte de γ -secretasa aumentando la proporción de $A\beta_{42}$ respecto a la de $A\beta_{40}$.

En este trabajo buscamos entender:

- ¿cómo algunas de estas mutaciones en γ -secretasa afectan la estabilidad y el funcionamiento de la enzima?
- ¿de qué forma las mutaciones en aminoácidos distantes al sitio activo influyen en la catálisis de la enzima?
- ¿existen redes de correlación entre los aminoácidos de PS1?
- ¿es posible clasificar las mutaciones de acuerdo al tipo de efecto que tienen en ésta?

3. Hipótesis

El uso de herramientas de simulación molecular multi-escala permitirá obtener información sobre el impacto de mutaciones relacionadas con la enfermedad de Alzheimer en la estabilidad del componente catalítico de γ -secretasa.

4. Objetivos

4.1 Objetivo general

Utilizando metodologías computacionales de simulación y dinámica molecular se busca caracterizar la estructura y la dinámica del componente catalítico de γ -secretasa así como las alteraciones provocadas por mutaciones involucradas en la aparición de la enfermedad de Alzheimer.

4.2 Objetivos particulares

- Realizar simulaciones de dinámica molecular atomística de dos modelos de γ -secretasa, cada uno embebido en una membrana lipídica al interior de una caja de moléculas de agua.
- Analizar las trayectorias de las simulaciones para identificar correlaciones en el movimiento de los aminoácidos del componente catalítico de γ -secretasa.
- Identificar los aminoácidos con mayor número de correlaciones altas y con mutaciones reportadas relacionadas con la enfermedad de Alzheimer.
- Construir un modelo de γ -secretasa para cada mutación elegida así como para la proteína nativa y realizar simulaciones de dinámica molecular de grano grueso para cada uno de ellos.
- Caracterizar la dinámica de cada simulación molecular de grano grueso e identificar los cambios causados por las mutaciones estudiadas.

5. Métodos

5.1 Metodología general

En el siguiente diagrama se presenta la metodología empleada en este trabajo. Se emplearon dos estructuras cristalográficas para construir dos modelos atomísticos consistentes cada uno en una de las proteínas embebida en una bicapa de POPC, agua e iones. Se realizó una dinámica molecular atomística de $1\mu\text{s}$ para cada modelo y se analizó la correlación dinámica entre todos los aminoácidos de PS1 durante los últimos 500 ns de cada trayectoria. Se eligieron los aminoácidos con mayor frecuencia de aparición en pares de correlación con separación espacial y correlación significativas. Posteriormente, éstos fueron filtrados escogiendo aquellos aminoácidos que presentaron alta correlación en ambos modelos y además que tuvieran mutaciones reportadas como causantes de FAD. Utilizando estas mutaciones se construyeron 13 modelos de γ -secretasa con una mutación sencilla cada uno y un modelo extra para la proteína nativa. Con estos modelos se construyeron 14 sistemas proteína-membrana de grano grueso, cada uno de los cuáles fue simulado con 20 réplicas durante $1\mu\text{s}$. De nueva cuenta, se analizó la correlación entre aminoácidos durante los últimos 500 ns de cada trayectoria; filtrando los residuos con mayor frecuencia de aparición en pares de correlación con distancias grandes y correlaciones altas. Se compararon los aminoácidos seleccionados de esta forma entre los modelos mutados y la proteína nativa.

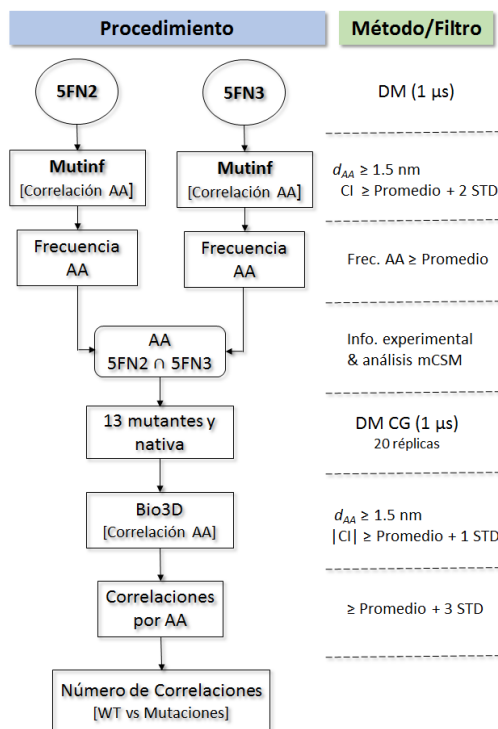


Figura 5. Metodología general empleada.

5.2 Preparación de los modelos

Se utilizaron dos estructuras de γ -secretasa (PDBs: 5FN2 y 5FN3) resueltas por criomicroscopía electrónica a 4.2 y 4.1 Å respectivamente, como coordenadas iniciales para realizar las dinámicas moleculares multi-escala. Las cadenas laterales faltantes se añadieron a la estructura con el servidor web WHAT IF.³⁴ Para asignar el arreglo espacial de los modelos de γ -secretasa en la membrana de POPC se utilizó el servidor web Orientations of Proteins in Membranes (OPM).³⁵

5.3 Dinámicas moleculares atomísticas

Cada uno de los modelos obtenidos fue embebido en una bicapa lipídica utilizando el módulo de CHARMM-GUI, Membrane Builder.^{36,37} Con el fin de comparar los resultados con datos experimentales disponibles en la literatura, se utilizó una membrana de palmitoil oleoil fosfatidilcolina (POPC).³⁸ El sistema derivado del modelo de 5FN2 consiste en la proteína reconstruida, 320 lípidos de POPC, 46,078 moléculas de agua, 143 iones de cloro y 148 iones de sodio. El sistema derivado del modelo de 5FN3 consiste en la proteína reconstruida, 326 lípidos de POPC, 46,827 moléculas de agua, 145 iones de cloro y 150 iones de sodio.

Los sistemas fueron minimizados y equilibrados energéticamente con dinámicas NVT y NPT, y simulados durante 1 μ s con un paso de tiempo de 2 fs. La temperatura fue fijada en 310 K utilizando el termostato de Nose-Hoover y la presión fue fijada a 1.0 bar mediante el algoritmo semi-isotrópico Parinello-Rahman. El potencial de Lennard-Jones se truncó mediante una función de desplazamiento entre los 0.9 y los 1.2 nm. Todas las interacciones electrostáticas se calcularon para vecinos a distancias iguales o menores a 1.2 nm, después de los cuales se calcularon utilizando el método de partículas de malla Ewald (PME). Las listas de vecinos se actualizaron cada 20 pasos y los enlaces que involucraban hidrógenos se restringieron mediante el algoritmo LINCS (LINear Constraint Solver).³⁹

Las simulaciones fueron realizadas con el programa GROMACS 5.0.6,⁴⁰ el campo de fuerza CHARMM36,⁴¹ y el modelo de agua TIP3P. La contribución apolar a la energía de γ -secretasa, asignada a la interacción de PS1 con los componentes adyacentes, fue calculada con el script `g_mmpbsa.py`.⁴² Finalmente, las simulaciones de DM fueron analizadas con las herramientas de GROMACS y las bibliotecas de MDAAnalysis⁴³ para Python.

5.4 Cálculo de la información mutua

La identificación de los movimientos correlacionados entre aminoácidos es clave para caracterizar efectos alostéricos en una proteína.⁴⁴ Por lo tanto, estudiar las redes de correlación en la dinámica de una proteína resulta esencial para obtener información sobre la regulación alostérica de aminoácidos con mutaciones reportadas.

Se evaluaron los movimientos correlacionados entre residuos de PS1 durante los últimos 500 ns de las simulaciones de los modelos de 5FN2 y 5FN3. La información mutua fue calculada para todos los ángulos del esqueleto de la proteína (Ω , Ψ y Φ) y de sus cadenas laterales (χ_1 a χ_5) utilizando el programa Mutinf.⁴⁵ Mutinf utiliza términos de segundo orden de la expansión de entropía configuracional (la información mutua) para identificar pares de residuos con conformaciones correlacionadas. Éste método usa coordenadas internas y se enfoca en los ángulos diedros para capturar la correlación altamente anarmónica de las cadenas laterales y eliminar correlaciones potencialmente espurias.

Se calculó la distancia promedio entre cada par de aminoácidos y se eligieron los pares con distancias grandes (≥ 1.5 nm) y correlaciones altas (\geq promedio + 2STD). Posteriormente se cuantificó la frecuencia de aparición de cada residuo en los pares de aminoácidos seleccionados y se escogieron aquellos con una frecuencia de aparición mayor al valor promedio. Finalmente, se generaron 13 modelos con cada una de las mutaciones reportadas para los aminoácidos seleccionados y presentes en ambos modelos. Junto con un decimocuarto modelo para la proteína nativa.

5.5 Dinámicas moleculares de grano grueso

Se construyeron catorce sistemas de grano grueso a partir de los trece modelos con mutaciones puntuales y la proteína nativa (fig. 6). Antes de preparar los sistemas para las simulaciones de dinámica molecular, se confirmó que el número y/o tipo de partículas de grano grueso en el aminoácido mutado era distinto que en la proteína nativa (tabla 1). Cada modelo fue embebido en una bicapa de POPC utilizando el módulo de CHARMM-GUI, Martini Bilayer Maker^{46,47} con el campo de fuerza Martini v2.2⁴⁸. La temperatura fue fijada en 310 K utilizando el acoplamiento *V-rescale* y la presión se fijó en 1.0 bar con el acoplamiento semi-isotrópico de Berendsen. Se realizaron las trayectorias para 20 réplicas de cada sistema durante 1 μ s con dinámicas NPT utilizando GROMACS 5.0.⁴⁰

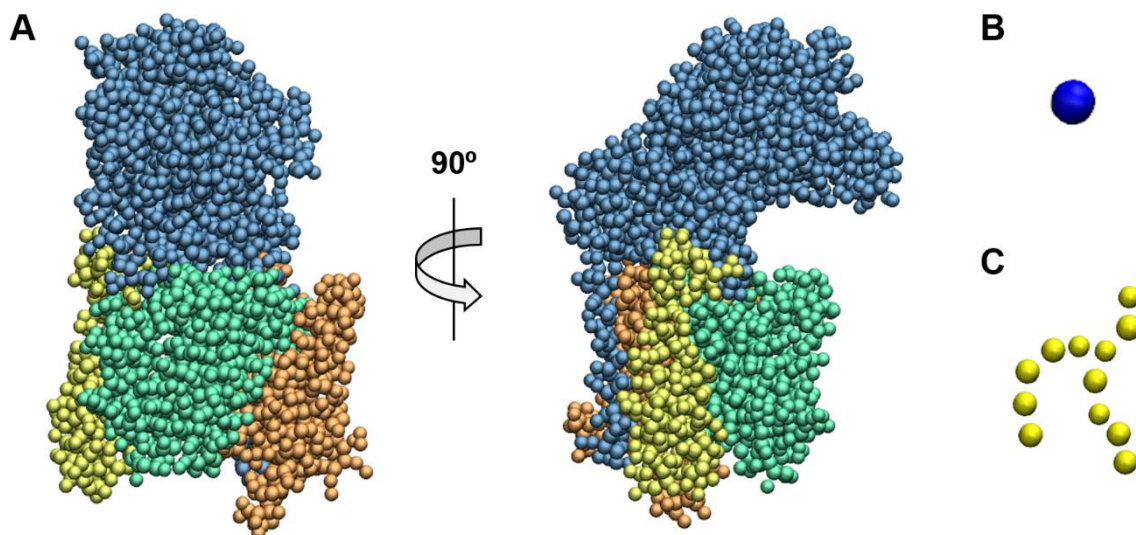


Figura 6. Representación de grano grueso de: (A) γ -secretasa, (B) molécula de agua y (C) POPC.

Tabla 1. Topología de grano grueso		
Mutación	Proteína nativa	Mutante
I83T	P5, AC1	P5, P1
R108Q	P5, N0, Qd	P5, P4
W165G	P5, SC4, SNd, SC5, SC5	P5
I213F	P5, AC1	P5, SC5, SC5, SC5
I213T	P5, AC1	P5, P1
H214Y	P5, SC4, SP1, SP1	P5, SC4, SC4, SP1
K239N	P5, C3, Qd	P5, P5
A260V	P4	P5, AC2
V261F	P5, AC2	P5, SC5, SC5, SC5
C263R	P5, C5	P5, N0, Qd
C263F	P5, C5	P5, SC5, SC5, SC5
I408T	P5 AC1	P5, P1
L435F	P5, AC1	P5, SC5, SC5, SC5

Partículas utilizadas para representar los aminoácidos de las mutaciones elegidas para construir sistemas de grano grueso y realizar simulaciones de dinámica molecular. Existen cuatro tipos de partículas (P, polar; N, intermedia; C, apolar; Q cargada) divididos en subniveles (0, no forma puentes de hidrógeno; a, algunas capacidades de aceptor de hidrógeno; d, algunas capacidades de donador de hidrógeno; ad, capacidades de donador y aceptor) o (1, 2, 3, 4 o 5) donde el subtipo 5 es más polar que el 1. Los tipos AC1 y AC2 se utilizan para evitar choques al interior de las proteínas. La S indica que la partícula forma parte de un anillo.

5.6 Correlación cruzada de las simulaciones de grano grueso

Se calculó la matriz de correlación cruzada para todos los carbonos alfa ($C\alpha$) en cada trayectoria de grano grueso utilizando el paquete de R, Bio3D.^{49,50} Se obtuvo el valor total de la frecuencia de aparición de cada aminoácido en las parejas correlacionadas, tomando en cuenta todas las réplicas de cada modelo. Éstos fueron filtrados seleccionando únicamente aquellos con alta correlación (\geq promedio en todas las réplicas + 1STD) y larga distancia entre aminoácidos (≥ 1.5 nm). El umbral obtenido para la correlación (aprox. 0.32) se encuentra dentro del intervalo usado por los autores del paquete Bio3D. Subsecuentemente, se seleccionaron los aminoácidos con mayor número de correlaciones. El límite utilizado se escogió en base a los resultados del modelo de la proteína nativa (\geq promedio + 3STD). Se compararon las correlaciones entre los modelos mutados y el modelo de la proteína nativa para análisis posteriores.

6. Resultados

Como primer paso se analizó la estabilidad de la proteína en las trayectorias de simulación atomística. Los sistemas alcanzaron estados de equilibrio después de 300 ns de simulación, con estructuras alrededor de 0.30 nm de RMSD de la estructura inicial (fig. 7A). De acuerdo con los valores calculados de RMSFs de PS1, las regiones más móviles de esta proteína son el lazo entre TM1 y TM2 y la región carboxilo terminal de TM6. Todos los lazos presentan un RMSF de al menos 1.5 nm.

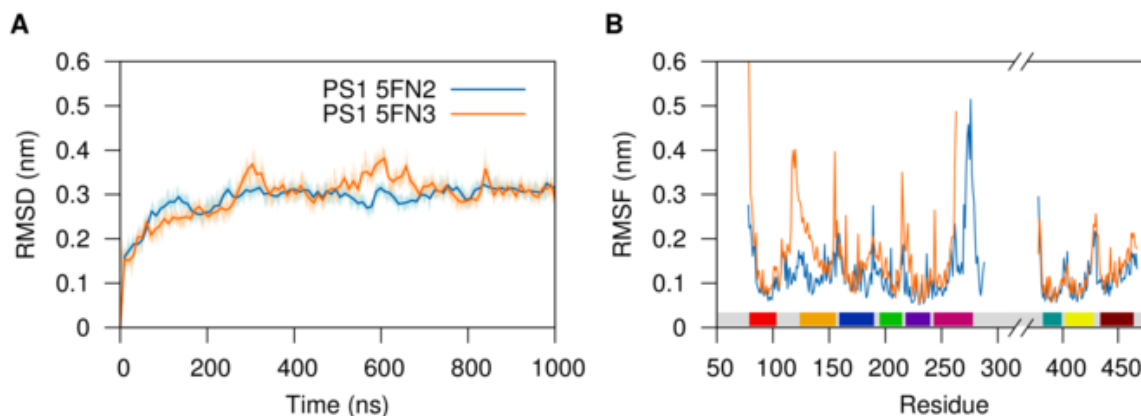


Figura 7. (A) Desviación cuadrática media (RMSD) del esqueleto de PS1 y B) fluctuación cuadrática media (RMSF) durante las simulaciones atomísticas de 5FN2 (azul) y 5FN3 (naranja).

6.1 Correlación entre residuos de PS1

Se utilizaron los últimos 500 ns de las simulaciones de dinámica molecular atomística de 1 μ s de los modelos de γ -secretasa 5FN2 y 5FN3 para calcular la información mutua de todos los ángulos diedros del esqueleto de la proteína (Ω , Ψ y Φ) y de sus cadenas laterales (χ). Los cálculos se hicieron utilizando el software Mutinf.⁴⁵

Se evaluaron las redes de fluctuaciones correlacionadas en cinco intervalos de 100 ns cada uno con el fin de capturar movimientos de aminoácidos correlacionados y no lineales de forma imparcial y estadísticamente robusta. Se obtuvieron un total de 44,253 y 37,128 pares correlacionados para cada intervalo de tiempo de las simulaciones de DM de 5FN2 y 5FN3 respectivamente (fig. 8). Los pares de correlación fueron clasificados en cuatro regiones (fig. 8A) de acuerdo a la distancia entre aminoácidos y al coeficiente de correlación: correlaciones pequeñas a distancias cortas (región I), correlaciones pequeñas a distancias largas (región II), grandes correlaciones a distancias cortas (región III) y grandes correlaciones a distancias largas (región IV).

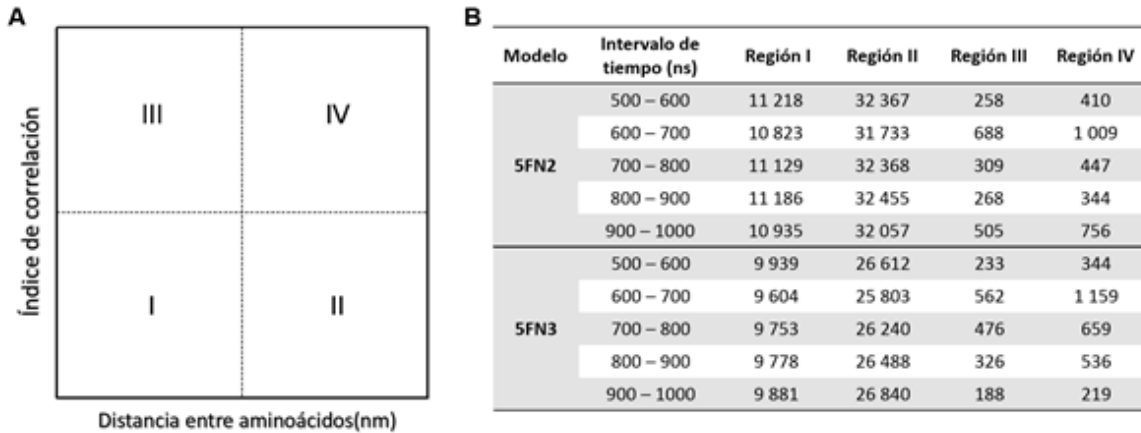


Figura 8. (A) Representación esquemática de las regiones en la gráfica empleada para la selección de aminoácidos con base en su índice de correlación y la distancia entre pares. (B) Número de pares de aminoácidos en cada región de la gráfica durante los diferentes intervalos de tiempo.

Primero se descartaron todas las correlaciones entre aminoácidos con distancias promedio menores a 1.5 nm ya que las correlaciones entre aminoácidos cercanos no contienen información sobre la regulación alostérica. Posteriormente, considerando que es menos probable que correlaciones altas entre residuos distantes provengan de fluctuaciones térmicas, se seleccionaron aquellos aminoácidos con correlaciones altas utilizando como valor umbral el valor promedio de correlación más dos desviaciones estándar (0.19472 para 5FN2 y 0.19568 para 5FN3).

La figura 9 muestra todos los datos divididos en las cuatro regiones mostradas en el esquema de la figura 9A. La región superior derecha contiene los pares de correlación encontrados entre aminoácidos con valores más altos y con distancias más grandes. Los pares de correlación de esta región (2,966 para 5FN2 y 2917 para 5FN3) fueron utilizados para cuantificar la frecuencia de aparición en parejas fuertemente correlacionadas. Se identificaron los aminoácidos con alta frecuencia de aparición (mayor al promedio, 20 para 5FN2 y 21 para 5FN3) y con mutaciones reportadas (fig. 9B).

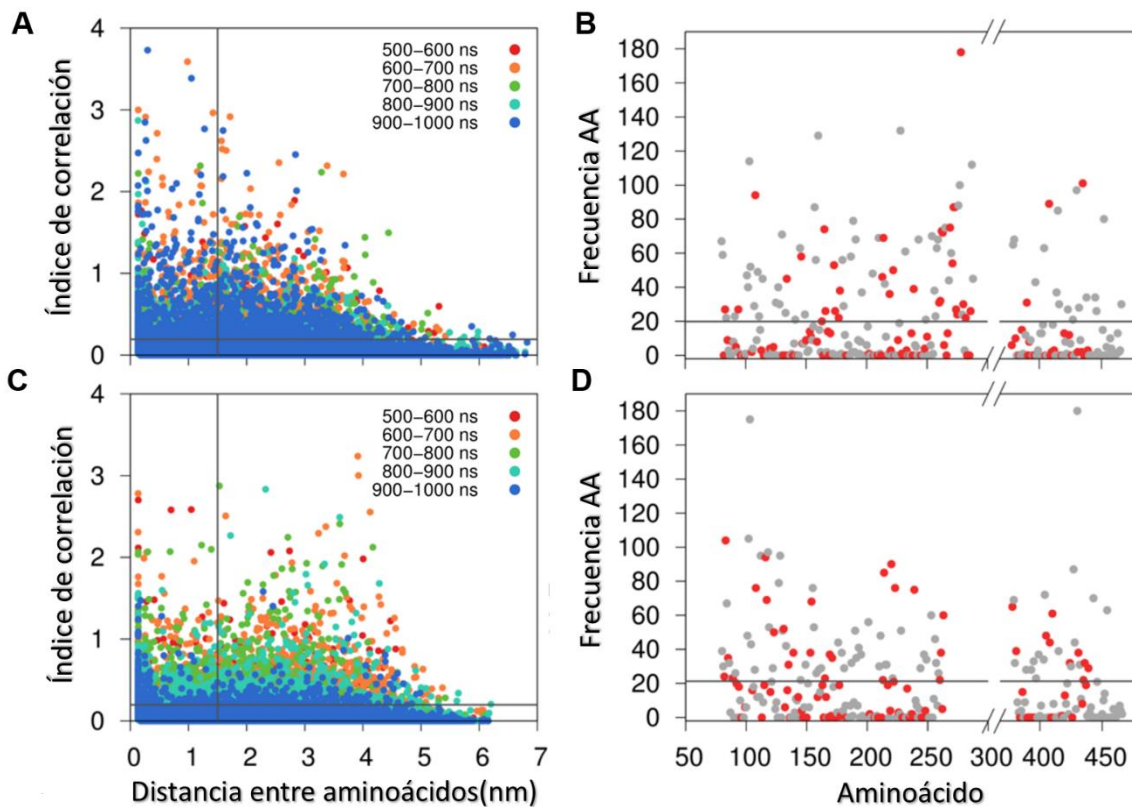


Figura 9. Análisis de correlación con información mutua. Pares correlacionados (izquierda) y frecuencia de aparición de los aminoácidos en los pares correlacionados (derecha) para 5FN2 (A, B) y 5FN3 (C, D). Los aminoácidos con mutaciones reportadas se encuentran resaltados en color rojo (B, D).

Se observó que un gran número de aminoácidos sin mutaciones reportadas exhiben un alto número de correlaciones, sugiriendo su importancia en la actividad o estabilidad de PS1. Es notable que un gran número de correlaciones relevantes del modelo de 5FN2 están localizadas en el fragmento C-terminal de TM6, el cuál no está presente en el modelo de 5FN3. Finalmente, se escogieron aquellos aminoácidos con alta frecuencia de aparición y mutaciones reportadas presentes en ambos modelos para realizar un análisis posterior (fig. 10A). La figura 10B muestra la ubicación espacial en la proteína de los aminoácidos seleccionados. La amplia distribución de los aminoácidos con movimientos altamente correlacionados en las nueve TMs de PS1 exhibe su amplia regulación alostérica, particularmente entre los aminoácidos localizados en la región citoplasmática de la proteína.

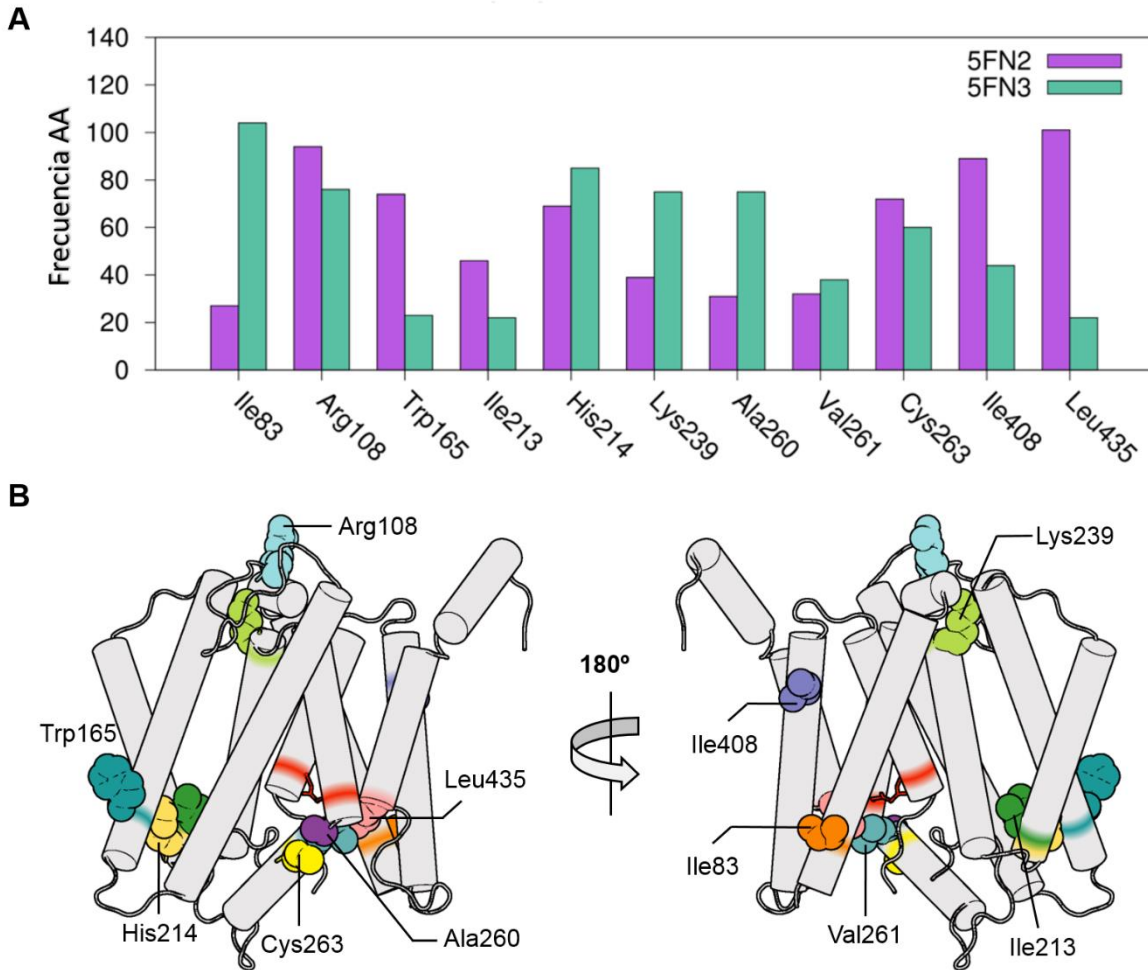


Figura 10. Aminoácidos seleccionados en base al estudio de correlación y a datos experimentales. (A) Frecuencia de aparición en los pares de correlación de los aminoácidos con mutaciones reportadas identificados en ambos modelos. (B) Ubicación en la estructura 3D de PS1 de los aminoácidos seleccionados.

Adicionalmente al análisis de correlación, se predijo el efecto de 210 mutaciones sencillas en la estabilidad de PS1 utilizando el método basado en teoría de grafos, mCSM (fig. 11).⁵¹ Cabe notar que este método no toma en cuenta las interacciones de PS1 con los otros componentes de γ -secretasa. De acuerdo con este análisis, se predijo que casi todas las mutaciones reportadas desestabilizan a PS1. Una de las excepciones, que mostró un aumento en la estabilidad de esta subunidad, corresponde a uno de los aminoácidos seleccionados con el análisis de correlación (H214Y).

El efecto de las mutaciones seleccionadas en la estabilidad de PS1 también puede ser descrito mediante un análisis dinámico de correlación cruzada, como se muestra en la siguiente sección.

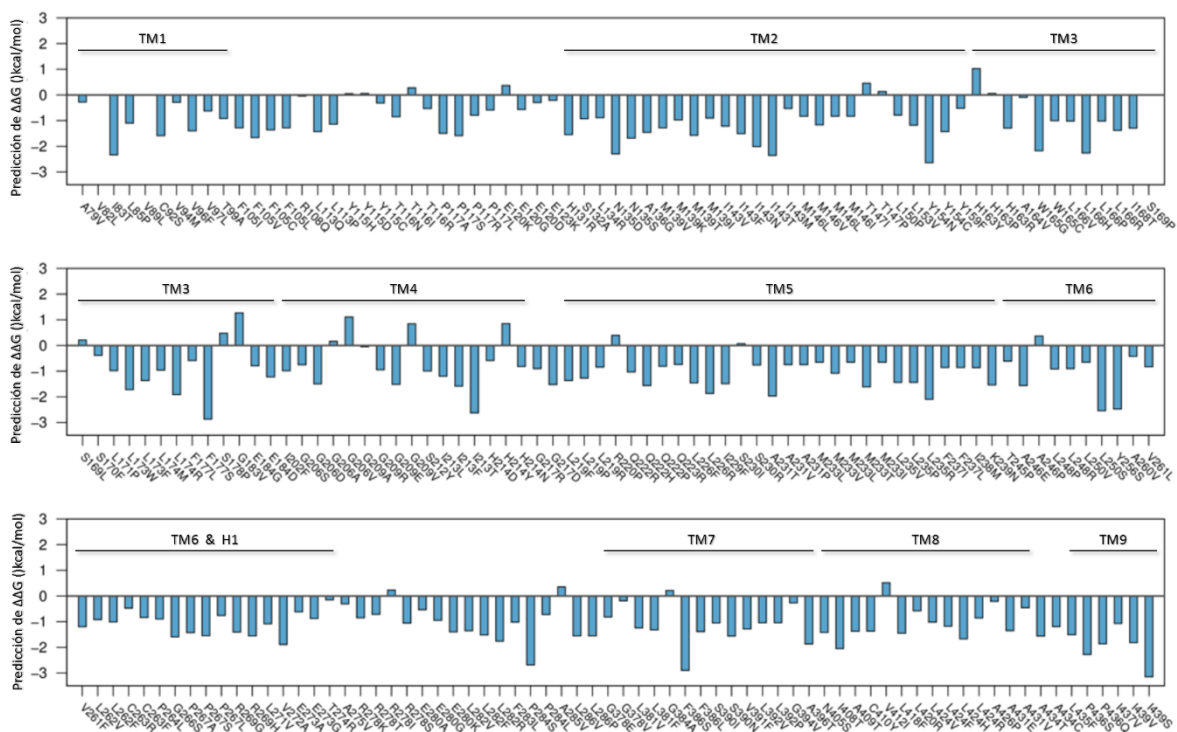


Figura 11. Predicción de estabilidad de la proteína para 210 mutaciones reportadas de PS1 utilizando el método mCSM.

6.2 Perturbaciones en PS1 introducidas por mutaciones sencillas

El estudio de correlación efectuado permitió analizar el comportamiento general del sistema con un modelo de PS1 nativa, sin embargo, es necesario evaluar el efecto de mutaciones sencillas para poder estimar su impacto en la estabilidad de la estructura. Con base en la información proporcionada por el estudio de correlación y resultados experimentales, se seleccionaron las siguientes mutaciones: I83T, R108Q, W165G, I213F/T, H214Y, K239N, A260V, V261F, C263R/F, I408R y L435F.

Se utilizó el modelo de 5FN2 para construir modelos de grano grueso de la proteína nativa y de 13 proteínas con una mutación sencilla, cada una utilizando el módulo Martini Maker del servidor CHARMM-GUI.⁴⁶ Antes de preparar los sistemas para su simulación con DM, se confirmó que el número y/o tipo de partículas CG en el aminoácido mutado era distinto al de la proteína nativa. Se realizaron un total de 20 réplicas de cada sistema, cada una con una trayectoria de 1 μ s.

Debido a que el método de información mutua no es aplicable a modelos de grano grueso, se calculó la matriz de correlación cruzada dinámica (DCCM) de los C α en los últimos 500 ns de cada réplica con el paquete Bio3D.⁴⁹ De manera similar al análisis del estudio de

información mutua, se identificaron los pares de correlación con un índice de correlación en valor absoluto mayor a 0.32 (Promedio + 1STD) y con una distancia entre pares de aminoácidos de al menos 1.5 nm. En este caso fue necesario utilizar el valor absoluto del índice de correlación debido a que Bio3D le asigna un signo dependiendo de si los aminoácidos correlacionados se están acercando (índice positivo) o alejando (índice negativo). Es importante mencionar que el umbral elegido para el índice de correlación fue definido en base a estudios anteriores realizados con este método.⁵⁰

Posteriormente se seleccionaron los aminoácidos presentes en al menos cuatro (Promedio + 3STD) de las veinte réplicas de la proteína nativa. Este límite se utilizó para filtrar los pares de correlación en las réplicas del resto de los modelos de PS1. Por último se cuantificó el número restante de correlaciones de los aminoácidos mutados para estudiar la red de correlación entre aminoácidos y hacer un estudio comparativo entre las correlaciones de la proteína nativa y las de los modelos mutados.

La figura 12 muestra los valores promedio de correlación para el movimiento de cada aminoácido de PS1 en el modelo nativo, con los residuos coloreados de menor (azul) a mayor (rojo) correlación. Se observa que TM2, TM4, la región N-terminal de TM7 y la región C-terminal de TM9 tienen los aminoácidos con mayor grado de correlación. En el caso de TM4 y TM9, esta alta correlación puede deberse a la interacción de estas regiones con las subunidades PEN-2 y APH1-A respectivamente, poniendo en realce su papel en la estabilidad de las regiones N-terminal (TM1 a TM6) y C-terminal (TM7 a TM9) de PS1. La alta correlación observada en TM2 podría estar relacionada con el mecanismo de entrada del sustrato, pues ha sido sugerido que esta hélice actúa como puerta de entrada en este proceso.⁵² Por último, la proximidad de los aminoácidos altamente correlacionados en TM7 al residuo catalítico Asp385 sugiere que éstos podrían estar involucrados en la actividad enzimática.

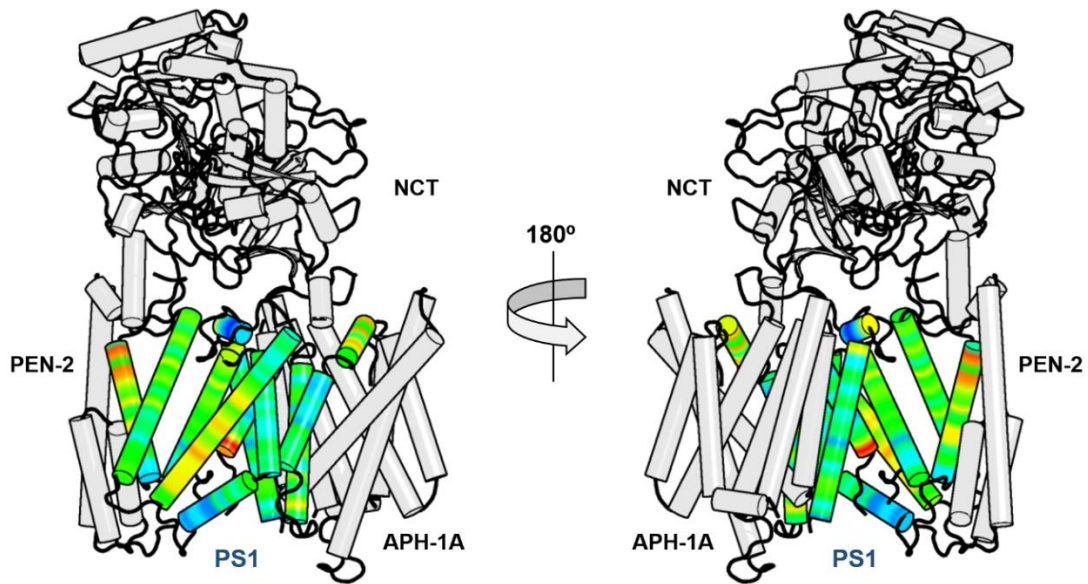


Figura 12. Análisis de correlación cruzada dinámica del modelo de la enzima γ -secretasa nativa coloreada de acuerdo a la correlación total por residuo, de menor (azul) a mayor (rojo) correlación.

Se compararon las correlaciones del modelo de la enzima γ -secretasa nativa con las correlaciones correspondientes a los aminoácidos modificados en los modelos mutados (fig. 13). Con excepción del aminoácido en la posición 263, todos los residuos mutados aumentaron o cuando menos mantuvieron su número de correlaciones al ser comparados con la proteína nativa. Esto confirma que pequeñas modificaciones estructurales pueden alterar el patrón de correlación en el sistema e interferir con la función y estabilidad de la enzima.

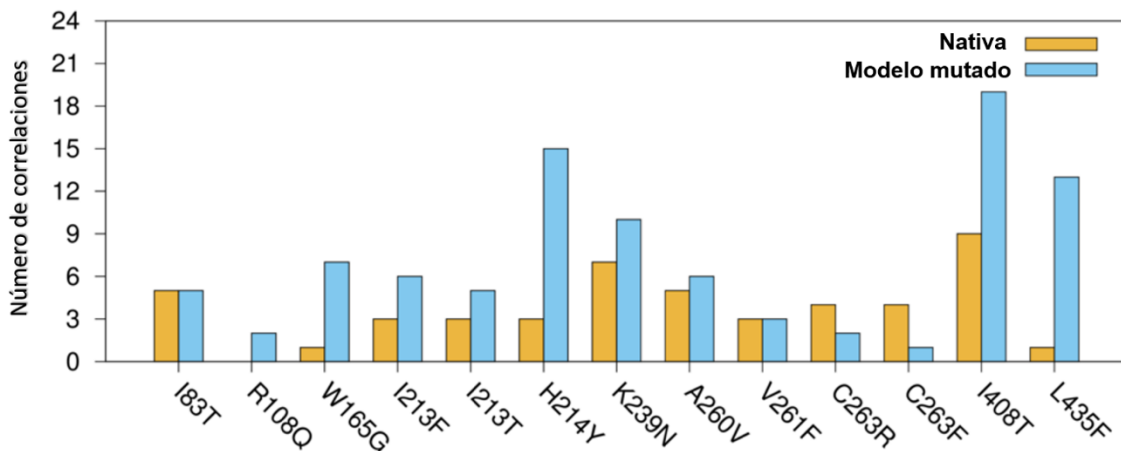


Figura 13. Comparación del número de correlaciones conservadas entre la proteína nativa y los modelos mutados.

La red de correlación en la figura 14A muestra que las correlaciones del modelo nativo no persisten en los modelos con mutaciones reportadas, con excepción de aquellos que se encuentran muy cercanos (213 y 214 o 260, 261 y 263). Por otro lado, el aumento en el número de correlaciones en los modelos mutados se reflejó en un mayor número de nodos en las redes de correlación. Éstos formaron dos redes principales, compartiendo correlaciones con uno o más residuos con mutaciones reportadas (fig. 14B). Es interesante notar que la red de correlación más grande conecta el segmento N-terminal (Y214), el sitio catalítico (N239) y la región C-terminal (T408) de PS1; mientras que la red más pequeña involucra aminoácidos que constituyen la interfaz entre PS1 y APH1-A.

Con base en las observaciones mencionadas, se clasificó el efecto alostérico de las mutaciones de acuerdo a su red de correlación con aminoácidos en las interfaces (1) PS1-APH1-A y (2) PS1-PEN2, (3) el sitio catalítico, (4) el sitio de entrada del sustrato y (5) el sitio de reconocimiento del sustrato (fig. 14C).

Las mutaciones I83T, I408T y L435F mostraron un gran número de aminoácidos correlacionados en TM8 y TM9, sugiriendo una posible alteración de la interacción de PS1 con APH1-A; mientras que las mutaciones que involucran correlaciones en TM3 y TM4 podrían perturbar la interacción entre PS1 y PEN-2 (I213F, I213T, H214Y y W165G). La correlación de algunas mutaciones podría afectar el modo de unión del sustrato al sitio catalítico (K239N y W165G), ya que los aminoácidos con los que están correlacionados se encuentran al centro de la cavidad de PS1. Asimismo, se observaron movimientos correlacionados entre aminoácidos mutados y las TMs 2 y 6 (A260V, V261F, C263F y C263R), que constituyen la posible puerta de entrada del sustrato. La alteración de aminoácidos en esta región podría alterar el mecanismo de entrada del sustrato o su unión para alcanzar el sitio catalítico.⁵² Por último, la mutación I408T mostró un gran número de correlaciones con aminoácidos en las TMs 6 y 9. Se ha propuesto que esta región es el primer sitio de acoplamiento del sustrato, lo que sugiere que esta mutación podría interferir en el reconocimiento del sustrato.⁵³ La mutación R108Q no fue clasificada en ninguna de las categorías propuestas debido a que no presentó correlaciones con aminoácidos cuya función resulte clara.

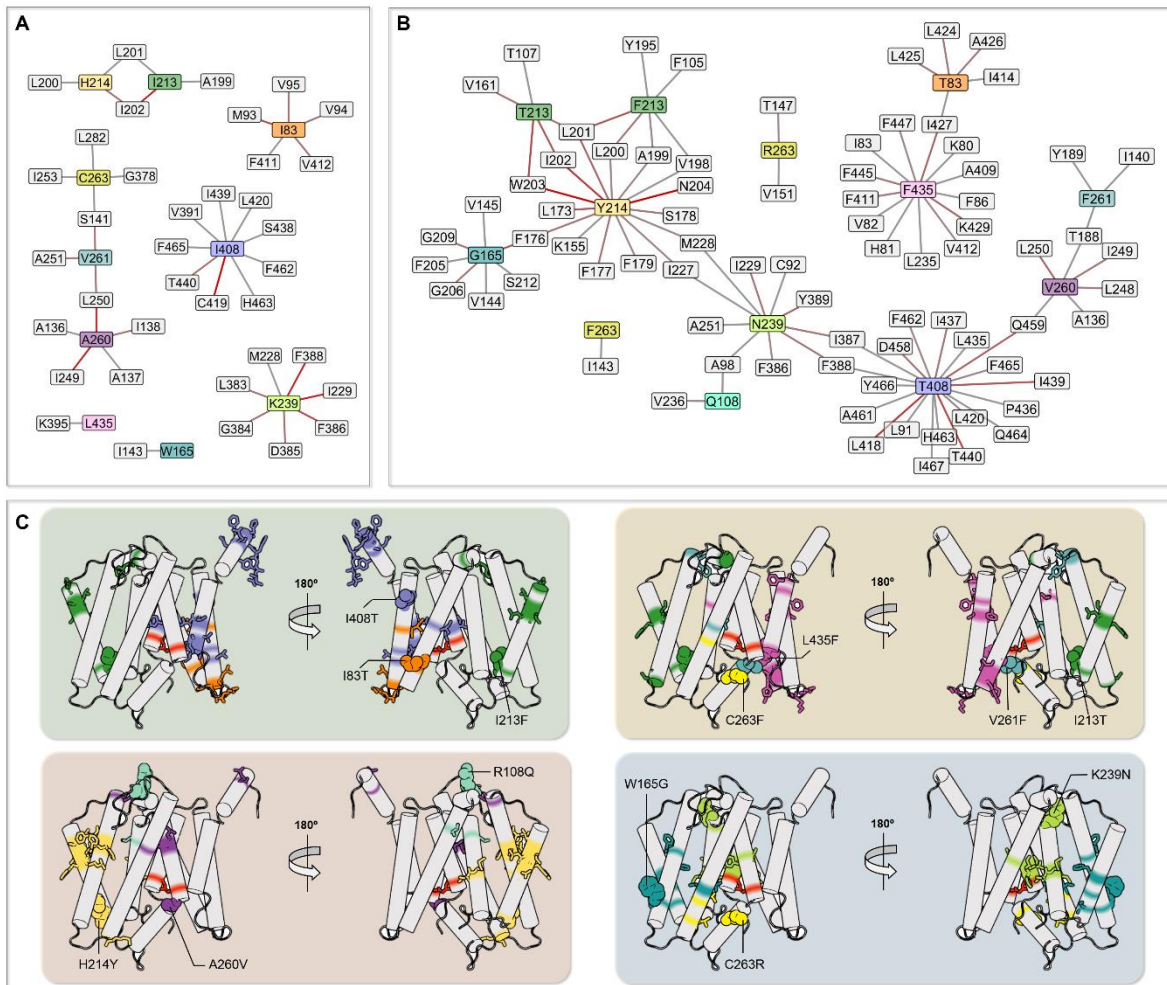


Figura 14. Redes de correlación entre aminoácidos en (A) la proteína nativa y (B) los modelos mutados. Las líneas que unen los pares correlacionados van del gris (correlación promedio) al rojo (alta correlación). (C) Ubicación espacial en la estructura de PS1 de los aminoácidos mutados y los residuos con los que están correlacionados.

6.3 Alteraciones en la estabilidad de PS1

Se analizó la estabilidad de la proteína en las trayectorias de grano grueso de los modelos mutados y de la proteína nativa (fig. 15). En promedio, todas las simulaciones de grano grueso se estabilizaron después de 300 ns con estructuras a 0.44 nm de RMSD respecto a la estructura inicial.

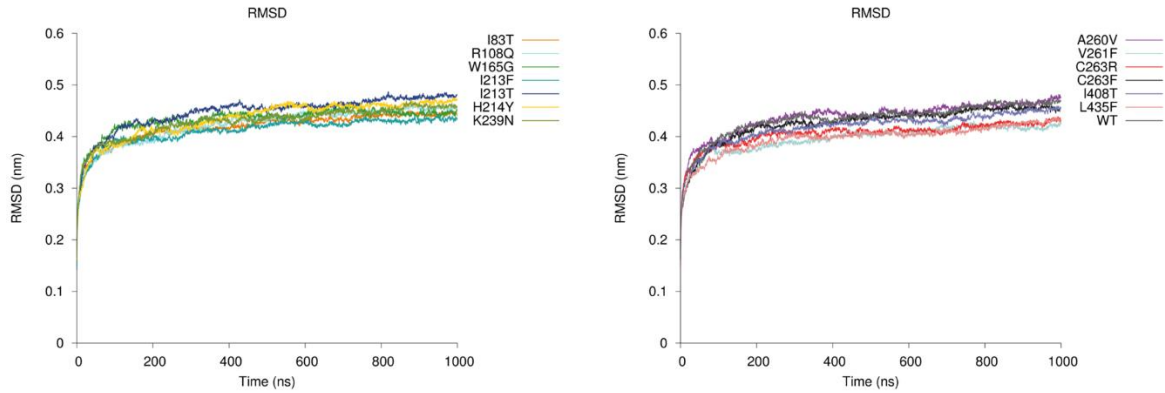


Figura 15. Promedio de la desviación cuadrática media (RMSD) de las 20 réplicas de cada modelo de grano grueso.

La figura 16 muestra el valor absoluto de la diferencia de RMSF (medida de la movilidad promedio por residuo) entre el modelo de PS1 nativa y los modelos con mutaciones sencillas. Las mutaciones que involucran un cambio de isoleucina a treonina (I83T, I213T e I408T), a pesar de que implican un cambio de polaridad, no afectaron la estabilidad de PS1. Similarmente, la mutación H214Y, consistente en un cambio menor ya que ambos aminoácidos son aromáticos, tampoco alteró la estabilidad del componente catalítico de γ -secretasa. En contraste, al mutar I213 a fenilalanina (un aminoácido de mayor tamaño que la isoleucina) se desestabiliza la región carboxilo terminal de PS1. Las mutaciones de los aminoácidos A260, V261 y C263, ubicados en TM5, modificaron el movimiento de los aminoácidos de la zona carboxilo terminal de la proteína, incluyendo a TM6, hélice transmembranal que contiene uno de los residuos catalíticos.

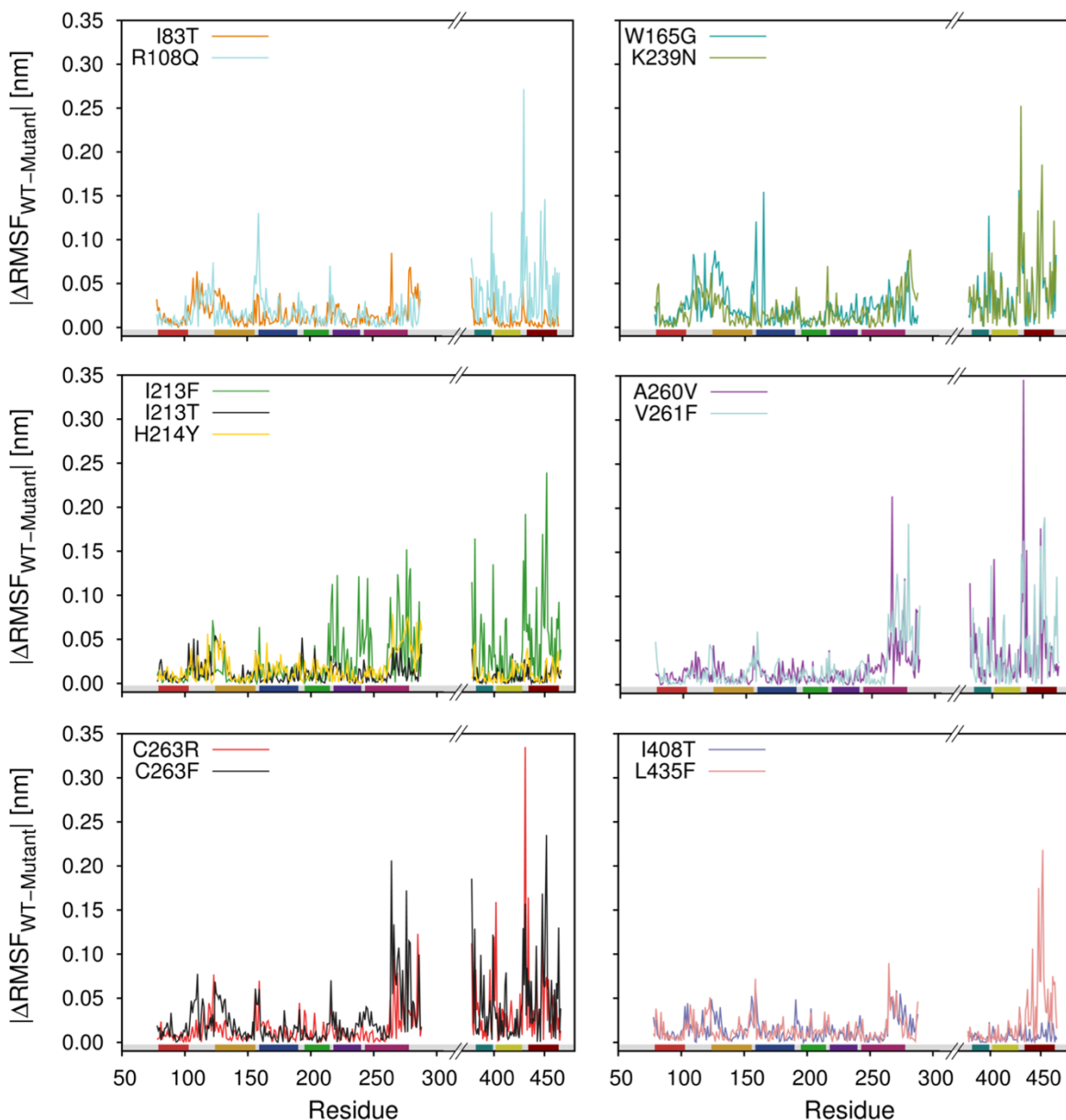


Figura 16. Valor absoluto de la diferencia de RMSF entre el modelo de proteína nativa y cada modelo mutado. En la parte inferior de cada gráfica se encuentra una barra que indica la ubicación de las TMs en la secuencia de PS1.

La mutación W165G, ubicada al inicio de TM3, altera el ensamble conformacional de la región en la que se encuentra, probablemente debido a que la glicina modifica la estabilidad de la hélice. R108Q podría estar desestabilizando la proteína debido a una pérdida de contactos, ya que la glutamina tiene un menor tamaño que la arginina y al contrario de ésta, no tiene carga. La mutación K239 interactúa con el lazo entre TM1 y TM2,

al comparar el modelo de PS1 mutada con el modelo de la proteína nativa se observa que tanto ese lazo como la región carboxilo terminal se desestabilizan. L435F desestabiliza TM9, que es precisamente la hélice en la que se encuentra.

La distribución de conformeros para cada inclinación de las TMs de PS1 en las trayectorias de los modelos con mutaciones sencillas y de la proteína nativa se muestra en la figura 18. Un cambio en la inclinación de las hélices puede modificar directamente la distancia entre los residuos catalíticos de la enzima. El ángulo de inclinación, que se calculó para cada hélice, se definió en función de la recta normal al plano de la interfaz membrana-agua (fig. 17).

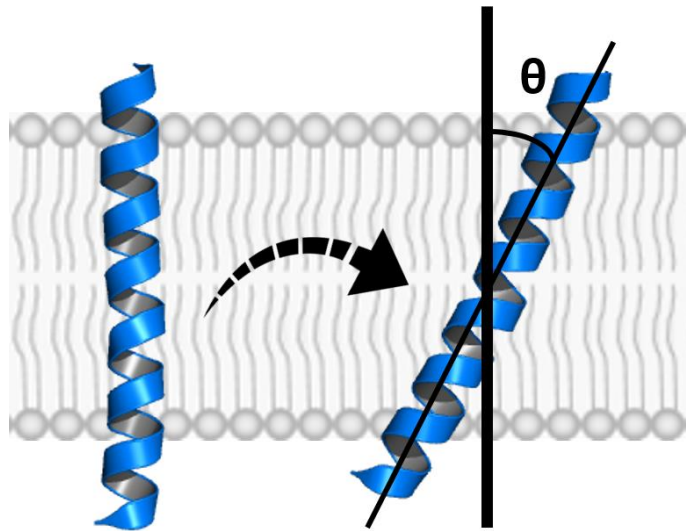


Figura 17. Inclinación de una hélice transmembranal. El ángulo θ se define en función a la recta normal al plano de la interfaz membrana-agua.

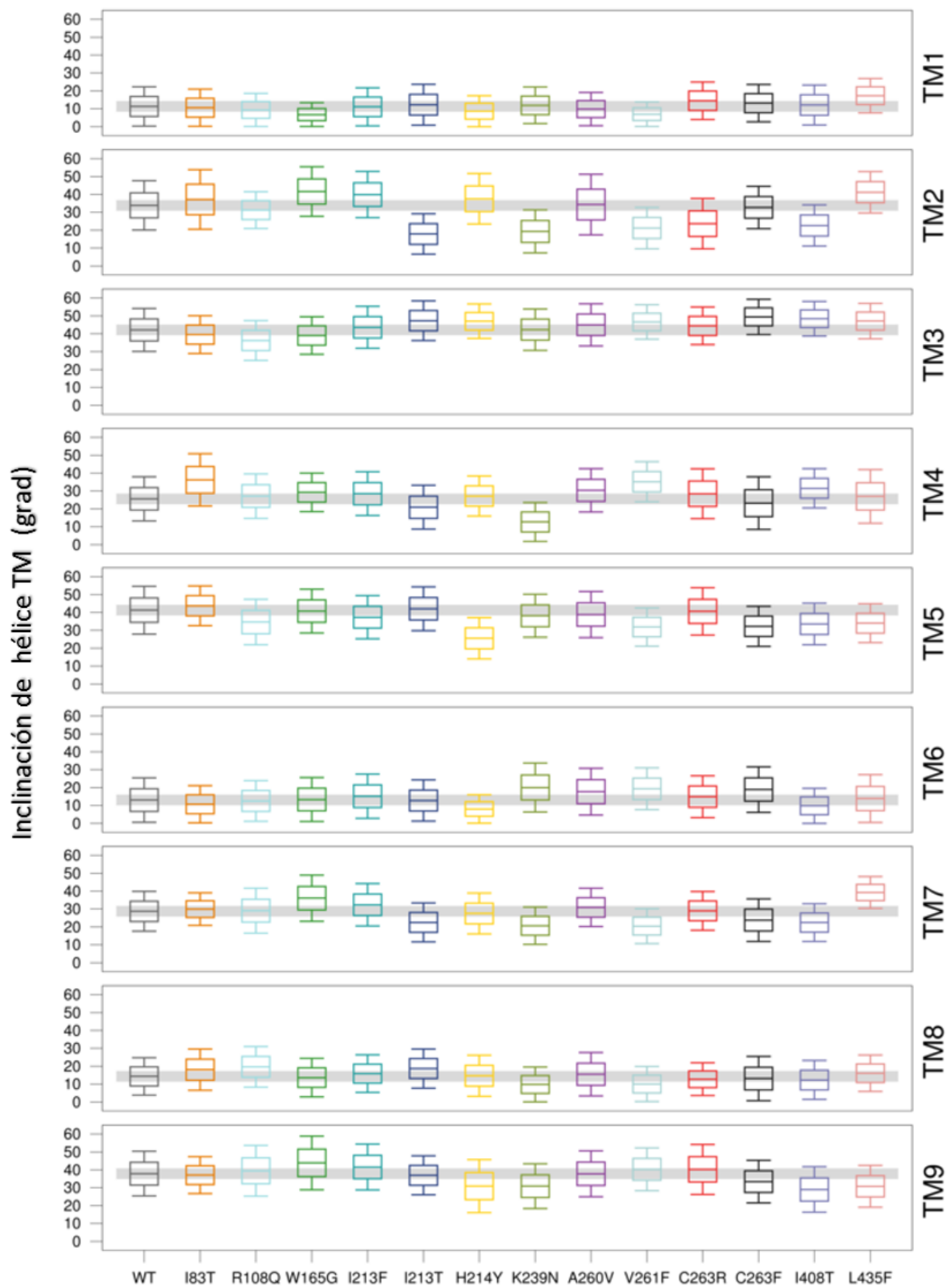


Figura 18. Distribución del ángulo de inclinación de las hélices transmembranales de PS1 de las 20 réplicas para la nativa y cada modelo de mutante obtenido mediante las simulaciones de grano grueso.

En las hélices TM1, TM3 y TM8 se observa la menor varianza en sus ángulos de inclinación. TM2 cubre un amplio espectro de inclinaciones en las trayectorias de las distintas mutantes, lo cual puede deberse a la alta movilidad de esta hélice. La mutación H214Y ubicada en TM4, altera la inclinación de TM5, mientras que la mutación K239 ubicada en TM5, modifica la inclinación de TM4. Ambos comparten correlación con dos aminoácidos (fig. 14B). Las mutaciones que más afectan las inclinaciones de las hélices que forman el sitio catalítico son H214Y, K239N, V261F y C263F (TM6) y W165G, I213T, K239N, V261F, I408T y L435F (TM7). A pesar de que las dos mutaciones C263R y C263F afectan al mismo residuo, el cambio de cisteína a fenilalanina afecta más las inclinaciones de las hélices que el cambio a arginina. Debido a que la arginina tiene carga, podría estar interactuando con Asp257 mientras que la fenilalanina, al ser un residuo hidrofóbico y estar ubicado en la cara exterior de TM6, podría preferir interactuar con los lípidos.

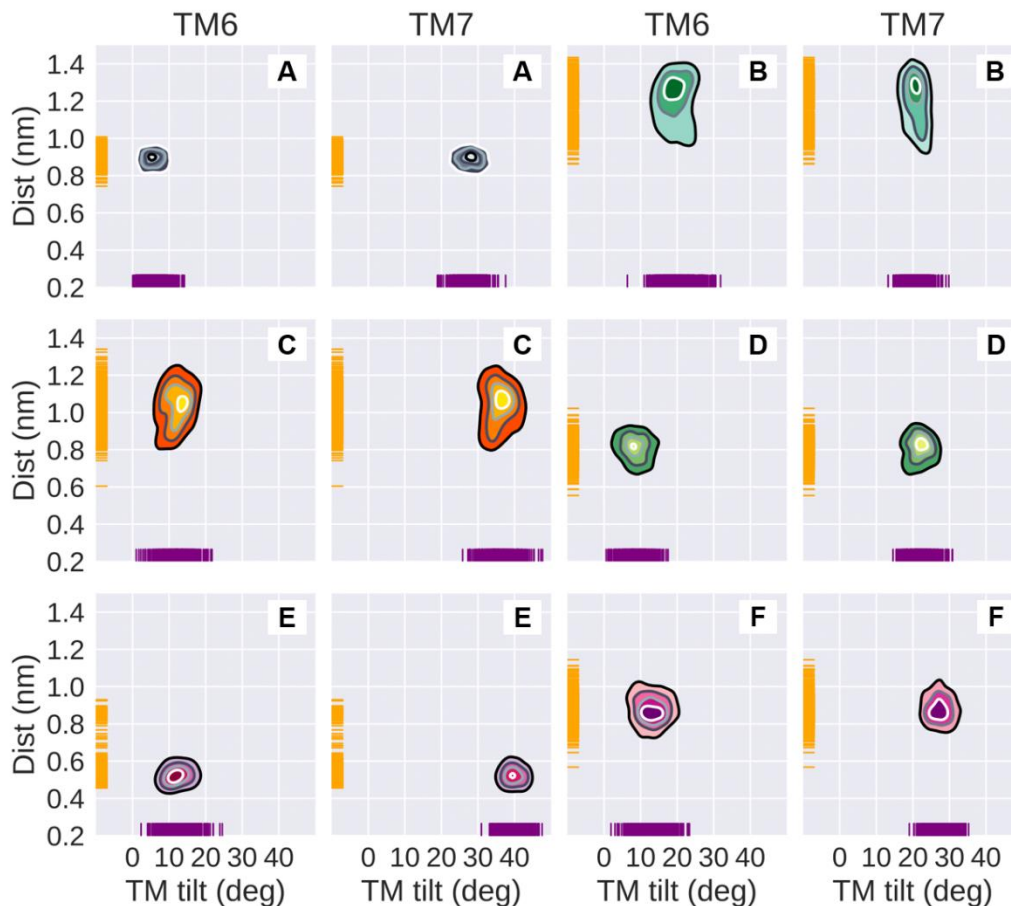


Figura 19. Histograma 2D de la distancia entre los aspartatos catalíticos y las inclinaciones de TM6 y TM7 en los modelos: A) H214Y, B) K239N, C) W165G, D) I408T, E) L435F y F) proteína nativa.

La figura 19 muestra un histograma 2D de las inclinaciones de las hélices TM6 y TM7 de las mutantes que más alteraron estas inclinaciones, así como la distancia entre los residuos catalíticos durante las simulaciones. Las mutantes que se incluyeron en este análisis se eligieron debido a que nuestros resultados anteriores mostraron que sus mutaciones se encuentran en los nodos más conectados de las redes de correlación (fig. 14B).

Las mutantes H214Y e I408T mantuvieron la misma separación entre los residuos catalíticos que la proteína nativa, mientras que las mutantes K239N y W165G la aumentaron. Para que la enzima pueda llevar a cabo la ruptura de los enlaces peptídicos de sus sustratos, es necesario que los aspartatos catalíticos se encuentren a una distancia de puente de hidrógeno (~ 0.2 nm). La proteína nativa se encuentra en una configuración inactiva debido a que sus dos residuos catalíticos se encuentran desprotonados. En la mutación L435F se observa una disminución en la distancia entre los dos aspartatos. La mutación podría estar provocando que la proteína se mantenga en una conformación “cerrada”, lo que podría modificar el proceso de corte del sustrato al sitio catalítico.

7. Conclusiones

Fue posible identificar redes de correlación en PS1 mediante herramientas de dinámica molecular multi-escala. Los aminoácidos con correlaciones grandes y largas distancias entre pares de correlación se encontraron distribuidos a lo largo de toda la proteína, lo que coincide con la amplia distribución de mutaciones reportadas causantes de FAD.

8. Perspectivas

- Estudiar y calcular con metodologías computacionales el estado de protonación de los aminoácidos aspartatos catalíticos de la enzima γ -secretasa.
- Realizar simulaciones multi-escala de los modelos mutados de PS1 con uno de los dos aspartatos catalíticos protonado.
- Evaluar cómo la composición de los lípidos y los componentes de la membrana (por ejemplo la presencia de colesterol y esfingolípidos) influyen o modifican la dinámica y las redes de correlación de la enzima γ -secretasa.
- Estudiar con metodologías de dinámica molecular el mecanismo de entrada del sustrato al sitio activo de la enzima.

Bibliografía

- 1 Instituto Nacional de Neurología y Neurocirugía, <http://www.innn.salud.gob.mx/interior/atencionapacientes/padecimientos/alzheimer.html>, (accessed 2 May 2017).
- 2 L. Mucke, *Nature*, 2009, **461**, 895–897.
- 3 F. M. LaFerla, K. N. Green and S. Oddo, *Nat. Rev. Neurosci.*, 2007, **8**, 499–509.
- 4 ALZFORUM, <http://www.alzforum.org/>, (accessed 12 April 2017).
- 5 D. J. Selkoe, J. Hardy, D. Selkoe and J. Hardy, *EMBO Mol Med*, 2016, **8**, 595–608.
- 6 R. Vassar, *J. Mol. Neurosci.*, 2004, **23**, 105–114.
- 7 H. Cai, Y. Wang, D. McCarthy, H. Wen, D. R. Borchelt, D. L. Price and P. C. Wong, *Nat. Neurosci.*, 2001, **4**, 233–234.
- 8 J. Kim, L. Onstead, S. Randle, R. Price, L. Smithson, C. Zwizinski, D. W. Dickson, T. Golde and E. McGowan, *J. Neurosci.*, 2007, **27**, 627–633.
- 9 K. A. Ball, A. H. Phillips, D. E. Wemmer and T. Head-Gordon, *Biophys. J.*, 2013, **104**, 2714–2724.
- 10 R. Singh, A. Barman and R. Prabhakar, *J. Phys. Chem. B*, 2009, **1**, 2990–2999.
- 11 A. J. Beel and C. R. Sanders, *Cell. Mol. Life Sci.*, 2008, **65**, 1311–1334.
- 12 G. Struhl and A. Adachi, *Mol. Cell*, 2000, **6**, 625–636.
- 13 W. T. Kimberly, M. J. LaVoie, B. L. Ostaszewski, W. Ye, M. S. Wolfe and D. J. Selkoe, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 6382–7.
- 14 B. De Strooper, *Neuron*, 2003, **38**, 9–12.
- 15 et al. E. I. Rogaev, R. Sherrington, E. A. Rogaeva, *Nature*, 1995, **376**, 775–778.
- 16 et al. R. Sherrington, E. I. Rogaev, Y. Liang, *Nature*, 1995, **375**, 754–760.
- 17 X. Bai, C. Yan, G. Yang, P. Lu, D. Ma, L. Sun, R. Zhou, S. H. W. Scheres and Y. Shi, *Nature*, 2015, **525**, 212–7.
- 18 O. Holmes, S. Paturi, D. J. Selkoe and M. S. Wolfe, *Biochemistry*, 2014, **53**, 4393–4406.
- 19 X.-C. Bai, E. Rajendra, G. Yang, Y. Shi and S. H. Scheres, *Elife*, 2015, **4**, e11182.

- 20 R. Aguayo-Ortiz and L. Dominguez, *Isr. J. Chem.*, 2016, 1–13.
- 21 G. Zhao, J. Tan, G. Mao, M. Z. Cui and X. Xu, *J. Neurochem.*, 2007, **100**, 1234–1246.
- 22 Y. Qi-Takahara, *J. Neurosci.*, 2005, **25**, 436–445.
- 23 C. Marc, AD&FTDMDDB, <http://www.molgen.ua.ac.be/ADmutations/>, (accessed 18 April 2017).
- 24 R. O'Brien and P. Wong, *Annu. Rev. Neurosci.*, 2011, **1987**, 185–204.
- 25 B. De Strooper, *EMBO Rep.*, 2007, **8**, 141–6.
- 26 D. Rapaport, *The art of molecular dynamics simulation*, Cambridge University Press, New York, Second., 2004.
- 27 D. W. Borhani and D. E. Shaw, *J. Comput. Aided. Mol. Des.*, 2012, **26**, 15–26.
- 28 S. V. Bennun, M. I. Hoopes, C. Xing and R. Faller, *Chem. Phys. Lipids*, 2009, **159**, 59–66.
- 29 S. J. and T. D. P. Marrink, *Chem Soc Rev*, 2013, 6801–6822.
- 30 G. van Meer, D. R. Voelker and G. W. Feigenson, *Nat. Rev. Mol. Cell Biol.*, 2008, **9**, 112–124.
- 31 P. Heftberger, B. Kollmitzer, F. A. Heberle, J. Pan, M. Rappolt, H. Amenitsch, N. Kučerka, J. Katsaras and G. Pabst, *J. Appl. Crystallogr.*, 2014, **47**, 173–180.
- 32 E. Winkler, F. Kamp, J. Scheuring, A. Ebke, A. Fukumori and H. Steiner, *J. Biol. Chem.*, 2012, **287**, 21326–21334.
- 33 Mathsisfun, <https://www.mathsisfun.com/data/correlation.html>, (accessed 5 May 2017).
- 34 G. Chinae, G. Padron, R. W. W. Hooft, C. Sander and G. Vriend, *Proteins Struct. Funct. Bioinforma.*, 1995, **23**, 415–421.
- 35 M. A. Lomize, A. L. Lomize, I. D. Pogozheva and H. I. Mosberg, *Bioinformatics*, 2006, **22**, 623–625.
- 36 J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. B. Iii, A. D. M. Jr, B. Klauda and W. Im, *J. Chem. Theory Comput.*, 2016, **12**, 405–413.
- 37 E. L. Wu, X. Cheng, S. Jo, H. Rui and K. C. Song, *J. Comput. Chem.*, 2014, **35**, 1997–2004.
- 38 P. Osenkowski, W. Ye, R. Wang, M. S. Wolfe and D. J. Selkoe, *J. Biol. Chem.*, 2008, **283**,

22529–22540.

- 39 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
- 40 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindah, *SoftwareX*, 2015, **1–2**, 19–25.
- 41 R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and J. MacKerell, Alexander D., *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.
- 42 R. Kumari, R. Kumar, O. Source, D. Discovery and A. Lynn, *J. Chem. Inf. Model.*, 2014, **54**, 1951–1962.
- 43 N. Michaud-agrawal, E. J. Denning, T. B. Woolf and O. Beckstein, *J. Comput. Chem.*, 2011, **32**, 2319–2327.
- 44 K. Gunasekaran, B. Ma and R. Nussinov, *Proteins Struct. Funct. Genet.*, 2004, **57**, 433–443.
- 45 C. L. McClendon, G. Friedland, D. L. Mobley, H. Amirkhani and M. P. Jacobson, *J. Chem. Theory Comput.*, 2009, **5**, 2486–2502.
- 46 Y. Qi, H. I. Ingólfsson, X. Cheng, J. Lee, S. J. Marrink and W. Im, *J. Chem. Theory Comput.*, 2015, **11**, 4486–4494.
- 47 S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. De Vries, *J. Phys. Chem. B*, 2007, **111**, 7812–7824.
- 48 S. O. Yesylevskyy, L. V. Schäfer, D. Sengupta and S. J. Marrink, *PLoS Comput. Biol.*, 2010, **6**, 1–17.
- 49 B. J. Grant, A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon and L. S. D. Caves, *Bioinformatics*, 2006, **22**, 2695–2696.
- 50 L. Skjærven, X.-Q. Yao, G. Scarabelli and B. J. Grant, *BMC Bioinformatics*, 2014, **15**, 399.
- 51 D. E. V Pires, D. B. Ascher and T. L. Blundell, *Bioinformatics*, 2014, **30**, 335–342.
- 52 A. Fukumori and H. Steiner, 2016, **35**, 1628–1643.
- 53 S. Takagi-Niidome, T. Sasaki, S. Osawa, T. Sato, K. Morishima, T. Cai, T. Iwatsubo and T. Tomita, *J. Neurosci.*, 2015, **35**, 2646–56.



Cite this: DOI: 10.1039/c7sc00980a

Characterizing the structural ensemble of γ -secretase using a multiscale molecular dynamics approach†

Rodrigo Aguayo-Ortiz,^a Cecilia Chávez-García,^a John E. Straub^b
and Laura Dominguez^{†*}

γ -Secretase is an intramembrane-cleaving aspartyl protease that plays an essential role in the processing of a variety of integral membrane proteins. Its role in the ultimate cleavage step in the processing of amyloid precursor protein to form amyloid- β ($A\beta$) peptide makes it an important therapeutic target in Alzheimer's disease research. Significant recent advances have been made in structural studies of this critical membrane protein complex. However, details of the mechanism of activation of the enzyme complex remain unclear. Using a multiscale computational modeling approach, combining multiple coarse-grained microsecond dynamic trajectories with all-atom models, the structure and two conformational states of the γ -secretase complex were evaluated. The transition between enzymatic state 1 and state 2 is shown to critically depend on the protonation states of the key catalytic residues Asp257 and Asp385 in the active site domain. The active site formation, related to our γ -secretase state 2, is observed to involve a concerted movement of four transmembrane helices from the catalytic subunit, resulting in the required localization of the catalytic residues. Global analysis of the structural ensemble of the enzyme complex was used to identify collective fluctuations important to the mechanism of substrate recognition and demonstrate that the corresponding fluctuations observed were uncorrelated with structural changes associated with enzyme activation. Overall, this computational study provides essential insight into the role of structure and dynamics in the activation and function of γ -secretase.

Received 2nd March 2017
Accepted 5th June 2017

DOI: 10.1039/c7sc00980a

rsc.li/chemical-science

Introduction

γ -Secretase is a membrane-embedded aspartyl protease that cleaves different integral membrane proteins within the lipid bilayer including amyloid precursor protein (APP), Notch, N-cadherin and ErbB4.¹ APP is initially cleaved by β -secretase to release the APP C-terminal fragment (APP-C99), which is subsequently processed by γ -secretase to yield a variety of amyloid- β ($A\beta$) peptides of different lengths. The pattern of cleavage of APP-C99 results in a primary isoform, $A\beta_{40}$ consisting of 40 amino acids, in addition to minor isoforms including $A\beta_{38}$ and $A\beta_{42}$.² Aberrant accumulation of $A\beta_{42}$ over $A\beta_{40}$ has been associated with the formation of amyloid- β plaques in the brain of Alzheimer's disease (AD) patients.^{2,3} As such, developing an understanding of the mechanism of cleavage of APP-C99 by γ -secretase is an important goal for the field.

The mature γ -secretase consists of four components (Fig. 1A): presenilin 1 (PS1), presenilin enhancer 2 (PEN-2),

anterior pharynx-defective 1A (APH-1A) and nicastrin (NCT).⁴⁻⁶ PS1 is the catalytic component of γ -secretase. It contains nine transmembrane helices (TMs) organized into a horseshoe-shaped structure with two catalytic Asp residues located at TM6 (Asp257) and TM7 (Asp385) within the convex side of the

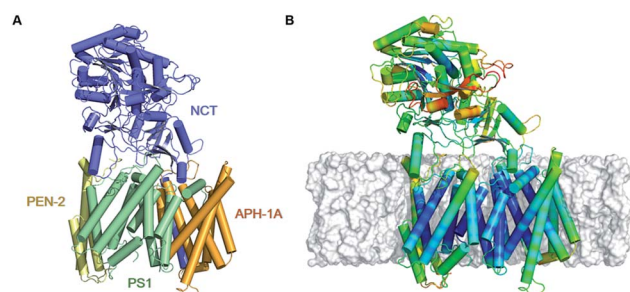


Fig. 1 (A) Depiction of the all atom model of γ -secretase, derived from the 5FN2 PDB structure, colored by its subunits: PS1, NCT, PEN-2, and APH-1A in green, blue, yellow, and orange, respectively. (B) Fluctuation analysis of 5FN2-derived atomistic model of γ -secretase in POPC bilayer (gray) color-coded by the normalized per-residue root mean square fluctuation (RMSF) from more flexible (red) to less flexible (blue). The analysis was performed during the last 500 ns of the simulation.

^aDepartamento de Físicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, Mexico City, 04510, Mexico. E-mail: lauradd@unam.mx

^bDepartment of Chemistry, Boston University, Boston, Massachusetts 02215, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7sc00980a



protein surface.^{7–10} Autoproteolysis of the intracellular loop connecting TM6 and TM7 leads to the formation of the **PS1** N-terminal (NTE, TMs 1–6) and C-terminal fragments (CTF, TMs 7–9) that interact with **PEN-2** and **APH-1A**, respectively.⁹ **PEN-2** consists of three TMs, of which TM1 and TM2 form a re-entrant loop extending halfway through the membrane from the intracellular side.^{7,11} The association of **PEN-2** with **PS1** has been related to the autocatalytic maturation of **PS1** and γ -secretase activity. **APH-1A** consists of seven TMs helices and a C-terminal juxtamembrane region.^{12,13} Several experimental studies suggest that this component is required for proper γ -secretase assembly.^{14,15} **NCT** is a transmembrane glycoprotein with a large N-terminal extracellular domain (ECD) and a single TM helix located in its C-terminal region.^{16,17} The globular ECD is comprised of a large and a small lobe composed of α -helices and β -strands.^{12,18,19} It has been proposed that Glu333 and Tyr337, located at **NCT**'s large lobe, interact with the substrate's extracellular domain and play a critical role in substrate recognition.^{17,20} These residues are buried in a hydrophilic pocket covered by a lid formed by the small lobe.¹² It is thought that a rotation of **NCT**'s large lobe around a central pivot may cause lid opening, exposing the substrate-binding site.¹⁷ However, mutagenesis studies involving Glu333 contradict this theory, suggesting that Glu333 is involved in the maturation assembly of the complex rather than in substrate recognition.^{21,22}

Novel high-resolution cryo-EM structures of human γ -secretase have been resolved.^{7,12,19} In the first proposed structure, Shi and coworkers⁷ reported the overall architecture of the complex at 4.5 Å resolution (PDB ID: 4UIS). Nevertheless, due to the limited resolution, partial assignment of the TMs side chains was only possible through sequence homology modeling of **PS1** using a presenilin homologue (PSH) (PDB ID: 4HYG).⁸ Subsequently, Bai, *et al.*¹² obtained an atomic-level γ -secretase structure with 3.4 Å resolution employing cryo-EM single particle analysis (PDB ID: 5A63). In this structure, the main-chain connectivity and side-chains were mostly displayed, facilitating identification of new interactions between the TMs, the presence of the juxtamembrane region of **APH-1A**, and a complete structure of the **NCT** ECD.

The structure of TM2 and the N-terminal region of TM6 in **PS1** are typically not resolved due to the high flexibility of these domains. Importantly, four new γ -secretase structures exhibiting varying conformational states in this critical region have been resolved (PDB IDs: 5FN2, 5FN3, 5FN4 and 5FN5).¹⁹ The first was obtained in complex with the dipeptidic inhibitor *N*-[*N*-(3,5-difluorophenacetyl)-*L*-alanyl]-*S*-phenylglycine *t*-butyl ester (DAPT); however, the structure of the inhibitor was not assigned. The second and third γ -secretase structures were solved in complex with peptide fragments. Interestingly, the authors suggest that the peptide fragment found in 5FN3, located between **PS1** TM2, TM3 and TM4, belongs to the N-terminal region of APP. Finally, the last structure did not appear to be complexed with any external agent, and as in the case of PDB ID: 5A63, it was not possible to resolve TM2.

A previous molecular dynamics (MD) study examined the dynamic properties and activation of a single human **PS1**

subunit embedded in a variety of membrane lipid compositions, with an initial **PS1** conformation derived from homology modeling.²³ However, in the absence of the **PEN-2**, **APH-1A**, and **NCT** subunits, the simulations exhibited large instabilities and flexibility in the TMs. More recently, Han and coworkers²⁴ explored the initial substrate binding site of the transmembrane region of γ -secretase using a multiscale MD approach. In that work, the authors demonstrated the importance of using atomistic and coarse grained models to assess the behaviour of the TMs of forming the complex. Moreover, their results suggest that TM2/6/9 (including the PAL motif) constitute the initial APP-C99 binding site.

Despite these recent advances, many essential aspects of the structure and dynamics of the multicomponent γ -secretase complex remain undescribed. In this study, we employed a multiscale simulation approach that combines atomistic and coarse-grained models starting from a variety of cryo-EM structures to explore the dynamic structural ensemble of γ -secretase embedded in a POPC lipid bilayer. Conformational changes are analyzed using order parameters that characterize the essential dynamics of the enzyme complex, including the transition between active and inactive functional states. Our study provides critical insight into (1) the nature of large-scale conformational transitions in the γ -secretase complex, (2) the identification of two conformational states of **PS1**, constituting a mechanism for its activation, and (3) the characterization of **NCT** ECD mobility and its relationship to the observed **PS1** states.

Methods

Model preparation

Two cryo-EM structures of γ -secretase (PDB IDs: 5FN2 and 5FN3) resolved at 4.2 and 4.1 Å resolution, respectively, were used as initial 3D coordinates to perform the multiscale MD simulations of the γ -secretase complexes. The missing side-chains of the structures were completed using the WHAT IF web server.²⁵ The spatial arrangement of the γ -secretase models in the POPC lipid bilayer was assigned with the Orientation of Proteins in Membranes (OPM) web server.²⁶

Molecular dynamics simulations

We used a multiscale computational approach (Table 1) to characterize the structure and dynamics of the γ -secretase enzyme complex embedded in a POPC bilayer, combining all-atom and CG representations, using CHARMM36 and MARTINI force field models, respectively, for the protein, membrane, and solvent environments. We simulated two distinct γ -secretase models, each in three different protonated states: two models with one of the two catalytic aspartic acid residues protonated (Asp257 and Asp385) and a third model with both catalytic residues unprotonated.

All-atom molecular dynamic simulations. Each protein was embedded in a palmitoyl oleoyl phosphocholine (POPC) membrane using the CHARMM-GUI membrane builder.^{27,28} The model of the 5FN2 derived system consisted of the



Table 1 Overview of all-atom and CG simulations employed in this study, color-coded according to the system model

System	Force field	S. time	# replicas
5FN2	CHARMM-36	1 μ s	1
5FN3	CHARMM-36	1 μ s	1
5FN2	MARTINI v2.2	1 μ s	50
5FN2 Asp-H ₂₅₇	MARTINI v2.2	1 μ s	50
5FN2 Asp-H ₃₈₅	MARTINI v2.2	1 μ s	50
5FN3	MARTINI v2.2	1 μ s	50
5FN3 Asp-H ₂₅₇	MARTINI v2.2	1 μ s	50
5FN3 Asp-H ₃₈₅	MARTINI v2.2	1 μ s	50
5FN2	MARTINI v2.2	5 μ s	1
5FN2	MARTINI + ElnDyn	5 μ s	1
Active Asp-H ₃₈₅	CHARMM-36	100 ns	1
Inactive Asp-H ₃₈₅	CHARMM-36	100 ns	1

reconstructed protein complex, 320 POPC lipids, 46 078 water particles, and 143 Cl⁻ and 148 Na⁺ ions; the 5FN3 derived system contained the reconstructed protein complex, 326 POPC lipids, 46 827 water particles, and 145 Cl⁻ and 150 Na⁺ ions. The systems were energy minimized and equilibrated with NVT and NPT dynamics. Equilibration of each simulation was followed by a 1 μ s production run with a time step of 2 ps. The temperature was set to 310 K using the Nose–Hoover coupling thermostat algorithm and the pressure was set to 1.0 bar using the semi-isotropic Parrinello–Rahman barostat algorithm. The Lennard-Jones potential was truncated using a shift function between 0.9 and 1.2 nm. Full electrostatic interactions were calculated between 0 and 1.2 nm, after which the electrostatic interactions were calculated using the Particle Mesh Ewald (PME) approach. Neighbor lists were updated every 20 steps and bonds involving hydrogens were constrained using the Linear Constraint Solver (LINCS) algorithm.²⁹ The simulations were performed using GROMACS 5.0.6 (ref. 30) with the CHARMM36 force field³¹ and the TIP3P water model. The apolar contribution to the energy assigned to interaction of PS with adjacent components was calculated with the *g_mmpbsa.py*³² script. Finally, the MD simulations were analyzed using the inbuilt GROMACS tools and the MAnalysis libraries³³ for python and the images were generated using Bendix,³⁴ VMD v1.9.2 (ref. 35) and PyMOL v0.9 (ref. 36).

Coarse-grained (CG) model simulations. Three coarse-grained systems were built for each of the 5FN2 and 5FN3 derived γ -secretase models: two models with one of the two catalytic aspartic acid residues protonated (Asp257 and Asp385) and a third model with both catalytic residues in an unprotonated state. Each model protein complex was embedded in a POPC bilayer using the CHARMM-GUI Martini bilayer maker^{37,38} with the polarizable water model of the Martini v2.2 force field.³⁹ The temperature was set to 310 K using V-rescale coupling and the pressure was set at 1.0 bar with a semi-isotropic Berendsen coupling. Trajectories for 50 replicas of the protonated and unprotonated states of the 5FN2 and 5FN3

derived models were performed for 1 μ s of NPT dynamics for a total of 150 μ s of dynamics for the membrane-embedded γ -secretase complex, time evolution of the CG trajectories suggested a broad sampling of the conformational space on each simulation (Fig. S4[†]).

Additionally, in order to evaluate the stability of our simulations and validate our methodology and results, two longer simulations of 5 μ s were performed for the unprotonated γ -secretase complex using our CG Martini model and a CG Martini with an Elastic Network in Dynamics (CG-ElnDyn).⁴⁰ The systems were prepared following the same methodology described above and results are presented in the ESI.[†]

The most representative structures of the state 1 and state 2 **PS1** conformers from the 5FN2 derived CG simulations were selected using the GROMACS cluster tool with a root mean square deviation (RMSD) threshold of 2.0 Å. These structures were back-mapped from coarse-grained (MARTINI force field) models to all-atom (CHARMM36 force field) models using the script *backward.py*.⁴¹ The final systems were simulated for 100 ns of MD following the previously described all-atom simulation protocols.

Finally, three structures were selected from the 5FN2 derived CG simulations with protonated Asp385 to represent the compact, intermediate, and extended conformations of the γ -secretase state models. The density maps of the structures were generated with the *molmap* command of UCSF Chimera v1.11 (ref. 42) with a resolution of 8.0 Å.

Results

We employed multiscale molecular dynamics (MD) simulations, combining coarse-grained (CG) and all-atom models, to investigate different conformational states of the γ -secretase complex. The initial models were derived from the available experimental information (PDB IDs: 5FN2 and 5FN3). Our initial all-atom models were simulated for 1 μ s in order to relax the structure into a POPC lipid bilayer environment and derive secondary and tertiary structural information required to build valid CG models. Using the constructed CG model of the full enzyme complex in a POPC lipid bilayer, 50 independent 1 μ s trajectories were simulated in order to thoroughly sample the dynamics on a time scale sufficient to observe critical large-scale structural transitions. Subsequently, all-atom models of the γ -secretase complex were constructed based on conformational distributions derived from the extensive CG simulations. Finally, the all-atom models of the bilayer-embedded γ -secretase complex were simulated for 100 ns each, in order to assess the overall behavior of the complex at the atomic scale.

In the following sections, we provide a detailed picture of (1) the relative roles of the enzyme subunits in stabilizing the structure of the **PS1** catalytic subunit, (2) the activation mechanism of the enzyme resulting from protonation of the catalytic aspartic acid residues, and (3) the characterization of the principal motions of **NCT** ECD related to transitions between two conformational states of the enzyme complex. Overall, this study provides the first complete picture of the γ -secretase



complex and the relative role of the enzyme subunits in the activation mechanism.

Protein dynamics and inter-subunit interactions of γ -secretase

γ -Secretase is a highly stable protein complex, the maturation and activation of which is related to the assembly of its four components.¹⁸ To develop a detailed atomistic description and evaluate the general dynamics of γ -secretase and its subunit interactions, we performed all-atom MD simulations of 5FN2 and 5FN3 derived models in a POPC lipid bilayer using the CHARMM36 force field (Fig. S1 and S2†). The initial structures display two important differences: (1) 5FN3 lacks the carboxyl terminal fragment of **PS1** TM6 (264–288) while 5FN2 contains the carboxyl terminal fragment of **PS1** TM6 and (2) 5FN3 presents a larger distance between the catalytic Asp residues (5.06 Å) compared to a shorter distance (3.89 Å) observed in 5FN2. Fig. 1B shows the all-atom fluctuation analysis of the 5FN2 structure model, depicting the high mobility of the **NCT** extracellular fragment and relatively low mobility of its TMs. **PEN-2**, TM5–7 of **APH1**, and TM2 of **PS1** exhibited higher mobility due to their exposure to lipids and fewer protein–protein contacts. In contrast, the **PS1** TMs in close contact with **PEN-2** and **APH-1A** exhibited low mobility and stable protein–protein interactions during the all-atom simulations. These findings are consistent with previous experimental observations suggesting that **PEN-2** and **APH-1A** play key roles in the catalytic subunit stabilization and activation.⁴³ Our results also provide an explanation for the structural instabilities observed in an earlier simulation study of the isolated **PS1** subunit.²³

Analysis of the apolar contributions to the inter-subunit binding free energies of the four γ -secretase components suggests that **PS1**-TM1, **PS1**-TM8 and **PS1**-TM9 are involved in critical interactions with **APH-1A**, contributing to the low observed energy values (Fig. S3A†). Similarly, we observed favorable interaction between **PS1**-TM4 and the first and third TMs of **PEN-2**. It is worth noting that we found significant interactions between the **PS1** TM3–TM4 loop and the N-terminal fragment of **PEN-2**, which was involved in significant contacts with a major lobe helix of **NCT**. This network of interactions may play an essential role in the communication pathway between these subunits and the **PS1** catalytic site. Fluctuation analysis reveals that the γ -secretase complex, including the catalytic subunit, remains structurally stable throughout the 1 μ s all-atom simulations, consistent with previous experimental studies.^{7,12} An equivalent analysis of our 5FN3 model is shown in the ESI (Fig S1–S3†).

Presenilin structural ensemble

To complement the atomic-scale fluctuation analysis, we explored in detail the structure, flexibility, and orientation of the **PS1** TMs in the POPC bilayer. Furthermore, we analyzed the dependence of the complex structure on the protonation state of the catalytic aspartic acid residues and the impact of protonation on the orientation of TMs defining the **PS1** structure in the active and inactive states of the enzyme.

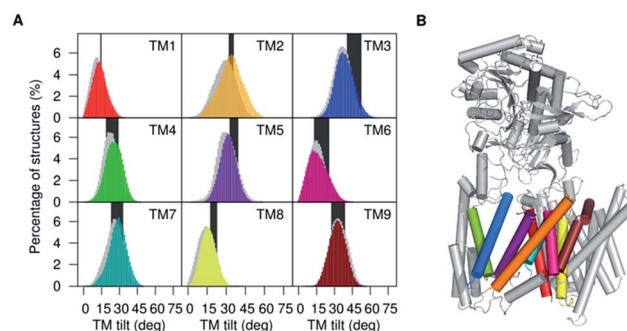


Fig. 2 (A) Distribution of tilt angles of the **PS1** TMs calculated for simulations of 5FN2 (color coded by TM helix number) and 5FN3 (in gray shade) and compared with TMs tilt angle ranges obtained from available experimental structures of γ -secretase (PDB IDs: 5A63, 4UIS, 5FN2, 5FN3, 5FN4 and 5FN5) (black bars). (B) Depiction of the 3D structure of γ -secretase (color coding **PS1** TMs as (A)).

PS1-TMs dynamics. The overall small changes in **PS1** structure observed during the all-atom simulations of the 5FN2 and 5FN3 derived models (Fig. S3B†) demonstrate that the initial structures of the complex remained stable throughout the simulations. Computed RMSF, which provides insight into residue mobility relative to the average simulation structure, was used to explore the TM mobility and loop flexibility in **PS1** (Fig. S3C†). The secondary structure of **PS1** was monitored and compared with the secondary structure derived from cryo-EM structural data demonstrating that the helicity of the TMs was preserved throughout the all-atom simulations (Fig. S3D†). The greatest flexibility was observed in loop regions separating relatively stable TMs. Due to the high flexibility of the TM1–TM2 loop (hydrophilic loop, HL1), Tomita and coworkers proposed that the HL1 loop may play a key role in substrate recognition.⁴⁴ Similarly, Wolfe and coworkers proposed that HL1 together with the **PS1**-CTF, which comprise a large fragment of the extracellular/luminal side of **PS1**, contribute to the “initial” substrate-binding conformation before substrate reaches the **PS1** γ -site.⁴⁵ Our simulation results support the plausibility of these conjectures.

The TM6 cytosolic fragment also showed high flexibility during the MD simulations. It has been proposed that alterations in the distance between cytosolic sides of TM6 and TM7 correlate with A β 42 production, suggesting that the TM6 conformation is a critical regulator of **PS1** catalytic activity.⁴⁶ The large RMSF values and highly conserved helicity observed in TM2 and TM6 suggest the presence of collective motion modulating the relative orientation of these helices. The mobility of TM2 and TM6 has also been associated with the significant plasticity of the active site suggested by cryo-EM structural analysis.¹² In Fig. 2 we present the computed tilt angle distribution of TMs, which agrees well with distributions derived from available γ -secretase structures. In particular, the wide tilt angle distribution of TM2 confirms the high mobility of this helix associated with the **PS1** subunit plasticity.

PS1 populates two principal conformational states

PS1 is an aspartyl protease that contains two catalytic Asp residues located in the intramembrane regions of TM6 and



TM7.^{8,47} Similar to other aspartyl proteases, the mechanism of catalysis is believed to involve formation of an initial intramolecular hydrogen-bond between the catalytic aspartic residues in coordination with a water molecule.^{48,49} In this acid-base mechanism, one Asp residue should be deprotonated (as aspartate, Asp⁻) to activate the water molecule, while the second Asp residue (as aspartic acid, Asp-H) donates a proton to the carbonyl group of the substrate.⁵⁰ The initial hydrogen-bond formed between the aspartates is believed to be essential to the proton transfer step necessary to the overall enzymatic reaction.^{51,52} These observations suggest that the distance between the catalytic Asp residues should be less than 0.5 nm in order to form the crucial hydrogen-bond required for the formation of the **PS1** active state.⁸ Previous structural studies of γ -secretase indicate that the relative proximity and orientation of the catalytic residues in the **PS1** active site depend on the tilt angle orientation of different TMs of **PS1**.^{12,18} Inspired by this idea, large-scale conformational changes of γ -secretase were simulated for three coarse grained (CG) Martini models varying protonation states of the catalytic residues with either or both charged. Trajectories of 50 replicas were each simulated for 1 μ s MD on each model. A pK_a calculation with the PROPKA v3.1 (ref. 53) module of PDB2PQR server⁵⁴ provides estimates of Asp385 (pK_a = 9.91) and Asp257 (pK_a = 5.12), suggesting these residues should be protonated and unprotonated, respectively, at pH = 7.0. As a consequence of the many replicas used and long simulation time, we expected to observe no dependence on the initial conditions (Fig. S4†). In order to validate our Martini CG model, we simulated 5 μ s of dynamics for a CG Martini model and a CG Martini model including an Elastic-Network in Dynamics (ElNeDyn) (Fig. S5 and S6†).⁴⁰ The CG-ElNeDyn simulation exhibited smaller RMSD values (\sim 0.2 nm) throughout the simulation. However, the RMSF showed that our CG-ElNeDyn system restricted motion of the HL1 loop, the CTF of TM6, and the loop between TM8 and TM9 (PAL motif). Importantly, the distance between the catalytic Asp residues in the CG-ElNeDyn model remained practically fixed. As mentioned above, experimental studies show that mobility of these regions is essential to conformational changes in the complex required to form the active and inactive states of the catalytic subunit. It is important to note that despite these differences, measurements of key distances between the TMs led to similar results in 5 μ s simulations of a Martini CG model and CG-ElNeDyn model. Finally, the curvature and kinks of the **PS1** TMs helices are consistent in all CG and AA simulation models (Fig. S7–S10†). Taken together these results validate our CG Martini models for the study of conformational changes of the γ -secretase enzyme complex.

Order parameters used to characterize the **PS1** ensemble include the TM tilt angles relative to the membrane normal (T_{TM}) and the distance between the catalytic Asp (dd_{Asp}). These order parameters are highly effective in differentiating the conformational state 1 (inactive state, long dd_{Asp} and proper T_{TM} angles) and state 2 (active state, short dd_{Asp} and proper T_{TM} angles) of γ -secretase. Short dd_{Asp} conformations involve Asp conformations suitable to form the crucial hydrogen bond required for the catalytic reaction.

In order to explore the correlation between the TMs orientation and proximity of the catalytic residues, the set of 50 1 μ s trajectories for the 5FN2 and 5FN3 derived CG models, varying Asp257 and Asp385 protonation states, were analyzed by projecting the simulated structural ensemble onto the dd_{Asp} distance and T_{TM} angles for TM2, TM6, TM7 and TM9. Fig. 3 displays the probability distributions for the dd_{Asp} distance and the T_{TM} angles for 5FN2 model of γ -secretase with unprotonated Asp and protonated Asp385 (similar results for 5FN3 and the two systems with the protonated Asp257 are available in Fig. S11†). There are notable changes in the relative probabilities of the conformational states of γ -secretase characterizing the structural ensemble of **PS1**: (1) the state 1, predominant in the unprotonated 5FN2 and 5FN3 derived CG models, is characterized by an inactive **PS1** catalytic conformation with a dd_{Asp} distance larger than 0.5 nm; (2) the state 2 observed in the protonated 5FN2 and 5FN3 derived CG models is characterized by a short dd_{Asp} distance consistent with active site formation. The T_{TM} angle distributions for the identified states indicate that TM6 and TM7 undergo an important T_{TM} angle modification that is strongly correlated with modulation of the dd_{Asp} distance. Moreover, the principal component analysis (PCA) of the **PS1** catalytic subunit suggests that a concerted motion of all TMs is required for the state 1 to state 2 transition (Fig. S12†).

Similarly, a slight change in the T_{TM} angles of TM2 and TM9 is observed to be correlated with a conformational change in T_{TM} angles of TM6 and TM7. Importantly, previous experimental studies have provided evidence that TM2 and TM9 constitute the initial substrate binding site in **PS1**, suggesting two different mechanisms of substrate entry into the active site: between TM6 and TM9 or between TM2 and TM6.^{43,45} Given the high flexibility and broad tilt angle distribution observed in our study, which is consistent with the conformational ensemble derived from cryo-EM studies, we specifically conjecture that TM2 acts as a gateway for substrate entry through TM2 and TM6.⁵⁵

In order to obtain a detailed atomistic description of the identified conformational states of **PS1** structures, we

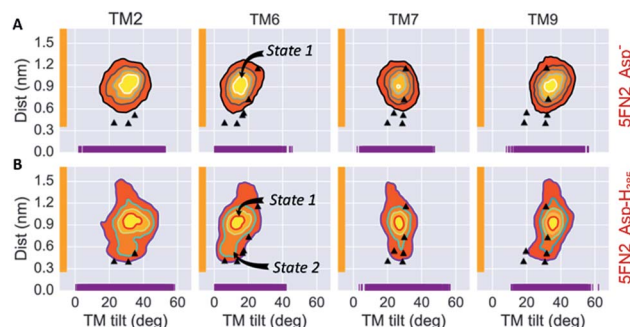


Fig. 3 Simulated distributions of 5FN2 derived CG model in POPC bilayer projected onto (1) the distance between the catalytic residues (Asp257 and Asp385) and (2) the calculated TM2, TM6, TM7 and TM9 tilt angles in the (A) unprotonated and (B) Asp385 protonated states. The black triangles depict the values of dd_{Asp} and T_{TM} angles obtained from the experimental structures of γ -secretase (PDB IDs: 5A63, 4UIS, 5FN2, 5FN3, 5FN4 and 5FN5).



performed 100 ns all-atom MD simulations (Fig. 4) of the γ -secretase complex employing the CHARMM force field in a POPC lipid bilayer. The most probable state 1 and state 2 γ -secretase conformations were selected using the GROMACS cluster tool with a RMSD threshold of 2.0 Å from the CG structural ensemble. As observed in the CG studies, the state 2 conformation is showed to fluctuate between more and less active conformations (dd_{ASP} distance does not exceed 0.75 nm). This state 2 (active/less-active) equilibrium is associated with Asp dihedral angle fluctuation with unaltered inclination of the key TMs. Equivalent fluctuations between state 1/state 2 (inactive/active) states were observed in the CG model simulations, indicating that the state 2 (active) conformation may require the substrate to be bound to the catalytic site in order to stabilize the active conformation. On the other hand, data derived from atomistic MD simulation of the **PS1** state 1

(inactive) conformations display substantial structural fluctuations with a 1.0 nm dd_{ASP} distance. This observation is consistent with the previously characterized inactive CG simulations. Furthermore, the analysis of water molecules at the active site showed the presence of intramolecular hydrogen-bonds between both catalytic aspartic residues and a coordinated water molecule within 3.0 Å distance and 150-degree angle (Fig. 4C). This hydrogen bond network was present 43.92% of the time of the last 50 ns of the state 2 all-atom simulation. Interestingly, this water coordination is a crucial step for **PS1** activation and proteolytic processing of γ -secretase substrates.

Nicastrin shows dramatic up-down movement and left-right rotation of the NCT extracellular domain

The large conformational modifications observed during cryo-EM studies of the complex have led to several hypotheses that seek to define the substrate recognition mechanism of γ -secretase.^{20,22,56} The observed structural changes in γ -secretase complex can be related to the large-scale motion of the NCT extracellular domain. Employing the 50 1 μ s CG MD simulations of our 5FN2 and 5FN3 derived models, we performed PCA to identify the most significant collective motions of the complex, for both protonated and unprotonated states. The PCA analysis was performed with the GROMACS inbuilt tools (covar and ana eig).³⁰

The first two eigenvectors, involving the NCT ECD and accounting for the majority (23%) of overall γ -secretase motion, were selected to analyze the NCT movement. Analysis of conformations projected onto the first two PC eigenvectors revealed that the first eigenvector is related to NCT “up/down” movement while the second eigenvector corresponds to NCT-ECD “left/right” rotation. Fig. 5 depicts motion along the first and second principal components (PCs) using a porcupine representation. Interestingly, we found the same two principal component motions during the 5 μ s simulation of our CG-ElNeDyn model with a slower dynamics and a smaller set of collective vectors (Fig. S6†).

In order to further analyze structural fluctuations of NCT, we constructed two order parameters based on the first and second PCs obtained from the CG structural ensemble (capturing up/down and left/right rotational motion of NCT, Fig. S13†). The first order parameter describes the distance between the center of geometry of **PS1** and the center of geometry of the large lobe of NCT ($dd_{\text{PS-NCT}}$). Significant variation in the distribution of the $dd_{\text{PS-NCT}}$ distances indicates that a percentage of the γ -secretase structures simulated using the CG model exhibit an “up” conformation (with $dd_{\text{PS-NCT}}$ distances larger to 5.0 nm) while a fraction of the NCT domain adopts a “down” conformation (with a $dd_{\text{PS-NCT}}$ distance shorter than 4.5 nm). The second order parameter characterizes the rotational motion of NCT (dih_{NCT}) and involves a dihedral angle formed by four points of the NCT structure: (1) the intracellular and (2) extracellular amino acids of the single TM helix of NCT, (3) the center of mass of the large lobe of NCT, and (4) the residues Arg583–Asp588 (located at the distal end of the NCT large lobe). Previous experimental structural studies of γ -secretase have proposed

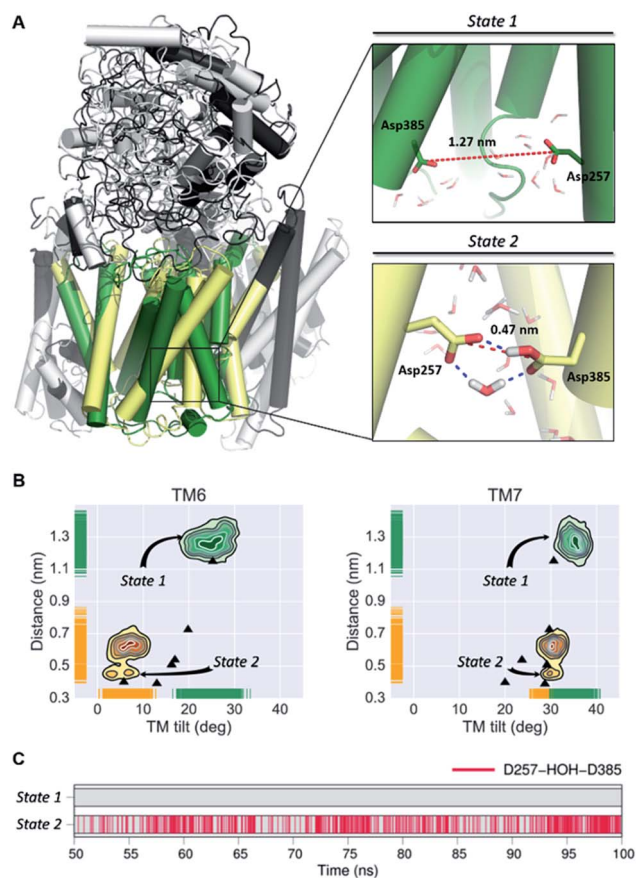


Fig. 4 (A) Most representative structures of the state 1 (inactive) and state-2 (active) conformations of the **PS1** subunit of γ -secretase obtained from all-atom MD simulations. The dotted red line represents the distance between the gamma carbons of Asp257 and Asp385. (B) Simulated distribution of both conformational state models projected onto the distance between the catalytic aspartic acid residues and the **PS1** tilt angles of TM6 and TM7. The black triangles depict dd_{ASP} distances and T_{TM} angles obtained from the experimental structures of γ -secretase (PDB IDs: 5A63, 4UIS, 5FN2, 5FN3, 5FN4 and 5FN5). (C) Time evolution of hydrogen bonds between a coordinated water molecule and both catalytic aspartic residues through the last 50 ns of the state 1 and state 2 simulations.



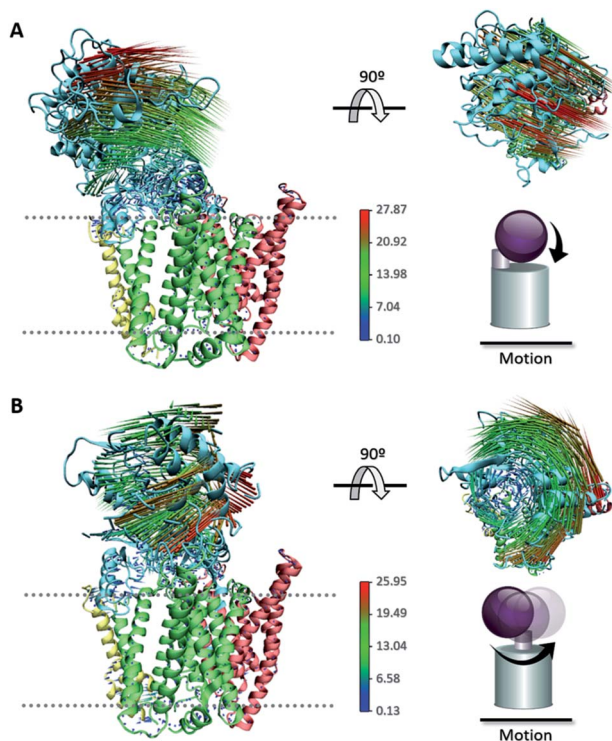


Fig. 5 Porcupine representation of the (A) "up/down" movement and (B) "left/right" rotation of the NCT ECD obtained from PCA.

a possible interaction of the substrate N-terminal region with exposed Glu333 located at a lid opening.¹⁷ Taken together, those observations and our simulation analysis support the conjecture that NCT rotation (dih_{NCT}) is essential to the γ -secretase substrate binding mechanism. The simulated dih_{NCT} distribution indicates that the NCT large lobe rotates in the left/right directions, with preference for leftward rotation (dih_{NCT} values of -50 deg). However, further analysis demonstrates that the identified NCT motion, described by the dd_{PS-NCT} distance and the dih_{NCT} rotation, does not modify the position of the NCT-lid (Fig. S14 and S15[†]). This suggests that Glu333 might not be involved in the substrate recognition mechanism. Furthermore, Wolfe and coworkers⁵⁷ proposed that the NCT-ECD acts as a steric gatekeeper for substrate entry into the PS1 active site, instead of having a specific interaction with the substrate. This steric block mechanism could be associated with ECD up-down movement, left-right rotation, and interaction with the transmembrane subunits, as shown in the following section.

The structural ensemble of γ -secretase is characterized by three unique conformational states. The simulated γ -secretase structural ensemble was compared with the experimentally derived conformational distribution of the γ -secretase complex.²² We measured the major axis length of our structures using the distance between the amino acids located at the lower intracellular and upper extracellular domains (Fig. 6).²² In an experimental single-particle EM study, Chávez-Gutiérrez and coworkers²² found that wild-type γ -secretase exists in three different structural states: compact, intermediate, and extended conformations. In agreement with their experimental observations, three structural states

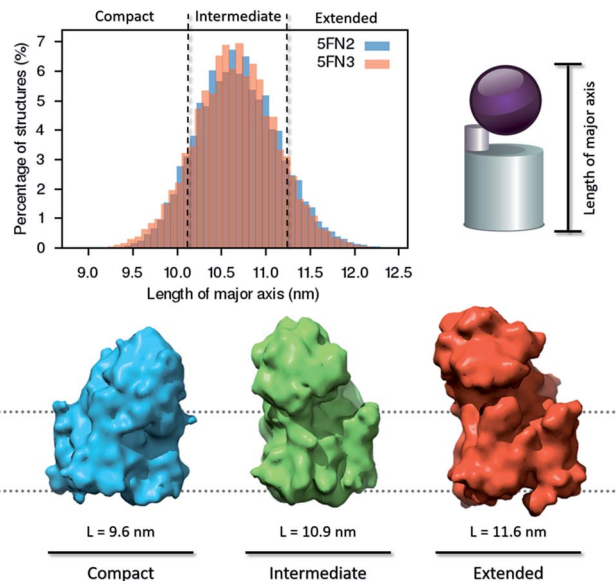


Fig. 6 Distribution of major axis length and density map representation of the compact, intermediate, and extended conformations of the γ -secretase complex derived from 50 $1 \mu s$ CG replica simulations of the γ -secretase complex.

of the γ -secretase complex were identified in the simulated length distribution. A dynamic equilibrium was observed, characterized by a greater abundance of the intermediate form. On the other hand, our CG-ElNeDyn simulation model only sampled the extended and intermediate groups due to restrictions imposed by the elastic network (Fig. S6C[†]).

In further analysis, we extracted representative structures characterizing the three distinct conformational states in order to generate and analyze three density maps (8.0 \AA resolution), employing UCSF Chimera software.⁴² The structural state transitions are observed to involve three distinctive interactions between the extracellular domain of NCT and key γ -secretase subunits: (1) in the extended state, the small lobe of NCT interacts with the C-terminal region of PEN2, (2) in the intermediate state, the large lobe of NCT interacts with the HL1 (hydrophilic TM1–TM2 loop) of PS1, and (3) in the compact state, the large lobe of NCT interacts with the APH-1A and PS1 extracellular domains. The broad distribution of structures in the intermediate state indicates that the second interaction, between the large lobe of NCT and HL1 of PS1, is essential for NCT ECD stabilization. Moreover, this conformation could be essential for NCT to act as a selective "gatekeeper" during peptide entry into the catalytic site.⁵⁸ Experimental studies have demonstrated that when substrate or other agents (*e.g.* inhibitors and modulators) are bound to γ -secretase, the compact conformation of the complex is the most stable and favored state.^{20,22,56} Importantly, we found that although the NCT ECD interact with the hydrophilic PS1 HL1, APH-1A, and PEN2, no correlated motion was observed between the large NCT ECD and the PS1 TMs helices. These results suggest that the large NCT ECD motion is not critical to the activation or inactivation mechanism of the catalytic PS1 subunit (Fig. S16[†]).



Conclusions

Our multiscale simulation approach has elucidated the first detailed atomistic description of the structure and dynamics of the complete γ -secretase complex. Analysis of the structure and dynamics of our simulated γ -secretase models supports the hypothesis that **PEN-2** and **APH-1A** play essential roles in stabilizing the **PS1** catalytic subunit through interaction with the **PS1** N-terminal and C-terminal fragments, respectively.

Overall, the TMs of γ -secretase form a stable structural complex of relatively low mobility. However, we specifically observed that higher flexibility and correlated motion of **PS1-TM2** and **PS1-TM6** impacts the dd_{Asp} distance essential for protease catalytic activity. Taken together, these findings suggest that correlated motion of key TMs helices, dependent on the dd_{Asp} distance and TM tilt angles, is essential to the transitions between inactive to active states of **PS1**.

We have further demonstrated that the equilibrium between the state 1 (“inactive”) and state 2 (“active”) **PS1** conformers is modulated by the protonation states of the catalytic residues Asp257 and Asp385. State 1 **PS1** conformations are predominant in proteins having unprotonated catalytic Asp residues, while state 2 **PS1** conformations are sampled when either catalytic Asp residue is protonated.

Analysis of global conformational changes in the γ -secretase complex identified significant mobility in the **NCT ECD**, characterized as up/down motion and left/right rotation of the large lobe of **NCT ECD**. Similar movement of the **NCT ECD** has been previously inferred from electron microscopy images.²² Based on these observations, it has been conjectured that γ -secretase function is controlled through modulation of relative populations of **NCT ECD** conformational states in the γ -secretase structural ensemble. Importantly, our data suggests that these movements are not correlated with **NCT** lid motion and the state 1/state 2 equilibrium of **PS1**, supporting that **NCT-ECD** only acts as steric gatekeeper for substrate entry into the active site, as proposed first by Wolfe and coworkers.⁵⁷

This multiscale simulation analysis provides a detailed picture of the global structure and dynamics of the γ -secretase complex. The insight provided into the nature of the active and inactive state conformation, as well as the mechanism of transition between state 1 and state 2 of the **PS1** catalytic domain, provides a foundation for future studies of the catalytic mechanism of substrate recognition and cleavage. It is our hope that the results of this study will contribute to a mechanistic understanding of the cleavage of APP-C99 by γ -secretase in the genesis of A β , critical to the structure-based design of AD therapeutics.

Conflict of interest

The authors declare no competing financial interest.

Acknowledgements

R. Aguayo-Ortiz (No. 510728/288862) and C. Chávez-García (No. 663225/583333) are grateful to CONACyT for the fellowships

granted. The authors gratefully acknowledge the support of the Programa de Apoyo a la Investigación y el Posgrado (PAIP 5000-9155), Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT IA204716) and the Dirección General de Cómputo y de Tecnologías de Información (LANCAD-UNAM-DGTIC-306). JES acknowledges the generous support of the National Institutes of Health (R01 GM107703). The authors are very grateful to Sac N. I. Morales Betanzos for contributions to figures in this manuscript.

Notes and references

- 1 B. De Strooper and W. Annaert, *Annu. Rev. Cell Dev. Biol.*, 2010, **26**, 235–260.
- 2 F. M. LaFerla, K. N. Green and S. Oddo, *Nat. Rev. Neurosci.*, 2007, **8**, 499–509.
- 3 L. Mucke, *Nature*, 2009, **461**, 895–897.
- 4 B. De Strooper, *Neuron*, 2003, **38**, 9–12.
- 5 W. T. Kimberly, M. J. LaVoie, B. L. Ostaszewski, W. Ye, M. S. Wolfe and D. J. Selkoe, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 6382–6387.
- 6 C. A. Hansson, S. Frykman, M. R. Farmery, L. O. Tjernberg, C. Nilsberth, S. E. Pursglove, A. Ito, B. Winblad, R. F. Cowburn, J. Thyberg and M. Ankarcrona, *J. Biol. Chem.*, 2004, **279**, 51654–51660.
- 7 L. Sun, L. Zhao, G. Yang, C. Yan, R. Zhou, X. Zhou, T. Xie, Y. Zhao, S. Wu, X. Li and Y. Shi, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 6003–6008.
- 8 X. Li, S. Dang, C. Yan, X. Gong, J. Wang and Y. Shi, *Nature*, 2013, **493**, 56–61.
- 9 T. Tomita and T. Iwatsubo, *J. Biol. Chem.*, 2013, **288**, 14673–14680.
- 10 S. Dang, S. Wu, J. Wang, H. Li, M. Huang, W. He, Y.-M. Li, C. C. L. Wong and Y. Shi, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 201502150.
- 11 X. Zhang, C. J. Yu and S. S. Sisodia, *Mol. Neurodegener.*, 2015, **10**, 39.
- 12 X. Bai, C. Yan, G. Yang, P. Lu, D. Ma, L. Sun, R. Zhou, S. H. W. Scheres and Y. Shi, *Nature*, 2015, **525**, 212–217.
- 13 Y. Li, L. S. Y. Liew, Q. Li and C. Kang, *Sci. Rep.*, 2016, **6**, 19522.
- 14 R. Francis, G. McGrath, J. Zhang, D. A. Ruddy, M. Sym, J. Apfeld, M. Nicoll, M. Maxwell, B. Hai, M. C. Ellis, A. L. Parks, W. Xu, J. Li, M. Gurney, R. L. Myers, C. S. Himes, R. Hiesch, C. Ruble, J. S. Nye and D. Curtis, *Dev. Cell*, 2002, **3**, 85–97.
- 15 N. Takasugi, T. Tomita, I. Hayashi, M. Tsuruoka, M. Niimura, Y. Takahashi, G. Thinakaran and T. Iwatsubo, *Nature*, 2003, **422**, 438–441.
- 16 D. M. Bolduc and M. S. Wolfe, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 14643–14644.
- 17 T. Xie, C. Yan, R. Zhou, Y. Zhao, L. Sun, G. Yang, P. Lu, D. Ma and Y. Shi, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 13349–13354.
- 18 P. Lu, X. C. Bai, D. Ma, T. Xie, C. Yan, L. Sun, G. Yang, Y. Zhao, R. Zhou, S. H. Scheres and Y. Shi, *Nature*, 2014, **512**, 166–170.



- 19 X.-C. Bai, E. Rajendra, G. Yang, Y. Shi and S. H. Scheres, *eLife*, 2015, **4**, e11182.
- 20 S. Shah, S. F. Lee, K. Tabuchi, Y. H. Hao, C. Yu, Q. LaPlant, H. Ball, C. E. Dann, T. Südhof and G. Yu, *Cell*, 2005, **122**, 435–447.
- 21 L. Chávez-Gutiérrez, A. Tolia, E. Maes, T. Li, P. C. Wong and B. De Strooper, *J. Biol. Chem.*, 2008, **283**, 20096–20105.
- 22 N. Elad, B. De Strooper, S. Lismont, W. Hagen, S. Veugelen, M. Arimon, K. Horre, O. Berezovska, C. Sachse and L. Chávez-Gutiérrez, *J. Cell Sci.*, 2015, **128**, 589–598.
- 23 R. Kong, S. Chang, W. Xia and S. T. C. Wong, *J. Struct. Biol.*, 2015, **191**, 120–129.
- 24 S. Li, W. Zhang and W. Han, *Initial Substrate Binding of γ -Secretase: The Role of Substrate Flexibility*, 2017.
- 25 M. L. Hekkelman, T. A. H. te Beek, S. R. Pettifer, D. Thorne, T. K. Attwood and G. Vriend, *Nucleic Acids Res.*, 2010, **38**, 719–723.
- 26 M. A. Lomize, A. L. Lomize, I. D. Pogozheva and H. I. Mosberg, *Bioinformatics*, 2006, **22**, 623–625.
- 27 J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. B. Iii, A. D. MacKerell Jr, B. Klauda and W. Im, *J. Chem. Theory Comput.*, 2016, **12**, 405–413.
- 28 E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda and W. Im, *J. Comput. Chem.*, 2014, **35**, 1997–2004.
- 29 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
- 30 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.
- 31 R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, J. MacKerell and D. Alexander, *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.
- 32 R. Kumari, R. Kumar, O. Source, D. Discovery and A. Lynn, *J. Chem. Inf. Model.*, 2014, **54**, 1951–1962.
- 33 N. Michaud-agrawal, E. J. Denning, T. B. Woolf and O. Beckstein, *J. Comput. Chem.*, 2011, **32**, 2319–2327.
- 34 A. C. E. Dahl, M. Chavent and M. S. P. Sansom, *Struct. Bioinf.*, 2012, **28**, 2193–2194.
- 35 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
- 36 W. L. DeLano, *DeLano Scientific*, San Carlos, CA, 2007, <http://www.pymol.org>.
- 37 S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. De Vries, *J. Phys. Chem. B*, 2007, **111**, 7812–7824.
- 38 Y. Qi, H. I. Ingólfsson, X. Cheng, J. Lee, S. J. Marrink and W. Im, *J. Chem. Theory Comput.*, 2015, **11**, 4486–4494.
- 39 S. O. Yesylevskyy, L. V. Schäfer, D. Sengupta and S. J. Marrink, *PLoS Comput. Biol.*, 2010, **6**, 1–17.
- 40 X. Periole, M. Cavalli, S. J. Marrink and M. A. Ceruso, *J. Chem. Theory Comput.*, 2009, **5**, 2531–2543.
- 41 T. A. Wassenaar, K. Pluhackova, R. A. Böckmann, S. J. Marrink and D. P. Tieleman, *J. Chem. Theory Comput.*, 2014, **10**, 676–690.
- 42 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
- 43 N. Watanabe, S. Takagi, A. Tominaga, T. Tomita and T. Iwatsubo, *J. Biol. Chem.*, 2010, **285**, 19738–19746.
- 44 S. Takagi-Niidome, T. Sasaki, S. Osawa, T. Sato, K. Morishima, T. Cai, T. Iwatsubo and T. Tomita, *J. Neurosci.*, 2015, **35**, 2646–2656.
- 45 A. Y. Kornilova, F. Bihel, C. Das and M. S. Wolfe, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 3230–3235.
- 46 A. Tominaga, T. Cai, S. Takagi-Niidome, T. Iwatsubo and T. Tomita, *J. Neurosci.*, 2016, **36**, 1362–1372.
- 47 H. Steiner, M. Kostka, H. Romig, G. Basset, B. Pesold, J. Hardy, a. Capell, L. Meyn, M. L. Grim, R. Baumeister, K. Fechteler and C. Haass, *Nat. Cell Biol.*, 2000, **2**, 848–851.
- 48 D. B. Northrop, *Acc. Chem. Res.*, 2001, **34**, 790–797.
- 49 E. Erez, D. Fass and E. Bibi, *Nature*, 2009, **459**, 371–378.
- 50 A. Tolia, L. Chávez-Gutiérrez and B. De Strooper, *J. Biol. Chem.*, 2006, **281**, 27633–27642.
- 51 M. Bentahir, O. Nyabi, J. Verhamme, A. Tolia, K. Horré, J. Wiltfang, H. Esselmann and B. De Strooper, *J. Neurochem.*, 2006, **96**, 732–742.
- 52 F. Sussman, M. C. Villaverde, J. L. Dominguez and U. E. Danielson, *Curr. Pharm. Des.*, 2013, **19**, 4257–4275.
- 53 M. H. M. Olsson, R. S. Chresten, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **2**, 525–537.
- 54 T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Res.*, 2004, **32**, 665–667.
- 55 A. Fukumori and H. Steiner, *EMBO J.*, 2016, **35**, 1628–1643.
- 56 Y. Li, S. H. J. Lu, C. J. Tsai, C. Bohm, S. Qamar, R. B. Dodd, W. Meadows, A. Jeon, A. McLeod, F. Chen, M. Arimon, O. Berezovska, B. T. Hyman, T. Tomita, T. Iwatsubo, C. M. Johnson, L. A. Farrer, G. Schmitt-Ulms, P. E. Fraser and P. H. St George-Hyslop, *Structure*, 2014, **22**, 125–135.
- 57 D. M. Bolduc, D. R. Montagna, Y. Gu, D. J. Selkoe and M. S. Wolfe, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E509–E518.
- 58 S. Urban, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 201524151.



Characterizing the structural ensemble of γ -secretase using a multiscale molecular dynamics approach

Rodrigo Aguayo-Ortiz^a, Cecilia Chávez-García^a, John E. Straub^b, and Laura Dominguez^{a,*}

^aDepartamento de Fisicoquímica, Facultad de Química, Universidad Nacional Autónoma de México, Mexico City, 04510, Mexico; ^bDepartment of Chemistry, Boston University, Boston, Massachusetts 02215, United States

*Corresponding Author: Phone: +52-5622-3773. E-mail: lauradd@unam.mx

Table of contents

Contents	Page
Fig. S1 RMSD of 5FN2 and 5FN3 derived all atom model components	1
Fig. S2 RMSF of 5FN2 and 5FN3 derived all atom model components	2
Fig. S3 Presenilin contribution energy, RMSD, RMSF and % helicity (%) for all-atom MD simulations.	3
Fig. S4 Projection of the Asp257-Asp385 distance and TM helix tilt of the CG models	4
Fig. S5 Time evolution of RMSD, RMSF, and key PS1 TMs distances calculated for 5 μ s of CG-Martini and CG-EINeDyn systems.	5
Fig. S6 PCA of the NCT ECD for the 5 μ s simulation of the CG-Martini and CG-EINeDyn systems as well as the distribution of the major axis length of the complex	6
Fig. S7 Curvature of the PS1 TMs during the last 500 ns of the 50 CG replicas	7
Fig. S8 Measurement of the per-residue kink angle for each PS1 TM of the 5FN2-derived models	8
Fig. S9 Measurement of the per-residue kink angle for each PS1 TM of the 5FN3-derived models	9
Fig. S10 Schematic representation of the kink angles of the 5FN2 PS1-derived AA and CG models	10
Fig. S11 Time-colored (ns) projection of the Asp257-Asp385 distance and TM helix tilt of the Asp257 protonated system	11
Fig. S12 PCA of the 5FN2 PS1-derived AA and CG models and the porcupine representation of PC1	12
Fig. S13 Simulated distribution for CG models projected onto dd_{PS-NCT} distance and dih_{NCT} rotation	13
Fig. S14 Projection of NCT lid and major lobe distance and NCT rotation	13
Fig. S15 Projection of NCT lid and Glu333 distance and NCT rotation	14
Fig. S16 Projection of γ -secretase <i>state 1</i> and <i>state 2</i> conformations and length of the major axis	14

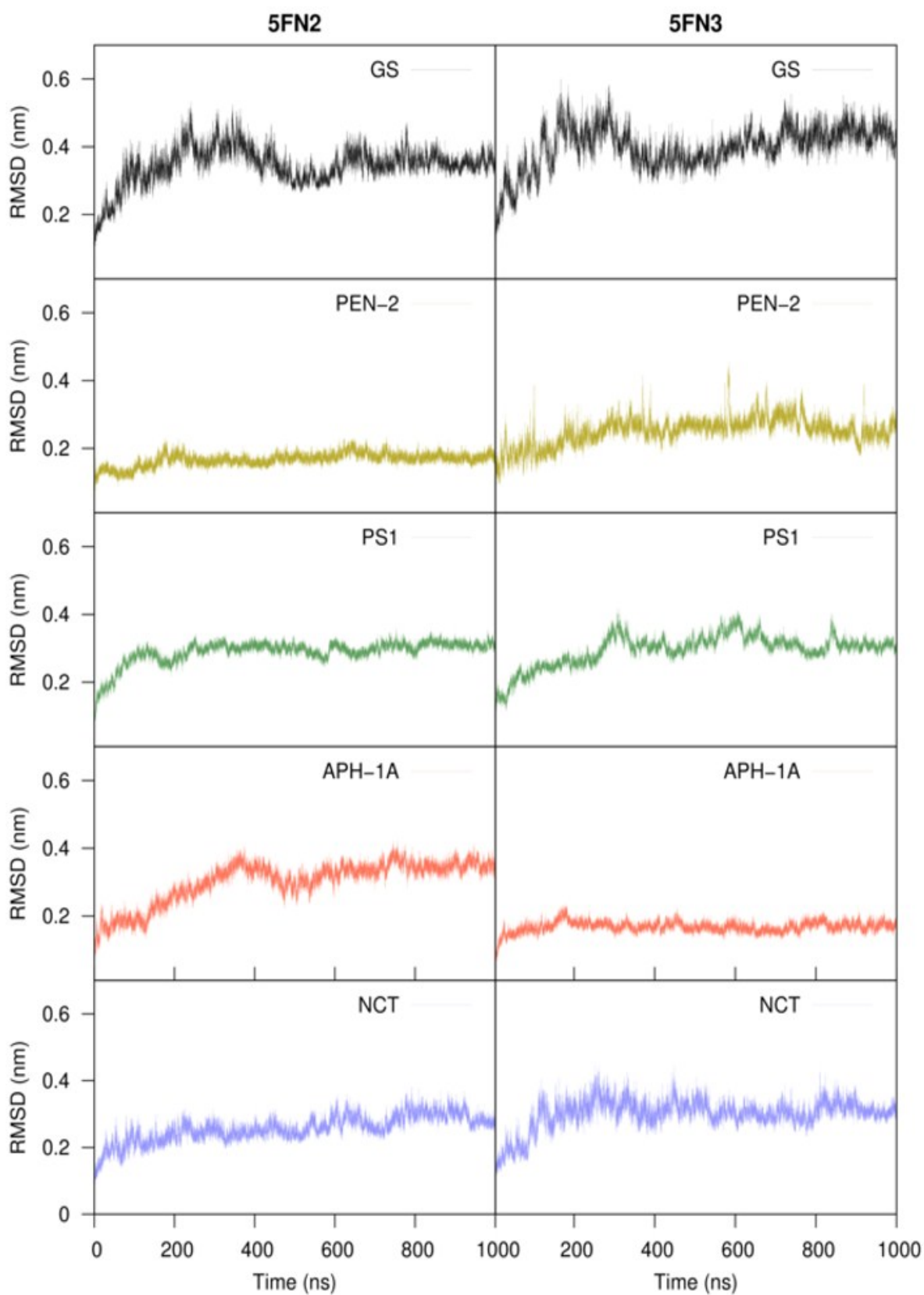


Fig. S1. Root mean square deviation (RMSD) of 5FN2 and 5FN3 derived all atom models of γ -secretase (GS) components: presenilin enhancer 2 (PEN-2), presenilin 1 (PS1), anterior pharynx-defective 1A (APH-1A), and Nicastrin (NCT).

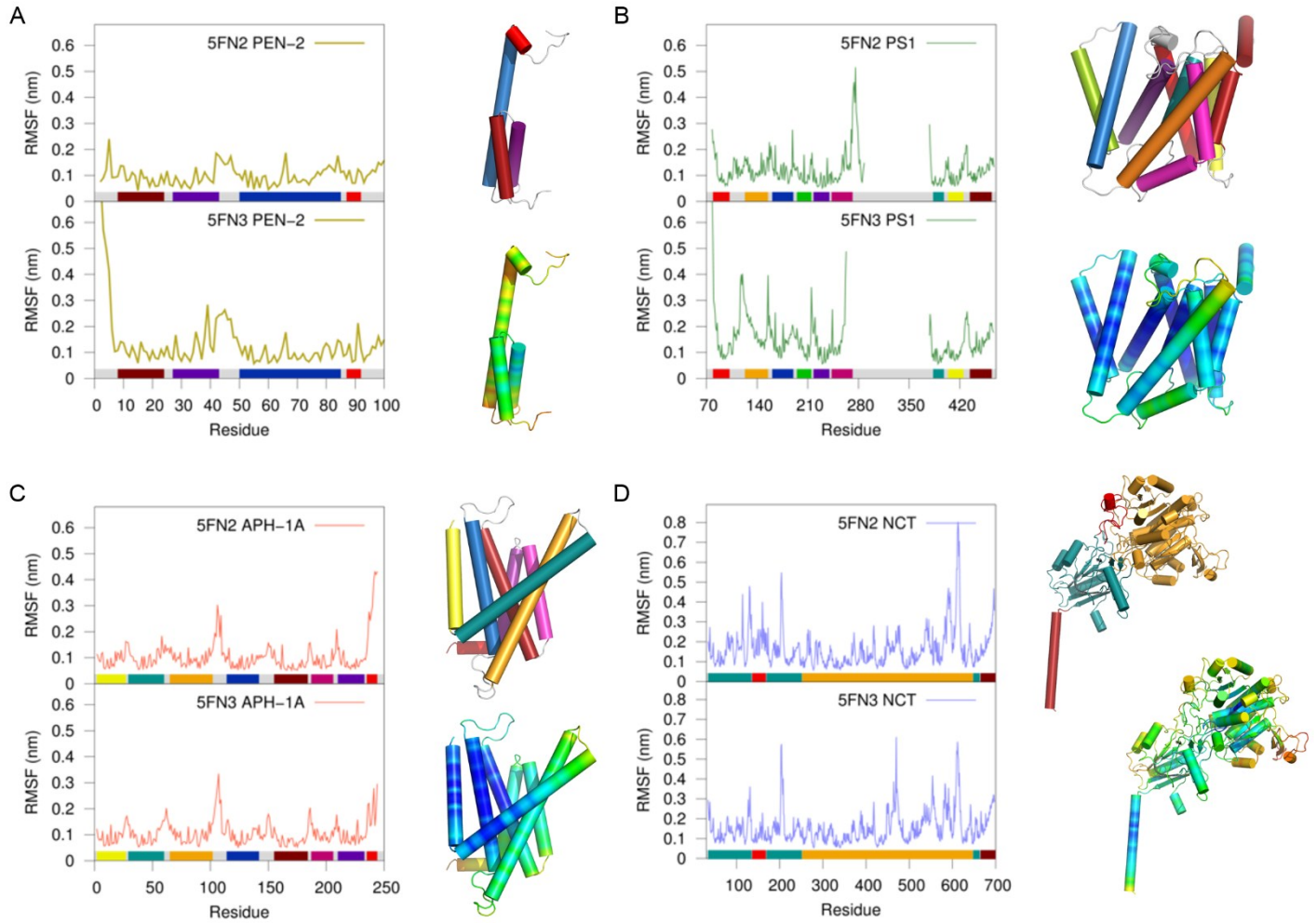


Fig. S2. Root mean square fluctuation (RMSF) of 5FN2 and 5FN3 derived all-atom models: (A) PEN-2, (B) PS1, (C) APH-1A and (D) NCT. The 3D structures depict the region represented in the RMSF plots (top) and the normalized per-residue RMSF (bottom) colored by the RMSF value from more flexible (red) to less flexible (blue).

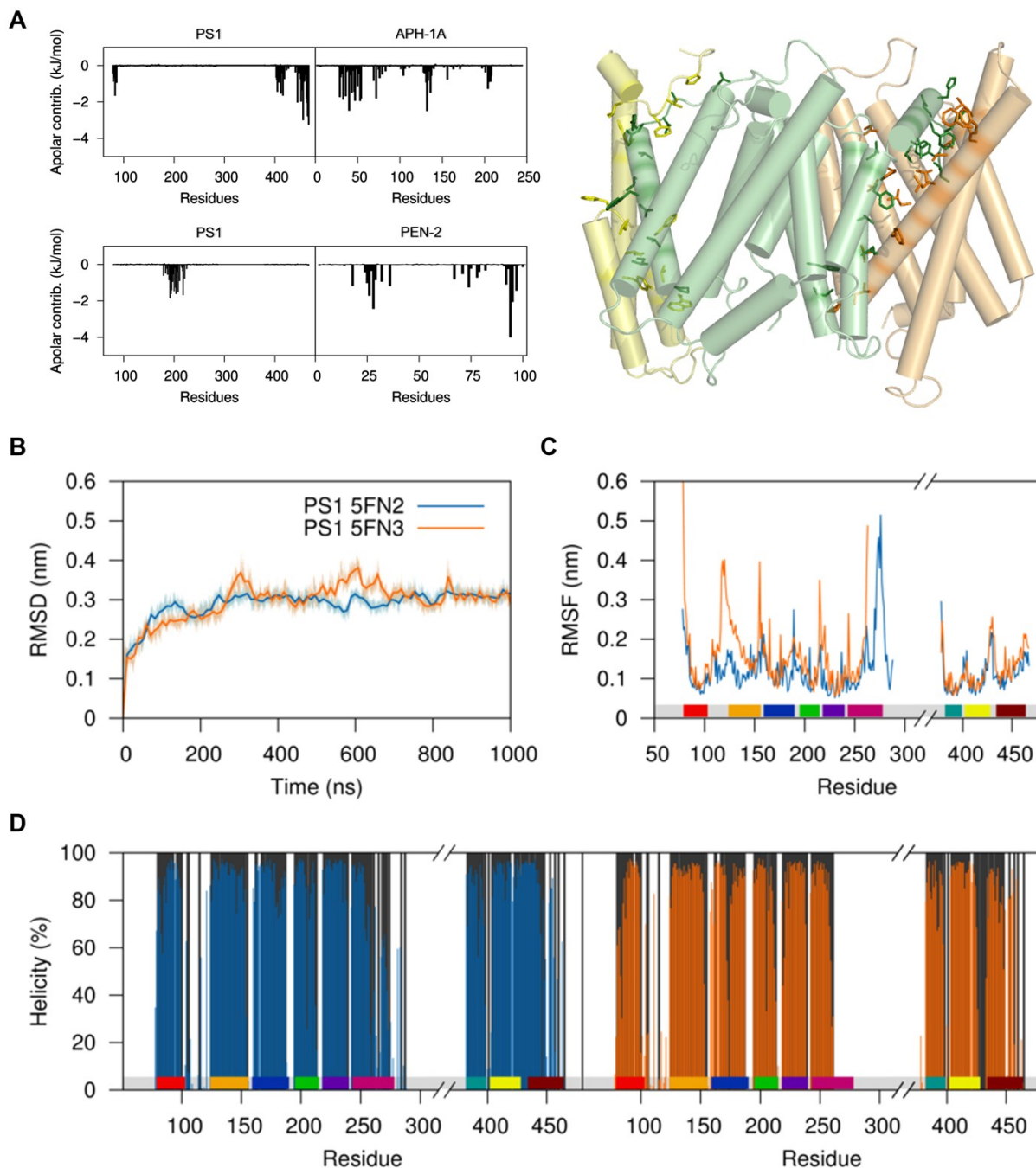


Fig. S3. (A) Calculated apolar free energy contribution of each amino acid by domain. Key residues making dominant contributions to apolar binding free energy are shown (color coded as in Figure 1A). (B) PS1 backbone root mean-square deviation (RMSD) during all-atom MD simulations of the 5FN2 (blue) and 5FN3 (orange) derived γ -secretase complex. The equilibrium states of the systems were reached after 300 ns of simulation, with resulting structures within approximately 0.30 nm RMSD of the initial structure. (C) Root mean-square fluctuation (RMSF) and (D) percent helicity of residues constituting the PS1 TMs.

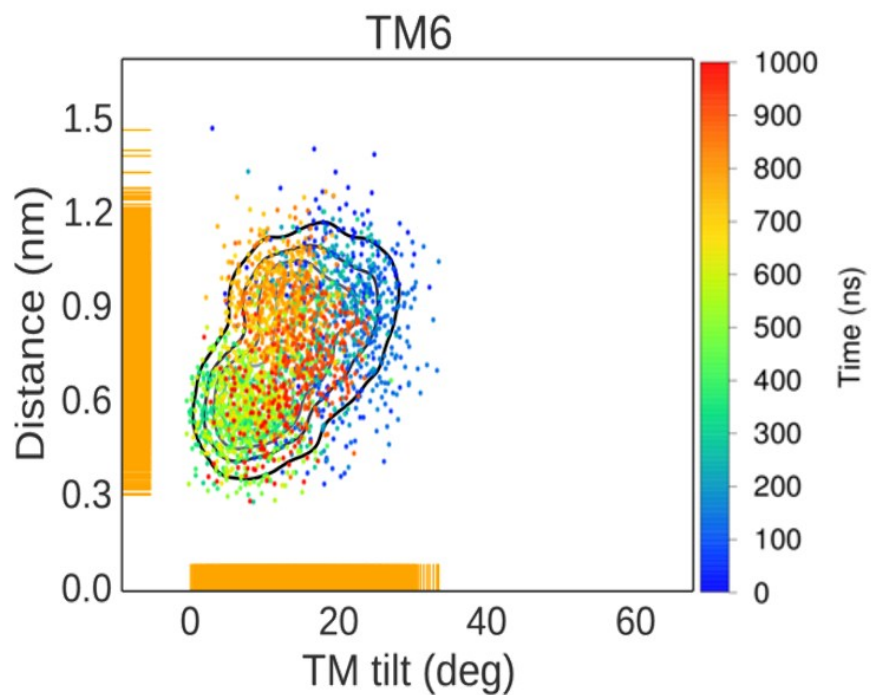


Fig. S4. Time evolution of one replica (out of fifty) of a CG 5FN2 derived model in a POPC bilayer projected onto the distance between the catalytic residues (Asp257 and Asp385) and the calculated TM6 tilt angle in the Asp257 protonated state, colored by time.

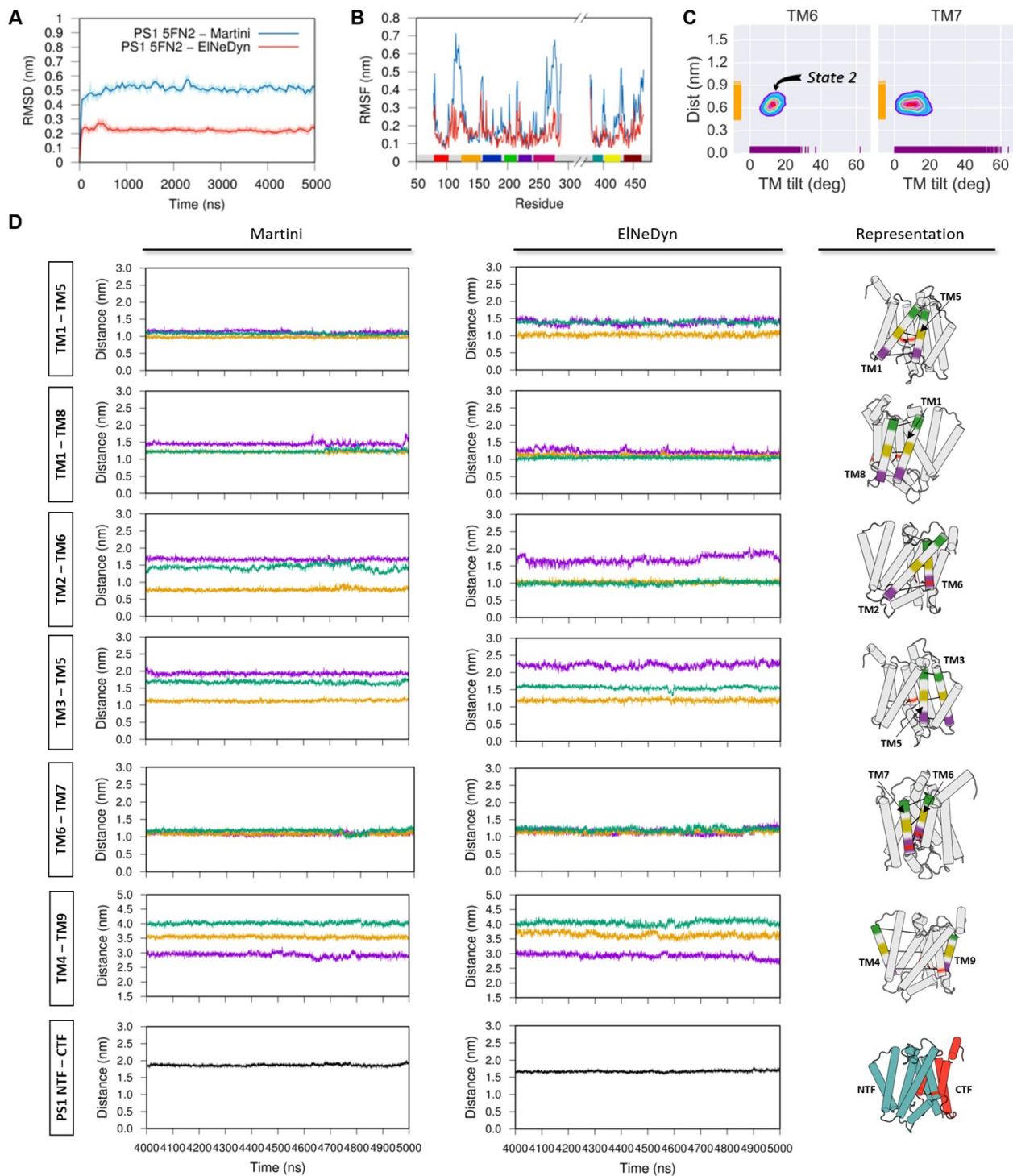


Fig. S5. (A) RMSD and (B) RMSF of the 5 μ s CG Martini and CG-EIneDyn PS1 derived models in a POPC bilayer. (C) Simulated distributions of our Martini CG-EIneDyn 5FN2-derived simulation model in a POPC bilayer projected onto (1) the distance between the catalytic residues (Asp257 and Asp385) and (2) the calculated TM6 and TM7 tilt angles for the unprotonated state. (D) Time evolution of six key TM-TM distances obtained from the catalytic PS1 subunit during the last μ s of the 5 μ s 5FN2-derived CG Martini model trajectory and the last μ s of the 5 μ s 5FN2-derived CG-EIneDyn model trajectory, together with a PS1 3D model to depict the measured distances and the color code segments in the plot. The last plot (PS1 NTF-CTF) depicts the distance between the centers of geometry of the PS1 C-terminal (CTF red) and N-terminal (NTF blue) fragments.

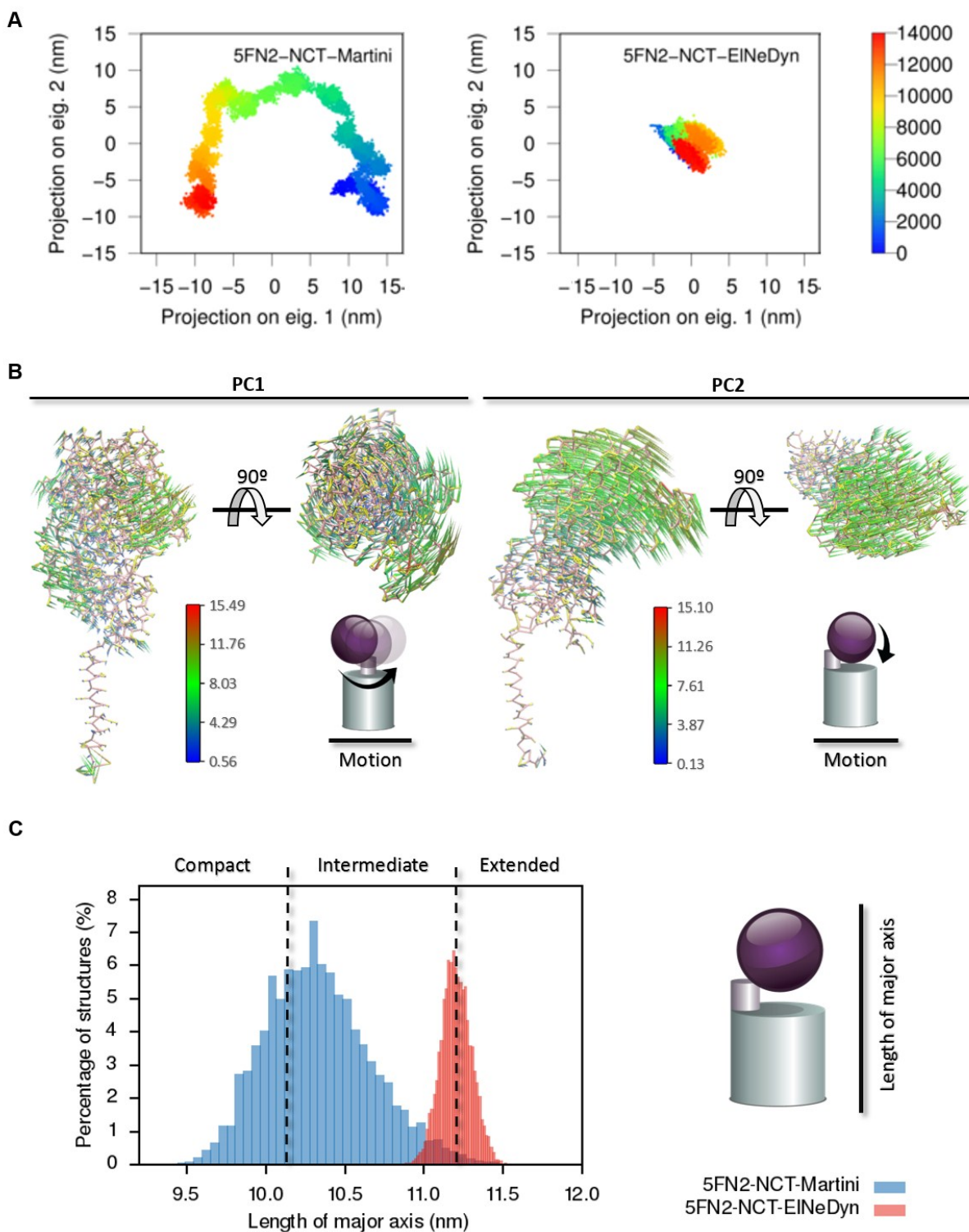


Fig. S6. (A) Projection of the first and second eigenvectors obtained from the PCA of the last 500 ns of the Martini CG trajectory (left) and the projection of the 5FN2 CG-EIeDyn trajectory into the same first and second eigenvectors derived from the Martini CG trajectory (right). (B) Porcupine representation of the NCT ECD “left/right” rotation (PC1) and “up/down” movement (PC2) obtained from the PCA of CG-EIeDyn 5FN2-derived model. (C) Distribution of major axis length of the compact, intermediate, and extended conformations of the γ -secretase complex derived from 5 μ s Martini CG-EIeDyn model simulations.

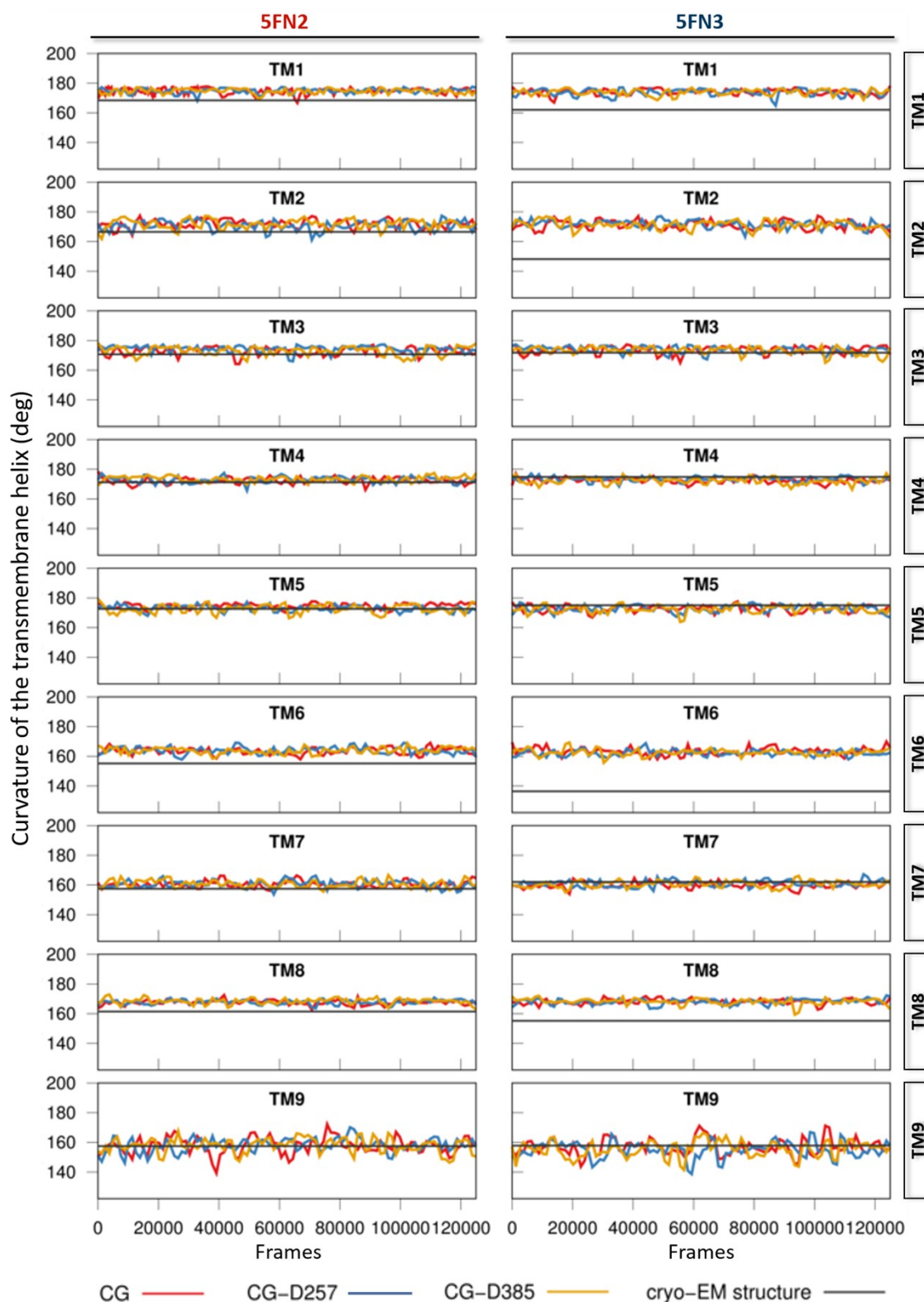


Fig. S7. PS1 TMs helix curvatures during the last 500 ns of the 50 CG replicas (125,000 frames) of the distinct γ -secretase 5FN2-derived (left) and 5FN3-derived (right) models in the unprotonated state (red) and the protonated D257 state (blue) and protonated D385 state (yellow) models, together with the reference curvature value for the cryo-EM structure (grey).

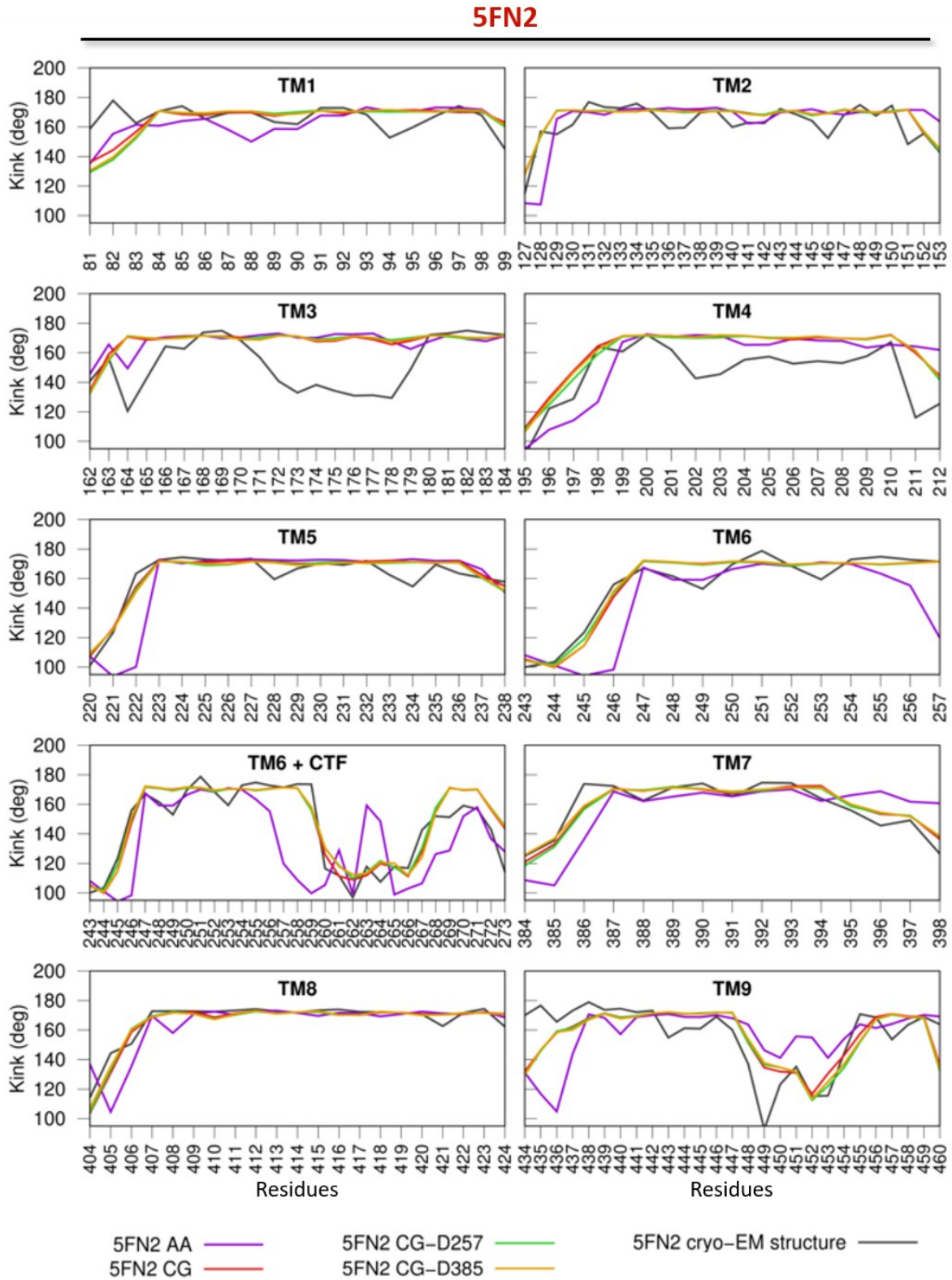


Fig. S8. Per-residue kink angles calculated for each PS1 helix in the AA (dark-violet), CG (red), protonated CG-D257 (green), and protonated CG-D385 5FN2-derived simulation models compared with the per-residue 5FN2-derived cryo-EM structure reference value (grey).

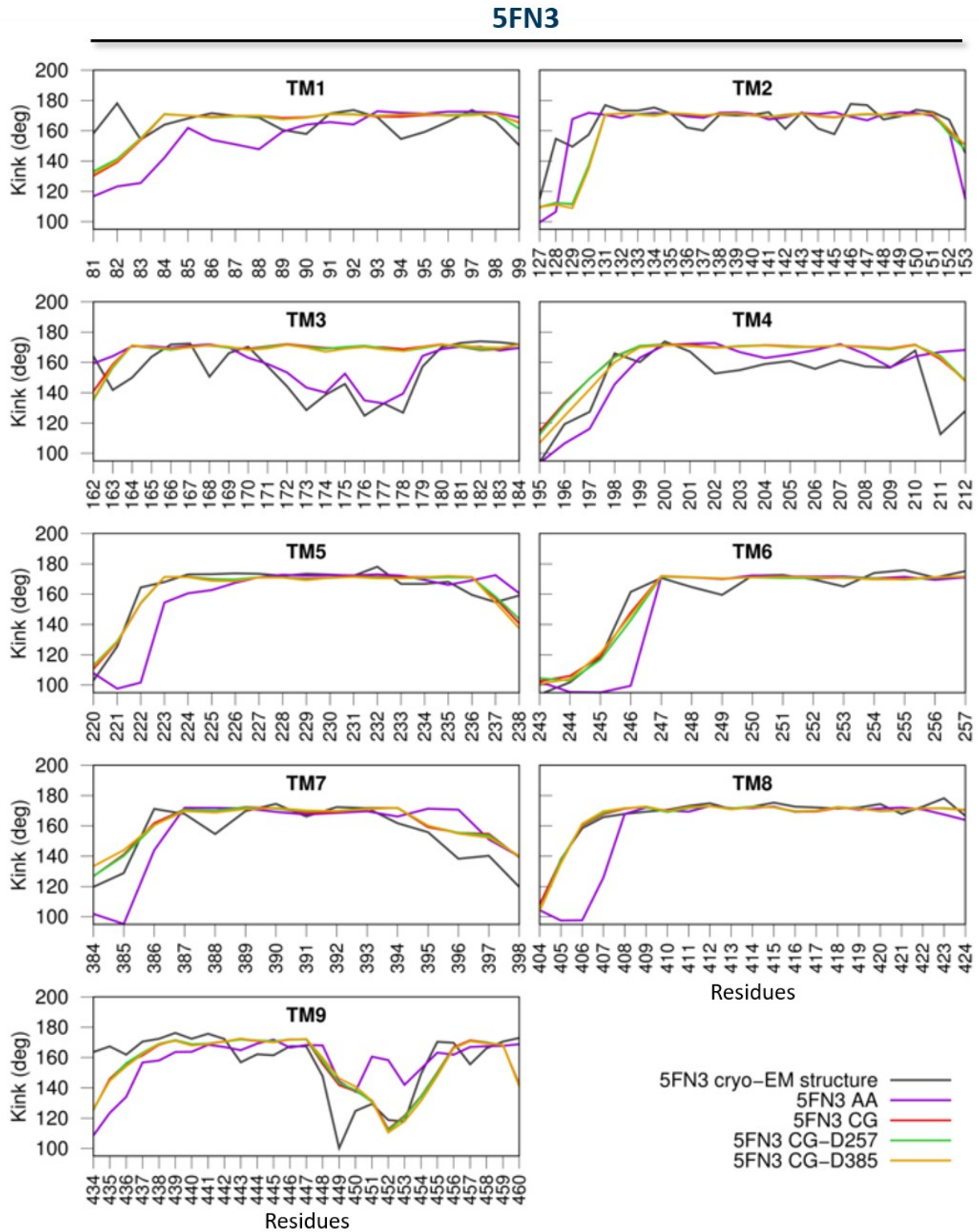


Fig. S9. Per-residue kink angles calculated for each PS1 TM helix in the AA (dark-violet), CG (red), protonated CG-D257 (green), and protonated CG-D385 5FN3-derived simulation models compared with the per-residue 5FN3 cryo-EM structure reference value (grey).

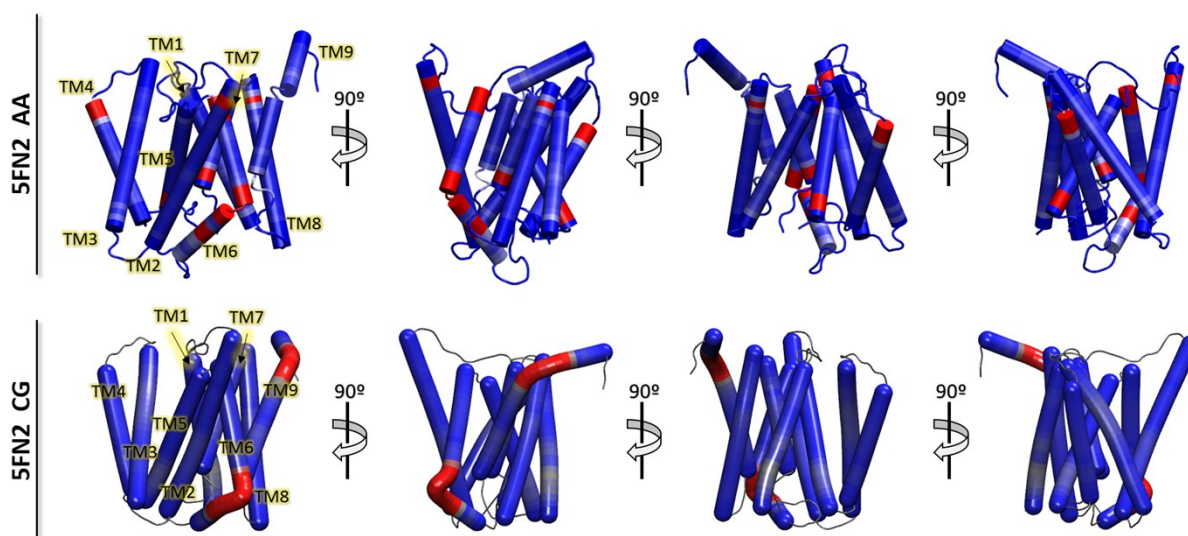


Fig. S10. Schematic representations of 5FN2 PS1-derived atomistic (top) and CG Martini (bottom) models of γ -secretase color-coded by our calculated kink angles, from more kinked (red) to less kinked (blue) values. The CG Martini model representation was generated with Bendix.

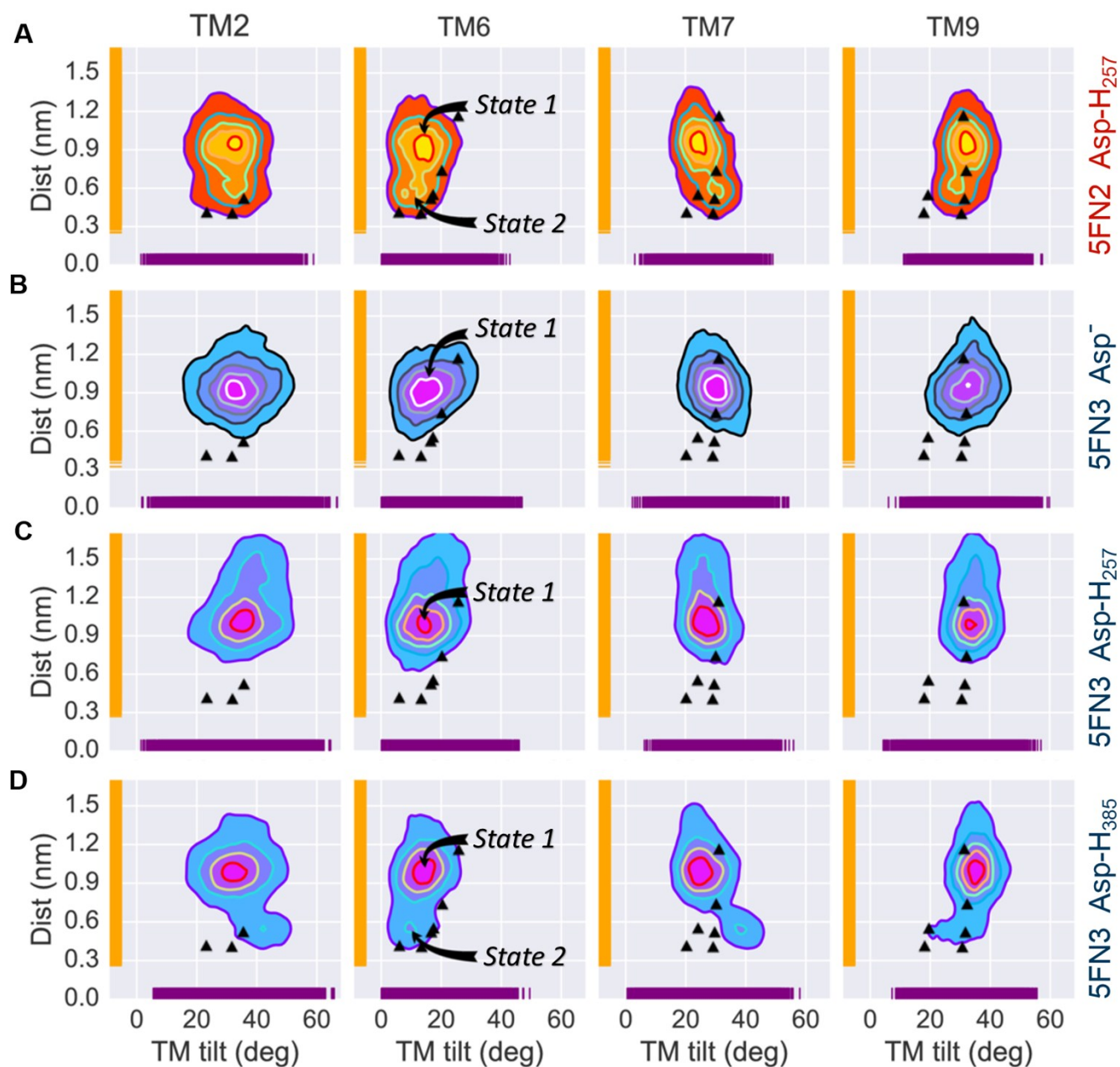


Fig. S11. Simulated distributions of 5FN2 (orange) and 5FN3 (blue) derived CG models in a POPC bilayer projected onto (1) the distance between the catalytic residues (Asp257 and Asp385) and (2) the calculated TM2, TM6, TM7 and TM9 tilt angles in the unprotonated and Asp257 and Asp385 protonated states. The black triangles depict the values of dd_{Asp} and T_{TM} angle obtained from the experimental structures of γ -secretase (PDB IDs: 5A63, 4UIS, 5FN2, 5FN3, 5FN4 and 5FN5).

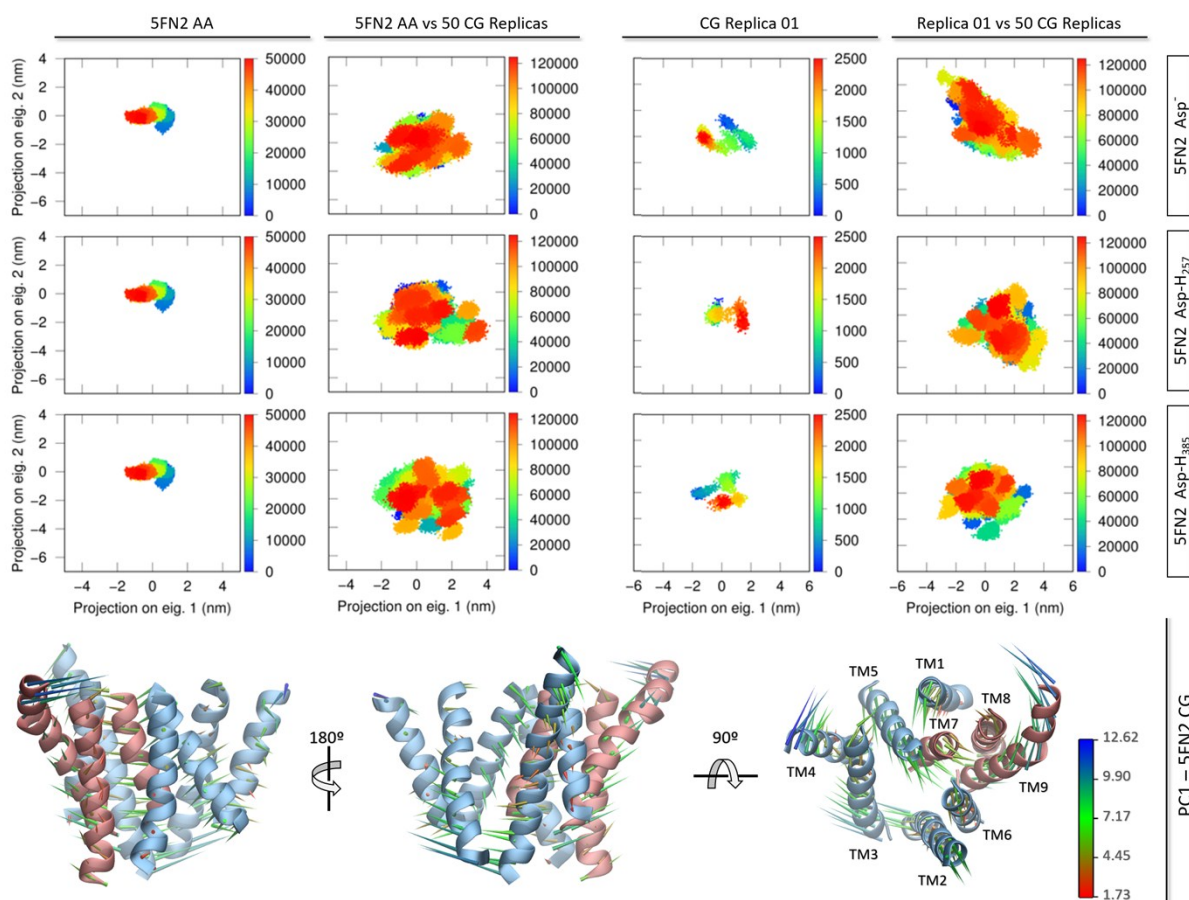


Fig. S12. First column, 2D projection of the first and second eigenvectors obtained from the PCA of the last 500 ns of the AA simulated model (5FN2 AA column) for each of the protonated states models: 5FN2 unprotonated, (top), 5FN2 Asp-H257 (middle), and 5FN2 Asp-H385 (bottom). Second column, the fifty CG trajectories projected onto the 5FN2 AA first and second eigenvectors (5FN2 AA vs 50 CG replicas) for the same protonated and unprotonated models. Third column, 2D projection of the first and second eigenvectors obtained from the PCA of one of the Martini CG replicas (CG Replica 01) (out of fifty) for the different protonated and unprotonated models. Fourth column, the projection of the 50 CG trajectories onto the CG Replica 01 first and second eigenvectors (Replica 01 vs 50 CG Replicas) for each of the protonated and unprotonated models. Bottom, PS1 3D-structure models and a porcupine representation of the first eigenvector associated with the transition between *state-1* (*inactive*) and *state-2* (*active*) conformations of PS1.

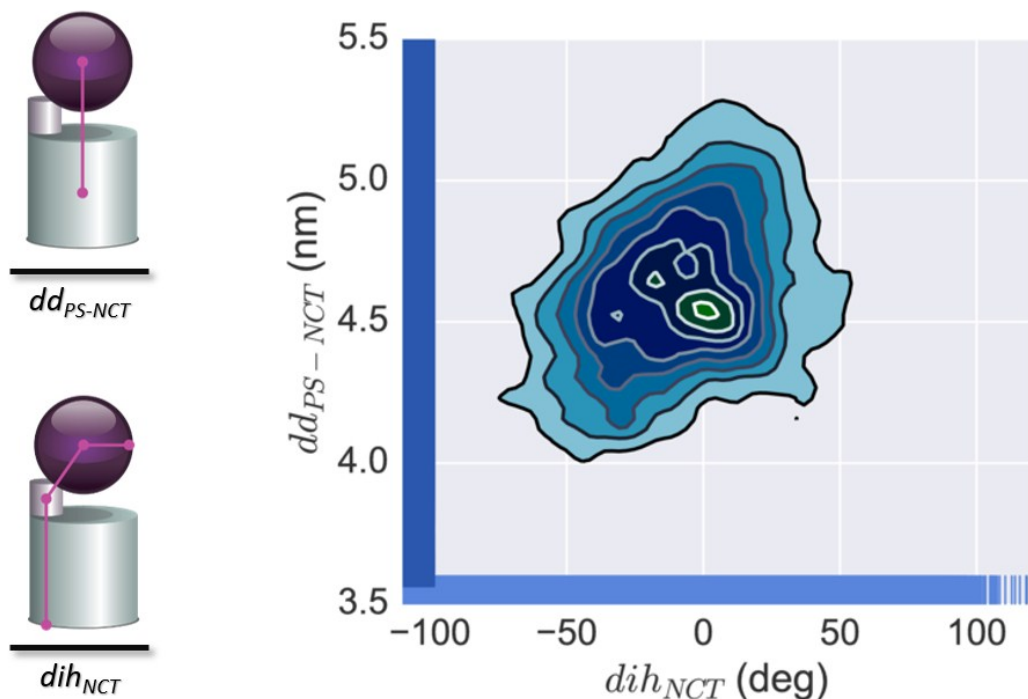


Fig. S13. Simulated distribution for 5FN2 and 5FN3 derived CG models in a POPC bilayer projected onto dd_{PS-NCT} distance and dih_{NCT} rotation.

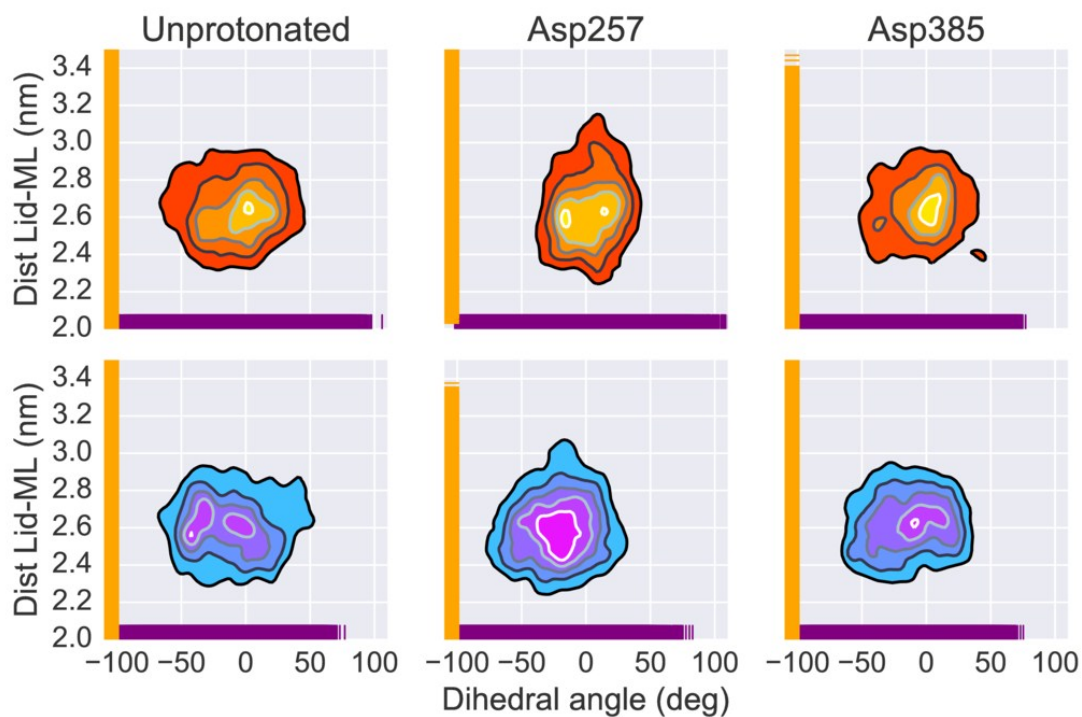


Fig. S14. Simulated distributions of 5FN2 (orange) and 5FN3 (blue) derived CG models in a POPC bilayer projected onto (1) the distance between center of geometries of the NCT lid and the NCT major lobe and (2) the NCT rotation (dihedral angle) in the unprotonated and protonated (Asp257 and Asp385) states.

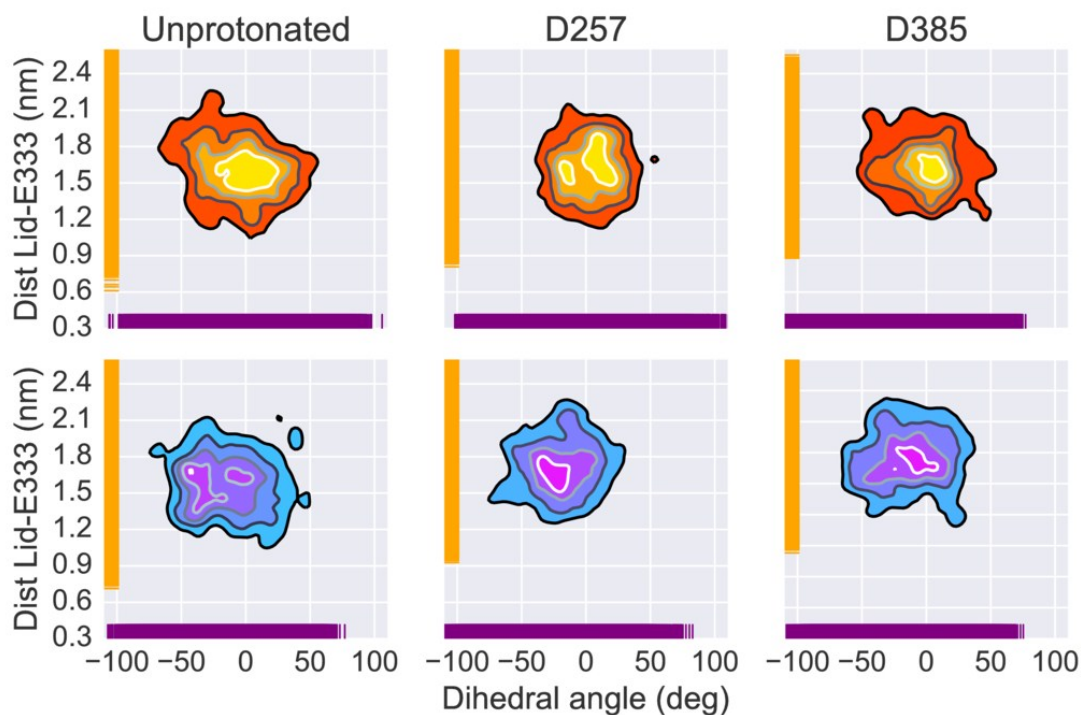


Fig. S15. Simulated distributions of 5FN2 (orange) and 5FN3 (blue) derived CG models in a POPC bilayer projected onto (1) the NCT rotation (dihedral angle) and (2) the distance between the NCT lid center of geometry and Glu333, shown for the unprotonated and two protonated (Asp257 and Asp385) states.

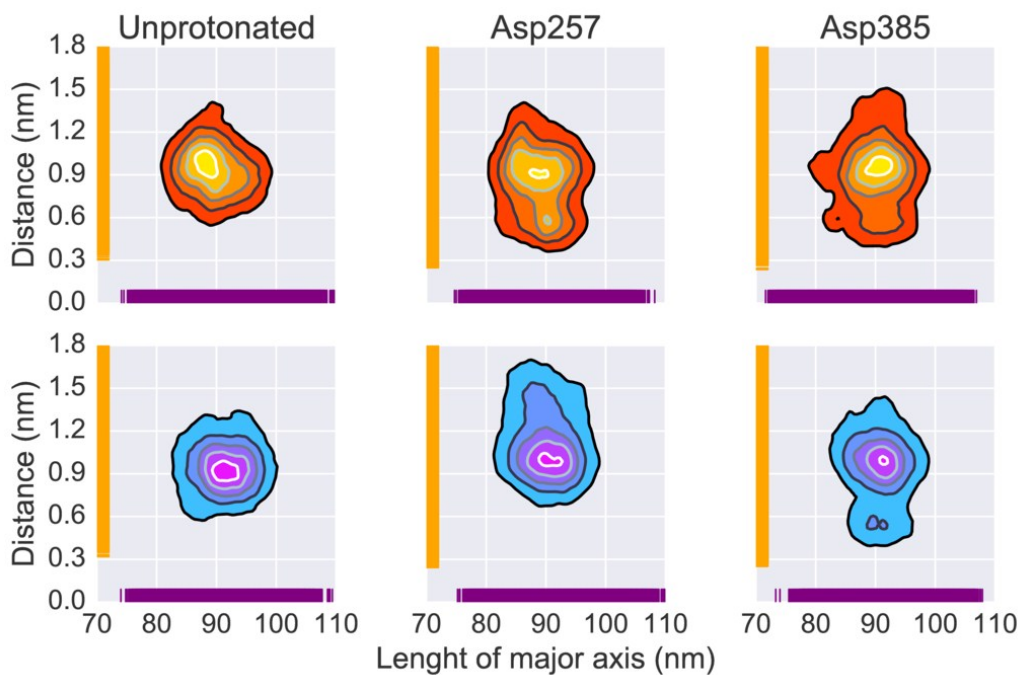


Fig. S16. Simulated distributions of 5FN2 (orange) and 5FN3 (blue) derived CG models in POPC bilayer projected onto (1) the length of the major axis of the γ -secretase complex and (2) the distance between the catalytic residues (Asp257 and Asp385) in the unprotonated and protonated (Asp257 and Asp385) states.