



UNIVERSIDAD DE TALCA
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA EN BIOINFOMÁTICA

**CAMBIOS CONFORMACIONALES EN
PROTEÍNAS TRANSPORTADORAS DE
MEMBRANA DEFINIDOS POR REPETICIONES
DE TOPOLOGÍA INVERTIDA**

MARIELA GONZÁLEZ AVENDAÑO

Profesor Tutor: DRA. ARIELA VERGARA JAQUE
Profesor Informate: DR. JULIO CABALLERO RUIZ

Memoria de título
Talca - Chile
Enero, 2020

CONSTANCIA

La Dirección del Sistema de Bibliotecas a través de su unidad de procesos técnicos certifica que el autor del siguiente trabajo de titulación ha firmado su autorización para la reproducción en forma total o parcial e ilimitada del mismo.



Talca, 2021

Contenidos

1. Introducción	11
1.1. Proteínas transportadoras de membrana	11
1.2. Transportadores activos secundarios	12
1.3. Mecanismo de transporte de acceso alternado	14
1.4. Repeticiones de topología invertida	15
1.5. Modelado por homología usando repeticiones de intercambio (<i>Repeat-swap homology modeling</i>)	17
1.6. Caso de estudio: estructura y función de transportadores de glucosinolatos (GTR)	19
2. Hipótesis	22
3. Objetivos	22
3.1. Objetivo general	22
3.2. Objetivos específicos	22
4. Materiales y Métodos	23
4.1. Refinamiento de la estructura cristalográfica de NRT1.1	23
4.2. Modelado de la estructura de NRT1.1 en una conformación orientada al medio extracelular	25
4.3. Modelado por homología de isoformas de GTR en distintos estados conformacionales	27
4.4. Análisis estructurales de los distintos modelos	28

4.5. Búsqueda de motivos funcionales	29
5. Resultados y Discusión	32
5.1. Construcción del modelo de largo completo del transportador de nitrato NRT1.1	32
5.2. Modelado de la estructura de NRT1.1 en una conformación orientada al medio extracelular	35
5.2.1. Identificación de las unidades repetidas	35
5.2.2. Alineamiento estructural y de secuencia de las unidades repetidas	37
5.2.3. Preparación del archivo PDB invertido	39
5.2.4. Construcción del modelo orientado hacia el medio extracelular . .	40
5.3. Modelado por homología de isoformas de GTR en distintos estados conformacionales	44
5.4. Análisis estructurales de los distintos modelos	47
5.5. Búsqueda de motivos de secuencia	52
6. Conclusiones	54
7. Referencias	56

Lista de Figuras

1.1.	Mecanismo de acceso alternado para un transportador de iones/sustratos.	15
1.2.	Pliegues característicos en transportadores activos secundarios.	16
1.3.	Protocolo de modelado por homología de intercambio repetido.	18
1.4.	Estructura base de glucosinolatos.	20
4.1.	Modelado de bucles de cierre cinemático (Rosetta KIC).	24
4.2.	Protocolo de modelado por homología usando repeticiones de intercambio.	27
4.3.	Etapas para la descripción de motivos funcionales en los transportadores GTRs.	30
5.1.	Modelo de largo completo del transportador de nitrato NRT1.1.	33
5.2.	Gráfico de las puntuaciones de ProQM para estimar la calidad estructural del modelo.	34
5.3.	Gráfico de Ramachandran.	35
5.4.	Identificación de las unidades repetidas.	37
5.5.	Alineación estructural de las unidades repetidas.	38
5.6.	Alineamiento de secuencia de las unidades repetidas.	39
5.7.	Esquema del archivo PDB invertido.	39
5.8.	Modelo de intercambio repetido.	41
5.9.	Gráfico de Ramachandran.	42
5.10.	Gráfico de las puntuaciones obtenidas con ProQM para evaluar la calidad estructural del modelo.	42

5.11. Comparación de las estructuras del transportador de nitrato NRT1.1 en una conformación orientada hacia el medio interno y externo.	44
5.12. Alineamientos de secuencias entre las isoformas de GTRs y el transportador de nitrato NRT1.1 usados para la generación de modelos.	46
5.13. Ruta de transporte de sustratos en la estructura de transportadores GTRs.	48
5.14. Residuos sugeridos para estar implicados en la unión a sustrato y acoplamiento de protones en las tres isoformas de transportadores GTRs. . . .	50
5.15. Análisis estructural de las tres isoformas de los transportadores de glucosinolatos.	51
5.16. Conservación de residuos en transportadores GTRs.	53

Lista de Tablas

5.1. Residuos claves en el acoplamiento de protones y transporte de sustrato en transportadores pertenecientes a la familia NPF y POT	49
---	----

Agradecimientos

Al concluir esta etapa de mi vida quiero extender un profundo agradecimiento, a todas aquellas personas que me acompañaron en este camino, a quienes estuvieron ahí para alentarme y apoyarme en todo momento.

Primero a mi familia, que siempre están brindándome consejos y demostrándome que no importa que suceda en el futuro, ellos estarán ahí para impedir que me rinda y logre mis objetivos. Ellos son la fuerza que me impulsa a seguir.

También quiero agradecer a mis amigos y compañeros de carrera por los buenos momentos vivos durante este ciclo, anécdotas y risas que no olvidare.

A la Dra. Ariela Vergara quien no solo guio esta tesis, sino que se ha convertido en un pilar fundamental para mí, otorgándome sus consejos y palabras de aliento, además de ser una inspiración constante en mi carrera.

A cada uno de ellos, muchas gracias

“Esta tesis esta dedicada a ti, mamá”

RESUMEN

Los transportadores activos secundarios son proteínas transmembranales que median el paso de una amplia gama de solutos a través de la membrana lipídica. El proceso de transporte se basa en un mecanismo de acceso alternado que alcanza al menos tres estados distintivos: una conformación orientada al medio interno, otra conformación intermedia ocluida y una conformación orientada al medio externo. Pese a estos tres posibles estados conformacionales, las proteínas transportadoras tienden a cristalizar en un solo estado. El protocolo denominado “modelado por homología usando repeticiones de intercambio” permite actualmente estimar la conformación opuesta de un transportador usando como templado su conformación inversa. En el presente estudio se utilizó este protocolo para modelar el transportador de nitrato NRT1.1 de *Arabidopsis thaliana* perteneciente a la familia NPF en una conformación orientada hacia el medio externo. Posteriormente, sobre la base de la estructura cristalográfica de NRT1.1 y el modelo generado en una conformación inversa, se construyeron modelos comparativos de tres simportadores de H⁺/glucosinolato (GTR1, GTR2 y GTR3). A través de un análisis estructural de estas proteínas se logró identificar elementos claves para el transporte de nutrientes en proteínas transportadoras de plantas. Adicionalmente, se identificaron residuos implicados en la especificidad y translocación de sustratos. Los resultados obtenidos a lo largo de esta investigación sugieren un sitio de unión a glucosinolatos dependiente de su cadena lateral y predicen mutaciones claves para un transporte selectivo de nutrientes.

ABSTRACT

Secondary active transporters are transmembrane proteins that allow the passage of a wide range of solutes across the lipid membrane. The transport process is based on an alternate access mechanism that reaches at least three distinctive states: inward-facing, occluded and outward-facing conformation. Despite these three possible conformational states, membrane protein transporters tend to crystallize in a single state. The protocol called Repeat swap homology modelling currently allows estimating the opposite conformation of a transporter using its inverse conformation as template. In the present study, this protocol was used to model the nitrate transporter NRT1.1 of *Arabidopsis thaliana* belonging to the NPF family in an inward-facing conformation. Subsequently, on the basis of the crystallographic structure of NRT1.1 and the model generated in an inverse conformation, comparative models of three H⁺/glucosinolate symporters (GTR1, GTR2 and GTR3) were constructed. Through a structural analysis of these proteins, it was possible to identify key elements for the transport of nutrients in plant transport proteins. Additionally, residues involved in the specificity and translocation of substrates were identified. The results obtained in this investigation suggest a substrate binding site in the three transporters dependent on the glucosinolate's side chains and predict key mutations for selective nutrient transport.

1. Introducción

1.1. Proteínas transportadoras de membrana

Las proteínas de membrana se encuentran mayormente insertas en la bicapa lipídica y permiten el transporte de componentes esenciales para la célula o la eliminación de compuestos indeseados. En este sentido, la membrana celular constituye una barrera altamente selectiva que limita la difusión de ciertas sustancias, por lo cual deben existir proteínas que regulen el paso selectivo de moléculas entre el citoplasma y el medio extracelular. La estructura básica de la membrana plasmática está determinada por una bicapa lipídica dividida en tres secciones: un núcleo interno hidrófobo y dos regiones polares [1]. Dado estas características, la bicapa lipídica actúa como un filtro relativamente impermeable al flujo de moléculas hidrosolubles, haciendo necesario un mecanismo de transporte mediado por proteínas de membrana.

Las proteínas transportadoras de membrana constituyen entre el 20 y el 30% de los genes en la mayoría de los organismos [2] y desempeñan funciones claves para la homeostasis celular. La clasificación de dichas proteínas se encuentra determinada por la localización que estas presentan en la membrana lipídica, permitiendo así diferenciar dos grandes grupos: (1) proteínas periféricas que se encuentran fuera de la membrana, pero ancladas a ella por interacciones tipo enlace de hidrógeno, van der Waals o iónicas; y (2) proteínas integrales que se encuentran embebidas en la membrana de forma permanente, cuyos segmentos atraviesan una o más veces la bicapa lipídica [3].

Al igual que la membrana lipídica, las proteínas transmembranales presentan características anfipáticas. La región hidrófoba se compone de una serie de segmentos de aproximadamente 20 aminoácidos de largo, formando mayormente α -hélices que atraviesan la membrana e interactúan con las colas de los fosfolípidos situados al interior de la bicapa [4]. Por el contrario, las regiones hidrofílicas o polares se encuentran expuestas al solvente en ambos lados de la membrana. Residuos de tirosinas y triptófanos sirven como elementos de anclaje para estas proteínas, interactuando simultáneamente con el interior hidrófobo de la membrana y el medio exterior acuoso [5].

Un número importante de proteínas de membrana realizan su función una vez que se asocian a otros polipéptidos formando una estructura oligomérica [5]. Las funciones

que desempeñan estas proteínas son muy variadas, destacando: el intercambio de iones de calcio, sodio o potasio a través de bombas y canales iónicos; la adhesión a la matriz extracelular o a otras células llevada a cabo por integrinas, cadherinas, selectinas y otras proteínas para formar andamiajes tridimensionales; el transporte de moléculas que participan en la comunicación celular y que actúan como receptores de señales para factores de crecimiento, neurotransmisores y hormonas; y la translocación de nutrientes como nitrógeno, fósforo, potasio, entre otros, esenciales en plantas para la síntesis de sus componentes celulares o utilizados como fuente de energía [6].

1.2. Transportadores activos secundarios

Las proteínas transportadoras de membrana median el paso de solutos a través de una serie de cambios conformacionales, los cuales permiten unir y liberar un sustrato ya sea desde el medio extracelular al citoplasma o viceversa. Específicamente, los transportadores activos secundarios corresponden a una forma de transporte activo que acopla el movimiento de un ion (típicamente Na^+ , Cl^- , H^+) [7] respecto a su gradiente electroquímico. De manera general, el método de transporte secundario depende de los gradientes electroquímicos generados por transportadores activos primarios [8]. Las proteínas que median este tipo de transporte son llamadas cotransportadores o transportadores acoplados y se dividen en dos clases: (1) simportadores, que transportan el sustrato y las moléculas de acoplamiento en la misma dirección; y (2) antiportadores, que transportan las dos moléculas en direcciones opuestas.

La clase más representativa de los transportadores activos secundarios es la superfamilia de facilitadores principales (major facilitator superfamily, MFS) presente en las células de todos los reinos [9]. Hasta ahora, más de 15,000 genes han sido identificados como codificadores de transportadores MFS [9]. Por ejemplo, el 25% de los transportadores procarióticos pertenecen a la familia MFS; mientras que, en el genoma humano se han identificado genes que codifican al menos 110 proteínas MFS [9]. Muchos de estos transportadores utilizan la fuerza protón-motriz (PMF) para impulsar el proceso de transporte [10]. En tales casos, un transportador MFS a menudo consume un protón para transportar una molécula de sustrato electroneutral. Por lo tanto, los transportadores MFS parecen ser más eficientes energéticamente en términos de la relación estequiométrica de sustrato a protones, en comparación con la otra clase principal de transportadores activos, los transportadores de unión a ATP (ATP-binding cassette transporters, ABC).

1.3. Mecanismo de transporte de acceso alternado

Para llevar a cabo el transporte de sustratos, los transportadores activos secundarios utilizan un mecanismo de acceso alternado (Figura 1.1). En este mecanismo, los cambios conformacionales secuenciales de la proteína de transporte otorgan una accesibilidad alternativa al sitio de unión [11]. Tal accesibilidad se logra debido al cambio estructural que sufren las puertas extracelulares e intracelulares en un transportador. De esta forma, las transiciones estructurales globales son un medio necesario para la translocación de sustratos entre ambos lados de la proteína [12]. Cabe señalar que la tasa de rotación de un transportador es relativamente lenta, alcanzando entre 200 a 50,000 moléculas de sustrato translocadas por segundo [13].

El mecanismo de acceso alternado se asocia a muchos tipos de transporte, entre los que destacan el mecanismo *rocking-bundle* y el mecanismo de transporte *elevator* [14]. En el primero, la unión del sustrato entre dos dominios estructuralmente distintos cataliza el movimiento acoplado de las puertas exteriores e interiores alrededor de un sitio de unión al sustrato ubicado centralmente [15]. Un ejemplo de un transportador que utiliza este mecanismo es el transportador de leucina bacteriana (LeuT) [16]. Por otra parte, en el mecanismo *elevator*, el sitio de unión al sustrato se limita en gran parte, o completamente, a un solo dominio rígido e inmóvil. De esta forma, la barrera de transporte permanece en una posición fija y el sustrato se mueve a través de ella, desde el lado extracelular al intracelular. Se han propuesto transiciones estructurales tipo *elevator* para transportadores de sacáridos acoplados a procesos de fosforilación [17], un transportador de vitamina C [18], el transportador de dicarboxilato de sodio vINDY [19] y el transportador de citrato [20].

Los elementos estructurales específicos que producen cambios conformacionales con accesibilidad diferente en las proteínas de transporte son en gran parte desconocidos. A diferencia de los transportadores acoplados a ATP, como la bomba *sarco/endoplasmic reticulum* $Ca^{2+}ATPase$ que cuenta con múltiples estructuras cristalinas representando los distintos estados conformacionales que constituyen el ciclo de transporte [21], las estructuras actualmente disponibles de transportadores secundarios proporcionan una conformación única de la proteína, con el sitio de unión al sustrato expuesto ya sea a un lado u otro de la membrana.

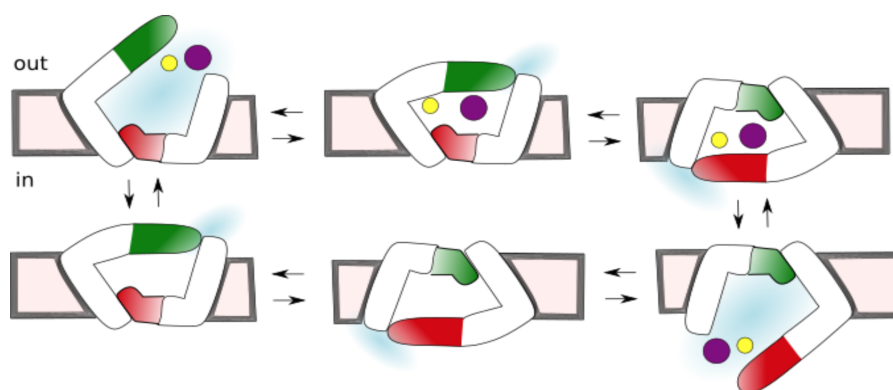


Figura 1.1: El ciclo comienza con un transportador vacío orientado hacia el medio extracelular, al cual se le une un sustrato y el ion acoplado. Esto genera una reacción que induce el cierre de la puerta extracelular (verde). El transportador entonces experimenta un cambio conformacional que expone el sitio de unión a sustrato hacia el medio intracelular. La puerta intracelular (rojo) queda habilitada y se abre, permitiendo la liberación del sustrato y el ion en el citoplasma. Finalmente, se produce un estado del transportador vacío, el cual experimenta nuevamente cambios conformacionales para exponer el sitio de unión al sustrato hacia el medio externo.

1.4. Repeticiones de topología invertida

Una característica estructural común en la mayoría de los transportadores activos secundarios es la presencia de repeticiones estructurales que comprenden varios segmentos transmembrana (TM) repetidos dos o más veces (Figura 1.2) [21]. Dichas repeticiones son relativamente comunes y ocurren en al menos el 14 % de todas las proteínas [17].

Curiosamente, los segmentos transmembranales invertidos muestran una pseudosimetría, cuyo eje puede ser paralelo o perpendicular a la membrana. Cuando el eje de simetría es paralelo a la membrana normal, los extremos N- y C-terminales de los protómeros se ensamblan en la membrana en la misma dirección [22]. Estos homooligómeros generalmente siguen la regla positiva interna, que establece que los residuos cargados positivamente de arginina y lisina tienden a mirar hacia la sección interna de la membrana [23].

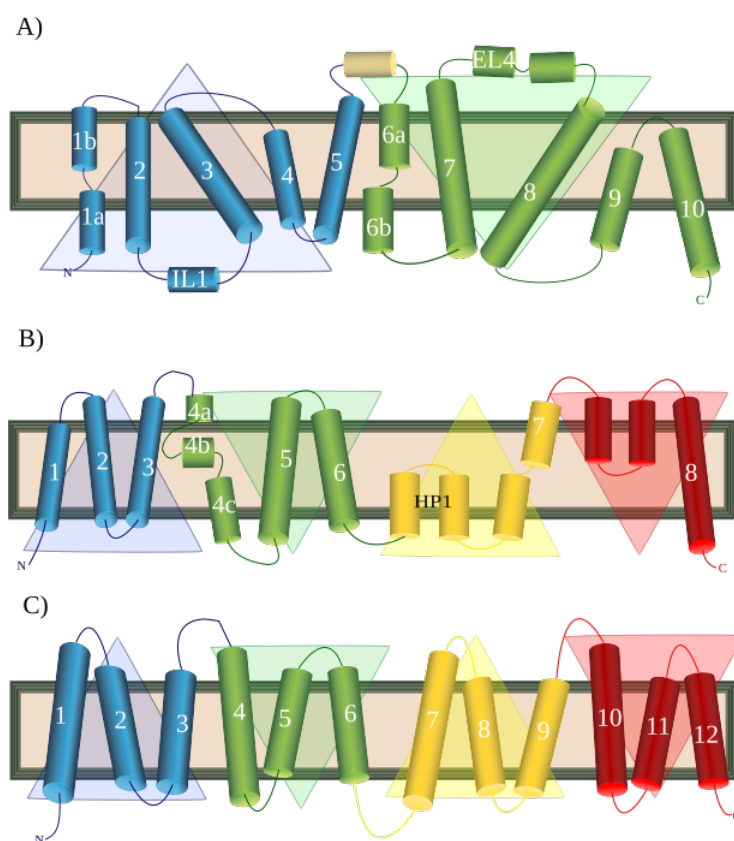


Figura 1.2: (A) Pliegue LeuT: comprende 10 TMs que se organizan en dos repeticiones estructurales invertidas, cada una de las cuales contiene 5 TMs consecutivas. (B) Pliegue GltPh: contiene 8 TMs, así como un motivo de hélice-giro-hélice, que apuntan hacia el centro de la membrana desde lados opuestos [24]. (C) Pliegue de la superfamilia facilitadora principal (MFS): contiene 12 TMs que se organizan en dominios N- y C-terminal relacionados. Cada repetición de topología invertida se resalta mediante triángulos.

Cuando el eje de simetría es perpendicular a la membrana normal, las unidades asimétricas básicas son repeticiones invertidas, cada una de las cuales comprende segmentos transmembranales poco empaquetados. Los protómeros se insertan en la membrana en direcciones opuestas [5], por lo cual sus dominios N- y C-terminal se organizan en extremos opuestos [25].

La pseudosimetría interna está presente no solo en las proteínas transportadoras, sino también en algunos canales como las acuaporinas y el canal FocA [26], en el que dos segmentos transmembrana de repetición invertida se unen para formar un protómero. La mayoría de los transportadores muestran poca similitud de secuencia entre repeticiones; por el contrario, dos repeticiones en una acuaporina comparten una homología de secuen-

cia considerable, lo que sugiere un evento de fusión génica durante la evolución [24]. Por tanto, tomando como referencia la conservación en los plegamientos estructurales de las proteínas de membrana, resulta interesante entender cómo estas proteínas han evolucionado para proporcionar funcionalidades ampliamente divergentes [21].

1.5. Modelado por homología usando repeticiones de intercambio (*Repeat-swap homology modeling*)

Como se mencionó anteriormente, los transportadores activos secundarios tienden a ser cristalizados en uno de sus principales estados conformacionales (*inward-facing* y *outward-facing*). Esto dificulta la caracterización de sus estructuras en otras conformaciones y por ende limita la comprensión de sus mecanismos de transporte [28].

Tomando en consideración la pseudosimetría doble presente en las repeticiones de transportadores activos secundarios, Forrest et al. en el año 2008 [16] desarrollaron un protocolo computacional (Figura 1.3) para crear un estado asimétrico de un transportador en una conformación opuesta, utilizando la estructura de rayos X de la proteína en la conformación complementaria como templado.

Para construir un modelo de intercambio repetido es necesario en primera instancia identificar las unidades repetitivas estructurales en la proteína que se utilizará como templado. Posteriormente, se genera un alineamiento estructural y de secuencia en base a las unidades repetidas. Luego, se elabora un templado con la estructura de las unidades de repetición intercambiadas y se utiliza esta estructura como plantilla para construir un modelo de intercambio de la proteína correspondiente a la conformación estructural opuesta.

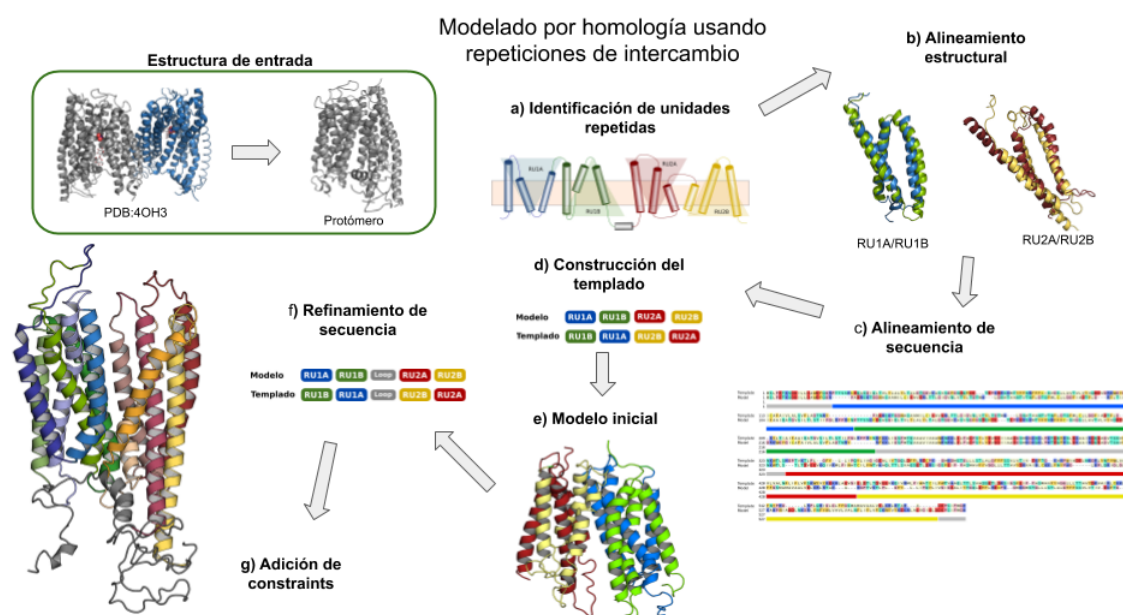


Figura 1.3: En detalle se muestran los pasos necesarios para obtener la conformación opuesta de un transportador usando su conformación complementaria como templado. Inicialmente, se (a) identifican las unidades repetidas en el templado, para luego (b) generar una superposición estructural de las repeticiones y (c) extraer un alineamiento de sus secuencias. Posteriormente, se (d) construye un ensamble estructural del templado invertido, el cual es usado para (e) generar un modelo en estado opuesto de la proteína de interés. Finalmente, para optimizar el modelo se (f) realiza un refinamiento del alineamiento usado en el proceso de modelado y se (g) incorporan ciertas restricciones estructurales para forzar el cambio conformacional.

La técnica de modelado de intercambio repetido fue aplicada por primera vez en el transportador de aminoácidos bacterianos LeuT de *Aquifex aeolicus* [28]. La comparación de la estructura de rayos X del transportador orientada hacia el exterior con el modelo orientado hacia el interior obtenido por medio de esta técnica, predijo con precisión la ruta de transporte de sustratos en transportadores NSS (neurotransmitter sodium symporter) [16]. En la actualidad esta metodología ha sido utilizada con éxito para obtener estructuras de varios transportadores, incluyendo: el de dicarboxilato bacteriano (VcINDY) [19], intercambiadores de aniones [29], transportadores de nucleósidos concentrativos, PepT1 Y Oct2 [30], proteínas pertenecientes a la familia SLC [31], además de cotransportadores de fosfato de sodio NaPi-II [32], entre otros.

1.6. Caso de estudio: estructura y función de transportadores de glucosinolatos (GTR)

La relevancia fisiológica del nitrógeno para las plantas está determinada por sus efectos en el crecimiento de las hojas, la senescencia, la arquitectura del sistema radicular, el tiempo de floración, entre otros aspectos [33]. La mayoría de los organismos como plantas, algas, hongos filamentosos, algunas bacterias y levaduras, son capaces de incorporar nitrógeno a partir de los iones de nitrato, nitrito y amonio como principales fuentes de nitrógeno inorgánico. Las características químicas tanto del nitrito como del nitrato hacen necesaria la existencia de sistemas de transporte para movilizar estas moléculas al interior de las células [34]. El transporte de nitrato y nitrito constituye un paso clave en el control de la cantidad de nitrógeno incorporado en las células y su subsecuente almacenamiento, reducción y exportación.

En plantas, un grupo de proteínas de membrana pertenecientes a la familia NRT1/PTR (*nitrate transporter/proton-dependent oligopeptide transporter*), más tarde denominada familia NPF, fueron identificadas como transportadores de nitrato o di/tri-péptido. Estudios recientes revelaron que esta familia de transportadores también permite la translocación de hormonas vegetales como auxinas (ácido indol-3-acético), ácido abscísico (ABA) y giberelinas (GA), así como metabolitos secundarios [35].

El primer miembro descubierto de la familia NPF es conocido como AtNRT1.1 o CHL1, identificado en *Arabidopsis thaliana* [36]. Esta proteína funciona como un transportador de doble afinidad, es decir, cambia de un estado de alta afinidad a otro de baja afinidad (HATS y LARS) por nitrato en función de la concentración de dicho nutriente en el medio [34]. En condiciones de bajas concentraciones de nitrato, un residuo Treonina en la posición 101 (T101) en AtNRT1.1 es fosforilado por una proteína kinasa CIPK23 [35], dejando al transportador en un estado de alta afinidad. A altas concentraciones de nitrato, sin embargo, el transportador es desfosforilado, convirtiéndose en un transportador de baja afinidad.

El transportador NRT1.1 adopta un plegamiento MFS, que se caracteriza por 12 segmentos transmembranales organizados en repeticiones guiadas por un pseudo eje simétrico como se nombró anteriormente. Esta proteína está conformada por 590 aminoácidos y se encuentra altamente conservada entre los ortólogos de NRT1.1 de plantas [36].

Un grupo de metabolitos secundarios transportados por proteínas pertenecientes a la

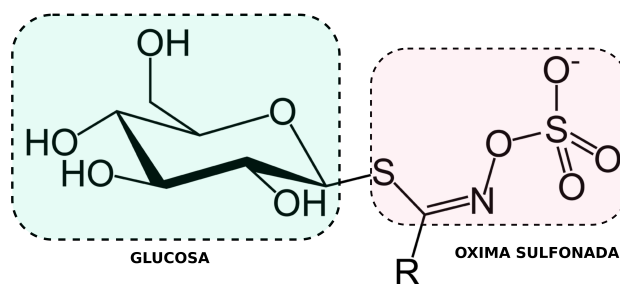


Figura 1.4: La imagen muestra los fragmentos que constituyen los glucosinolatos incluyendo: un fragmento de β -tioglucosa, un fragmento de oxima sulfonada y una cadena lateral de aglicona variable derivada de un α -aminoácido.

familia de NPF son los glucosinolatos [33]. Estos metabolitos son químicamente estables en condiciones normales [37], pero cuando las células y los tejidos de las plantas sufren algún daño, son hidrolizados por la enzima miosinasa, lo que conduce a la degradación de varios productos, como los isotiocianatos [38] que están involucrados en la defensa de las plantas contra patógenos y herbívoros. Por otra parte, los glucosinolatos en humanos son ingeridos a través del consumo de vegetales como el brócoli, col rizada y coles de Bruselas, y confieren propiedades anticancerígenas [6].

Los glucosinolatos se sintetizan en casi todas las partes de las plantas, a excepción de las semillas. Sin embargo, el contenido de glucosinolatos en semillas es mayor que en otros tejidos y órganos, lo cual se debe a un sistema de transporte asociado a H^+ /glucosinolatos. Actualmente, se han identificado tres genes que codifican proteínas transportadoras de glucosinolatos: AtGTR1 (NPF2.10), AtGTR2 (NPF2.11) y AtGTR3 (NPF2.9) [22].

El presente proyecto consistió en la implementación del protocolo denominado “modelado por homología usando repeticiones de intercambio” para inicialmente modelar el transportador de nitrato NRT1.1 de *Arabidopsis thaliana* en una conformación orientada hacia el exterior, utilizando como templado su estructura cristalográfica en una conformación orientada hacia el interior. Ambas conformaciones estructurales se utilizaron posteriormente para modelar los tres simportadores de H^+ /glucosinolato encontrados en *Arabidopsis thaliana* (GTR1, GTR2 y GTR3). Una vez obtenidas las estructuras de todos los transportadores en conformaciones orientadas al medio intra y extracelular, se realizó un análisis comparativo a nivel estructural y de secuencias con el propósito de identificar divergencias evolutivas y funcionales entre todos los transportadores.

2. Hipótesis

Es posible identificar los dominios estructurales involucrados en el mecanismo de transporte de sustratos y aquellos residuos que participan en la translocación de protones en simportadores de H^+ /glucosinolato, a través de un modelado comparativo de las proteínas GTR1, GTR2 y GTR3 basado en las conformaciones orientadas al medio interno y externo del transportador NRT1.1 de plantas.

3. Objetivos

3.1. Objetivo general

Modelar la estructura de proteínas transportadoras de H^+ /glucosinolato en plantas, con el propósito de identificar los motivos de secuencia y elementos estructurales que contribuyen al transporte de sustratos.

3.2. Objetivos específicos

- Refinar la estructura del transportador de nitrato NRT1.1 en una conformación orientada al medio intracelular.
- Construir un modelo por homología del transportador de nitrato NRT1.1 en una conformación orientada al medio extracelular usando el método de modelado de intercambio.
- Modelar los simportadores de H^+ /glucosinolatos GTR1, GTR2, y GTR3 sobre la base estructural de NRT1.1 en sus conformaciones orientadas hacia el medio intra y extracelular.
- Caracterizar la estructura de los modelos generados e identificar motivos de secuencias que determinen sus divergencias evolutivas y funcionales.

4. Materiales y Métodos

4.1. Refinamiento de la estructura cristalográfica de NRT1.1

Para el desarrollo de esta tesis se utilizó como referencia la estructura del transportador de nitrato NRT1.1 que se encuentra disponible en la base de datos Protein Data Bank (PDB) bajo el código de acceso 4OH3 [36]. El cristal presenta una resolución de 3.25 Å y contiene los residuos desde la posición 9 a la 581; por consiguiente, la proteína no presenta los dominios N- y C-terminal, además de los loops entre los residuos 282-295, 311-324 y 451-458.

Para refinar la estructura de NRT1.1 y adicionar los fragmentos faltantes, se implementó un protocolo en base a modelado *ab initio*, que permite predecir el plegamiento de los fragmentos ausentes en la estructura de una proteína.

El modelo de la secuencia de largo completo del transportador de nitrato (Código Uniprot Q05085) se obtuvo utilizando el software Rosetta [39] que implementa un algoritmo de predicción de estructura de *novo*, basado en el ensamble de fragmentos cortos de proteínas conocidas para producir conformaciones de tipo nativas mediante una estrategia de Monte Carlo [40]. Los fragmentos fueron proporcionados a través del servidor Robetta, que genera una librería de fragmentos a partir de una secuencia aminoacídica, la cual utiliza para hacer una búsqueda de estructuras homólogas en el PDB. De esta forma, se generan fragmentos de 3 y 9 aminoácidos de longitud [41], que se ensamblan y se evalúan a través de una serie de iteraciones de modelado para estimar el plegamiento más probable.

Para el refinamiento de loops se implementó un protocolo basado en el modelado de bucles de cierre cinemático (Rosetta KIC) (Figura 4.1). Este método consiste en la división de la cadena de aminoácidos faltantes en dos segmentos incluyendo tres “puntos de pivote”: uno al comienzo de la cadena, otro en el centro de los dos segmentos (pivote medio) y uno al final [42]. El objetivo es calcular los valores de torsión del pivote medio orientando cada segmento de modo que conserve la forma original y rígida solo en los

extremos del bucle [42]. Este protocolo permite la construcción y refinamiento de bucles ausentes en una estructura proteica con el propósito de encontrar conformaciones de mínima energía a través de un muestreo de al menos 10.000 modelos.

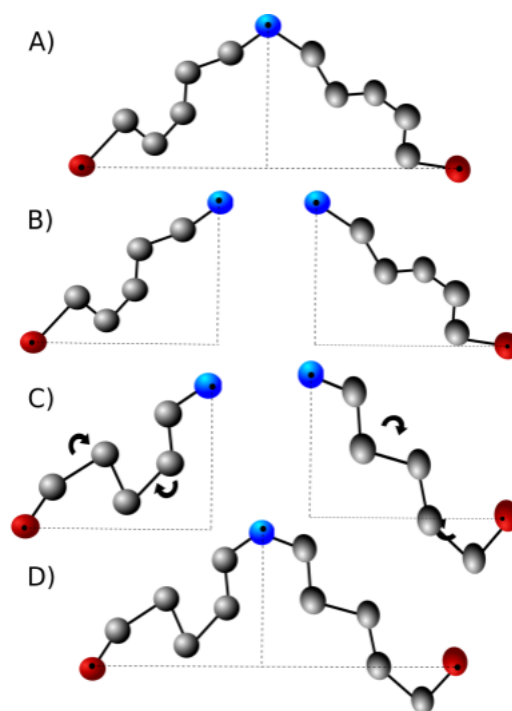


Figura 4.1: (A) Se definen tres “puntos de pivote”: uno al comienzo y al final de la cadena (rojo), y otro en el centro de los dos segmentos (pivote medio, azul). (B) La cadena se divide en dos segmentos respecto a su pivote medio, mantenido rígida la posición de cada pivote. (C) Se calculan los valores de torsión de modo que conserve la forma original y rígida solo en sus extremos. (D) Se unen los dos segmentos para formar la cadena de largo original [42].

El mejor modelo de NRT1.1 fue seleccionado en base al puntaje que asigna el software Rosetta a través de una serie de potenciales estadísticos. Las funciones de puntuación en Rosetta son sumas ponderadas de términos de energía, algunos de los cuales representan fuerzas físicas como la electrostática y las interacciones de van der Waals, mientras que otros representan términos estadísticos como la probabilidad de encontrar los ángulos de torsión en el espacio de Ramachandran [43].

Para la evaluación de la calidad de los modelos se utilizaron softwares orientados a predecir la calidad de las estructuras de proteínas incluyendo: (1) ProQM que utiliza una serie de parámetros en combinación con características de proteínas de membrana

para pronosticar la calidad de cada residuo [44] y (2) PROCHECK que verifica la calidad estereoquímica de una estructura proteica, produciendo una serie de gráficos de Ramachandran que analizan la geometría general de la proteína considerando cada uno de los residuos [45].

4.2. Modelado de la estructura de NRT1.1 en una conformación orientada al medio extracelular

Para construir un modelo del transportador de nitrato en una conformación orientada hacia el medio extracelular se utilizó la técnica de modelado por homología usando repeticiones de intercambio invertidas (*Repeat-swap homology modeling*) [28].

Esta metodología (Figura 4.2) crea un nuevo estado asimétrico de un transportador en una conformación opuesta, utilizando la estructura de rayos X de la proteína en la conformación complementaria como templado. Para ello, se considera la pseudosimetría de fragmentos repetidos invertidos presente en los transportadores activos secundarios [16]

La primera etapa de este protocolo consistió en la identificación de las unidades estructurales repetidas de un protómero de la proteína a modelar, en este caso NRT1.1. La información referente a la pseudosimetría interna dentro de la estructura fue evaluada a través de herramientas como SymD[46] y CE-symm[47], además se utilizó las definiciones estructurales proporcionadas por los autores que cristalizaron la estructura de NRT1.1 disponible en el Protein Data Bank (PDB ID 4OH3).

Posteriormente se construyó un alineamiento de secuencias basado en una superposición estructural de las repeticiones invertidas, el cual fue obtenido empleando el software TM-Align [48]. Utilizando este alineamiento, la primera repetición de la proteína se modela en base a la segunda repetición, mientras que ésta última se modela en base a la primera repetición. De esta forma, se obtiene un modelo del transportador en la conformación opuesta.

Inicialmente, se utilizó un alineamiento preliminar como guía para construir un modelo por homología del transportador NRT1.1 en una conformación orientada hacia el medio externo, basado en una plantilla "intercambiada" de su estructura cristalográfica. Esta plantilla se obtuvo cambiando el orden de las repeticiones en el archivo de coordenadas

en formato PDB, de manera que la primera parte del archivo de coordenadas contenga los residuos de la segunda repetición, seguida por los residuos que constituyen la primera repetición. Los fragmentos de la proteína que no son parte de las repeticiones estructurales mantienen sus coordenadas cristalográficas durante el proceso de modelado. Para obtener una amplia búsqueda conformacional, la generación de modelos se realizó mediante el uso del software MODELLER [49], el cual implementa un enfoque automatizado para el modelado comparativo de estructuras de proteínas mediante el uso de restricciones espaciales [50].

Una vez obtenido un modelo preliminar de NRT1.1 en la conformación opuesta, el alineamiento inicial de secuencias fue optimizado, adicionando descripciones de estructura secundaria y patrones de conservación evolutivos, generados con los softwares DSSP [51] y ConSurf [52], respectivamente. Luego, se generó un nuevo conjunto de modelos utilizando un alineamiento optimizado con el fin de identificar las regiones “andamio” y los dominios de transporte en la estructura de NRT1.1.

Para conservar la disposición de estructura secundaria nativa y el empaquetamiento intramolecular de los residuos que componen la proteína a modelar, se implementaron restricciones de distancias en el protocolo de MODELLER [49] construyendo 2.000 modelos. Dichas restricciones corresponden a las posiciones de los átomos de $C\alpha$ respecto a la estructura cristalográfica.

Finalmente, la calidad de los modelos fue evaluada, seleccionando el mejor modelo como aquel que cumplió con la mayoría de las predicciones de ProQM y parámetros estereoquímicos calculados con Procheck.



Figura 4.2: Se muestran las etapas principales para llevar a cabo el modelado de NRT1.1 en una conformación orientada hacia el medio externo, indicando además las herramientas computacionales utilizadas para llevar a cabo cada objetivo.

4.3. Modelado por homología de isoformas de GTR en distintos estados conformacionales

El modelado comparativo o por homología de proteínas es una técnica predictiva que permite obtener una aproximación de la estructura atómica de una cadena polipeptídica en base a estructuras conocidas de proteínas relacionadas, en su mayoría homólogas, denominados moldes o templados [49]. El modelado comparativo consta de cuatro pasos principales: (i) asignación de pliegues, que identifica la similitud entre la proteína *target* y al menos una estructura templado conocida; (ii) alineamiento de la secuencia *target* y el o los templado(s); (iii) construcción de un modelo basado en el alineamiento generado con el/los templado(s) seleccionado(s); y (iv) evaluación de la calidad del modelo.

Los simportadores de H^+ /glucosinato encontrados en *Arabidopsis thaliana* GTR1 (Uniprot Código:Q944G5), GTR2 (Uniprot Código: Q9LV10), y GTR3 (Uniprot Código: Q9M9V7), actualmente no cuentan con una estructura cristalográfica disponible en el PDB. Por esta razón, se realizó un protocolo de modelado por homología empleando el software MODELLER [49] a fin de obtener estructuras tridimensionales representativas de las tres isoformas de GTR.

Utilizando el modelo de largo completo refinado de NRT1.1 en una conformación orientada al medio intracelular y el modelo generado a través de repeticiones invertidas en una conformación orientada al medio externo, se construyeron 3.000 modelos para cada isoforma de los transportadores GTR en ambos estados conformacionales.

Inicialmente se generaron alineamientos de secuencias empleando el servidor AlignMe [53], los cuales fueron posteriormente refinados en base a predicciones de estructura secundaria con PsiPred [54] y predicciones de segmentos transmembrana con Topcons [55].

El mejor modelo de cada isoforma de GTR fue seleccionado en base al puntaje que asigna el software MODELLER por medio de MolPDF, el cual describe qué tan bien un modelo satisface las restricciones obtenidas a partir del templado, el alineamiento de secuencia y cualquier restricción adicional aplicada. Por lo tanto, este es un valor arbitrario que depende de tales características. El modelo con el puntaje MolPDF de menor valor satisface mejor todas las restricciones.

Para la selección de los modelos de mejor calidad se utilizaron también las evaluaciones de ProQM y Procheck.

4.4. Análisis estructurales de los distintos modelos

La caracterización bioquímica y estructural de las proteínas integrales de membrana presenta grandes desafíos, debido a que durante su purificación estas proteínas son aisladas de la bicapa lipídica mediante el uso de detergentes, solventes no polares o, en casos determinados, por medio de agentes desnaturizantes [3]. El principal problema reside en que muchas proteínas de membrana tienden a desnaturizarse si se solubilizan con estos detergentes, por lo cual pierden su estructura terciaria lo que hace extremadamente difícil llevar a cabo estudios funcionales, espectroscópicos o ensayos de cristalización. La estructura de las proteínas está estabilizada por diferentes tipos de enlaces,

tales como enlaces covalentes (enlace peptídico, enlace por puentes disulfuro), enlaces de hidrógeno (interacciones dipolo-dipolo), interacciones hidrofóbicas, enlaces salinos (interacciones electrostáticas) o las fuerzas de los contactos de Van der Waals. Todos estos tipos de enlaces juegan un importante papel en el correcto plegamiento de la estructura tridimensional de las proteínas.

Las interacciones inter e intramoleculares de los modelos generados para los distintos estados conformacionales de GTR1, GTR2 y GTR3, se analizaron mediante los programas VMD [56] y PyMOL [57]. Con el propósito de evaluar los elementos estructurales que determinan los cambios conformacionales requeridos para el transporte de sustrato en cada proteína, se realizaron alineamientos estructurales de los dominios andamio en las estructuras proteicas, y luego se cuantificó el cambio conformacional de los dominios de transporte en términos de traslación, ángulo de rotación y el desplazamiento de las hélices.

Para identificar el posible dominio de transporte de cada una de las isoformas de GTR se empleó la herramienta HOLE v2.2 [58] que permite estimar las dimensiones de una cavidad, en este caso la ruta de transporte que atraviesa un modelo estructural. Además, se realizó una búsqueda de aminoácidos que pudieran permitir el paso de protones e interacciones que regulen la selectividad por ciertos sustratos. Por otra parte, mediante el uso de Bendix [59], un complemento de VMD, se analizó la flexibilidad de las α -hélices que juegan un rol esencial para el transporte de sustratos.

4.5. Búsqueda de motivos funcionales

Las secuencias conservadas son secuencias biológicas similares o idénticas que pueden encontrarse en proteína homólogas ya sea cercanas o distantes a nivel evolutivo. Estas secuencias son utilizadas para agrupar proteínas en familias, que normalmente realizan funciones similares y comparten un ancestro evolutivo común. Un motivo es un elemento conservado en la secuencia de aminoácidos o nucleótidos, que habitualmente se asocia con una función concreta. Tales motivos se incluyen en diferentes bases de datos de secuencias de proteínas para asignar funciones putativas a proteínas desconocidas. En los ortólogos de NRT1.1 en plantas se ha identificado un motivo conservado, ExxER, que juega un papel importante en el acoplamiento de protones. Este motivo se encuentra cercano al bolsillo de unión a nitrato y sus residuos se orientan al túnel de transporte de la proteína [60]. Considerando la homología entre las estructuras GTRs y NRT1.1, se realizó un análisis de secuencia entre todas las proteínas incluidas en este estudio para detectar

motivos asociados al transporte de protones (Figura 4.3).

Los motivos de patrones o las descripciones de motivos cualitativos utilizan un formato de consenso, mientras que los perfiles o las descripciones de motivos cuantitativos, utilizan puntuaciones para cada posición de un aminoácido. Un motivo ideal identifica a todos los miembros de una familia de proteínas. Sin embargo, debido al alto grado de divergencia entre algunos miembros de familias de proteínas, a menudo es difícil definir un motivo que identifique con exactitud un grupo de proteínas relacionadas.

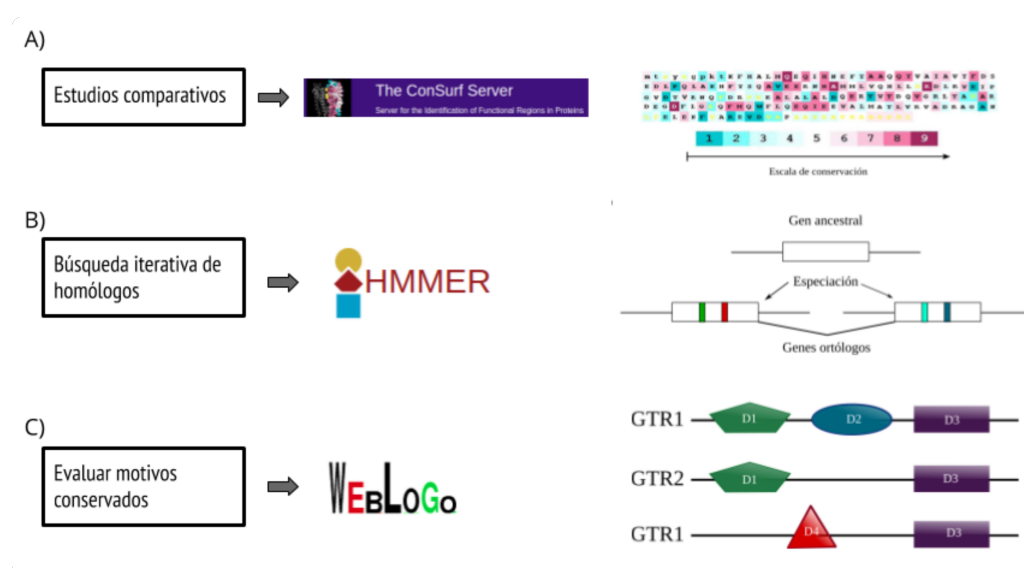


Figura 4.3: Se detallan los programas computacionales usados en cada etapa del análisis de las secuencias, incluyendo un esquema del funcionamiento del software en cada caso.

Para identificar divergencias evolutivas y funcionales entre los transportadores descritos se implementó un protocolo definido en tres fases:

1. Búsqueda iterativa de homólogos con HMMER. Este programa busca en distintas bases de datos secuencias homólogas para realizar alineamientos, implementando un método que utiliza modelos probabilísticos llamados modelos ocultos de Markov (perfil HMM). HMMER a menudo se usa en conjunto con una base de datos de perfiles como Pfam, aunque también puede trabajar con secuencias de consulta, no solo con perfiles, como BLAST. Por ejemplo, puede buscar una secuencia de consulta de proteínas en una base de datos con phmmer, o hacer una búsqueda iterativa con jackhmmmer [61].
2. Estudios comparativos para evaluar la conservación de residuos. El servidor Con-

Surf [52] es una herramienta para estimar la conservación evolutiva de aminoácidos entre secuencias homólogas. El grado en que la posición de un aminoácido se conserva evolutivamente (es decir, su tasa de evolución) depende en gran medida de su importancia estructural y funcional. Por lo tanto, el análisis de conservación de las posiciones entre los miembros de la misma familia a menudo puede revelar la importancia de cada posición para la estructura o función de la proteína.

3. Búsqueda de motivos conservados: Se realizaron búsquedas en las secuencias de proteínas de todos los miembros conocidos de la familia de los GTRs en la base de datos Uniprot. Las secuencias se ordenaron por longitud y agrupados al 60 % de identidad utilizando el algoritmo USEARCH v11. [62]. Luego, se construyó una alineación de secuencias múltiple con el servidor MAFFT v37 [63] y se analizó con Jalview v2.8[64]. La conservación de residuos se estimó como probabilidades de residuos para la alineación de secuencia múltiple completa usando el servidor Weblogo v3.4 [65].

5. Resultados y Discusión

5.1. Construcción del modelo de largo completo del transportador de nitrato NRT1.1

La estructura de los dominios N- y C-terminal, además de los *loops* entre los residuos 282-295, 311-324 y 451-458 de NRT1.1 fueron modeladas utilizando el software Rosetta, con el cual se construyeron 10.000 modelos de la estructura completa del transportador NRT1.1, implementando un algoritmo de predicción de estructura de novo que permite ensamblar los fragmentos faltantes de la proteína. La estructura cristalográfica de NRT1.1 posee dos protómeros que se pueden superponer con una desviación cuadrática media (RMSD) de 0.9 Å sobre 504 átomos de C α , lo que indica una estructura general común [36] y por ende solo se construyó el modelo en base a una sola subunidad. El mejor modelo fue seleccionado en base al puntaje asignado por Rosetta y que cumpliera las predicciones de estructura secundaria y de topología de proteínas de membrana realizadas con PSIPRED [54] y TOPCONS [55], respectivamente (Figura 5.1). El modelo seleccionado obtuvo un score de Rosetta de 2363.81 que corresponde a sumas ponderadas de términos de energía que permiten estimar que tan estable a nivel termodinámico es la estructura. Si bien, gran parte de la función energética en Rosetta se basa en principios fisicoquímicos que no tienen una conversión directa a unidades de energía física como kcal/mol, esta función incluye ciertos términos estadísticos para favorecer modelos que se relacionan con estructuras proteicas conocidas. Cabe destacar, que la naturaleza a menudo conserva los pliegues proteicos [40]; por lo tanto, es más probable que una estructura de puntuación más baja esté más cerca de la estructura nativa. Pese a ello, los cálculos de energía en Rosetta a veces se ven obstaculizados por la gran cantidad de grados de libertad a considerar y el rendimiento limitado de las funciones energéticas. En particular, el refinamiento de modelos y el modelado de *loops*, requieren búsquedas más precisas en una región del espacio conformacional [66].

Existen una serie de herramientas que ayudan a evaluar la fiabilidad de un modelo estructural; sin embargo, estos métodos predictivos tienden a mostrar resultados de menor calidad al momento de evaluar proteínas de membrana. Esto se debe a que durante la parametrización de dichas proteínas las propiedades hidrofóbicas que confiere la bicapa lipídica no son consideradas. Para corroborar los resultados obtenidos por Rosetta se

realizó un análisis mediante ProQM [44], que permite evaluar la calidad estructural global y local de un modelo seleccionado incluyendo parámetros para las propiedades de la membrana.

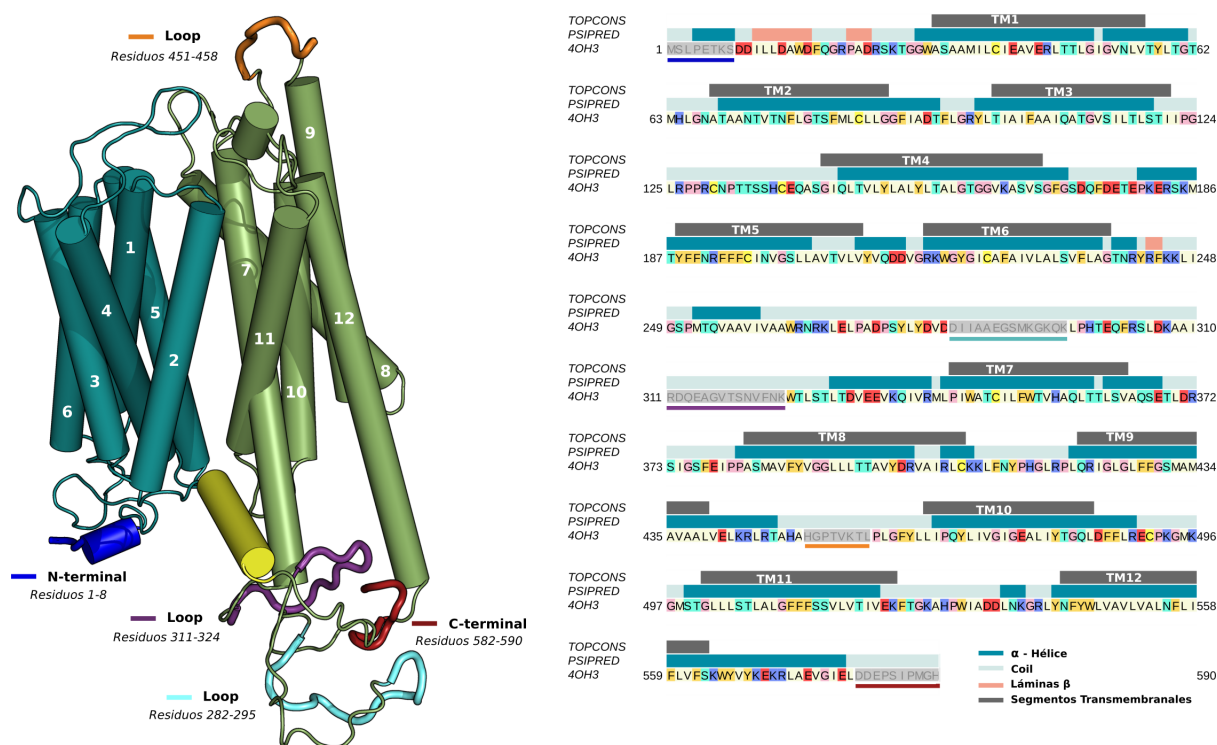


Figura 5.1: (A) Representación cartoon de la estructura del transportador. Se adicionaron los segmentos faltantes de la proteína que comprenden los dominios N- y C-terminal (azul, rojo), además de loops entre los residuos 282-295 (calipso), 311-324 (violeta) y 451-458 (naranja). (B) Predicciones de estructura secundaria y segmentos transmembranales para la secuencia de NRT1.1. Se aprecian 12 segmentos transmembranales (TMs) predichos por TOPCONS [55], lo que corrobora la topología MFS del transportador. Como se observa en la estructura, NRT1.1 posee un pseudo eje de simetría que relaciona los dominios N-terminal (TM1-6, azul) y C-terminal (TM7-12, verde)[48]. Las predicciones de estructura secundaria fueron realizadas con PSIPRED [54].

Existen una serie de herramientas que ayudan a evaluar la fiabilidad de un modelo estructural; sin embargo, estos métodos predictivos tienden a mostrar resultados de menor calidad al momento de evaluar proteínas de membrana. Esto se debe a que durante la parametrización de dichas proteínas las propiedades hidrofóbicas que confiere la bicapa lipídica no son consideradas. Para corroborar los resultados obtenidos por Rosetta se realizó un análisis mediante ProQM [44], que permite evaluar la calidad estructural global

y local de un modelo seleccionado incluyendo parámetros para las propiedades de la membrana.

Los puntajes de ProQM pueden variar de 0 a 1, siendo 1 el más parecido a una proteína de membrana, aunque las estructuras típicas tienen puntajes de ~ 0.7 [44]. El modelo seleccionado obtuvo una puntuación global de 0.664 que valida el modelo de largo completo generado por Rosetta. Pese a esto, los resultados muestran regiones de baja calidad estructural en los dominios N- y C-terminal al igual de un *loop* situado en el medio extracelular que comprende los residuos 243 al 328 (Figura 5.2). Dichas regiones no afectan el correcto funcionamiento del transportador debido a que los residuos claves implicados en la unión del sustrato y el acoplamiento de protones están situados en la región transmembranal de la proteína. Específicamente, los residuos descritos en la unión de protones y sustrato son: LYS164, HIS356 y GLU476, los cuales poseen puntajes de 0.841867, 0.854448 y 0.937315, respectivamente, validando la calidad estructural del modelo seleccionado. Por otra parte, el motivo ExxER conservado en todos los ortólogos de NRT1.1 y que desempeña un papel importante en el acoplamiento de protones muestra puntajes por sobre 0.6.

Sumado a los análisis anteriores, para validar la estructura seleccionada se realizó un análisis del modelo utilizando PROCHECK [45], el cual evalúa el grado de consistencia del modelo comparándolo con estructuras proteicas conocidas en términos de distancias de enlace, ángulos, diedros y átomos superpuestos [62]. El análisis mostró que el 99.4% de los residuos de NRT1.1 se encontraban en regiones estereoquímicamente permitidas del gráfico de Ramachandran (Figura 5.3).

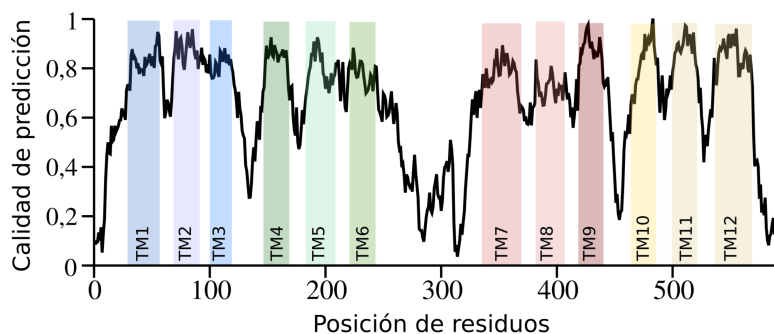


Figura 5.2: La predicción muestra que los residuos que componen cada uno de los segmentos transmembranales (TM) poseen puntajes de ~ 0.7 .

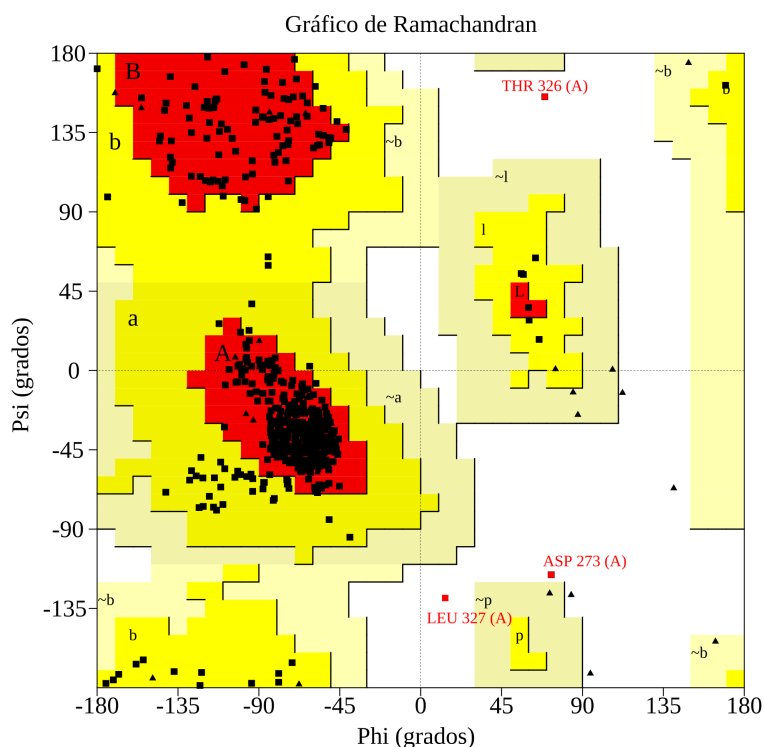


Figura 5.3: Se aprecia que solo tres residuos se encuentran en regiones no permitidas según su estereoquímica, estos son THR326, ASP273 y LEU327 los cuales se encuentran ubicados en coils extracelular alejados de la zona de interés de la proteína y por ende no afectan su funcionalidad.

5.2. Modelado de la estructura de NRT1.1 en una conformación orientada al medio extracelular

5.2.1. Identificación de las unidades repetidas

La superfamilia facilitadora principal (MFS) de transportadores constituye la familia más representativa dentro de los transportadores activos secundarios y es la responsable del transporte acoplado a iones de una amplia gama de solutos, desde monoaminas hasta azúcares y péptidos. El estudio cristalográfico del transportador de nitrato NRT1.1 (código de acceso 4OH3), reveló un pliegue perteneciente a la familia MFS de transportadores, que se caracteriza por tener 12 segmentos transmembrana, que se organizan en dos dominios N- y C-terminal. Ambos dominios, se encuentran relacionados entre sí por un eje de pseudosimetría perpendicular a la membrana. Los segmentos TM 1–3 y TM 4–6 en el do-

minio N-terminal, y los segmentos TM 7–9 y TM 10–12 en el dominio C-terminal, forman en total cuatro repeticiones compuestas de tres segmentos transmembrana consecutivos cada uno.

Existen varios métodos y herramientas computacionales para detectar y analizar repeticiones estructurales en proteínas. CE-Symm [47] utiliza un algoritmo basado en alineamientos estructurales. A través de esta herramienta fue posible corroborar lo expuesto en el estudio cristalográfico, revelando un eje simétrico entre los dos dominios N- y C-terminal que comprenden los residuos 1 al 312 y 313 al 590, respectivamente (Figura 5.4). Se cree que la simetría interna de las proteínas surge de las estructuras cuaternarias ancestrales fusionadas en una sola cadena de polipéptidos [67]. Sin embargo, dado que los pliegues simétricos de proteínas en teoría tienen una ventaja termodinámica, su simetría también podría haber surgido por la convergencia evolutiva [68].

La simetría entre ambos dominios en la estructura del transportador de nitrato NRT1.1 y el análisis de secuencias de la familia de transportadores MFS sugiere que la relación entre los dominios N- y C-terminal puede haberse originado a partir de un evento de duplicación génica [69]. Por otra parte, dado que CE-Symm se basa en una alineación de cuerpo rígido, esta herramienta no es capaz de detectar todas las repeticiones en estructuras multidominio [70]; por consiguiente, la estructura del transportador fue analizada tomando cada dominio de manera independiente. Los resultados de CE-Symm muestran que ambos dominios poseen dos unidades repetidas cada uno (Figura 5.4), las cuales se encuentran relacionadas por una doble pseudo-simetría. Esta característica otorga la capacidad al transportador de adoptar dos conformaciones principales durante el ciclo de transporte, un estado orientado hacia afuera de la membrana y otro hacia adentro. Específicamente, las repeticiones presentes en cada dominio adoptan una asimetría distintiva que al intercambiar sus conformaciones permite crear un nuevo estado asimétrico. De esta forma, es posible generar en los transportadores de topología invertida un mecanismo de transporte de acceso alternando.

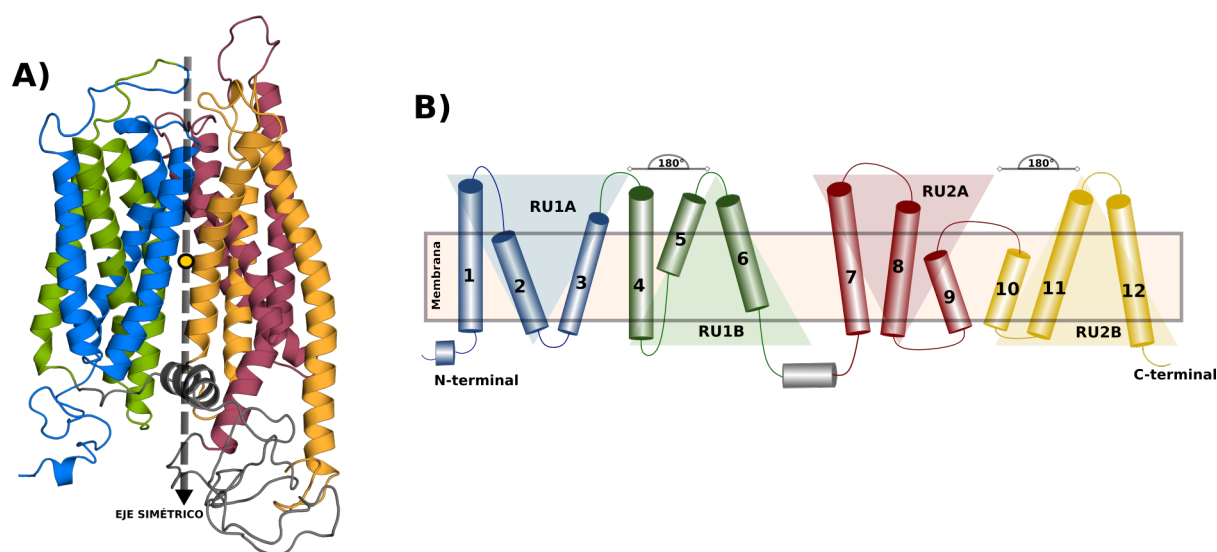


Figura 5.4: A) Identificación del eje de simetría en el transportador NRT1.1. Los segmentos transmembranales TM1-6 (hélices azules y verdes) componen la unidad repetida 1 (RU1), mientras que la unidad repetida 2 (RU2) se compone de los segmentos transmembrana TM7-12 (hélices rojas y naranja). La línea punteada gris denota la disposición del eje simétrico entre los dominios N- y C-terminal. B) Representación esquemática de la topología del transportador. Las hélices en los fondos triangulares azul y verde comprenden la unidad de repetición 1A (RU1A) y 1B (RU1B), respectivamente; mientras que la unidad de repetición 2A (RU2A) y 2B (RU2B) se componen de las hélices en los fondos triangulares rojo y amarillo. RU1A y RU2A están relacionados con RU1B y RU2B por pseudo-simetría doble (rotación de 180°).

5.2.2. Alineamiento estructural y de secuencia de las unidades repetidas

Una vez identificadas las unidades repetidas se realizó una superposición estructural de los segmentos mediante el programa TM-Align [48]. Para dos estructuras proteicas de equivalencia desconocida, TM-Align genera primero una alineación optimizada de residuo a residuo basada en la similitud estructural, utilizando iteraciones de programación dinámica heurística [48]. Una vez realizada la alineación, el software entrega una superposición óptima de las dos estructuras construidas sobre la alineación detectada, así como el valor de TM-score que escala la similitud estructural. TM-score tiene valores entre 0 y 1, donde 1 indica una coincidencia perfecta entre dos estructuras. Los puntajes por debajo de 0.2 corresponden a proteínas no relacionadas elegidas al azar, mientras que valores

mayores a 0.5 asumen generalmente el mismo pliegue en la base de datos SCOP [71] y CATH [72]. La superposición estructural de RU1A (residuos 1-131) con RU1B (residuos 132-242) y de RU2A (residuos 329-450) con RU2B (residuos 451-590) dieron como resultado un TM-score de 0.6 y 0.5, respectivamente; ambas por sobre el umbral de la función puntuación de similitud estructural de TM-align. Además, tras la superposición de las unidades repetidas, se aprecia claramente que comparten una arquitectura similar, con un valor de desviación cuadrática media (RMSD) para los átomos de C α de 3.08 Å para RU1 y de 3.56 Å para RU2 (Figura 5.5).

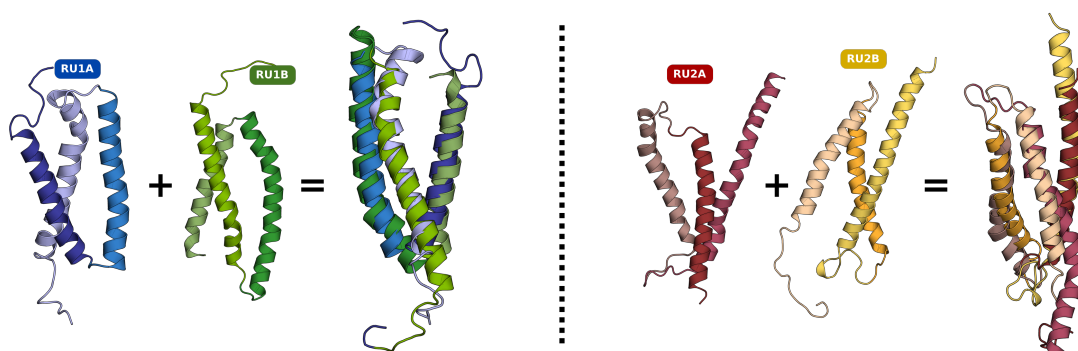


Figura 5.5: Representación en cartoon de la superposición de la unidades repetidas realizada con TM-Align. La coloración de las hélices corresponde a la topología descrita en la Figura 5.4.

Las alineaciones estructurales suelen ser confiables, especialmente cuando se comparan estructuras relacionadas de forma distante. Esto se debe a que las estructuras de las proteínas se encuentran más conservadas evolutivamente que su secuencia aminoacídica [73], y por ende las simetrías y pseudosimetrías que presentan a menudo las proteínas de membrana, tanto en niveles intramoleculares como cuaternarios, pueden ser usadas para predecir sitios activos, cambios conformacionales y mecanismos de transporte [74].

A partir de las superposiciones estructurales realizadas con TM-Align, se obtuvieron también los alineamientos de secuencias entre RU1 y RU2, mostrando una identidad menor al 9% (Figura 5.6). Esta alineación preliminar se ajustó posteriormente para eliminar gaps presentes en regiones transmembranales y agregar los residuos que no forman parte de las unidades repetidas, los cuales comprenden los aminoácidos desde la posición 243 al 328.

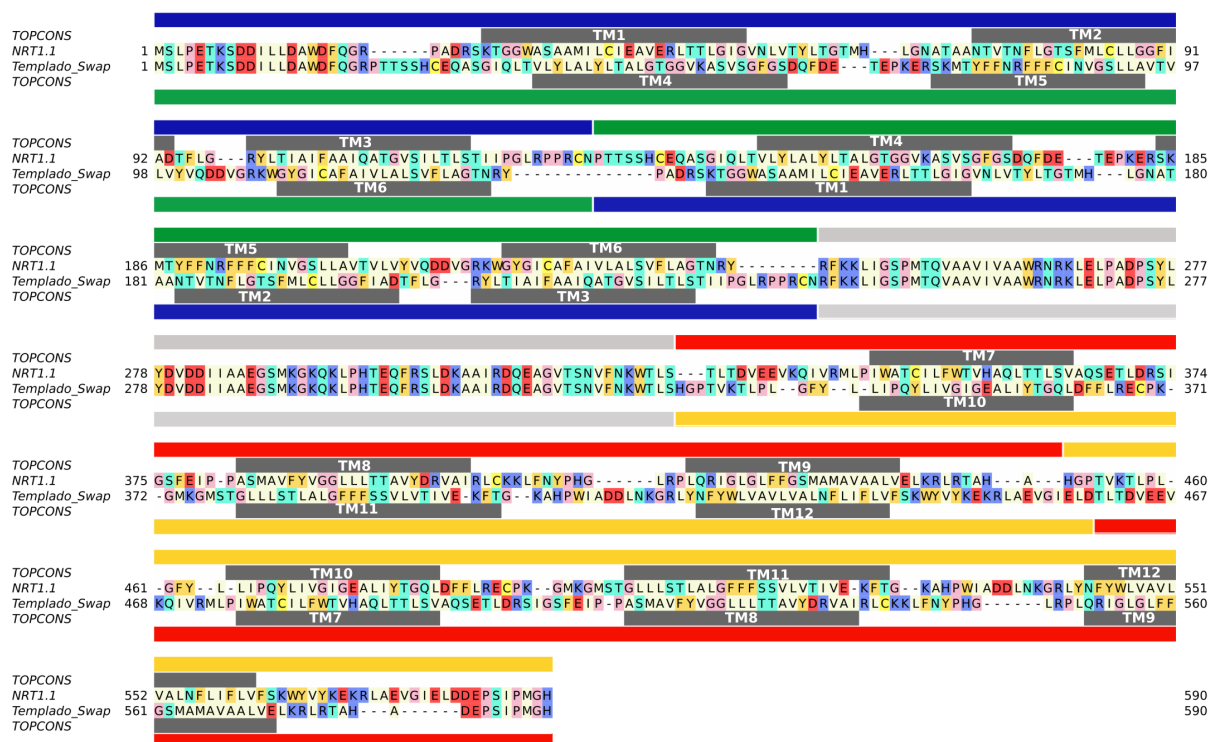


Figura 5.6: Se muestra el alineamiento de secuencia obtenido a partir de la superposición estructural de los segmentos repetidos. Los segmentos transmembrana corresponde a las regiones grises, mientras que las unidades repetidas se muestran coloreadas según la topología descrita en la Figura 5.4.

5.2.3. Preparación del archivo PDB invertido

Para generar la plantilla “intercambiada” de la estructura cristalográfica de un transportador, es necesario invertir el orden de las repeticiones en el archivo PDB. Posteriormente, usando esta estructura como templatado se modela la primera repetición RU1A en base a los residuos de la segunda repetición RU1B, seguido de los fragmentos de la proteína que no son parte de las repeticiones estructurales (243-328) y que por lo tanto mantienen sus coordenadas cristalográficas durante el proceso de modelado, para finalmente modelar RU2A en base a RU2B (Figura 5.7).



Figura 5.7: Se muestra el orden de las secuencias repetidas en el archivo PDB invertido, el cual será utilizado como templatado para construir un modelo de “intercambio”.

5.2.4. Construcción del modelo orientado hacia el medio extracelular

El modelo del transportador NRT1.1 orientado hacia el medio extracelular se generó a través del programa MODELLER v9.22[49]. Debido a la baja identidad de secuencia y a la divergencia estructural de las repeticiones fue necesario realizar una serie de ciclos de refinamiento para eliminar gaps dentro de los elementos estructurales repetidos, donde se asigna la estructura secundaria de acuerdo a los resultados obtenidos con DSSP [51]. Además, el modelo se ajustó a los patrones de conservación obtenidos con el servidor web ConSurf [52], con el fin de que los residuos más conservados se empaqueten dentro de la proteína, mientras que las posiciones variantes de secuencia se orienten hacia el exterior, imitando la disposición de los residuos en la estructura de rayos X (Figura 5.1). Por otra parte, se agregaron restricciones de distancia entre los átomos de carbono alfa tomadas desde la estructura cristalográfica. Estas restricciones permitieron conservar la disposición helicoidal nativa y empaquetamiento intramolecular dentro de los dominios, reduciendo así, las desviaciones que surgen debido a la divergencia de secuencia de las repeticiones. Las alineaciones de secuencia refinadas finales se usaron para construir 2.000 modelos de intercambio repetido del transportador NRT1.1, con el fin de aumentar las posibilidades de que el muestreo genere modelos que satisfagan todas las restricciones de entrada (Figura 5.8). El mejor modelo fue seleccionado en base a las puntuaciones de Modeller, ProQM [44] y Procheck [45] (Figura 5.9 5.10). El puntaje MolPDF describe qué tan bien el modelo satisface las restricciones de entrada, incluidas las creadas a partir de la plantilla, la alineación y cualquier restricción adicional aplicada.

El modelo final de intercambio repetido mostró $\sim 99\%$ de los residuos en las regiones favorecidas en el gráfico de Ramachandran, mientras que el puntaje ProQM fue 0.579.

Los análisis realizados con ProQM muestran que los residuos que componen cada uno de los segmentos transmembranales (TM) poseen puntajes de ~ 0.7 . Específicamente, las regiones de menor calidad corresponden a loops extracelulares que no influyen en el sitio de unión a sustrato. En el modelo de “intercambio” los residuos implicados en la unión a sustrato y el acoplamiento de protones LYS164, HIS356 y GLU476 poseen puntajes de 0.829, 0.834 y 0.883, respectivamente.

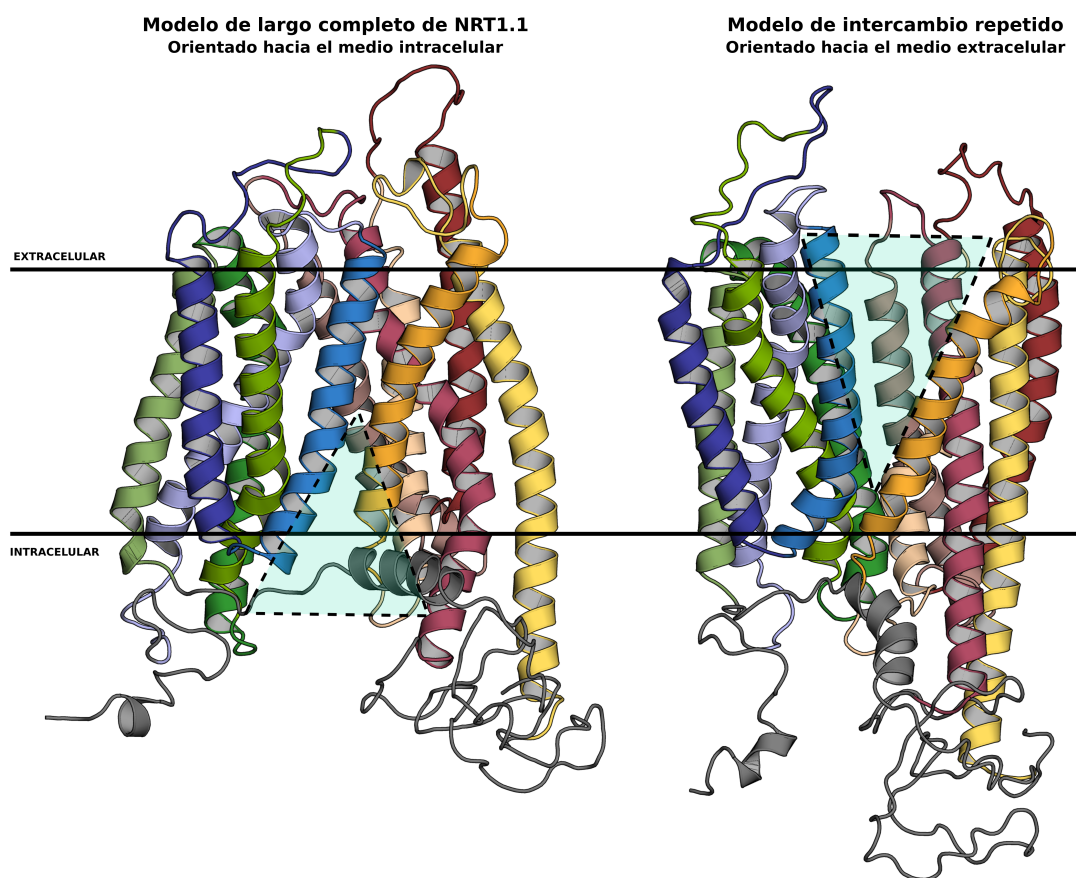


Figura 5.8: Modelo de largo completo refinado del transportador NRT1.1 en una conformación orientada al medio intracelular (figura izquierda) y modelo generado a través de repeticiones invertidas en una conformación orientada al medio externo (figura derecha). Las estructuras se muestran como hélices en cartoon, coloreadas en base a la topología descrita en la Figura 5.4. La posición de las proteínas en la membrana se determinó con el servidor OPM. La apertura de la ruta de unión a sustratos se muestra en triángulos transparentes.

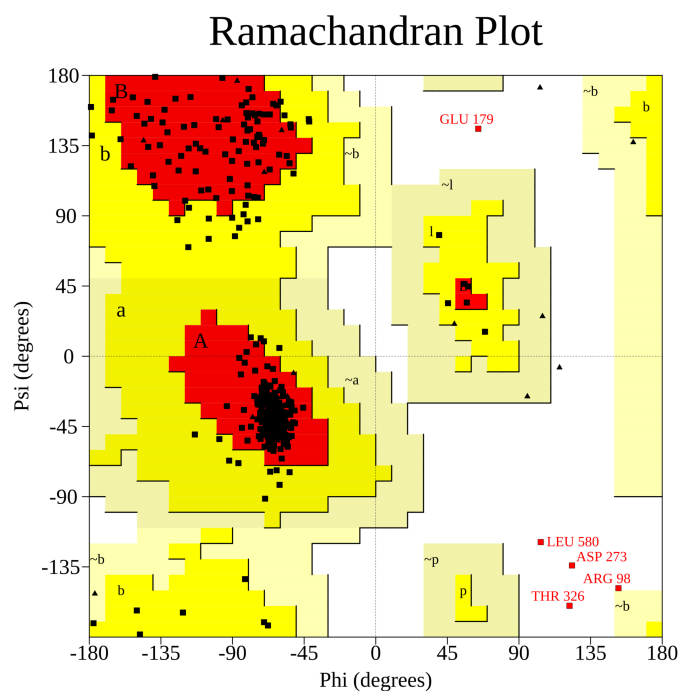


Figura 5.9: En la figura se aprecia que cinco residuos se encuentran en regiones no permitidas según su esteoquímica, estos son ARG98, GLU179, THR326, ASP273 y LEU580, los cuales se encuentran ubicados en coils extracelular alejados del sitio de unión a sustrato.

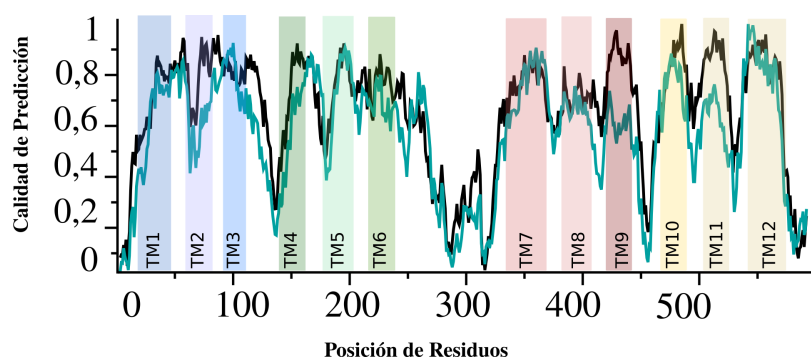


Figura 5.10: Se aprecian los puntajes de ProQM local del modelo de largo completo obtenido con Rosetta (negro) y del modelo en una conformación orientada hacia el medio extracelular obtenido mediante la técnica de modelado por homología usando repeticiones de intercambio invertidas (calipso).

Existen dos factores que determinan la precisión de un modelo de “intercambio”; primero, el grado en que las repeticiones son estructuralmente similares, lo que a su vez está

relacionado con su similitud de secuencia; y segundo, el grado de simetría que relaciona los dos estados. La similitud entre las estructuras de las repeticiones está relacionada con la precisión esperada de un modelo de homología típico, que disminuye en función de la identidad de secuencia decreciente [29]. El segundo factor que determina la precisión del modelo de intercambio repetido es el estado conformacional particular del templado. Por lo tanto, si la estructura de entrada está semiocclusa, entonces el modelo de salida estará en el mismo estado. De esta forma, el objetivo principal en el modelo de intercambio repetido es identificar los elementos involucrados en el cambio conformacional y la direccionalidad de su movimiento.

El modelo de “intercambio” de NRT1.1 se comparó con la estructura obtenida mediante Rosetta que se encuentra en una conformación orientada hacia el medio intracelular. Esta comparación reveló un cambio conformacional que consiste en una rotación de $\sim 35^\circ$ en el dominio de transporte alrededor de un eje perpendicular al eje de simetría y una translocación de $\sim 10 \text{ \AA}$, dando como resultado que el sitio de unión a sustrato quede expuesto al lado opuesto de la membrana (Figura 5.11A).

Estudios estructurales previos de varios miembros de la familia MFS han establecido un mecanismo *rocking-bundle* de transporte de sustrato, en el que los segmentos transmembrana de los transportadores se desplazan para generar conformaciones orientadas hacia afuera, ocluidas y hacia adentro [75]. El movimiento en los dominios se traduce a una variación de entre 4 a 10 \AA para los átomos de $C\alpha$, lo cual es consistente con la translocación de hélices observada en NRT1.1.

Análisis de conservación de residuos realizados con el server ConSurf [52] (Figura 5.11B) muestran los residuos conservados agrupados mayormente en el interior de la proteína en ambos estados conformacionales, mientras que los residuos variantes se localizan en las regiones periféricas que hacen contacto con la membrana lipídica.

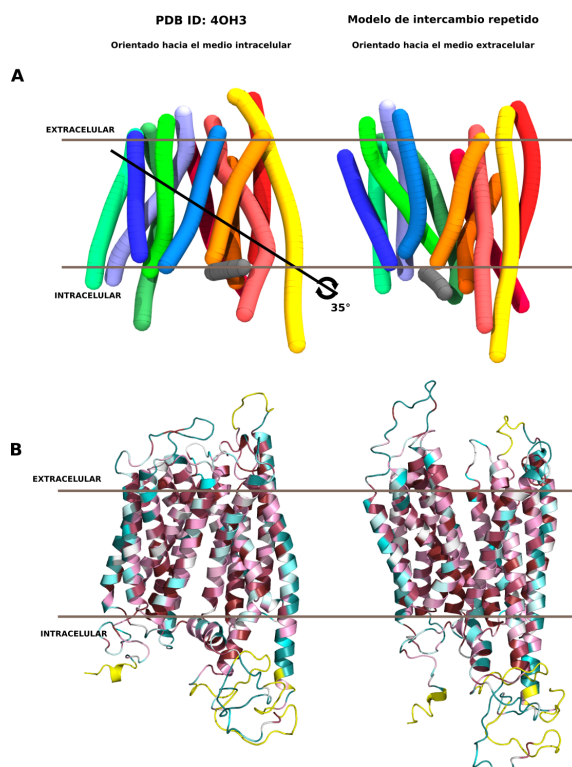


Figura 5.11: (A) Las estructuras en ambos estados fueron comparadas para cuantificar el cambio conformacional. La línea en negro representa el eje de rotación, que se encuentra aproximadamente perpendicular al eje de simetría. El complemento Bendix de VMD se utilizó para representar las hélices. (B) Comparación de los patrones de conservación de ConSurf para la estructura de NRT1.1 en ambos estados conformacionales. En cian se muestran los residuos variantes, en rosa oscuro los residuos conservados y en amarillo se indican posiciones con información de secuencia insuficiente para permitir una asignación de conservación.

5.3. Modelado por homología de isoformas de GTR en distintos estados conformacionales

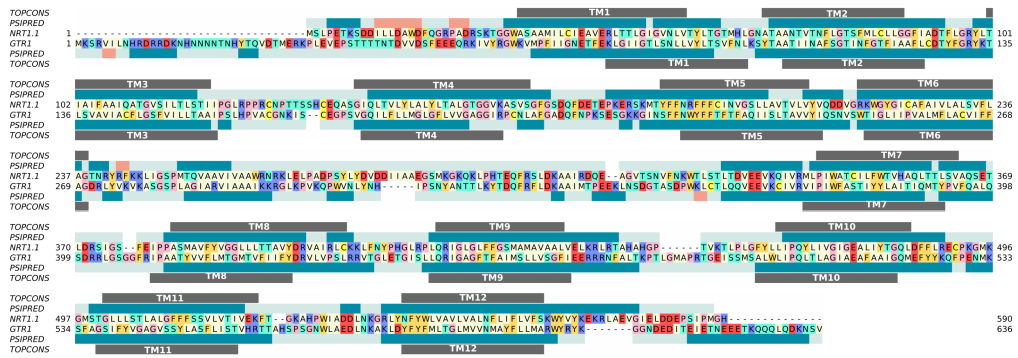
El modelo de largo completo refinado de NRT1.1 en una conformación orientada al medio intracelular y el modelo generado a través de repeticiones invertidas en una conformación orientada al medio externo, se utilizaron como plantilla para construir 3.000 modelos de cada isoforma de GTR en ambos estados conformacionales utilizando el programa Modeller v9.22. Los alineamientos de secuencias usados para el proceso de modelado se generaron empleando el servidor AlignMe [53]. Dichos alineamientos fueron

posteriormente refinados en base a predicciones de estructura secundaria con PsiPred [54] y predicciones de segmentos transmembrana con Topcons [55] (Figura 5.12). El modelo final de cada isoforma de GTR se seleccionó en base al puntaje MolPDF y las puntuaciones de Procheck y ProQM.

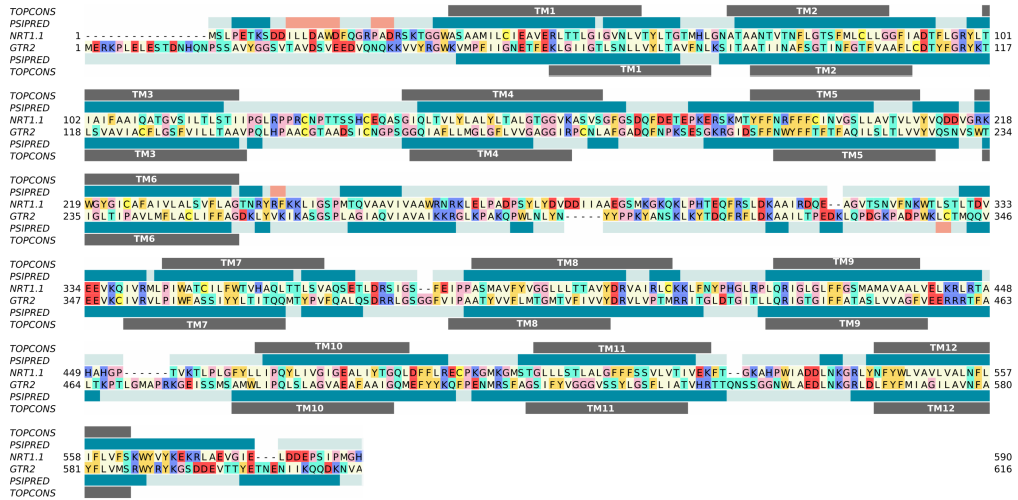
GTR1 (Uniprot Código:Q944G5) y GTR2 (Uniprot Código: Q9LV10) codifican dos proteínas transportadoras específicas de glucosinolato dependiente de protones de alta afinidad cuyas secuencias se componen de 636 y 616 aminoácidos, respectivamente [72]. Estos transportadores presentan un porcentaje de identidad de 32.49 % (GTR1) y 33.82 % (GTR2) respecto al transportador de nitrato NRT1.1 (Uniprot Código:Q05085) utilizado como templado para modelar ambas estructuras. Por otra parte, GTR3 (Uniprot Código: Q9M9V7) ha sido caracterizado principalmente como un transportador de nitrato de baja afinidad que facilita la carga de nitrato en el floema de la raíz, pero que tiene una actividad de transporte de glucosinolatos débil [73]. Este transportador se compone de 587 aminoácidos y presenta una identidad de 34.60 % con NRT1.1.

Dos factores importantes influyen en la capacidad de predecir modelos óptimos de proteínas: el grado de conservación estructural entre la proteína *target* y el templado, y la exactitud del alineamiento usado durante el proceso de modelado [74]. Los modelos basados en plantillas con más del 50 % de identidad de secuencia son generalmente muy precisos y pueden exhibir un RMSD de 1 Å respecto a la estructura experimental. Las proteínas con 30-50 % de identidad de secuencia, como es el caso de las isoformas de los GTRs, comparten al menos el 80 % de su estructura con el templado, generando modelos con un RMSD que no debería exceder los 4 Å respecto a una estructura experimental. Los errores en estos modelos se ubican principalmente en las regiones de bucles. Cabe destacar, que las alineaciones de secuencias suelen ser casi óptimas para targets con más del 30 % de identidad; por debajo de este umbral, la calidad de la alineación disminuye bruscamente e incluso la mitad de todos los residuos pueden estar desalineados cuando la identidad de secuencia es inferior al 20 % [75].

A) GTR1



B) GTR2



C) GTR3

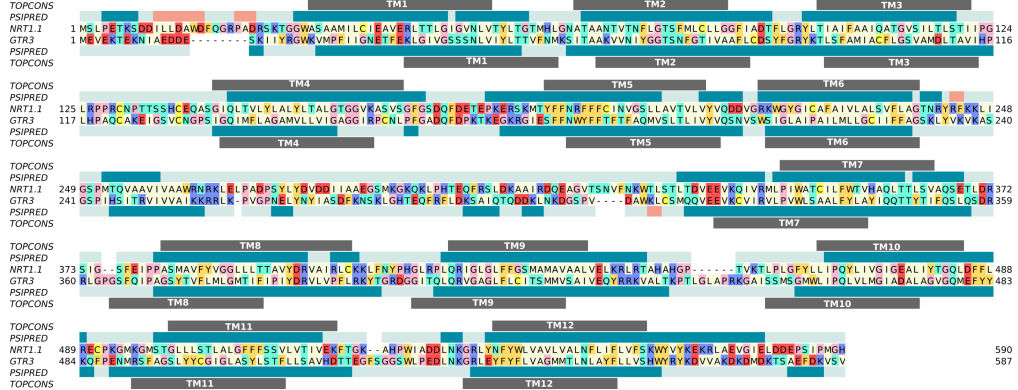


Figura 5.12: Las secuencias corresponden a (A) NRT1.1/GTR1, (B) NRT1.1/GTR2, y (C) NRT1.1/GTR3. Los alineamientos son coloreados en base a las propiedades químicas de los residuos. La asignación de estructura secundaria fue obtenida con PSIPRED y se muestra en rectángulos azul petroleo, mientras que en gris se muestran los segmentos transmembrana predichos con TOPCONS.

5.4. Análisis estructurales de los distintos modelos

Como se mencionó anteriormente las estructuras conocidas de transportadores MFS presentan dos dominios terminales con un eje de doble simetría perpendicular a la membrana lipídica. El cambio conformacional generado por las repeticiones de topología invertida presente en los dominios N- y C-terminal sugiere un mecanismo de transporte de tipo *rocking-bundle*, el cual se caracteriza por incluir dos dominios de las proteínas de tamaño comparable que se mueven uno respecto al otro. Ambos dominios, durante el proceso de transporte, también cambian su orientación con respecto a la membrana a medida que la proteína cambia de un estado orientado hacia el medio extracelular a una conformación expuesta al interior de la membrana. Para proporcionar acceso a estos sitios desde cualquier lado de la membrana, los dominios de la proteína se reorientan alrededor de los sitios de unión al sustrato, como un punto pivote de este movimiento, entre los dos dominios [61]. Las estructuras conocidas de los transportadores MFS describen un túnel de transporte situado entre los dominios terminales.

Para dilucidar el dominio de transporte asociado a glucosinolatos en las isoformas de GTRs se realizó un análisis con el programa HOLE v2.2[58], el cual permite conocer las dimensiones de una cavidad que atraviesa un modelo estructural. Este método utiliza un algoritmo basado en Monte Carlo para encontrar la mejor ruta en la que una esfera de radio variable, que simula el tamaño de un sustrato, pase a través del transportador. Los resultados muestran claramente la apertura del transportador en ambas conformaciones, con el túnel de transporte situado exactamente entre los dominios N- y C-terminal. Dicho dominio de transporte está delimitado por los segmentos transmembrana TM1, TM2, TM4, TM5, TM7, TM8, TM10 Y TM11 como se muestra en la Figura 5.13. Distintos estudios han descrito que los cambios conformacionales globales en el mecanismo de transporte *rocking-bundle* pueden ser causados por la formación y ruptura de puentes salinos intra e interdominios [76]. Esto implica que los aminoácidos cargados presentes en los segmentos transmembranales que limitan el túnel de transporte pudieran estar involucrados en el acoplamiento de iones y por consiguiente en el transporte de glucosinolatos.

A través de un análisis con HOLE en las estructuras de las tres isoformas de GTR incluidas en este estudio (Figura 5.13C), se observó que los transportadores en su conformación orientada hacia el medio extracelular disminuyen su radio de superficie a medida que el eje z se aproxima al medio intracelular, mientras que las conformaciones orientadas al medio intracelular muestran que el radio del túnel de transporte disminuye a medida que el eje z se aproxima al medio extracelular.

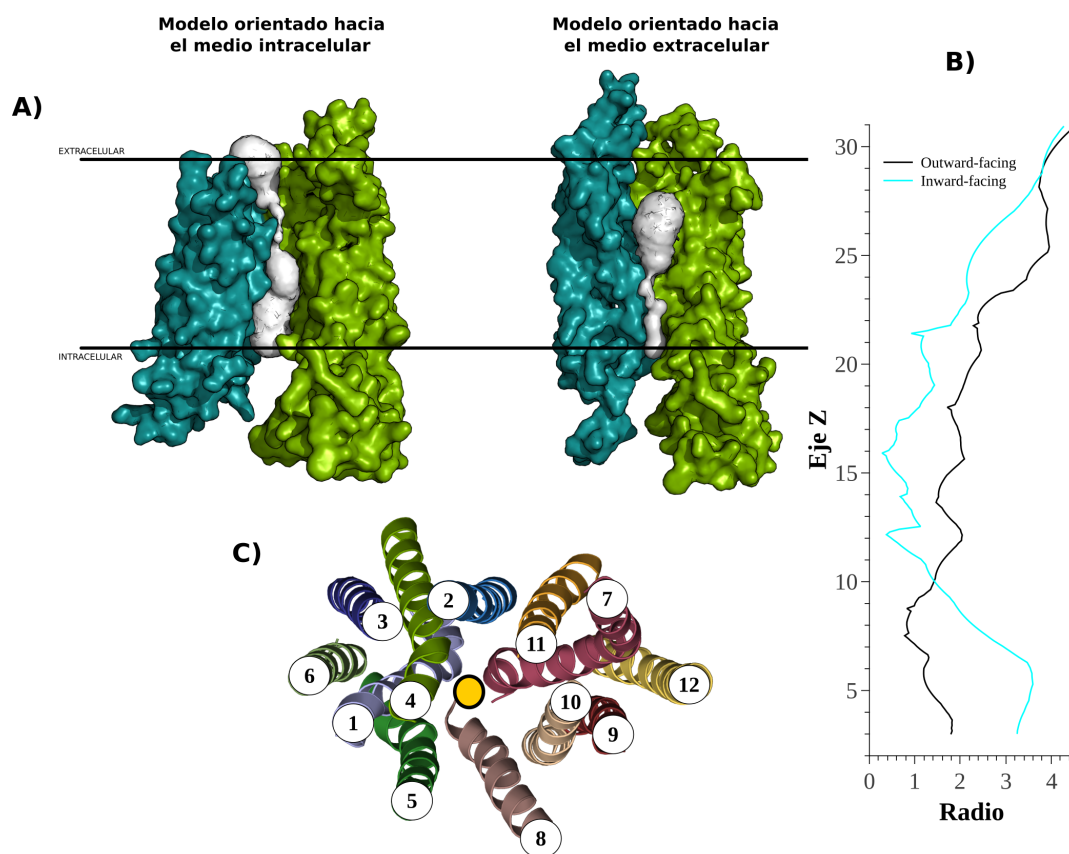


Figura 5.13: A) La imagen muestra el túnel de transporte de GTR1 obtenido mediante la técnica de HOLE (blanco) en las conformaciones intra y extracelular del transportador. El túnel de transporte está situado entre los dominios N- y C-terminal (calipso y verde, respectivamente). B) Gráfico obtenido con HOLE mostrando la variación del radio en la superficie del túnel de transporte de cada conformación del transportador respecto a sus coordenadas en el eje Z. C) Esquema de los 12 segmentos transmembrana del transportador GTR1 que se orientan formando el túnel de transporte (círculo amarillo) accesible desde ambos lados de la membrana dependiendo de la conformación del transportador.

Estudios estructurales del transportador de nitrato NRT1.1 y homólogos pertenecientes a la familia de transportadores de oligopéptidos dependientes de protones (POT) muestran un total de 13 residuos implicados en el transporte de sustrato y acoplamiento de protones [39]. Dichos aminoácidos fueron identificados en el alineamiento de secuencia utilizado para generar los modelos de las tres isoformas de GTRs, con el fin de encontrar una relación entre la secuencia templado y los modelos generados (Tabla 1). El análisis mostró que GTR1, GTR2 y GTR3 contienen residuos homólogos a aquellos descritos

para NRT1.1, los cuales se orientan a la cavidad central de los transportadores (Figura 5.14). Como consecuencia, se asume que dichos aminoácidos están involucrados en el acoplamiento de protones y el transporte de sustratos.

Posición	NRT1.1	GTR1	GTR2	GTR3
P1	E41	E75	E57	E33
P2	E44	E78	E60	E36
P3	R45	K79	K69	K37
P4	T48	I82	I64	I40
P5	L49	I83	I65	V41
P6	G52	I88	L68	S44
P7	F82	N116	N88	N74
P8	K164	R196	R180	R156
P9	N198	T230	T214	T190
P10	H356	I385	I369	Y343
P11	A357	T386	T370	I344
P12	Y388	L419	L403	L377
P13	E476	E513	E497	D471

Cuadro 5.1: Residuos claves en el acoplamiento de protones y transporte de sustrato en transportadores pertenecientes a la familia NPF y POT

. Residuos homólogos a los descritos para estas familias se identificaron en las tres isoformas de transportadores GTRs.

En NRT1.1 el acoplamiento de protones se encuentra mediado por el motivo conservado ExxER y los aminoácidos LYS164 y GLU476. El motivo ExxER mostró una sustitución de arginina por lisina en las tres isoformas de los GTRs, mientras que el residuo LYS164 fue sustituido por un residuo de arginina. Por otro lado, el residuo GLU476 se conserva tanto en GTR1 como en GTR2, en cambio en GTR3 presenta un ácido aspártico en la misma posición. Esto sugiere que el mecanismo de acoplamiento de protones se conserva en la familia de transportadores NPF, independientemente de la especificidad del sustrato [77]. Por otra parte, la HIS356 en NRT1.1 que estabiliza el sitio de unión a nitrato a través de una interacción carga-carga no muestra una conservación en los GTRs.

En cuanto a los glucosinolatos, estos están constituidos de tres partes: un fragmento de β -tioglucosa, un fragmento de oxima sulfonada y una cadena lateral de aglicona variable derivada de un α -aminoácido [78]. Los aminoácidos F82 y N198, implicados en la unión de sustratos en NRT1.1, corresponde a asparaginas y treoninas en los transportadores

GTRs y se conservan en las tres isoformas. Tales residuos han sido involucrados en la interacción con el fragmento de glucosa en los gluosinatos.

Existen tres tipos de glucosinatos capaces de ser transportados por GTRs: glucosinatos de origen alifático derivados de metionina, glucosinatos indol o indólico de fuente de triptófano y glucosinatos aromáticos formados a partir de fenilalanina [78]. Estudios previos han identificado a GTR1 y GTR2 como transportadores de glucosinatos alifáticos, mientras que GTR3 ha sido descrito como un transportador de glucosinatos de tipo indol [72]. A pesar de los avances recientes en la comprensión de la especificidad de sustratos por parte de estos transportadores, actualmente no se tiene conocimiento sobre los aminoácidos que determinan la especificidad desde una perspectiva evolutiva. Sin embargo, las mutaciones de aminoácidos observadas en las posiciones P4, P5, P6 de la Tabla 1 son consistente con los cambios en la especificidad del sustrato en GTRs y podrían contribuir a determinar la selectividad de los transportadores por los distintos tipos de glucosinatos.

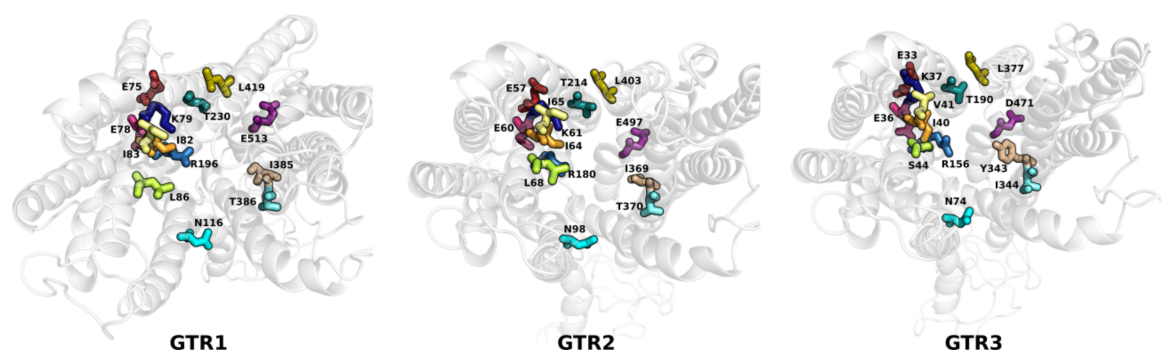


Figura 5.14: La figura muestra la disposición de los residuos descritos en la Tabla 1 en el túnel de transporte.

Para determinar si la orientación de los residuos identificados para participar en la unión de sustratos cambia respecto a los distintos estados conformacionales de los transportadores, se realizó un análisis estructural de las tres isoformas de GTR en su conformación orientada al medio intra y extracelular. A partir de estos estudios se observó una rotación en el túnel de transporte de $\sim 36^\circ$, 28° y 32° alrededor de un eje perpendicular al eje de simetría para GTR1, GTR2 y GTR3, respectivamente (Figura 5.15). Además, en los tres transportadores se estimó una translocación menor a 3\AA , dando como resultado que el sitio de unión a sustrato quede expuesto al lado opuesto de la membrana. Cabe destacar, que los valores de ángulo de rotación, desplazamiento y translocación se relacionan con los obtenidos para el transportador de nitrato NRT1 descrito anteriormente. Esto

implica que los dominios de transporte de NRT1.1 son homólogos a los que conforman las isoformas de los transportadores GTRs.

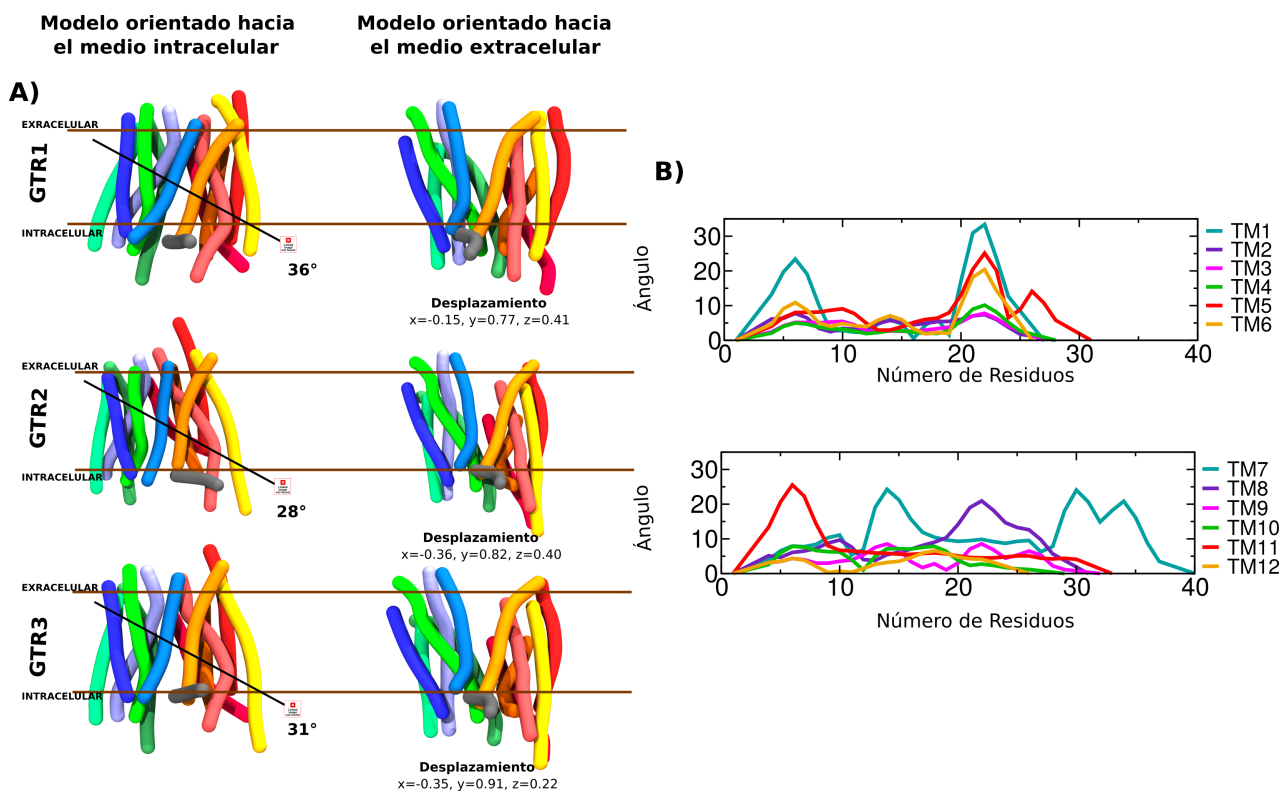


Figura 5.15: A) Las estructuras de GTRs en ambos estados conformacionales fueron comparadas para cuantificar el cambio estructural. Se evaluaron los ejes de rotación, ángulo de rotación y el desplazamiento de las hélices. Las líneas en negro representan el eje de rotación, que se encuentra aproximadamente perpendicular al eje de simetría. B) Gráfico obtenido con Bendix representando la curvatura versus número de residuos en los transportadores. Las hélices TM1, TM5, TM6, TM7, TM8 y TM11 son las que muestran una mayor curvatura en todas las isoformas de los GTRs.

5.5. Búsqueda de motivos de secuencia

Para identificar motivos conservados que pudieran estar implicados en el acoplamiento de protones y/o el transporte de glucosinolatos en las tres isoformas de GTRs, se realizó un análisis comparativo con ConSurf. Inicialmente se encontraron más de 10.000 homólogos utilizando HMMER, los cuales fueron procesados mediante USEARCHv11 seleccionando las secuencias cuyo porcentaje de identidad fuera mayor al 60 % y de un largo superior a 450 aminoácidos. Los homólogos que cumplieran esta restricción fueron usados para generar un alineamiento múltiple utilizando el server MAFFT. Posteriormente se generó un weblogo con la finalidad de ver qué aminoácidos se conservan en la familia de GTRs (Figura 5.16).

Los resultados obtenidos mediante el logotipo de secuencia revelan que existe una conservación predominante del motivo ExxE[R/K] en la familia NPF de transportadores. Esto implicaría que el acoplamiento de protones se conserva independiente del sustrato a transportar. Los residuos en la posición P8 y P13, descritos en la Tabla 1, muestran una alta conservación y han sido descritos como residuos claves al igual que el motivo ExxE[R/K] en el acoplamiento de protones. Por otra parte, los aminoácidos en las posiciones P4, P5 y P6 que podrían ser claves en el transporte de glucosinolatos muestran una conservación relativa; mientras que P4 y P5 se conservan predominantemente como residuos de isoleucina, P6 varía principalmente entre un residuo hidrófobo a un residuo de serina.

Por otra parte, los análisis con ConSurf (Figura 5.16) muestran que los residuos conservados se agrupan en el interior de la proteína entre los dominios N- y C-terminal, mientras que los residuos variantes se encuentran principalmente en las regiones periféricas que hacen contacto con la membrana lipídica. Específicamente, el segmento transmembrana TM1 muestra una alta conservación en todas las isoformas de los GTRs. Por el contrario, la región que comprende los aminoácidos en las posiciones P4, P5 y P6 (Tabla 1) muestra patrones de conservación variantes, lo que podría significar que dichos residuos están implicados en la especificidad de sustrato, y por ende interaccionan con la cadena lateral de los distintos tipos de glucosinolatos.

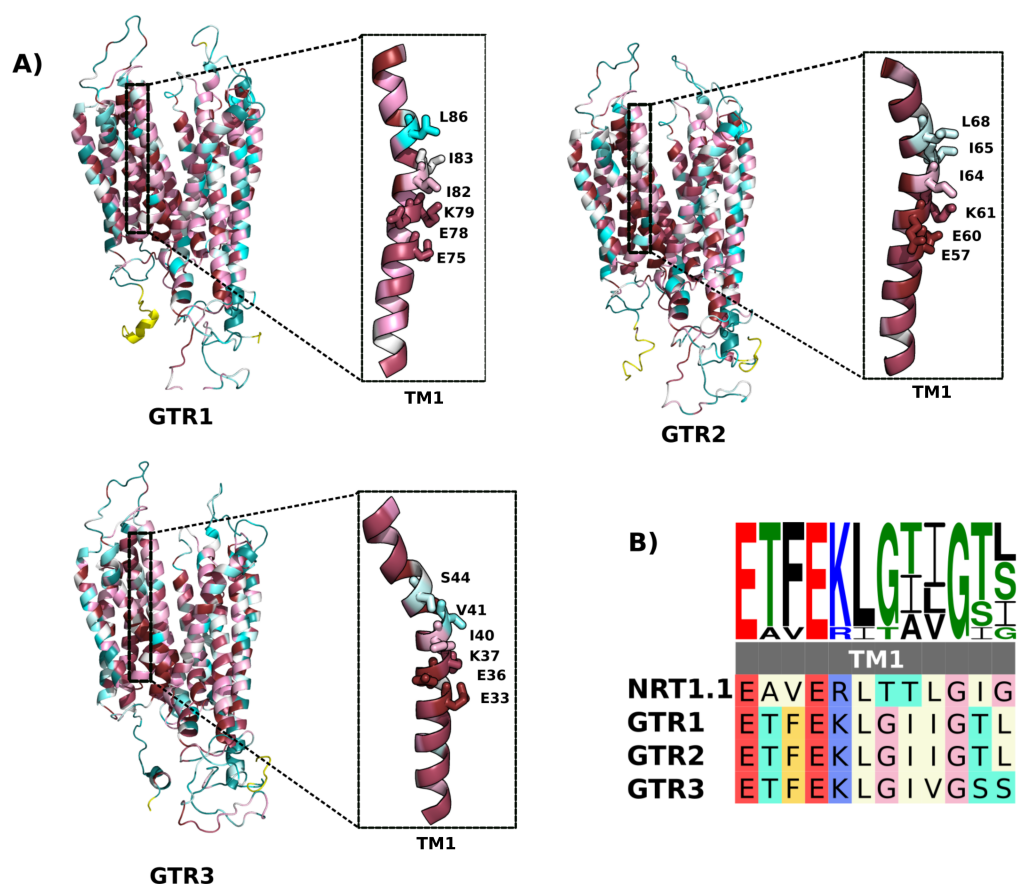


Figura 5.16: A) Patrones de conservación estimados con ConSurf para cada isoforma. Los residuos en TM1 se muestran coloreados en base a su conservación, utilizando colores desde calipso (la mayoría de los residuos variantes), hasta rosa oscuro (la mayoría de los residuos conservados). En amarillo se muestran las posiciones con información de secuencia insuficiente para permitir una asignación de conservación. (B) Logotipo de secuencia mostrando la conservación de los residuos implicados en el transporte y acoplamiento de protones. Se evaluó la conservación entre 400 miembros de la familia NPF.

6. Conclusiones

Diferentes técnicas computacionales permiten describir fenómenos biológicos a escala molecular. El modelado por homología usando repeticiones de intercambio (*Repeat-swap homology modeling*) ha demostrado ser útil para el análisis de los cambios conformacionales en diversos transportadores activos secundarios. Aunque los modelos generados pueden no ser de alta resolución para describir detalles a escala atómica, la identificación de conformaciones alternativas y mecanismos de intercambio conformacional es esencial para comprender características estructurales que pueden afectar la unión de sustratos en diferentes transportadores.

El refinamiento de la estructura cristalina de NRT1.1 orientado hacia el medio intracelular y el modelo de “intercambio” en una conformación expuesta hacia el citoplasma, generados a lo largo de esta tesis, proporcionan una descripción a nivel molecular de los cambios conformacionales en proteínas de la familia NRT1/PTR (NPF). En base a estas estructuras es posible determinar las características estructurales que sustentan la estequiometría y la unión de sustratos en transportadores de esta familia. En términos generales, el modelo de “intercambio” sugiere un mecanismo de transporte *rocking-bundle* para NRT1.1, en donde el sitio de unión a sustrato se expone en ambos lados de la membrana lipídica a través de un evento de rotación y traslación.

Los glucosinolatos son metabolitos secundarios en plantas característicos de las crucíferas y son responsables del sabor fuerte de plantas como el rábano picante, varias mostazas, berros y rábanos. Estos metabolitos, de los cuales se han identificado cerca de 200 tipos con diferentes sustituyentes, se sintetizan a partir de aminoácidos. Los glucosinolatos alifáticos se derivan de alanina, leucina, isoleucina, valina y metionina. Los glucosinolatos de indol y los glucosinolatos aromáticos se derivan de triptófano, fenilalanina o tirosina. Debido a la dificultad de aislar proteínas integrales de membrana, actualmente no se cuenta con una estructura cristalina de un transportador de glucosinolatos que permita comprender las diferencias en la selectividad de sustratos. Por lo tanto, en la presente tesis, se diseñaron modelos moleculares de tres isoformas de GTRs con el fin de lograr cierta comprensión de los sistemas de transporte asociado a la translocación de glucosinolatos y protones en estas proteínas.

A través de diversos análisis en la estructura de los modelos de GTRs se logró identificar el desplazamiento de hélices y ángulo de rotación que caracterizan el mecanismo

de transporte de estas proteínas. Además, se detectaron trece residuos implicados en el transporte de glucosinolatos y el acoplamiento de protones. Algunos de estos residuos son parte de patrones conservados en las distintas proteínas o reflejan ciertas mutaciones que pudieran explicar la especificidad por sustratos. Específicamente, se encontró un dominio conservado en el acoplamiento de protones, ExxER, sumado a tres residuos implicados en el transporte específico de sustrato. En GTR1 los residuos fueron I82, I83 y L86, GTR2 I64, I65 y L68, mientras que en GTR3 fueron I40, V41 y S44.

Cabe destacar que los resultados obtenidos a partir de métodos computacionales deben ser respaldados con experimentos que permitan comprobar las predicciones teóricas. De esta manera, para futuras investigaciones se propone la mutación de los residuos identificados como claves para el transporte de sustratos en las tres isoformas de los GTRs, con el fin de corroborar su participación en el mecanismo de transporte de sustratos y especificidad de cada proteína.

7. Referencias

1. E. Wallin, G. von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms., *Protein Sci. Publ. Protein Soc.* 7 (1998) 1029–1038.
2. K.R. Vinothkumar, R. Henderson, Structures of membrane proteins, *Q. Rev. Biophys.* 43 (2010) 65–158.
3. J.P. Whitelegge, Integral Membrane Proteins and Bilayer Proteomics, *Anal. Chem.* 85 (2013) 2558–2568.
4. O. Jardetzky, Simple allosteric model for membrane pumps, *Nature.* 211 (1966) 969–970.
5. S.S. Mansy, Membrane transport in primitive cells, *Cold Spring Harb. Perspect. Biol.* 2 (2010) a002188.
6. P. Nacry, E. Bouguyon, A. Gojon, Nitrogen acquisition by roots: physiological and developmental mechanisms ensuring plant adaptation to a fluctuating resource, *Plant Soil.* 370 (2013) 1–29.
7. J.V. Vermaas, N. Trebesch, C.G. Mayne, S. Thangapandian, M. Shekhar, P. Mahint-hichaichan, J.L. Baylon, T. Jiang, Y. Wang, M.P. Muller, E. Shinn, Z. Zhao, P.-C. Wen, E. Tajkhorshid, Microscopic Characterization of Membrane Transporter Function by In Silico Modeling and Simulation, *Methods Enzymol.* 578 (2016) 373–428.
8. U.K. Kar, M. Simonian, J.P. Whitelegge, Integral membrane proteins: bottom-up, top-down and structural proteomics, *Expert Rev. Proteomics.* 14 (2017) 715–723.
9. X.C. Zhang, Y. Zhao, J. Heng, D. Jiang, Energy coupling mechanisms of MFS transporters, *Protein Sci. Publ. Protein Soc.* 24 (2015) 1560–1579.
10. Y. Shi, Common folds and transport mechanisms of secondary active transporters, *Annu. Rev. Biophys.* 42 (2013) 51–72.
11. L.R. Forrest, Y.-W. Zhang, M.T. Jacobs, J. Gesmonde, L. Xie, B.H. Honig, G. Rudnick, Mechanism for alternating access in neurotransmitter transporters, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 10338–10343.

12. A. Accardi, C. Miller, Secondary active transport mediated by a prokaryotic homologue of ClC Cl⁻ channels, *Nature*. 427 (2004) 803–807.
13. O. Boudker, G. Verdon, Structural perspectives on secondary active transporters, *Trends Pharmacol. Sci.* 31 (2010) 418–426.
14. T.J. Crisman, S. Qu, B.I. Kanner, L.R. Forrest, Inward-facing conformation of glutamate transporters as revealed by their inverted-topology structural repeats, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 20752–20757.
15. W.D. Stein, W.R. Lieb, S.J.D. Karlish, Y. Eilam, A Model for Active Transport of Sodium and Potassium Ions as Mediated by a Tetrameric Enzyme*, *Proc. Natl. Acad. Sci. U. S. A.* 70 (1973) 275–278.
16. L.R. Forrest, Y.-W. Zhang, M.T. Jacobs, J. Gesmonde, L. Xie, B.H. Honig, G. Rudnick, Mechanism for alternating access in neurotransmitter transporters, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 10338–10343.
17. E. Ficici, J.D. Faraldo-Gómez, M.L. Jennings, L.R. Forrest, Asymmetry of inverted-topology repeats in the AE1 anion exchanger suggests an elevator-like mechanism, *J. Gen. Physiol.* 149 (2017) 1149–1164.
18. E. Screpanti, C. Hunte, Discontinuous membrane helices in transport proteins and their correlation with function, *J. Struct. Biol.* 159 (2007) 261–267.
19. C. Mulligan, C. Fenollar-Ferrer, G.A. Fitzgerald, A. Vergara-Jaque, D. Kaufmann, Y. Li, L.R. Forrest, J.A. Mindell, The bacterial dicarboxylate transporter VcINDY uses a two-domain elevator-type mechanism, *Nat. Struct. Mol. Biol.* 23 (2016) 256–263.
20. D. Drew, O. Boudker, Shared Molecular Mechanisms of Membrane Transporters, *Annu. Rev. Biochem.* 85 (2016) 543–572.
21. R.K. Henderson, K. Fendler, B. Poolman, Coupling efficiency of secondary active transporters, *Curr. Opin. Biotechnol.* 58 (2019) 62–71.
22. J. Reizer, A. Reizer, M.H. Saier, The MIP family of integral membrane channel proteins: sequence comparisons, evolutionary relationships, reconstructed pathway of evolution, and proposed functional differentiation of the two repeated halves of the proteins, *Crit. Rev. Biochem. Mol. Biol.* 28 (1993) 235–257.
23. A.V. Kajava, Review: proteins with repeated sequence–structural prediction and modeling, *J. Struct. Biol.* 134 (2001) 132–144.

24. L.R. Forrest, Structural Symmetry in Membrane Proteins, *Annu. Rev. Biophys.* 44 (2015) 311–337.
25. L.R. Forrest, Structural Symmetry in Membrane Proteins., *Annu. Rev. Biophys.* 44 (2015) 311–337.L.R. Forrest, Structural Symmetry in Membrane Proteins., *Annu. Rev. Biophys.* 44 (2015) 311–337.
26. Y. Wang, Y. Huang, J. Wang, C. Cheng, W. Huang, P. Lu, Y.-N. Xu, P. Wang, N. Yan, Y. Shi, Structure of the formate transporter FocA reveals a pentameric aquaporin-like channel, *Nature.* 462 (2009) 467–472.
27. N. Reyes, C. Ginter, O. Boudker, Transport mechanism of a bacterial homologue of glutamate transporters, *Nature.* 462 (2009) 880–885.
28. A. Vergara-Jaque, C. Fenollar-Ferrer, D. Kaufmann, L.R. Forrest, Repeat-swap homology modeling of secondary active transporters: updated protocol and prediction of elevator-type mechanisms, *Front. Pharmacol.* 6 (2015) 183.
29. E. Ficici, J.D. Faraldo-Gómez, M.L. Jennings, L.R. Forrest, Asymmetry of inverted-topology repeats in the AE1 anion exchanger suggests an elevator-like mechanism, *J. Gen. Physiol.* 149 (2017) 1149–1164.
30. D. Latek, Rosetta Broker for membrane protein structure prediction: concentrative nucleoside transporter 3 and corticotropin-releasing factor receptor 1 test cases, *BMC Struct. Biol.* 17 (2017).
31. H.R. Zhekova, M. Damergi, S.Y. Noskov, J. Jiang, Z.H. Zhou, A. Pushkin, I. Kurtz, Modelling of Ion Binding and Selectivity in SLC4 Transporters, *Biophys. J.* 116 (2019) 555a.
32. C. Fenollar-Ferrer, L.R. Forrest, Structural models of the NaPi-II sodium-phosphate cotransporters, *Pflugers Arch - Eur J Physiol.* 471 (2019) 43–52.
33. V.I. Shattuck, W. Wang, Nitrogen dioxide fumigation alters the glucosinolate and nitrate levels in pak choy (*Brassica campestris* ssp. *chinensis*), *Sci. Hortic.* 56 (1993) 87–100.
34. F.-Q. Guo, J. Young, N.M. Crawford, The nitrate transporter AtNRT1.1 (CHL1) functions in stomatal opening and contributes to drought susceptibility in *Arabidopsis*, *Plant Cell.* 15 (2003) 107–117.

35. Y. Chiba, T. Shimizu, S. Miyakawa, Y. Kanno, T. Koshiba, Y. Kamiya, M. Seo, Identification of *Arabidopsis thaliana* NRT1/PTR FAMILY (NPF) proteins capable of transporting plant hormones, *J. Plant Res.* 128 (2015) 679–686.
36. J. Sun, J.R. Bankston, J. Payandeh, T.R. Hinds, W.N. Zagotta, N. Zheng, Crystal structure of the plant dual-affinity nitrate transporter NRT1.1, *Nature.* 507 (2014) 73–77.
37. M. Cavaiuolo, A. Ferrante, Nitrates and Glucosinolates as Strong Determinants of the Nutritional Quality in Rocket Leafy Salads, *Nutrients.* 6 (2014) 1519–1538.
38. T. Kiba, A. Krapp, Plant Nitrogen Acquisition Under Low Availability: Regulation of Uptake and Root Architecture, *Plant Cell Physiol.* 57 (2016) 707–714.
39. S.D. Khare, T.A. Whitehead, Introduction to the Rosetta Special Collection, *PLOS ONE.* 10 (2015) e0144326.
40. K.W. Kaufmann, G.H. Lemmon, S.L. DeLuca, J.H. Sheehan, J. Meiler, Practically Useful: what the Rosetta protein modeling suite can do for you, *Biochemistry.* 49 (2010) 2987–2998.
41. D.E. Kim, D. Chivian, D. Baker, Protein structure prediction and analysis using the Robetta server, *Nucleic Acids Res.* 32 (2004) W526–W531.
42. G. Bhardwaj, V.K. Mulligan, C.D. Bahl, J.M. Gilmore, P.J. Harvey, O. Cheneval, G.W. Buchko, S.V.S.R.K. Pulavarti, Q. Kaas, A. Eletsky, P.-S. Huang, W.A. Johnsen, P.J. Greisen, G.J. Rocklin, Y. Song, T.W. Linsky, A. Watkins, S.A. Rettie, X. Xu, L.P. Carter, R. Bonneau, J.M. Olson, E. Coutsiias, C.E. Correnti, T. Szyperski, D.J. Craik, D. Baker, Accurate de novo design of hyperstable constrained peptides, *Nature.* 538 (2016) 329–335.
43. H. Park, P. Bradley, P. Greisen, Y. Liu, V.K. Mulligan, D.E. Kim, D. Baker, F. DiMaio, Simultaneous Optimization of Biomolecular Energy Functions on Features from Small Molecules and Macromolecules, *J. Chem. Theory Comput.* 12 (2016) 6201–6212.
44. A. Ray, E. Lindahl, B. Wallner, Model quality assessment for membrane proteins, *Bioinformatics.* 26 (2010) 3067–3074.
45. R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* 26 (1993) 283–291.

46. C.-H. Tai, R. Paul, D. Kc, J.D. Shilling, B. Lee, SymD webserver: a platform for detecting internally symmetric protein structures, *Nucleic Acids Res.* 42 (2014) W296–W300.
47. D. Myers-Turnbull, S.E. Bliven, P.W. Rose, Z.K. Aziz, P. Youkharibache, P.E. Bourne, A. Prlić, Systematic Detection of Internal Symmetry in Proteins Using CE-Symm, *J. Mol. Biol.* 426 (2014) 2255–2268.
48. Y. Zhang, TM-align: a protein structure alignment algorithm based on the TM-score, *Nucleic Acids Res.* 33 (2005) 2302–2309.
49. N. Eswar, B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M. Shen, U. Pieper, A. Sali, Comparative Protein Structure Modeling Using Modeller, *Curr. Protoc. Bioinforma.* 15 (2006) 5.6.1–5.6.30.
50. A. Šali, T.L. Blundell, Comparative Protein Modelling by Satisfaction of Spatial Restraints, *J. Mol. Biol.* 234 (1993) 779–815.
51. W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers.* 22 (1983) 2577–2637.
52. H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules, *Nucleic Acids Res.* 44 (2016) W344–W350.
53. M. Stamm, R. Staritzbichler, K. Khafizov, L.R. Forrest, AlignMe—a membrane protein sequence alignment web server, *Nucleic Acids Research.* 42 (2014) W246–W251.
54. L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, *Bioinformatics.* 16 (2000) 404–405.
55. A. Bernsel, H. Viklund, A. Hennerdal, A. Elofsson, TOPCONS: consensus prediction of membrane protein topology, *Nucleic Acids Res.* 37 (2009) W465–W468.
56. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, *J. Mol. Graph.* 14 (1996) 33–38.
57. S. Yuan, H.C.S. Chan, Z. Hu, Using PyMOL as a platform for computational drug design: PyMOL: platform for computational drug design, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 7 (2017) e1298.

58. HOLE: A program for the analysis of the pore dimensions of ion channel structural models - ScienceDirect, (n.d.).
59. Dahl, A. C., Chavent, M., Sansom, M. S. (2012). Bendix: intuitive helix geometry analysis and abstraction. *Bioinformatics* (Oxford, England), 28(16), 2193–2194.
60. M.E. Jørgensen, C.E. Olsen, D. Geiger, O. Mirza, B.A. Halkier, H.H. Nour-Eldin, A Functional EXXEK Motif is Essential for Proton Coupling and Active Glucosinolate Transport by NPF2.11, *Plant Cell Physiol.* 56 (2015) 2340–2350.
61. R.D. Finn, J. Clements, S.R. Eddy, HMMER web server: interactive sequence similarity searching, *Nucleic Acids Res.* 39 (2011) W29–W37.
62. SEARCH16S: A new algorithm for identifying 16S ribosomal RNA genes in contigs and chromosomes bioRxiv, (n.d.).
63. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, (n.d.).
64. A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview Version 2—a multiple sequence alignment editor and analysis workbench, *Bioinformatics.* 25 (2009) 1189–1191.
65. G.E. Crooks, WebLogo: A Sequence Logo Generator, *Genome Research.* 14 (2004) 1188–1190.
66. Forrest, S. Tavoulari, Y.-W. Zhang, G. Rudnick, B. Honig, Identification of a chloride ion binding site in Na⁺/Cl⁻-dependent transporters, *Proceedings of the National Academy of Sciences.* 104 (2007) 12761–12766.
67. J. Lee, M. Blaber, Experimental support for the evolution of symmetric protein architecture from a simple peptide motif, *Proc Natl Acad Sci USA.* 108 (2011) 126–130.
68. A. Broom, A.C. Doxey, Y.D. Lobsanov, L.G. Berthin, D.R. Rose, P.L. Howell, B.J. McConkey, E.M. Meiering, Modular Evolution and the Origins of Symmetry: Reconstruction of a Three-Fold Symmetric Globular Protein, *Structure.* 20 (2012) 161–171.
69. L.R. Forrest, R. Krämer, C. Ziegler, The structural basis of secondary active transport mechanisms, *Biochimica et Biophysica Acta (BBA) - Bioenergetics.* 1807 (2011) 167–188.

70. A. Andreeva, D. Howorth, C. Chothia, E. Kulesha, A.G. Murzin, SCOP2 prototype: a new approach to protein structure mining, *Nucl. Acids Res.* 42 (2014) D310–D314.
71. N.L. Dawson, T.E. Lewis, S. Das, J.G. Lees, D. Lee, P. Ashford, C.A. Orengo, I. Sillitoe, CATH: an expanded resource to predict protein function through structure and sequence, *Nucleic Acids Res.* 45 (2017) D289–D295.
72. A. Dettai, G. di Prisco, G. Lecointre, E. Parisi, C. Verde, Inferring Evolution of Fish Proteins: The Globin Case Study, in: *Methods in Enzymology*, Elsevier, 2008: pp. 539–570.
73. N. Solcan, J. Kwok, P.W. Fowler, A.D. Cameron, D. Drew, S. Iwata, S. Newstead, Alternating access mechanism in the POT family of oligopeptide transporters: Alternating access mechanism in the POT family, *The EMBO Journal.* 31 (2012) 3411–3421.
74. H.H. Nour-Eldin, T.G. Andersen, M. Burow, S.R. Madsen, M.E. Jørgensen, C.E. Olsen, I. Dreyer, R. Hedrich, D. Geiger, B.A. Halkier, NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds, *Nature.* 488 (2012) 531–534.
75. Y.-Y. Wang, Y.-F. Tsay, Arabidopsis Nitrate Transporter NRT1.9 Is Important in Phloem Nitrate Transport, *Plant Cell.* 23 (2011) 1945–1957.
76. M.A. Martí-Renom, A.C. Stuart, A. Fiser, R. Sanchez, F. Melo, Comparative Protein Structure Modeling of Genes and Genomes, (2000) 37.
77. I.N. Smirnova, V. Kasho, H.R. Kaback, Protonation and sugar binding to LacY, *PNAS.* 105 (2008) 8896–8901.
78. M. Ishida, M. Hara, N. Fukino, T. Kakizaki, Y. Morimitsu, Glucosinolate metabolism, functionality and breeding for the improvement of Brassicaceae vegetables, *Breed. Sci.* 64 (2014) 48–59.