



**Universidad Nacional Autónoma de México**  
**Maestría y doctorado en Ciencias Bioquímicas**

“Estudio estructural del asa L9 en la proteína M271 de *Solanum tuberosum*, un inhibidor de proteasas de la familia Kunitz-STI”

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Maestro en Ciencias

PRESENTA:

Alexis Omar Campuzano González

TUTOR PRINCIPAL

Dr. Enrique Rudiño Piñera, Instituto de Biotecnología, U.N.A.M.

MIEMBROS DEL COMITÉ TUTOR

Dra. Gloria Saab Rincón, Instituto de Biotecnología, U.N.A.M.

Dra. Carmen Nina Pastor Colón, Centro de Investigación en Dinámica Celular,  
U.A.E.M.

Ciudad de México. Marzo, 2019



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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

## Agradecimientos

Al Dr. Enrique Rudiño Piñera, por aceptarme en su laboratorio, confiar en mi y darme la libertad creativa para desarrollar mi proyecto. Así como también por los regaños, las llamadas de atención y unos buenos “estate quieto” cuando mi terquedad iba más allá de donde debía.

A las Dras. Carmen Nina Pastor Colón y Gloria Saab Rincón por ser parte de mi comité tutor, por el tiempo que dedicaron a revisar mi proyecto y también por ese tiempo extra que les robe cuando tenía alguna duda a medio semestre.

A los Drs. Agustin Lopez Munguia Canales, Adela Rodríguez Romero, Lorenzo Patrick Segovia Forcella, Laura Dominguez Dueñas y Marcela Ayala Aceves miembros de mi comité evaluador, por las correcciones, observaciones, pero sobre todo los consejos no solo para mejorar mi proyecto, sino también para mejorar y apoyar mi carrera en el maravilloso mundo de la ciencia.

Al Dr. Yasel Guerra Borrego por compartirme un pedazo de su gran proyecto y enseñarme todos los trucos necesarios para poder desenvolverme mejor en el proyecto y en el laboratorio.

A la Dra. Adelaida Diaz Vilchis y a la Biol. Sonia Patricia Rojas Trejo por todo lo que me enseñaron al llegar al laboratorio y también durante estos ya más de dos años. Por siempre darme un consejo cuando llevaba a cabo algún experimento nuevo o al explicarme algunas cosas que solo la experiencia es capaz de enseñarte.

Al Dr. Gilberto Valdés García por enseñarme no solo dinámica molecular, sino también muchos consejos que me ayudaron a mejorar mi rendimiento en el trabajo.

A mis amigos del laboratorio Alfredo, Angela, Aranza, Edith, Jorge y Ricardo quienes no solo me ayudaron en el laboratorio, sino también fuera de él y por todas esas fiestas que hicimos y haremos.

A mis amigos Brenda, Gabriela, Juan Carlos, Perla y Verónica que, aunque estos últimos años no hemos podido convivir mucho, siempre han estado ahí para escucharme, apoyarme y darme uno que otro golpe si era necesario.

A mis amigos de la banda Reliant: Adrian, Carlos, Fernando, Juan Carlos y Raúl. Llegamos lejos juntos y aunque tuve que alejarme para cumplir mi sueño de volverme científico loco, nunca me apartaron completamente de la banda.

Al Dr. Gustavo Rodriguez Alonso quien me enseñó el camino del mal, quien siempre me motivo a seguir adelante y quien me recomendó el gran laboratorio al que me integre, aunque me diga traidor. Y como olvidarlo, mi primer amigo del IBT.

A mi familia mi madre Elma González Gómez, por apoyarme durante todos estos años y nunca dejarme solo y a mi segunda madre, mi abuela Rosa Bahena Benítez, quien me aceptó en su hogar a los 12 años y siempre estuvo ahí para mí.

A mi pequeña familia Thania Quiroz Hernandez por estar a mi lado y por darme lo más maravilloso que tengo, a mi hijo **Evan Darío Campuzano Quiroz**. ¡Va por ti hijo!

Por último, al Instituto de Biotecnología, quien me abrió sus puertas hace ya casi 6 años. Al Consejo Nacional de Ciencia y Tecnología por apoyarme con una beca para la realización de la maestría. A la Universidad Nacional Autónoma de México. Al Laboratorio Nacional de Estructuras de Macromoléculas (LANEM) y al Laboratorio Nacional de Cómputo de Alto Desempeño (LANCAD) por el proyecto aprobado LAN-CAD-UNAM-DGTIC-266 en el supercomputo Miztli de la U.N.A.M.

¡Va por todos ustedes!

“Entre lo que pienso, lo que quiero decir, lo que creo decir, lo que digo,  
lo que quieres oír, lo que oyes, lo que crees entender, lo que entiendes,  
existen nueve posibilidades de no entenderse”

Frase popular

“La ciencia, está compuesta de errores,  
pero son errores que es útil cometer,  
porque conducen poco a poco a la verdad.”

Julio Verne

Para Evan Darío Campuzano Quiroz, mi hijo y mi mayor motivación.

Para Jorge Romay Mugica, Don Romy, quien nos dejó este año, pero aún así aprovecho su tiempo para enseñarnos lo más que pudo.

Unos llegan y otros se van.

## Contenido

Agradecimientos .....	2
Contenido .....	6
Lista de figuras .....	8
Listado de tablas.....	9
Abreviaturas y acrónimos.....	9
RESUMEN.....	10
ABSTRACT.....	10
1 INTRODUCCIÓN.....	11
1.1 <i>Solanum tuberosum</i> .....	11
1.2 Proteasas .....	11
1.3 Proteasas aspárticas .....	12
1.4 Inhibidores proteicos de proteasas .....	15
1.5 Inhibidores de proteasas de la familia Kunitz-STI .....	15
2 ANTECEDENTES.....	16
2.1 Mecanismo de acción de los inhibidores de proteasas de la familia Kunitz-STI.....	16
2.2 M271 .....	18
3 JUSTIFICACIÓN.....	19
4 HIPÓTESIS.....	20
5 OBJETIVOS.....	20
5.1 Objetivo general .....	20
5.2 Objetivos específicos.....	20
6 MATERIALES Y MÉTODOS.....	20
6.1 Obtención de la cepa x33 de <i>Pichia pastoris</i> con la clona de M271.....	20
6.2 Expresión heteróloga y purificación del inhibidor M271 .....	20
6.3 Ensayos de actividad inhibidora contra PlmII.....	21
6.4 Obtención de la estructura cristalográfica del inhibidor .....	22
6.5 Colecta y procesamiento de datos de la difracción de rayos X .....	22
6.6 Comportamiento anisotrópico en cristalografía.....	22
6.7 Simulaciones de Dinámica Molecular .....	23
7 RESULTADOS Y DISCUSIÓN.....	24
7.2 Expresión y purificación de M271 .....	24
7.2 Pruebas de inhibición de PlmII con M271 .....	25
7.3 Estudios cristalográficos de M271 .....	26
7.4 Estudio anisotrópico de la estructura a 1.7 Å de resolución de M271. ....	34

7.5 Simulaciones de Dinámica Molecular .....	37
8 CONCLUSIONES.....	51
9 PERSPECTIVAS.....	52
10 BIBLIOGRAFÍA.....	52



## Lista de figuras

Figura 1: Esquema de la nomenclatura Berger y Schechter .....	12
Figura 2: Esquema de la PIm II.....	14
Figura 3: Esquema general de la estructura terciaria de los inhibidores de la familia Kunitz-STI.....	16
Figura 4: Esquema de la estructura secundaria del inhibidor E3Ad .....	17
Figura 5: Comparación del asa L9 entre E3Ad y PDI.....	17
Figura 6: Comparación de secuencia a tres inhibidores de la familia Kunitz-STI M271, E3Ad y DPI.....	19
Figura 7: Análisis de SDS-PAGE.....	25
Figura 8: Gráfico de la velocidad residual de PImII .....	26
Figura 9: Cristales obtenidos . .....	27
Figura 10: Patrón de difracción de M271. ....	28
Figura 11: Densidad electrónica de la estructura de M271 con resolución máxima de 1.7 Å .....	29
Figura 12: Densidad electrónica de la manosa .....	30
Figura 13: Modelo de listón de M271 .....	31
Figura 14: Densidad electrónica del asa L9 .....	32
Figura 15: Comparación estructural de M271, E3Ad y PDI. ....	33
Figura 16: Comparación entre el asa L9 de M271, E3Ad y PDI .....	34
Figura 17: Gráfica de anisotropía media de cada residuo de M271 .....	35
Figura 18: Comportamiento anisotrópico del asa L9 en la cadena B de M271.....	36
Figura 19: Comportamiento anisotrópico del asa L9 en la cadena A de M271 .....	36
Figura 20: RMSD de la cadena principal de M271. El asa L9 esta resaltada en elipses rojas.....	38
Figura 21: RMSD de la cadena principal de PDI. El asa L9 esta resaltada en elipses rojas. ....	39
Figura 22: RMSD de la cadena principal de E3Ad. El asa L9 esta resaltada en elipses rojas.....	40
Figura 23: RMSD de M271 y el asa L9 de la misma. En negro M271 y en rojo el asa L9. ....	41
Figura 24: RMSD de PDI y el asa L9 de la misma. En negro PDI y en rojo el asa L9. ....	42
Figura 25: RMSD de E3Ad y el asa L9 de la misma. En negro E3Ad y en rojo el asa L9. ....	43
Figura 26: Comparación de los valores mínimos y máximos de RMSD de M271 .....	44
Figura 27: Comparación de los valores mínimos y máximos de RMSD de PDI.....	45
Figura 28: Comparación de los valores mínimos y máximos de RMSD de E3Ad. ....	46
Figura 29: Cálculo de RMSF de M271 .....	47
Figura 30: Cálculo de RMSF de PDI.....	47
Figura 31: Cálculo de RMSF de E3Ad. ....	48
Figura 32: Cálculo de estructura secundaria de M271. ....	49
Figura 33: Cálculo de estructura secundaria de PDI .....	50
Figura 34: Cálculo de estructura secundaria de E3Ad .....	51

## Listado de tablas

Tabla 1: Datos estadísticos de las 4 estructuras cristalográficas obtenidas durante este proyecto.....	28
Tabla 2: Comparación de RMSD entre M271, E3Ad y PDI. ....	34

## Abreviaturas y acrónimos

Å: Ångström

AOX1: Alcohol oxidasa 1

API-A: Arrowhead protease inhibitor A

BMGY: Buffered Glycerol-complex medium

BMMY: Buffered Methanol-complex medium

Da: Daltons

IP: Inhibidores Proteicos

nm: nano metros

ns: nano segundos

PDA: Parámetros de desplazamiento atómico

PDB: Protein Data Bank

PDI: Potato Cathepsin D Inhibitor

PEG: Polietilén glicol

Plm II: Plasmepsina II de *Plasmodium falciparum*

ps: pico segundos

RMSD: Root-Mean-Square Desviation

RMSF: Root-Mean-Square Fluctuation

SDS-PAGE: Sodium Dodecyl Sulphate-Polyacrylamide gel Electrophoresis

STI: Soybean Trypsin Inhibitor

## RESUMEN

Ritonja desde 1990 (Ritonja et al., 1990) propuso por medio de la comparación de la secuencia de un inhibidor de la familia Kunitz-STI que estos podrían inhibir proteasas aspárticas. Esto se ha comprobado con trabajos previos (Guerra, Valiente, Pons, Berry, & Rudiño-Piñera, 2016), pero no se ha descrito cómo se lleva a cabo esta inhibición.

En este trabajo se estudió el inhibidor M271 de la familia Kunitz-STI, extraído de *Solanum tuberosum*, el cual es una variante de los dos inhibidores previamente descritos, E3Ad (Guerra et al., 2016) y PDI (Guo, Erskine, Coker, Wood, & Cooper, 2015). Este inhibidor se estudió por medio de cristalografía de rayos X y se realizaron simulaciones de dinámica molecular de M271 y de los dos inhibidores previamente reportados.

El inhibidor M271, presentó una alta flexibilidad en la región del asa L9. Esto se observó analizando el comportamiento anisotrópico del modelo tridimensional obtenido con cristalografía de rayos X, y en simulaciones de dinámica molecular en las cuales se observó que la región del asa L9 es una región que presenta alta flexibilidad en la proteína.

Del mismo modo las simulaciones de dinámica molecular que se llevaron a cabo en los inhibidores E3Ad y PDI, muestran que en estos inhibidores L9 es una región que presenta alta flexibilidad en las proteínas.

## ABSTRACT

Since 1990 Ritonja (Ritonja et al., 1990) proposed by comparing the sequence of a Kunitz-STI family inhibitor that those could inhibit aspartic protease. This has been verified with previous works (Guerra, Valiente, Pons, Berry, & Rudiño-Piñera, 2016), but it hasn't been described how this inhibition is carried out.

In this work we studied the M271 inhibitor of the Kunitz-STI family, extracted from *Solanum tuberosum*, which is a variant of the two previously described inhibitors, E3Ad (Guerra et al., 2016) and PDI (Guo et al., 2015). This inhibitor was studied by X-ray crystallography and with the performance of molecular dynamics simulations of M271 and the two previously described inhibitors.

M271 inhibitor showed a high flexibility in the region of L9. This was observed by analysing the anisotropic behaviour of the three-dimensional model obtained with X-ray crystallography, and in molecular dynamics simulations in which it observed that L9 is a high flexibility region in the protein.

In this way, the simulations of molecular dynamics that were carried out in the E3Ad and PDI inhibitors show that in those inhibitors L9 is a region with high flexibility.

# 1 INTRODUCCIÓN

## 1.1 *Solanum tuberosum*

La especie *Solanum tuberosum*, una variante de papa, contiene contra una gran cantidad de factores para asegurar su supervivencia (Koiwa, Bressan, & Hasegawa, 1997). Estos factores incluyen la presencia de depredadores (Koiwa et al., 1998), parásitos (Williamson & Hussey, 1996), invasores (Dunaevsky et al., 1997) y plagas (Pannetier et al., 1997). Dentro de todos estos factores, los depredadores y las plagas son especialmente preocupantes ya que afectan negativamente la supervivencia de la especie. Debido a esto la especie ha desarrollado un mecanismo defensivo conformado por inhibidores proteicos (IP) de proteasas.

## 1.2 Proteasas

Las proteasas, también conocidas como peptidasas, son enzimas hidrolíticas involucradas en la digestión de proteínas. Estas enzimas se encuentran presentes en múltiples organismos, tales como plantas, animales y en muchos microorganismos; llevan a cabo eventos proteolíticos, los cuales sirven como mediadores para señales de iniciación, transmisión y terminación de eventos celulares, tales como la respuesta inflamatoria, apoptosis celular y vías de procesamiento hormonal, por lo cual están estrechamente involucradas en la correcta función de los organismos (González-Rábade, Badillo-Corona, Aranda-Barradas, & Oliver-Salvador, 2011).

Estas enzimas se clasifican en exoproteasas, cuando el aminoácido que escinde se ubica en uno de los extremos de la cadena peptídica o a pocos residuos del extremo y en endoproteasas cuando el aminoácido a hidrolizar se ubica en el interior de la cadena peptídica (Rawlings, Tolle, & Barrett, 2004). Otra clasificación se basa en las especies involucradas en la catálisis enzimática, tales como metales o aminoácidos como: ácido aspártico, cisteína, ácido glutámico, serina y treonina (Barrett, Rawlings, & Woessner, 2013).

Aquellas que realizan un ataque nucleofílico que escinde el enlace peptídico por medio de una catálisis ácido-base general, utilizan ácido aspártico, ácido glutámico o iones metálicos, por otro lado, la catálisis covalente donde se forma un intermediario covalente el mecanismo utiliza serina, treonina o cisteína (Rawlings & Barrett, 2013).

Esta clasificación general, igualmente puede subclasificarse en clanes y familias. Los clanes se agrupan en familias de proteínas homólogas, fundamentalmente con base en la similitud a nivel de la estructura terciaria y en la organización de los aminoácidos catalíticos y su entorno (Rawlings et al., 2004).

Las familias se organizan de acuerdo con la similitud entre la secuencia de aminoácidos de sus miembros, para realizar esta subclasificación el análisis de las secuencias se restringe únicamente a la sección de la proteína directamente

responsable de la actividad de la misma, la cual se conoce como “unidad proteolítica”(Rawlings & Barrett, 2013). La nomenclatura de los clanes utiliza letras acordes al grupo catalítico que presentan las proteasas; A para aspártico, S para serino y así sucesivamente.

También se desarrolló una nomenclatura para las posiciones de los aminoácidos en las regiones del centro activo de la proteasa y del sustrato de la enzima que interactúan con cada una de ellas. A esta nomenclatura se le conoce como “nomenclatura de Berger & Schechter” y nombra las posiciones hacia el extremo amino del enlace a hidrolizar como P1-Pn y a las posiciones hacia el extremo carboxilo como P1'-Pn' (Schechter & Berger, 1967). (Figura 1).

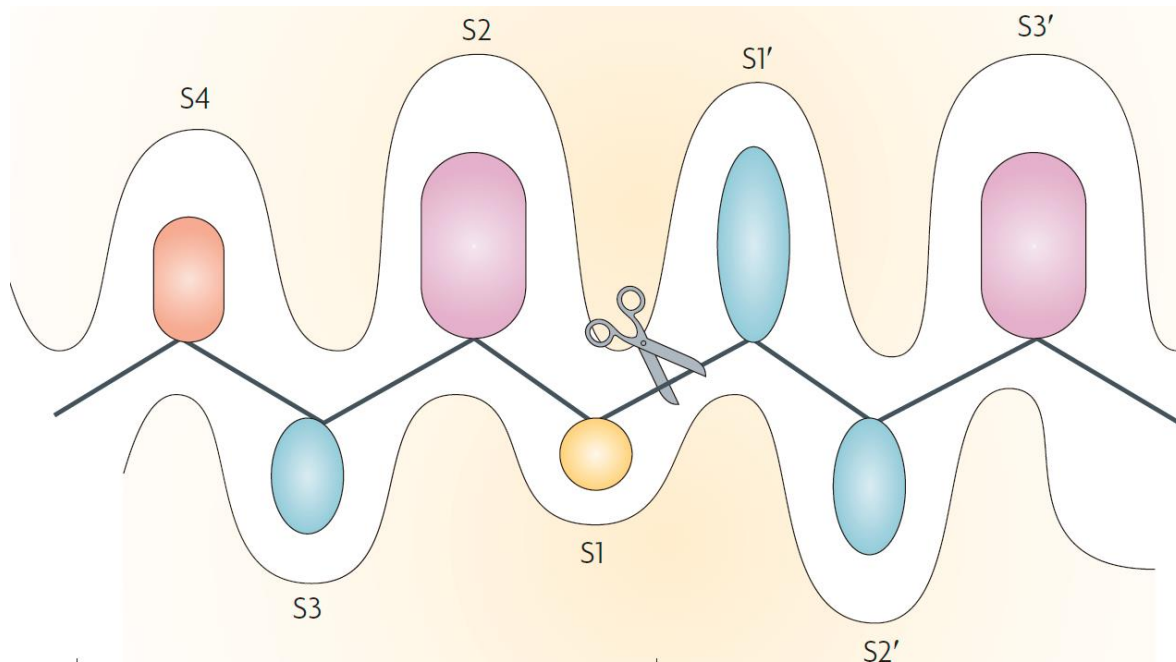
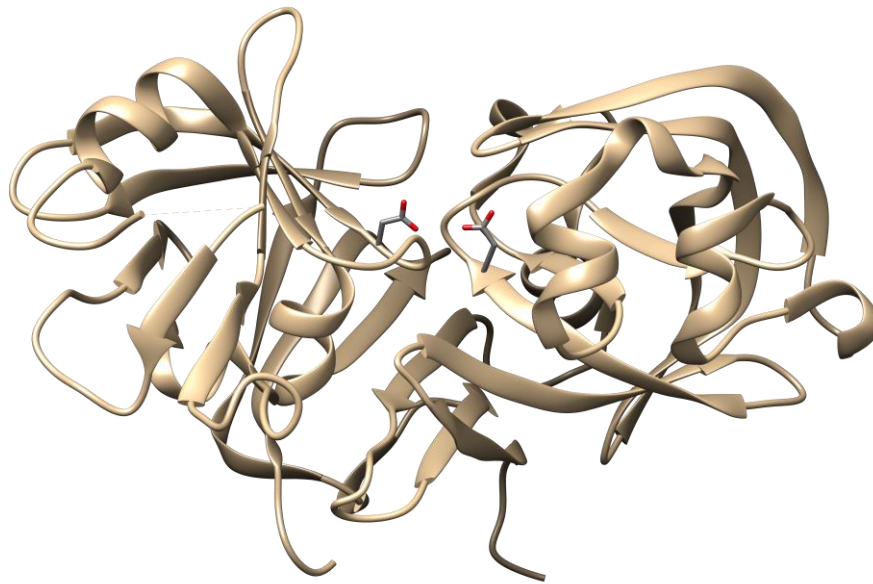


Figura 1: Esquema de la nomenclatura Berger y Schechter. Tomado y modificado de Drag & Salvesen, 2010.

### 1.3 Proteasas aspárticas

Las proteasas aspárticas son nombradas de esta forma porque un residuo de ácido aspártico está involucrado en la catálisis de la proteasa. En la mayoría de las proteasas de este tipo, se reporta un par de residuos de ácido aspártico los cuales llevan a cabo la catálisis. En el caso de la Plasmepsina II de *Plasmodium falciparum* (Plm II) estos residuos corresponden a los aspárticos 34 y 214. Pero también es posible que otro residuo participe junto con el aspártico en esta catálisis. También se ha descrito que estas enzimas únicamente presentan actividad de endopeptidasas (Tang & Wong, 1987). Las proteasas aspárticas forman parte de los clanes AA, AC, AD, AE y AF. Y se conoce la estructura terciaria de todas ellas con excepción de AC, estas proteasas presentan un plegamiento único, que no está relacionado con ninguna otra estructura de otro clan (Rawlings & Barrett, 2013).

El clan AA contiene la familia más grande de proteasas aspárticas, la familia A1, la cual tiene a la pepsina como su mejor ejemplo (Rawlings & Barrett, 2013). Estructuras cristalográficas de esta familia muestran que esta enzima es bilobulada, con el centro activo ubicado entre ambos lóbulos, este tipo de estructura es muy común en muchas proteasas, pero en la familia A1 un lóbulo evolucionó a partir del otro por medio de duplicación genética. Sin embargo, estos lóbulos presentan poca identidad entre ellos, pero conservan la estructura tridimensional, cada lóbulo provee un aspártico para llevar a cabo la catálisis y ambos lóbulos son homólogos (Tang & Wong, 1987). (Figura 2)



*Figura 2:* Esquema de la Plm II, arriba se observa la proteasa completa con sus dos lóbulos. Abajo se observa únicamente un lóbulo y se señala la distancia entre la entrada de la cavidad y uno de los aspárticos involucrados en la reacción, la cual es de 20 Å.

Sin embargo, debido a la naturaleza misma de las proteasas es necesario que éstas se encuentren finamente reguladas y controladas, para así evitar posibles daños a los organismos donde se encuentran presentes o en los procesos en los cuales se busca involucrarlas (Theron & Divol, 2014). Esta regulación se lleva a

cabo por medio de IP de proteasas, los cuales inhiben las reacciones de proteólisis llevadas a cabo por las proteasas (Valueva & Mosolov, 2004).

#### 1.4 Inhibidores proteicos de proteasas

Los IP de plantas son proteínas pequeñas que se presentan en altas concentraciones en tejidos de almacenaje. Contribuyen a más del 10% del total de proteína y también es posible encontrarlas en las hojas en respuesta al ataque de insectos o microorganismos patógenos (Ryan, 1990). Algunos IP han demostrado actuar como compuestos defensivos contra plagas por contacto directo o por expresión en cultivos transgénicos (Krattiger, 1996).

La primera evidencia de la protección de las plantas por medio de inhibidores de proteasas se observó con el desarrollo de las larvas de ciertos insectos los cuales eran incapaces de desarrollarse de manera normal en las plantas de soya. Posteriormente se demostró que los inhibidores de tripsina presentes en la soya eran tóxicos a las larvas de los escarabajos de harina (Lawrence & Koundal, 2002). Además de este estudio en otros trabajos se demostró su efectividad contra insectos (Koiwa et al., 1998; Pannetier et al., 1997; Zhao et al., 1996), nematodos patogénicos (Williamson & Hussey, 1996), hongos (Dunaevsky et al., 1997; Joshi, Sainani, Bastawade, Gupta, & Ranjekar, 1998).

La expresión de los genes de los IP se lleva a cabo en órganos específicos en fases particulares del crecimiento de la planta. Como la germinación (Botella et al., 1996), la senescencia temprana de la hoja (Huang et al., 2001) y durante sequías (Pernas, Sánchez-Monge, & Salcedo, 2000; Van der Vyver et al., 2003).

Durante algunos eventos de estrés hídrico, los tubérculos de la papa secretan dos tipos de IP de la familia Kunitz, esos IP están asociados con otros polipéptidos secretados y los protegen contra la degradación extracelular de la quimotripsina. Estos IP también interactúan con el sistema defensivo de la planta (Ledoigt et al., 2006; Valueva et al., 2001)

#### 1.5 Inhibidores de proteasas de la familia Kunitz-STI

La familia Kunitz, la cual se distingue por ser una familia proteica que inhibe principalmente proteasas serínicas, también es capaz de actuar ante otros mecanismos como proteasas aspárticas y proteasas cisteínicas (Pouvreau, Gruppen, Van Koningsveld, Van Den Broek, & Voragen, 2003). Esta familia presenta un plegamiento de  $\beta$ -*trefoil*, compuesto por 12 láminas beta, formando un triple pliegue en unidades seudo simétricas o subdominios con la presencia de hélices en algunos casos. Las unidades seudo simétricas están compuestas por cuatro láminas beta, dos de las cuales forman un barril beta conformado por seis láminas beta, las otras dos forman un  $\beta$ -*hairpin* el cual actúa como tapa del barril beta (Park et al., 2005) (Figura 3).



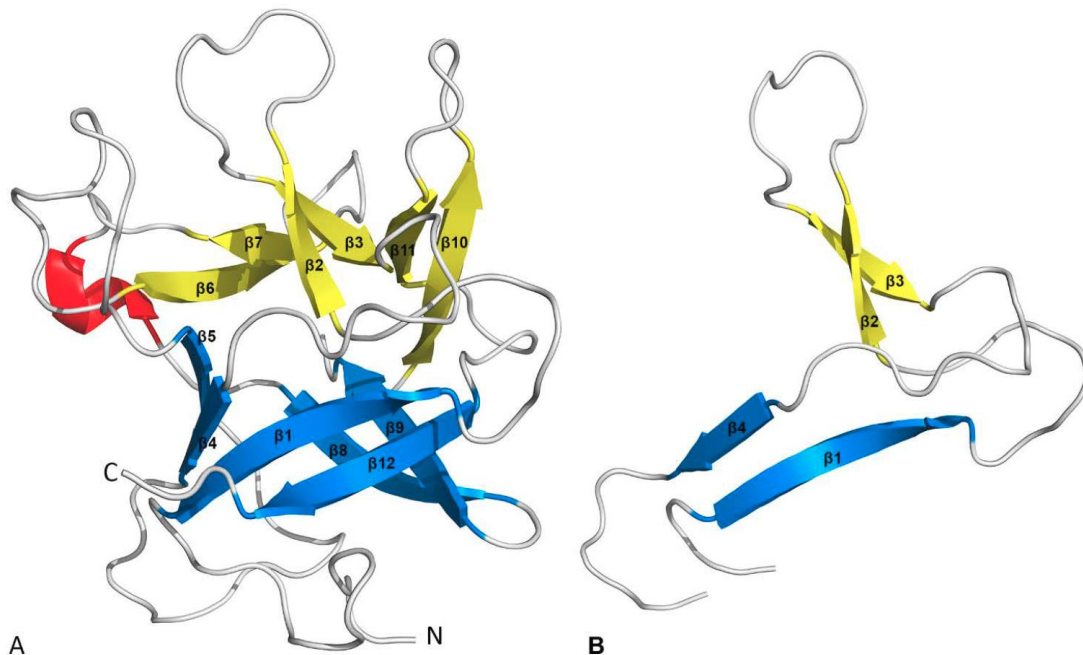


Figura 3: A) Esquema general de la estructura terciaria de los inhibidores de la familia Kunitz-STI. B) Esquema de cada sub-unidad que conforma la estructura terciaria por medio de un eje de pseudo simetría de orden 3.

Se ha reportado que las regiones estructurales involucradas en las interacciones proteasa-inhibidor incluyen al N- y al C- terminal, así como a las asas expuestas (Pouvreau et al., 2003). Estos últimos pueden presentar variación en su secuencia de aminoácidos y su longitud y aun así conservar su actividad (Renko, Sabotič, & Turk, 2012).

## 2 ANTECEDENTES

### 2.1 Mecanismo de acción de los inhibidores de proteasas de la familia Kunitz-STI

Los inhibidores de la familia Kunitz-STI poseen asas de unión a proteasas involucrados en la actividad del inhibidor por medio de la formación de un complejo proteasa-inhibidor, con secuencias y longitudes muy variadas, donde la mayoría de las asas presentan una conformación canónica en las posiciones P3-P3' (Bao et al., 2009). Sin embargo, no todos los miembros de la familia Kunitz-STI presentan esta conformación y conservan su función inhibitoria (Guerra et al., 2016)

Para la mayoría de los miembros de la familia Kunitz-STI previamente caracterizados estructuralmente se observó que el asa L4 lleva a cabo la inhibición para proteasas serínicas, localizado entre las láminas beta 4 y 5. Sin embargo, esto no se observa en todos los casos, debido a que también es posible encontrar esta actividad en diferentes asas, como el asa L2, ubicada entre las láminas beta 2 y 3, en el asa L5, entre las láminas beta 5 y 6 (Bao et al., 2009) y el asa L9, entre las láminas beta 9 y 10 (Figura 4), el cual previamente se propuso como la región involucrada en la inhibición de proteasas aspárticas (Ritonja et al., 1990), pero esto no ha sido demostrado ni descrito.

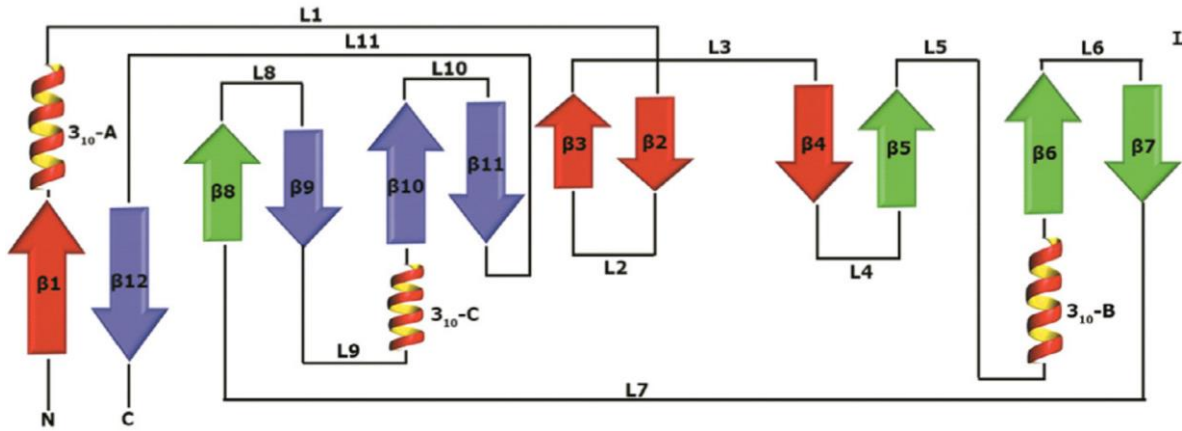


Figura 4: Esquema de la estructura secundaria del inhibidor E3Ad de la familia Kunitz-STI. Figura tomada y modificada de Guerra *et al.* 2016

En el asa L9, de dos inhibidores de la familia Kunitz-STI extraídos de *Solanum tuberosum* descritos previamente. Se observó en el inhibidor E3Ad (Guerra *et al.*, 2016), que la conformación estructural del asa L9 se encuentra extendida hacia el exterior del inhibidor, mientras que en otro trabajo (Guo *et al.*, 2015) se encontró que esta misma asa en PDI, el asa L9 se encuentra contraída hacia el núcleo de la proteína. (Figura 5).

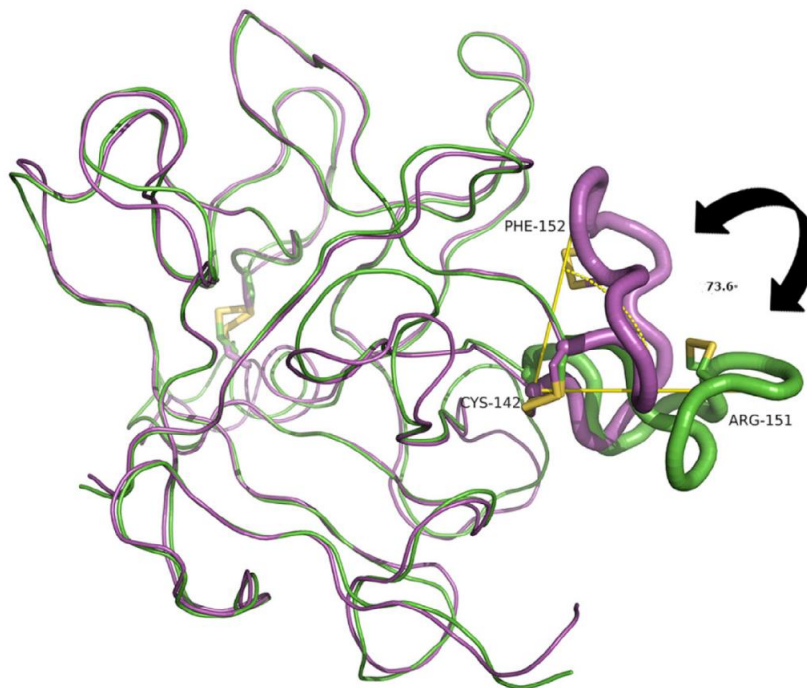


Figura 5: Comparación del asa L9 entre E3Ad (PDB:5fnw, en verde) y PDI (PDB:5dzu, en morado). Tomado de Guerra, *et al.*, 2016.

Igualmente en trabajos previos se predijo mediante diversas herramientas computacionales que un inhibidor de tomate de la familia Kunitz-STI podría utilizar

el asa L9 como región responsable de la inhibición de proteasas aspárticas y se propuso que la Arg154, ubicada en el asa L9 de un inhibidor de proteasa aspártica, lleva a cabo la inhibición de la proteasa (Cater et al., 2002; Guerra, Valiente, Berry, & Pons, 2012). Sin embargo, Guerra en 2016, mutó la arginina equivalente en el inhibidor E3Ad por una alanina y no se alteró la función inhibitoria del mismo (Guerra et al., 2016).

Es importante mencionar que el asa L9 se encuentra ubicado estructuralmente en una posición equivalente a la unión proteasa-inhibidor usado por el inhibidor API-A para interactuar con tripsina (Guerra et al., 2012). Considerando todo lo anterior, es posible que la actividad del inhibidor contra proteasas aspárticas se encuentra regulada por la conformación estructural del asa L9.

## 2.2 M271

Un inhibidor de proteasas, extraído de *Solanum tuberosum*, se nombró M271, el cual tiene una masa molecular de 21700 Da, presenta un punto isoeléctrico de 8.75, está compuesto por 189 aminoácidos y pertenece a la familia Kunitz-STI.

Este inhibidor presenta diferencias en la secuencia de aminoácidos con respecto a E3Ad, reportado por Guerra en 2016, presentando una identidad<sup>1</sup> del 87.3% y una similitud<sup>2</sup> del 94.7%. De los 189 aminoácidos del inhibidor 24 son diferentes respecto al inhibidor E3Ad, de estas 24 diferencias 10 se ubican en la región del asa L9, el cual presenta una identidad del 44.4% y una similitud del 72.2%.

En el caso de PDI, reportado por Guo en 2015, M271 y I.D.C.D. presentan una identidad del 92.1% y una similitud de 95.8%. De los 189 aminoácidos del inhibidor 15 son diferentes respecto a I.D.C.D., de los cuales 3 se ubican en la región del asa L9, el cual presenta una identidad del 83.3% y una similitud del 88.9%.

---

<sup>1</sup> En el alineamiento de secuencias se denomina identidad, a la comparación directa de residuos. Es decir, que el residuo es el mismo en las secuencias comparadas.

<sup>2</sup> Por otro lado, el término de similitud se refiere a la comparación de características. Es decir, dos residuos que no son idénticos, pero presentan propiedades similares. *E.g.*: dos residuos ácidos, un ácido aspártico y un ácido glutámico, no presentan identidad, pero sí similitud.

(Figura 6).



Figura 6: Comparación de secuencia de los tres inhibidores de la familia Kunitz-STI M271, E3Ad y DPI, en azul se resalta la zona del asa L9. Código de colores, arriba a la izquierda. Código de símbolos: asterisco, identidad. Dos puntos, similitud. Espacio, sin identidad ni similitud.

Por la alta similitud del IP M271 con los dos IP bifuncionales con estructura cristalográfica reportados previamente, se decidió utilizar a M271 como modelo adicional para el estudio de la dinámica de movimiento del asa L9, en el contexto de la posible inhibición de proteasas aspárticas.

### 3 JUSTIFICACIÓN

Debido a trabajos previos (Cater et al., 2002; Guerra et al., 2012, 2016; Guo et al., 2015) donde se encontró información que sugiere que la conformación del asa L9 de los inhibidores de la familia Kunitz-STI podría estar regulando la interacción proteasa-inhibidor, se decidió estudiar los cambios estructurales que presenta esta asa por el cambio en la secuencia de aminoácidos del asa L9.

Desde 1990 Ritonja sugirió la posibilidad de que el asa L9 sea la causante de la inhibición de proteasas aspártico, en especial se menciona a la catepsina D, aunque esto basado únicamente en la secuencia de aminoácidos del mismo (Ritonja et al., 1990), lo que nos lleva al hecho de que este problema lleva casi 30 años sin una respuesta.

## 4 HIPÓTESIS

El asa L9 de M271, un inhibidor de proteasas de la familia Kunitz-STI, posee una alta flexibilidad.

## 5 OBJETIVOS

### 5.1 Objetivo general

Estudiar los cambios conformacionales del asa L9 en la variante M271 aislado de *Solanum tuberosum*, un inhibidor de proteasas de la familia Kunitz-STI, en la inhibición de proteasas aspárticas.

### 5.2 Objetivos específicos

1. Obtención del inhibidor M271 con pureza y concentración suficientes para estudios de cristalografía.
2. Realizar pruebas de inhibición de M271 sobre proteasas aspárticas y compararla con E3Ad.
3. Determinación estructural de la variante M271 por cristalografía de rayos X.
4. Realizar simulaciones de dinámica molecular para estudiar los cambios estructurales del asa L9 en ausencia de vecinos cristalográficos y el probable impacto de estos cambios sobre la actividad inhibitoria frente a proteasas aspárticas.

## 6 MATERIALES Y MÉTODOS

### 6.1 Obtención de la cepa x33 de *Pichia pastoris* con la clona de M271

La secuencia codificante del inhibidor M271 se amplificó a partir del ADN genómico de *Solanum tuberosum*, en el laboratorio del Dr. Colin Berry de la Escuela de Biociencias de la Universidad de Cardiff en el Reino Unido, y fue donado al laboratorio para su estudio.

El inhibidor se expresó en la levadura *Pichia pastoris* X33, la construcción tiene integrada el factor de expresión por metanol AOX1 y el factor  $\alpha$  de secreción de *Saccharomyces cerevisiae* fusionado con su extremo amino, por lo cual el inhibidor fue secretado al medio de cultivo después de su expresión, esta construcción cuenta con una etiqueta de histidinas (6x) en su extremo carboxilo para facilitar su purificación con una columna de afinidad a  $Ni^{2+}$ .

### 6.2 Expresión heteróloga y purificación del inhibidor M271

Se inocula una cepa de *Pichia pastoris* previamente transformada conocida como la cepa M271, en 200 mL de medio BMGY (2% triptona, 1% extracto de levadura, 0.1 M de fosfato de potasio, pH 6.0, 1.34% de base nitrogenada de levadura, 40  $\mu$ M de biotina y 1% de glicerol) en frascos bafleados a 30 °C y 180 rpm por 24 horas. Se mide la densidad óptica del medio a 600 nm para medir indirectamente la concentración de biomasa, posteriormente el cultivo se centrifuga para separar la biomasa del medio BMGY a 3000 g por 10 minutos y posteriormente

se lava con medio BMMY (2% triptona, 1% extracto de levadura, 0.1 M de fosfato de potasio, pH 6.0, 1.34% de base nitrogenada de levadura, 40  $\mu\text{M}$  de biotín y 1% de metanol) y se centrifuga bajo las condiciones previamente mencionadas. La expresión del inhibidor se lleva a cabo resuspendiendo la biomasa recuperada previamente en medio BMMY hasta una  $\text{OD}_{600\text{nm}}$  de 1.0. Este medio se incuba con agitación continua a 28 °C y 180 rpm por 72 horas, metanol al 100% se agrega cada 24 horas hasta una concentración final de 1% para mantener la expresión del inhibidor. Este medio se centrifuga a 3000 g por 10 minutos y el sobrenadante es recuperado para realizar la purificación de la proteína.

El sobrenadante se lleva al doble del volumen obtenido y se agrega cloruro de sodio e imidazol, hasta tener concentraciones finales de 50 mM de fosfato de potasio, 200 mM de cloruro de sodio y 40 mM de imidazol y se ajusta el pH a 7.2. Esto se deja incubando toda la noche bajo las mismas condiciones donde inició la purificación, posteriormente se centrifuga a 10000 g por 30 minutos, se desecha el precipitado y se recupera el sobrenadante, se filtra con filtros de tamaño de poro de 0.45  $\mu\text{m}$  y se realiza una cromatografía de afinidad con una columna de níquel, que se equilibra con un buffer que tenga las mismas propiedades del sobrenadante y se realiza un gradiente de elución de 0 a 100 % de concentración de imidazol para separar al inhibidor de la columna. Las fracciones se analizan con SDS-PAGE y posteriormente se colecta la proteína y se concentra utilizando unidades de filtración en centrifuga de 10 kDa (Millipore). La muestra concentrada del inhibidor se aplica en una columna superdex 75 16/600 HR (*GE Healthcare Life Sciences*) equilibrada con 10 mM de formiato de sodio y 150 mM de cloruro de sodio a pH 4.0 acoplado a un sistema Äkta-pure (*GE Healthcare Life Sciences*). Las fracciones se analizan con SDS-PAGE y se concentra utilizando unidades de filtración en centrifuga de 10 kDa (Millipore).

### 6.3 Ensayos de actividad inhibidora contra PlmII

Para los ensayos de actividad inhibidora contra Plasmeptina II (PlmII), se utilizó el sustrato peptídico cromogénico KPFEFNphRL ( $K_M$ : 50.8  $\mu\text{mol L}^{-1}$ ; *Genscript*, EUA) a una concentración final de ensayo igual a 84.0-84.5  $\mu\text{mol L}^{-1}$  y usando el amortiguador NaAc 100  $\text{mmol L}^{-1}$  a pH 4.4, estas reacciones fueron monitoreadas de forma continua a 310 nm en un espectrofotómetro Cary60 (*Agilent Technologies*, EUA) a una temperatura constante de 37 °C, temperatura óptima de PlmII.

Para estos ensayos cinéticos, tanto la enzima como el inhibidor se incuban durante 5 minutos previos a la adición del sustrato. La velocidad inicial de la reacción se determinó con el programa Cary WinUV Kinetics v5.0 (*Agilent Technologies*, EUA), a diferentes concentraciones del inhibidor. La fracción de actividad residual ( $v_i/v_0$ ) se calculó a partir de la velocidad inicial de las reacciones en presencia de diferentes concentraciones de inhibidor ( $v_i$ ) respecto a la velocidad de la reacción en ausencia de inhibidor ( $v_0$ ).

#### 6.4 Obtención de la estructura cristalográfica del inhibidor

Se evaluaron condiciones de cristalización usando las matrices de condiciones de cristalización *Wizard I-IV* (*Emerald Biosciences*), *PEG/lon*, *Crystal Screen*, *Index*, *Crystal Screen Cryo* (*Hampton Research*, EUA) y el método de gota asentada. Se realizarán pruebas de cristalización a 4 y 18 °C. El método de gota apoyada utilizó un robot de cristalización *TTP Labtech's Mosquito® LCP* (*TTP Labtech*, Reino Unido) para preparar gotas compuestas por 100 nL de la solución de la proteína más 100 nL de la solución para cristalización. Se dispensaron 20 µL de la solución de cristalización en el pocillo correspondiente a cada gota.

También se realizó una matriz de optimización alrededor de la condición 6 (Cloruro de magnesio hexahidratado 0.2 M, TRIS 0.08 M, PEG 4000 24% p/v y glicerol 20% v/v) del *kit Crystal Screen Cryo* con el método de gota colgante a 4° C. En esta matriz, se realizaron ajustes a la concentración de la sal (Cloruro de magnesio hexahidratado) en un rango del 0.18 M al 0.24 M, con aumentos de 0.02 en molaridad. Igualmente, se ajustó la concentración del PEG 4000 en un rango del 18% al 28% p/v, con aumentos del 2%.

Los cristales obtenidos se remojaron con una solución crio-preservante (condición de cristalización con 30-40% glicerol, etilenglicol o PEG 200-400). Posteriormente los cristales fueron montados a -173 °C en una crío-asa de rayón.

#### 6.5 Colecta y procesamiento de datos de la difracción de rayos X

La difracción por rayos X de los cristales de M271 se realizó en la línea 19-ID del APS (Advanced Photon Source), con un detector ADSC Q315r CCD. Los patrones de difracción obtenidos fueron indexados e integrados con los programas XDS (Kabsch, 2010) y Mosflm (Leslie & Powell, 2007). El escalamiento, ordenamiento y empalme de los datos se realizó con XSCALE (Kabsch, 2010) y AIMLESS (Evans, 2006). Para la conversión de los valores de intensidad de las reflexiones, procesadas con XDS, en estimados de los valores de los factores de estructura, se utilizó el programa XDSCONV (Kabsch, 2010). Las fases fueron resueltas por reemplazo molecular con el programa PHASER (McCoy et al., 2007) y la estructura del inhibidor de papa PDI (PDB ID: 5DZU) que presenta una identidad del 92.1% con M271, esto para la primera estructura de esta tesis. Las tres estructuras obtenidas posterior a esta utilizaron el modelo generado con PDI para resolver fases por medio de reemplazo molecular. Los modelos obtenidos se afinaron en condiciones de cuerpo rígido con el programa REFMAC5 (Murshudov et al., 2011). Para la validación final del modelo, se realizaron ciclos de construcción manual y afinamiento restringido global con los programas Coot (Emsley & Cowtan, 2004) y PHENIX (Adams et al., 2010), respectivamente.

#### 6.6 Comportamiento anisotrópico en cristalografía

Se utilizó el servidor TLSMD (Translation/Libration/Screw Motion Determination) (Painter & Merritt, 2006b) el cual utiliza modelos matemáticos para predecir el desplazamiento de las posiciones locales de un grupo de átomos en una

estructura cristalográfica basándose en la suposición subyacente de que cada átomo actúa como miembro de un cuerpo rígido desplazado normalmente sobre una posición media. El desplazamiento neto es el resultado de la vibración de la molécula en el cristal y por el desorden estático dentro de la red cristalina por la obtención de varias micro conformaciones de la molécula en las diferentes celdas unitarias que forman el cristal (Painter & Merritt, 2006a).

Debido a la simplificación del análisis del comportamiento anisotrópico en el cristal por medio de grupos de átomos con un comportamiento anisotrópico similar, y que es posible emparejar, es que se decidió utilizar el servidor TLSMD.

Esto porque el cálculo del comportamiento anisotrópico en átomos individuales requiere de seis parámetros de desplazamiento atómico (PDA) y de tres parámetros de posición, lo que describiría una nube electrónica con forma elipsoide y con una dirección preferente de movimiento. Mientras que el cálculo del comportamiento isotrópico requiere de 3 parámetros de posición para generar una nube electrónica uniforme en todas direcciones, es decir, de forma esférica. Por lo tanto, se requiere de datos de difracción de alta resolución para tener una relación datos parámetros mayor a 2, para no sobre afinar los datos del modelo obtenido por cristalografía y poder validar el mismo.

TLSMD utiliza un algoritmo de optimización de dos pasos para seleccionar los  $n$  grupos óptimos en las cadenas proteicas para transformar en grupos TLS. El primer paso de la optimización agrupa segmentos consecutivos de la secuencia de residuos, para agregar a una subrutina que ajusta los parámetros de los modelos TLS con los parámetros de desplazamiento atómico de cada átomo. Esto genera varios modelos con grupos TLS de distinta longitud, pero con residuos subsecuentes, es decir en cadena y no saltados.

El segundo paso de la optimización selecciona los grupos TLS generados en el primer paso que mejor ajusten con los parámetros de desplazamiento atómico. Esto se lleva a cabo con la minimización de ruta más corta sobre un gráfico. Estos gráficos se construyen colocando vértices entre cada residuo y al principio y final de la cadena. Para cada curva del gráfico, la penalización de cada curva es el residuo de mínimos cuadrados en cada curva. Con estos datos se emplea un algoritmo de ruta corta (Restricción de Bellman-Ford) para encontrar las rutas con menor penalización, desde el inicio hasta el final de la cadena. Los grupos TLS con menor penalización son los mostrados al final por el servidor TLSMD.

La implementación del análisis del comportamiento anisotrópico por medio de TLS nos ayudó a mejorar los valores de  $R$ -work y  $R$ -free en el afinamiento de la estructura a 1.7, mejorando los  $R$ 's de 20, 24 a 17, 20.

## 6.7 Simulaciones de Dinámica Molecular

Las simulaciones de dinámica molecular se llevaron a cabo con el software GROMACS (Abraham et al., 2015). Los archivos de topología y estructura de los



inhibidores M271, E3Ad e I.D.C.D. se crearon con pdb2gmx (Abraham et al., 2015). Se utilizó el conjunto de parámetros del campo de fuerza Amberff99sb (Hornak et al., 2006). Los tres inhibidores se simularon a un pH=7. El estado de protonación de los aminoácidos ionizables de los tres inhibidores se calculó con el programa PROPKA (Li, Robertson, & Jensen, 2005) en el servidor PDB2PQR ([http://nbcrc-222.ucsd.edu/pdb2pqr\\_2.0.0/](http://nbcrc-222.ucsd.edu/pdb2pqr_2.0.0/)). Se generaron condiciones periódicas de entorno en una caja cúbica con una distancia de 1 nm (10 Å) de la superficie de la proteína a las paredes de la caja con la herramienta editconf (Abraham et al., 2015). Las moléculas de solvente se añadieron con el programa genbox (Abraham et al., 2015) con el modelo de agua TIP3P (Jorgense, Chandrasekjar, & Madura, 1983). Se neutralizaron cargas en el sistema añadiendo 0.1 M de NaCl, con la herramienta genion (Abraham et al., 2015). Se obtuvieron modelos con número de átomos igual a 53101 para M271, 41240 para E3Ad y 34338 para I.D.C.D.

La temperatura y presión de los sistemas se controló mediante el algoritmo de Berendsen (Berendsen, Postma, Van Gunsteren, Dinola, & Haak, 1984) y el algoritmo de Parrinello-Rahman (Parrinello & Rahman, 1981), respectivamente. A una temperatura de 296 K y una presión de 100 kPa. Las interacciones de largo alcance y de van der Waals, se calcularon con el método PME (del inglés "*Particle Mesh Ewald*") (Darden, York, & Pedersen, 1993) y la expresión de Lennard-Jones, respectivamente. El radio de corte de las interacciones electrostáticas de largo alcance se fijó en una distancia mayor a 1.0 nm, y la expresión de Lennard-Jones para las interacciones de corto alcance se fijó a un radio de corte menor de 1.0 nm.

Los sistemas se equilibraron con una minimización de energía en la cual se tomó como criterio de optimización que la energía no fuera mayor a 1000 KJ mol<sup>-1</sup> nm<sup>-1</sup>. Adicionalmente se realizaron equilibraciones del sistema en los ensambles NVT (número de partículas, volumen y temperatura constante), en el cual se estabilizó durante 100 ps (pico segundo) la temperatura del sistema y posteriormente se equilibró el sistema con un ensamble NPT (número de partículas, presión y temperatura constante), en el cual se estabilizó la presión del sistema durante 100 ps.

## 7 RESULTADOS Y DISCUSIÓN

### 7.2 Expresión y purificación de M271

La expresión de M271 en la cepa x33 de *Pichia pastoris* se optimizó a través de modificaciones en el tiempo de expresión, de 3 a 5 días. En la concentración final de metanol, de 1 a 2 % v/v, y en la concentración de glicerol, de 0 a 1% v/v en la fase de expresión. Con lo cual se obtuvieron concentraciones y volúmenes similares de proteína pura, 15 mg/ml en 400 µL de volumen, con el cambio de volumen de cultivo de 900 mL a 300 mL.

La purificación se llevó a cabo con una combinación de cromatografía de afinidad a Ni<sup>2+</sup> y exclusión molecular (Figura 7).

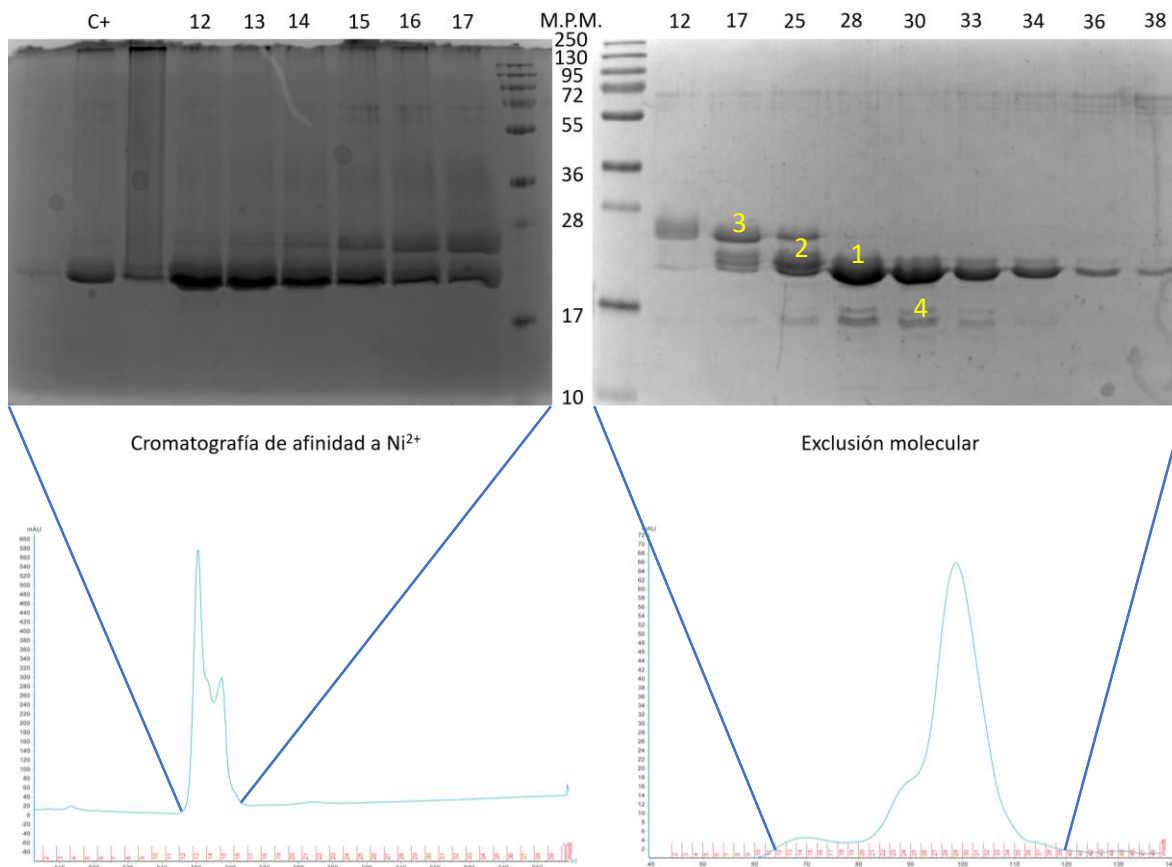


Figura 7: Análisis de SDS-PAGE de la cromatografía de afinidad a  $\text{Ni}^{2+}$  y la exclusión molecular de M271. Perfiles de elución de los pasos de afinidad y exclusión molecular.

En la purificación se obtuvieron cuatro especies principales: la primera (de concentración mayor a menor) corresponde al inhibidor M271 no glicosilado (1), la segunda al inhibidor glicosilado (2), el cual se ubica justo encima de la banda del inhibidor no glicosilado, la tercera especie podría corresponder a otra especie glicosilada (3), la cual no pudo comprobarse y la cuarta especie podría corresponder a un degradado de proteínas durante el proceso de purificación (4). En las fases de cromatografía de afinidad y en la exclusión molecular la proteína se concentró con unidades de filtración de 10 kDa.

Estas muestras fueron utilizadas para llevar a cabo pruebas de cristalización y de inhibición a la proteasa Plm II. La concentración del inhibidor fue de entre 11 a 18 mg/ml, con un rendimiento total de 4.5 mg de proteína por litro de cultivo.

## 7.2 Pruebas de inhibición de PlmII con M271

En las pruebas de inhibición de M271 se obtuvo una  $K_i$  de  $5.9 \pm 1.05$  nmol/L, esto con un  $R^2$  de 0.98, el ajuste de la curva se puede observar en la figura 8. En comparación con el valor de  $K_i$  reportado para E3Ad el cual es de  $10.5 \pm 2.5$  nmol/L (Guerra et al., 2016). Estos resultados se obtuvieron de un ajuste de regresión no

lineal con la ecuación de Morrison (Ecuación 1) para inhibidores con el régimen de unión fuerte.

$$\frac{v_i}{v_0} = 1 - \frac{([E]_0 + [I]_0 + Ki^{app}) - \sqrt{([E]_0 + [I]_0 + Ki^{app})^2 - 4[E]_0[I]_0}}{2[E]_0} \quad (\text{Ecuacion 1})$$

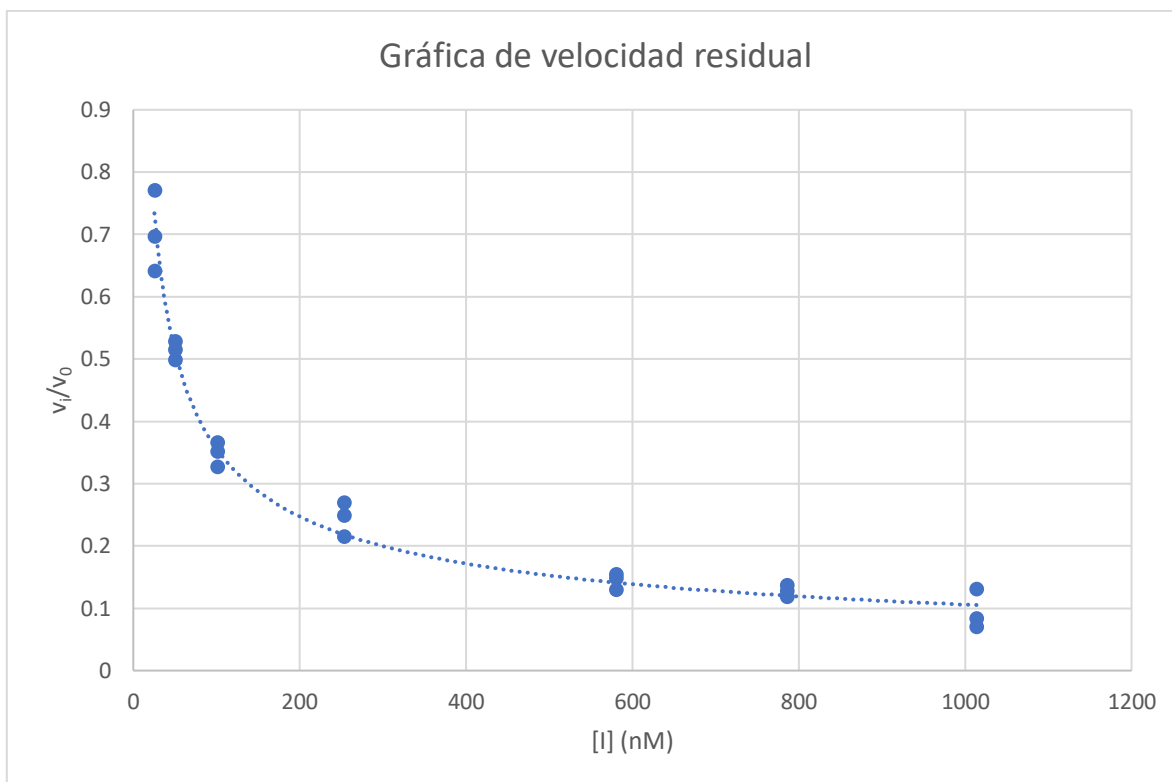


Figura 8: Gráfico de la velocidad residual de PlmII de acuerdo con la concentración del inhibidor M271.

Debido al régimen de inhibidores de unión fuerte las diferencias entre M271 y E3Ad, son prácticamente despreciables. Esto debido a que, a diferencia de otros inhibidores, donde los cambios en el valor directo de  $K_i$ , determina su afinidad a la enzima inhibida, los inhibidores de unión fuerte presentan cambios en órdenes de magnitud, esto quiere decir que requieren cambios de, por ejemplo, nmol a  $\mu$ mol para considerarse cambios significativos.

### 7.3 Estudios cristalográficos de M271

Se realizaron pruebas de cristalización con los kits *Wizard I-IV* (Emerald Biosciences), *PEG/lon*, *Crystal Screen*, *Crystal Screen Lite*, *Crystal Screen Cryo*, *Index* (Hampton Research, EUA). Se probaron 495 condiciones de cristalización diferentes, con dos concentraciones diferentes 5 y 10 mg/ml y a dos temperaturas diferentes 4 y 18 °C.

Se obtuvo un cristal en la condición 6 del kit *Crystal Screen Cryo* (Cloruro de magnesio hexahidratado 0.2 M, TRIS 0.08 M, PEG (Polietilenglicol) 4000 24% p/v y

glicerol 20% v/v), el cual se llevó al Laboratorio Nacional de Estructuras de Macromoléculas (LANEM) en el Instituto de Química de la U.N.A.M. en Ciudad Universitaria. Este cristal difractó a una resolución de 3.37 Å. Sin embargo, debido a la mala calidad del cristal fue imposible realizar una colecta completa de datos de este. Adicionalmente se obtuvo un nuevo cristal en la misma condición de cristalización, el cual fue llevado al APS (*Advanced Photon Source*) en la línea 19 - ID - D, el cual fue difractado a una resolución máxima de 1.8 Å y que dio lugar a una de las estructuras cristalográficas de esta tesis.

Con esta información se realizó una matriz de optimización alrededor de la condición 6 del *kit Crystal Screen Cryo*, en la cual se obtuvieron varios cristales de la proteína M271. De estos cristales obtenidos, dos fueron difractados en el APS en la línea 19 - ID - D. Estos dos cristales difractaron a una resolución máxima de 1.7 Å y 1.9 Å, y dieron lugar a dos nuevas estructuras cristalográficas.

También se obtuvo un cristal de M271 en la condición 12 del *kit Crystal Screen* (Cloruro de magnesio hexahidratado 0.2 M, HEPES sodio 0.1 M, 2-propanol 30% v/v), el cual igualmente fue difractado en el APS, en la línea 19 - ID - D. Este cristal difractó a una resolución máxima de 2.5 Å y, a pesar de la necesidad de usar simetrías no cristalográficas (NCS por sus siglas en inglés) al inicio del afinamiento, dio lugar a la cuarta estructura cristalográfica de esta tesis. (Figura 9)

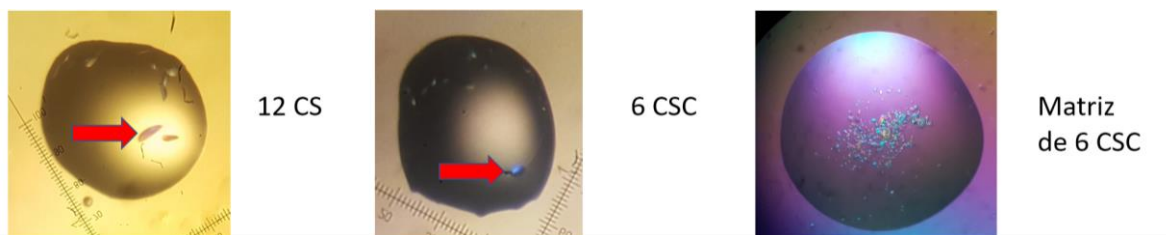


Figura 9: Cristales obtenidos en diferentes condiciones de cristalización y en una matriz de optimización.

De los cuatro paquetes de datos (imágenes de patrones de difracción) obtenidos de los cuatro cristales que difractaron (Figura 10), se logró determinar la estructura de M271 en cada una de ellas. Todos estos patrones de difracción fueron obtenidos con una longitud de onda de 0.9791 Å. Los cuatro cristales presentaron el mismo grupo espacial,  $P2_12_12$  y con valores en la celda unitaria con diferencias despreciables:  $A \approx 62$ ,  $B \approx 76$  y  $C \approx 82$  y  $\alpha = 90$ ,  $\beta = 90$  y  $\gamma = 90$ . Las principales consideraciones al momento de integrar los datos de difracción fueron los valores de:  $CC \frac{1}{2}$ ,  $I/\sigma$  y el *completeness* de los datos. Estas consideraciones fueron:  $CC \frac{1}{2}$  mayor a 80%,  $I/\sigma$  mayor a 2 y un *completeness* mayor al 99%. De las cuatro estructuras obtenidas durante este proyecto, únicamente se reportan dos, debido al hecho de que estas son las estructuras con los mejores datos. En la figura 11 vemos la densidad electrónica de la estructura a 1.7 Å. Tabla 1.

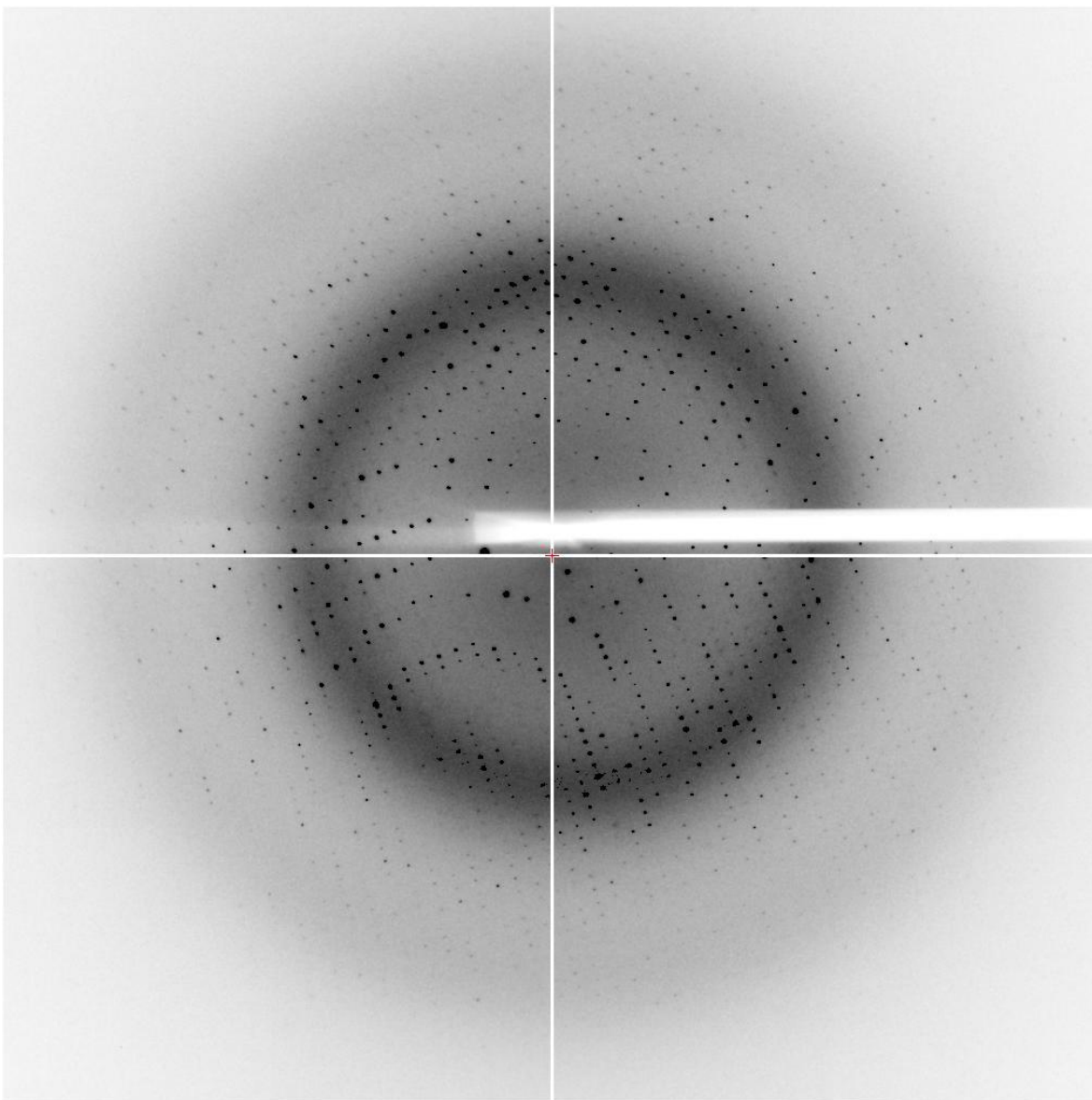


Figura 10: Patrón de difracción de M271.

Tabla 1: Datos estadísticos de las 4 estructuras cristalográficas obtenidas durante este proyecto.

Resolución	1.7	1.8
Longitud de onda (Å)	0.9791	0.9791
Grupo espacial	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2
Celda unitaria:		
a, b, c (Å)	62.39, 76.75, 82.37	62.11, 75.96, 82.73

$\alpha, \beta, \gamma$ (°)	90, 90, 90	90, 90, 90
<i>Mosaicity</i> (°)	0.14	0.32
Resolución (Å)	48.41 – 1.69	29.09 – 1.80
R <sub>merge</sub> (%)	10.5 (75.7)	6.5 (71)
CC ½ (%)	99.9 (85)	99.9 (83.3)
<i>Completeness</i> (%)	99.9 (99.8)	99.9 (99.9)
I/σ (I)	21.58 (3.36)	21.1 (2.66)
Redundancia	12.58	10.79
R-work (%)	17	19
R-free (%)	20	23

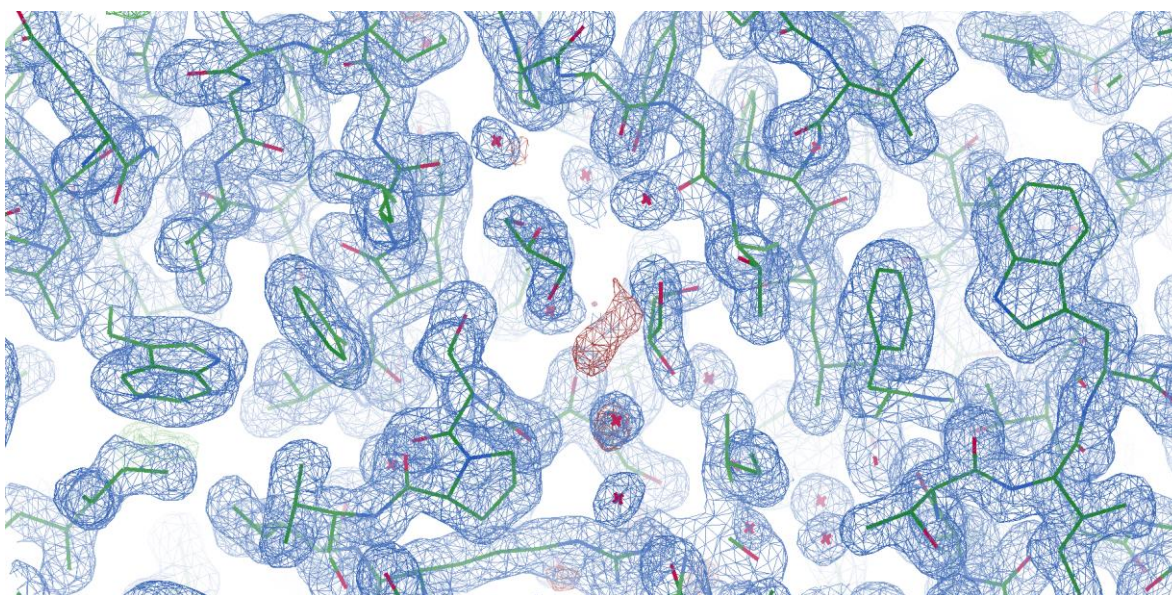


Figura 11: Densidad electrónica de la estructura de M271 con resolución máxima de 1.7 Å

Al momento de analizar las estructuras cristalográficas de M271 se observó que, aunque comparten el mismo grupo espacial y valores de la celda unitaria, las estructuras no son exactamente iguales, RMSD=0.12 Å. La cadena principal (carbonos alfas) de la proteína tiene un comportamiento sin diferencias apreciables.

Por otro lado, las cadenas laterales de los residuos de la proteína presentan mayores diferencias, debido principalmente al ambiente químico que les rodea.

Por ejemplo, la cadena lateral de la fenilalanina 136 de la cadena A de la estructura a resolución máxima de 1.7 Å, difiere de la misma fenilalanina de la cadena A en la estructura a resolución máxima de 1.8 Å. Este cambio en la estructura a 1.7 Å se debe a una manosa cercana a la fenilalanina (Figura 12), las cuales al estabilizarse mutuamente adoptaron una conformación diferente a la estructura a 1.8 Å, en la cual se observa en el ambiente de la fenilalanina una molécula de PEG.

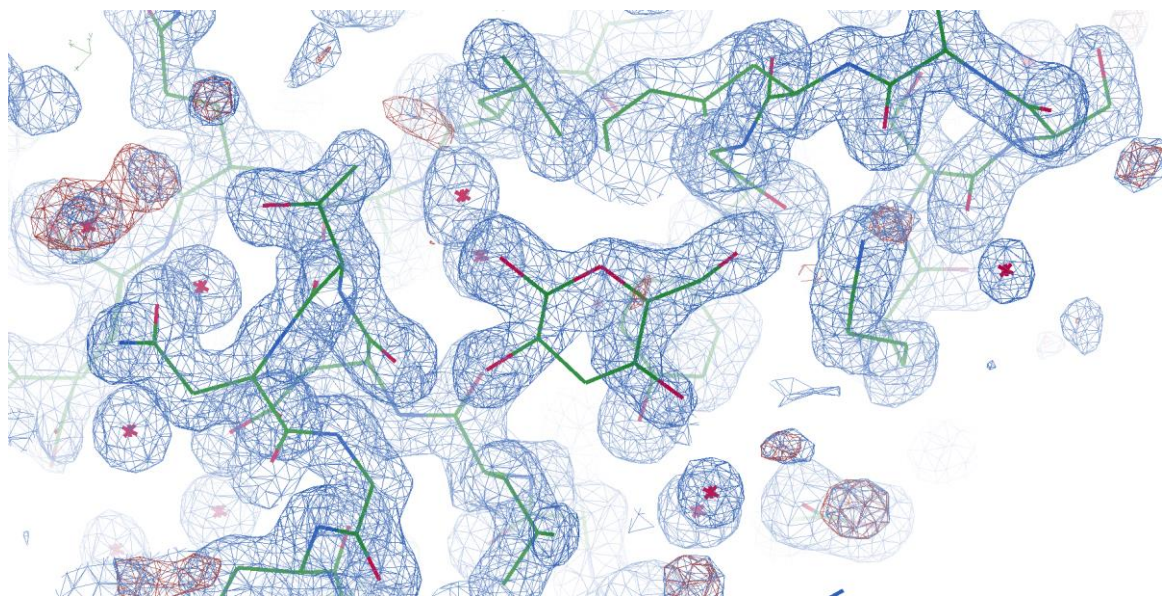


Figura 12: Densidad electrónica de la manosa encontrada en la estructura a resolución máxima de 1.7 Å

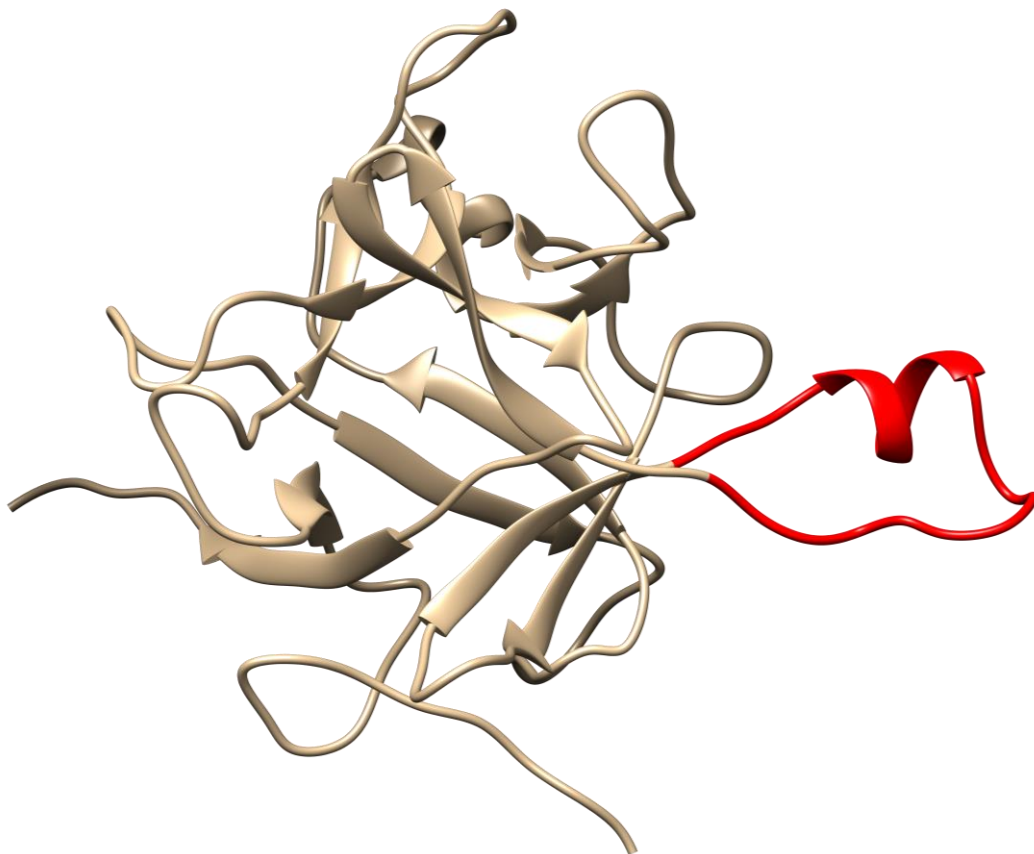
Y podemos encontrar algunos otros ejemplos de comportamientos dispares en la comparación de las estructuras. Recordemos que la estructura a 1.8 Å es la condición original del *kit Crystal Screen Cryo* y las estructuras a 1.7 Å y 1.9 Å son una variación de esta donde hay un cambio de 0.02 M de la sal (cloruro de magnesio hexahidratado) y 2% p/v de PEG 4000. Diferencias mínimas, que obtuvieron cristales diferentes.

También se pudo observar que en la estructura a 1.7 Å fue posible encontrar una n-glicosilación (la manosa previamente mencionada), cuya densidad electrónica no es posible observar en toda la cadena de glicosilación. Únicamente se observa la densidad de una hexosa a una distancia de 19 Å de la asparagina 19, sitio de glicosilación previamente descrito para los inhibidores Kunitz.

En esta estructura proponemos que la n-glicosilación que se observa pertenece a la manosa  $\alpha$ -1,3 (Choi et al., 2003), por la ausencia del oxígeno en enlace con el carbono 3 de la manosa en el mapa de densidad electrónica. Y no se observan perturbaciones en la región del carbono 1. Por lo que proponemos que la

manosa que se observa está enlazada en el carbono 3 y que es la última glicosilación de la cadena. Nosotros suponemos que el resto de la cadena posee un alto grado de flexibilidad, por lo que no es posible definir la densidad electrónica del resto de la cadena de glicosilación. Es importante mencionar que, aunque el modelar esta manosa ayuda a mejorar nuestros valores de R durante el afinamiento, las glicosilaciones no intervienen en la capacidad inhibitoria de los inhibidores de la familia Kunitz-STI. Guerra en el año 2016 cambió la asparagina 19 por un ácido glutámico, escindió la glicosilación del inhibidor, pero al realizar los ensayos de capacidad inhibitoria observó que la misma no cambiaba.

Sin embargo, el cambio más significativo se encuentra en la región del asa L9, donde se aprecia que la densidad electrónica de varios residuos del asa, sobre



*Figura 13:* Modelo de listón de M271 obtenido por cristalografía de rayos X. En rojo se señala el asa L9.



todo después del puente disulfuro entre las cisteínas 151 y 154, no está bien definida

(Figura 14).

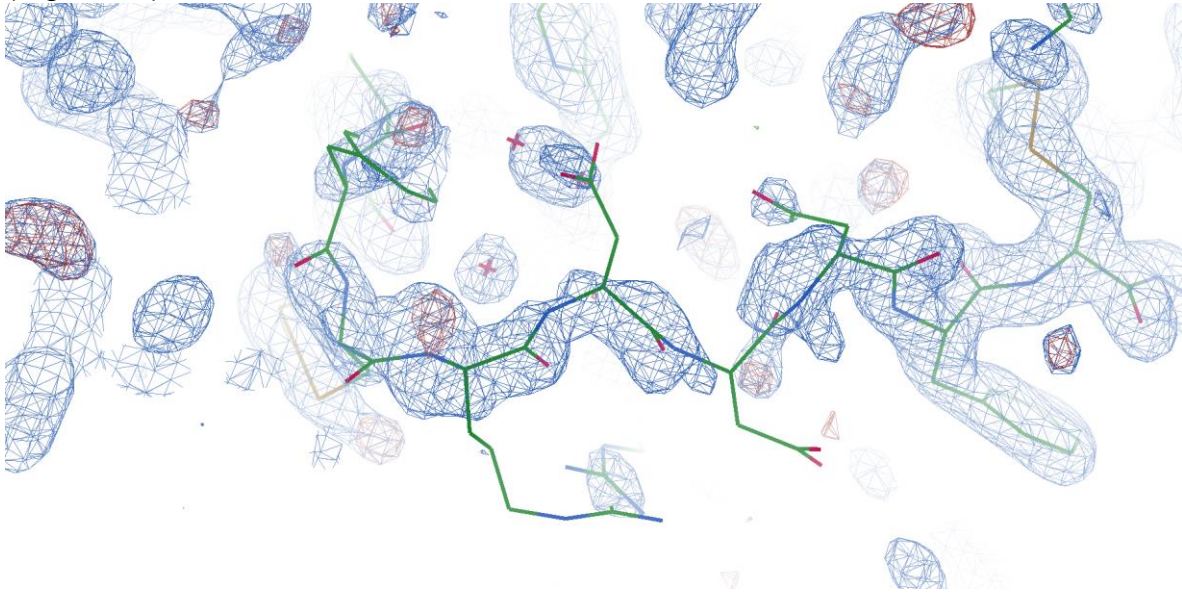


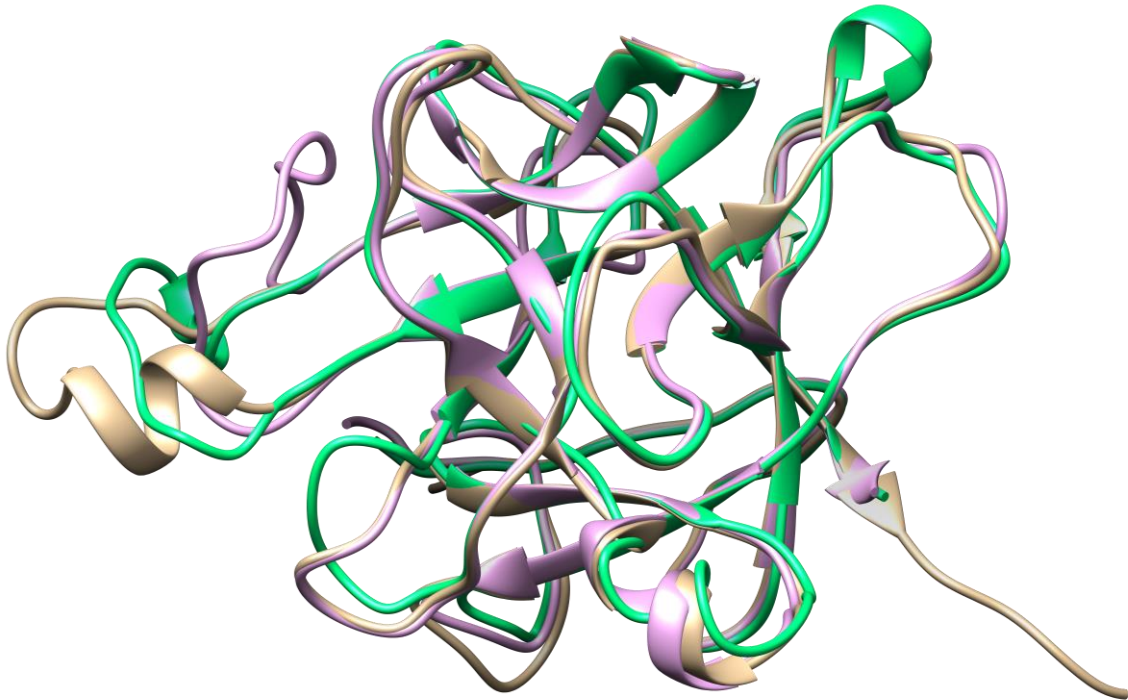
Figura 14: Densidad electrónica del asa L9, se observa que algunos residuos no presentan densidad electrónica en la región de las cadenas laterales.

Sin embargo, la falta de densidad electrónica solo afecta a la región de las cadenas laterales, la densidad electrónica de la cadena principal se aprecia sin problema alguno. Lo cual indica que esa zona posee un alto grado de flexibilidad, y aunque en muchos estudios de cristalografía se buscaría minimizar ese comportamiento, en el caso de nuestro estudio el no observar esa región apoya la teoría de la alta flexibilidad de la zona. (Figura 13).

Igualmente, al comparar la estructura de M271, E3Ad y PDI, es apreciable que la conformación estructural adoptada por la proteína cristalizada es diferente. Esto debido a los contactos que posee cada proteína en su forma cristalizada. Sin embargo, es posible decir que a menor número de contactos cristalinos, la conformación adoptada por la proteína es más cercana a la conformación de la proteína en solución. Por lo que, es importante analizar los contactos presentes en las estructuras de las 3 proteínas, para poder correlacionar los cambios en la conformación estructural del asa L9 con el número de contactos presentes en el cristal.

De la superposición de las 3 estructuras tridimensionales, se observó que el asa L9 posee una conformación diferente en las 3 estructuras. En el PDI se observa que el asa L9 se encuentra contraída hacia la proteína. En E3Ad se observa que la proteína se encuentra extendida hacia el solvente. Y en M271 se observa que el asa también se encuentra contraída hacia el solvente. La diferencia entre M271 y

E3Ad es que la zona del segundo puente disulfuro del asa L9 presenta una inclinación mayor. (Figura 15).



*Figura 15:* Comparación estructural de M271, E3Ad y PDI. A la izquierda se observa la región del asa L9. En dorado M271, en verde E3Ad y en morado PDI.

También, al comparar estas 3 estructuras tridimensionales en la región del asa L9 se observó que esta asa posee una hélice, la cual no está ubicada en la misma región en las 3 estructuras. En M271 esta hélice se encuentra antes del puente disulfuro entre las cisteínas 151 y 154, mientras que en E3Ad se ubica una hélice 3 10 en la región posterior al puente disulfuro, en la estructura de PDI no se logra observar una hélice, sin embargo, se logra apreciar cierta curvatura, la cual debido a la conformación estructural que adopta en el cristal no se logra reconocer como una hélice, en la región posterior al puente disulfuro de esta proteína. (Figura 16).

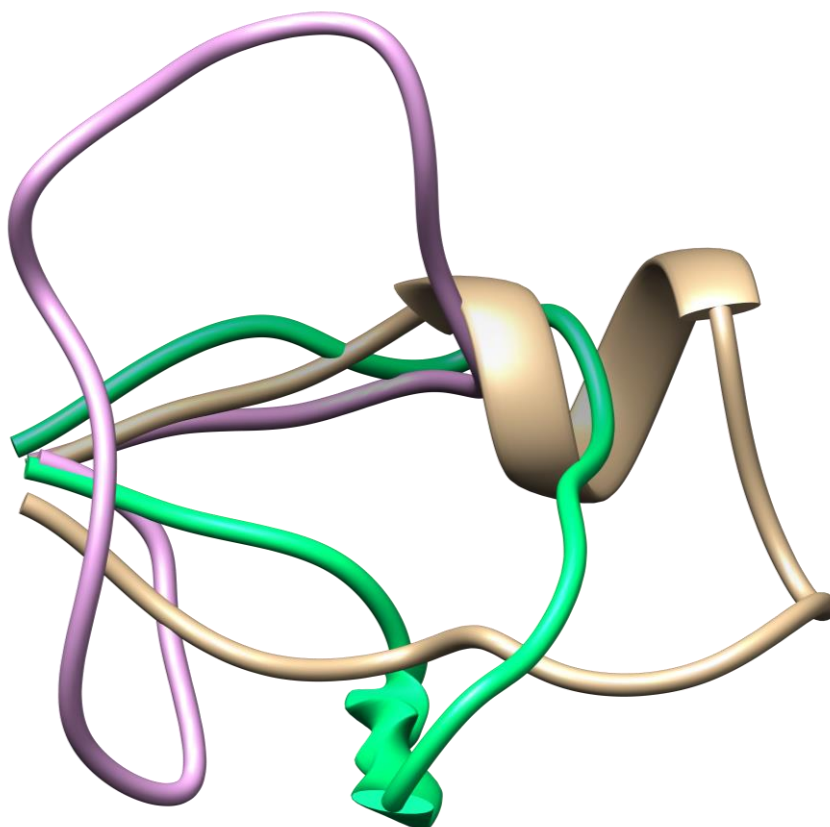


Figura 16: Comparación entre el asa L9 de M271, E3Ad y PDI. En dorado M271, en azul E3Ad y en morado PDI

Con excepción de los cambios observados e indicados en la región del asa L9, los 3 modelos tridimensionales no presentan mayores diferencias entre ellos. Los valores de RMSD de la superposición de las estructuras sin incluir al asa L9, se observan en la Tabla 2.

Tabla 2: Comparación de RMSD entre M271, E3Ad y PDI.

Proteína/RMSD	M271	E3Ad	I.D.C.D.
<b>M271</b>	0	0.69 Å	0.56 Å
<b>E3Ad</b>	0.69 Å	0	0.65 Å
<b>PDI</b>	0.56 Å	0.65 Å	0

#### 7.4 Estudio anisotrópico de la estructura a 1.7 Å de resolución de M271.

Para el estudio de átomos anisotrópicos se seleccionó la estructura con resolución máxima de 1.7 Å. El motivo de la selección de esta estructura se debió a

ser la de mayor resolución y por ende presentar una mayor relación datos – parámetros. Aun así, fue necesario realizar la determinación de TLS (Painter & Merritt, 2006a), debido a que los TLS separan la proteína en zonas que se calcularán en conjunto, disminuyendo los cálculos necesarios para afinar la estructura. El parámetro para determinar el número de zonas en el cálculo de TLS fue, que el número de zonas fuera el menor posible y que el asa L9 pudiera representarse en 3 zonas, los extremos del asa, sin importar que las zonas fueran más largas que el extremo mismo, y que la zona central del asa fuera una zona independiente. De esta forma se determinó, que la cadena A se dividiría en 6 secciones, y la cadena B en 7.

Estas zonas fueron para la cadena A: del residuo 4 al 40, del 41 al 45, del 46 al 146 (zona inicial del asa L9), del 147 al 159 (asa L9), del 160 al 178 (zona final del asa L9) y del 179 al 194. Y para la cadena B: del residuo 2 al 27, del 28 al 46, del 47 al 103, del 104 al 110, del 111 al 151 (zona inicial del asa L9), del 152 al 158 (asa L9) y del 159 al 194 (zona final del asa L9).

Con esto se observó que la estructura cristalográfica de M271 presenta poca flexibilidad en el núcleo de la proteína, y que las asas presentan una mayor flexibilidad. Se observó que el asa L9 es la que presenta mayor flexibilidad de todas las asas, en especial en la cadena B. Esto debido a que en la cadena B el asa L9 tiene un menor número de interacciones con sus vecinos cristalográficos, por lo que tiene un menor número de impedimentos estéricos, por lo tanto la zona presenta una mayor flexibilidad, en comparación con el resto de la proteína.

Este comportamiento es una prueba de que incluso en la matriz cristalina, M271 posee un comportamiento dinámico muy importante, especialmente en el asa L9.

Estos datos fueron analizados y validados con el servidor PARVATI (Merritt, 1999), el cual nos arrojó los siguientes datos. (Figura 17), (Figura 18 y 19)

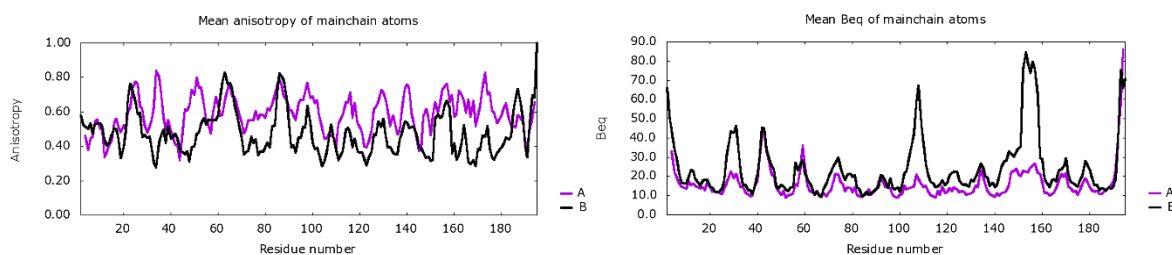


Figura 17: A la izquierda, gráfica de anisotropía media de cada residuo en cada cadena (A y B en este caso), este gráfico nos muestra el comportamiento anisotrópico individual de cada residuo. A la derecha, gráfica del comportamiento anisotrópico en un conjunto de cuerpo rígido, nos muestra secciones delimitadas de acuerdo con su comportamiento anisotrópico.

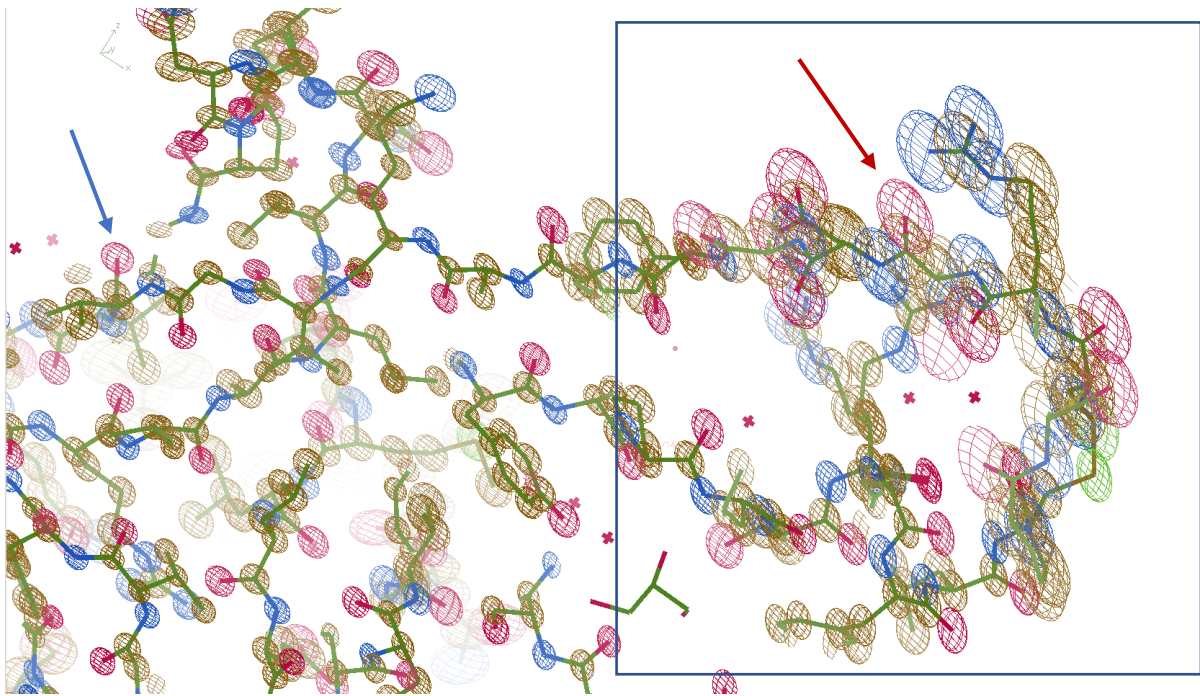


Figura 18: Comportamiento anisotrópico del asa L9, remarcada en un rectángulo, en la cadena B de M271 en el modelo tridimensional a 1.7 Å de resolución. Se observan elipsoides en los átomos que presentan un comportamiento anisotrópico (flecha roja), mientras que en los átomos con un comportamiento isotrópico se observan esferas (flecha azul).

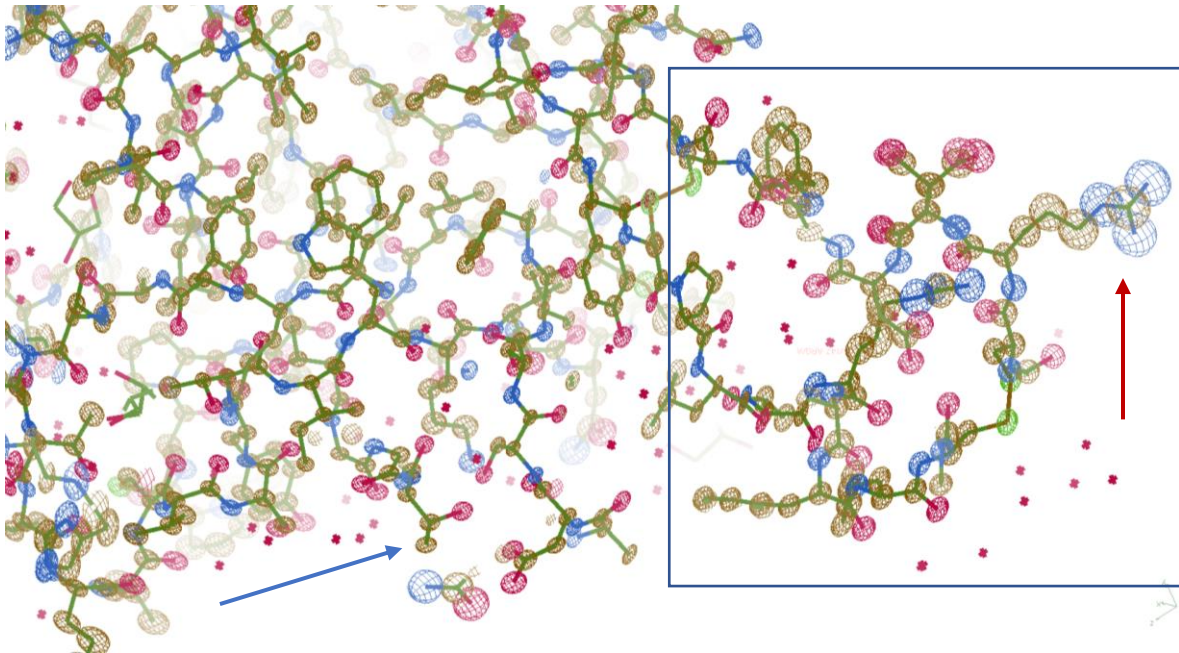


Figura 19: Comportamiento anisotrópico del asa L9 en la cadena A de M271 en el modelo tridimensional a 1.7 Å de resolución. Se observan elipsoides en los átomos que presentan un comportamiento anisotrópico (flecha roja), mientras que en los átomos con un comportamiento isotrópico se observan esferas (flecha azul).

Debido a los resultados de los estudios cristalográficos del inhibidor M271, en los cuales por medio del cálculo de átomos anisotrópicos se observó un comportamiento dinámico en la región del asa L9, es que se decidió explorar por medio de dinámica molecular el comportamiento de M271, y en particular del asa L9. Esto con el objetivo de estudiar el comportamiento de esta asa sin las restricciones de los vecinos cristalográficos.

### 7.5 Simulaciones de Dinámica Molecular

Las simulaciones de dinámica molecular se realizaron en la supercomputadora Miztli de la U.N.A.M. en el proyecto LAN-CAD-UNAM-DGTIC-266.

La simulación de dinámica molecular se llevó a cabo con el software GROMACS y se simularon 5000 ns (nano segundos). En esta simulación, se utilizaron los modelos de las estructuras obtenidas por cristalografía de rayos X de los inhibidores E3Ad, PDI y M271. En la simulación se asignó un estado de protonación equivalente a un pH 7, temperatura de 25° C y 1 atm de presión en las tres estructuras.

El cálculo de RMSD (Root-Mean-Square Deviation) sobre la cadena principal de M271 nos muestra que al inicio de la simulación es cuando se registra la mayor variación en la conformación del asa L9. En la figura 20, se observa que durante los primeros 500 nano segundos (representados en las figuras como pico segundos) la conformación del asa L9 cambia de expuesta al solvente a contraerse hacia el núcleo de la proteína. Y durante el resto de la simulación el asa L9 se mantiene en la conformación contraída.

## RMSD de M271

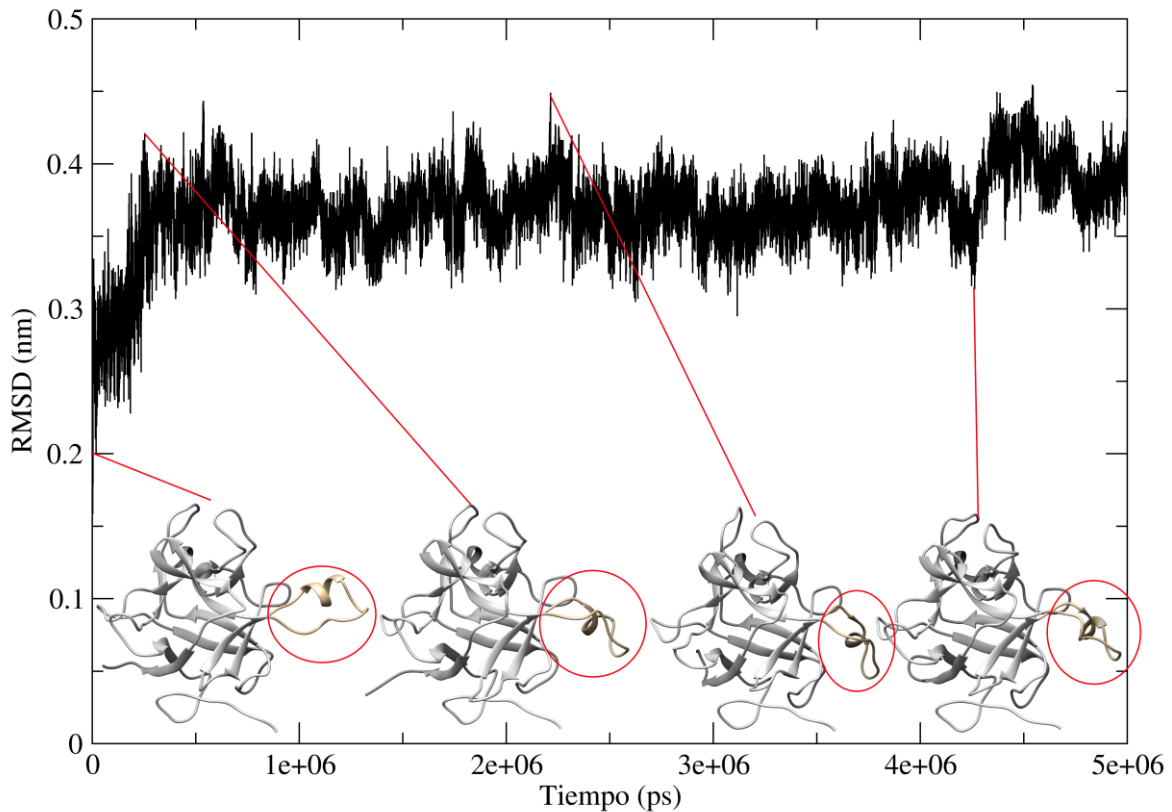


Figura 20: RMSD de la cadena principal de M271. El asa L9 está resaltada en elipses rojas.

En la simulación de PDI, se observa en la figura 21 que esta presenta una mayor variación, presentando varios picos a lo largo de la simulación. Se observa también que el asa L9 visita varias conformaciones durante la simulación. Iniciando contraída hacia el núcleo de la proteína, para extenderse hacia el solvente. Visitando estos estados y sus intermedios durante toda la simulación.

Es importante mencionar que el dinamismo del asa L9 en el inhibidor PDI es la más activa de los 3 inhibidores estudiados. Por lo menos en los estudios de dinámica molecular.

## RMSD de PDI

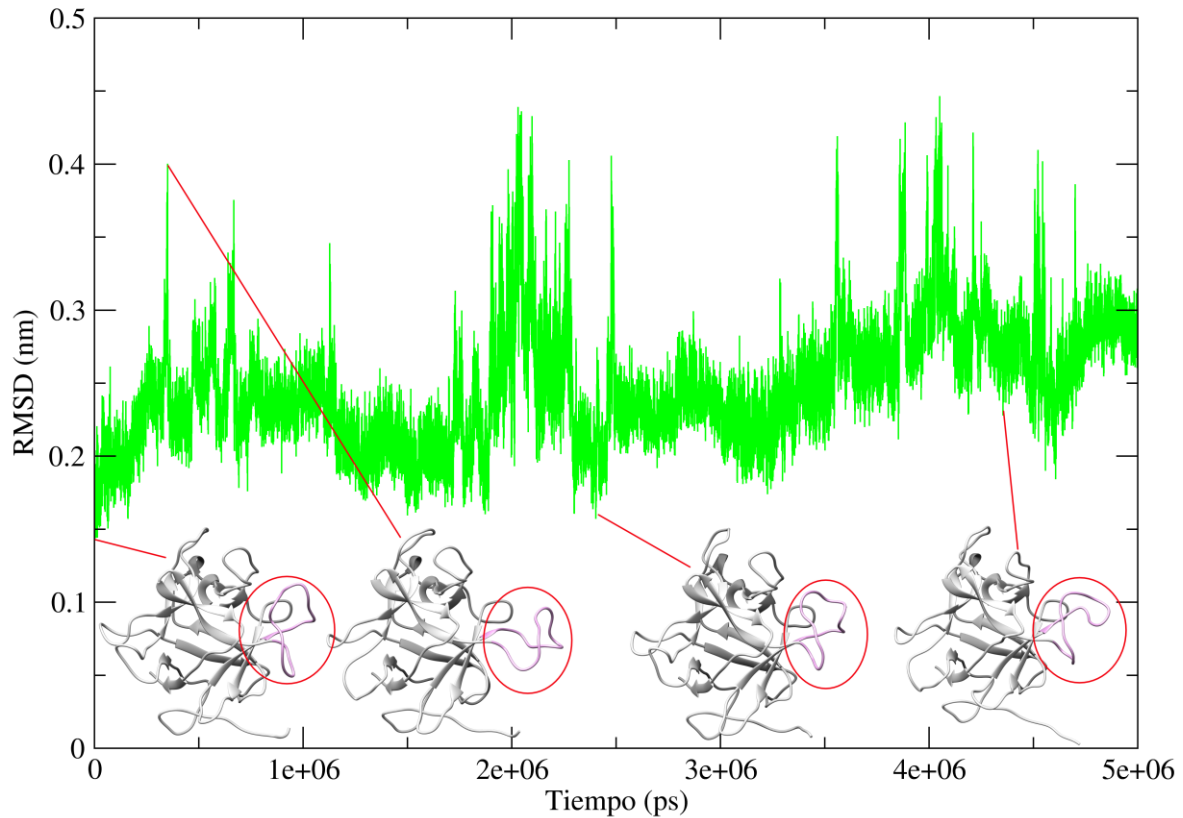


Figura 21: RMSD de la cadena principal de PDI. El asa L9 esta resaltada en elipses rojas.

En la simulación de E3Ad, se observa en la figura 22 que de los 3 inhibidores estudiados este presenta la menor variación. Aún así es posible observar que al igual que M271 y PDI, E3Ad presenta varios picos durante la simulación. Sin embargo, estas variaciones son menores, indicando principalmente cambios en la



región central del asa L9. Aún así no es posible atribuir las variaciones en el cálculo de RMSD de E3Ad únicamente al asa L9.

### RMSD de E3Ad

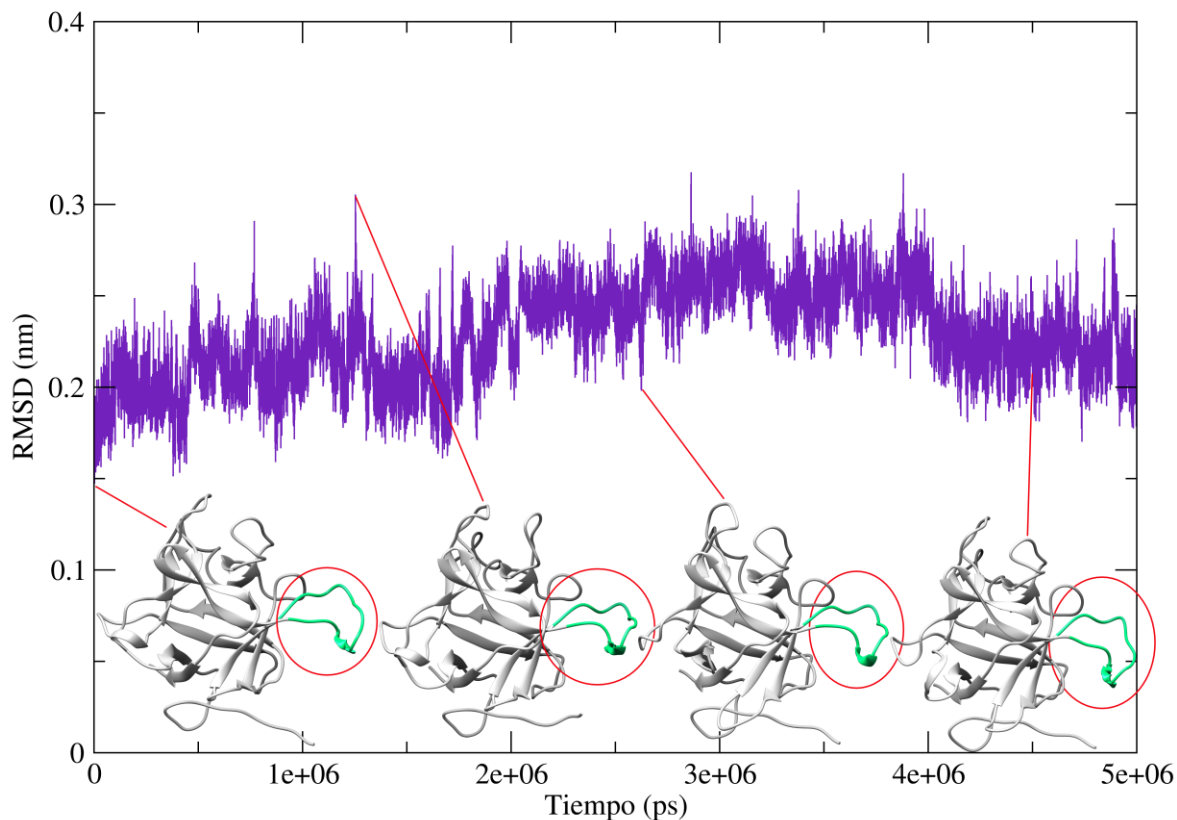


Figura 22: RMSD de la cadena principal de E3Ad. El asa L9 esta resaltada en elipses rojas.

En la figura 23 se muestra la comparación entre el RMSD general de la proteína (eje "Y" izquierda) contra el RMSD del asa L9 (eje "Y" derecha) de M271. Se observa que la variación del RMSD de la proteína es propiciado por la variación

en la conformación del asa L9, sin embargo, no es posible atribuir todos los cambios del RMSD de la proteína únicamente a estas variaciones.

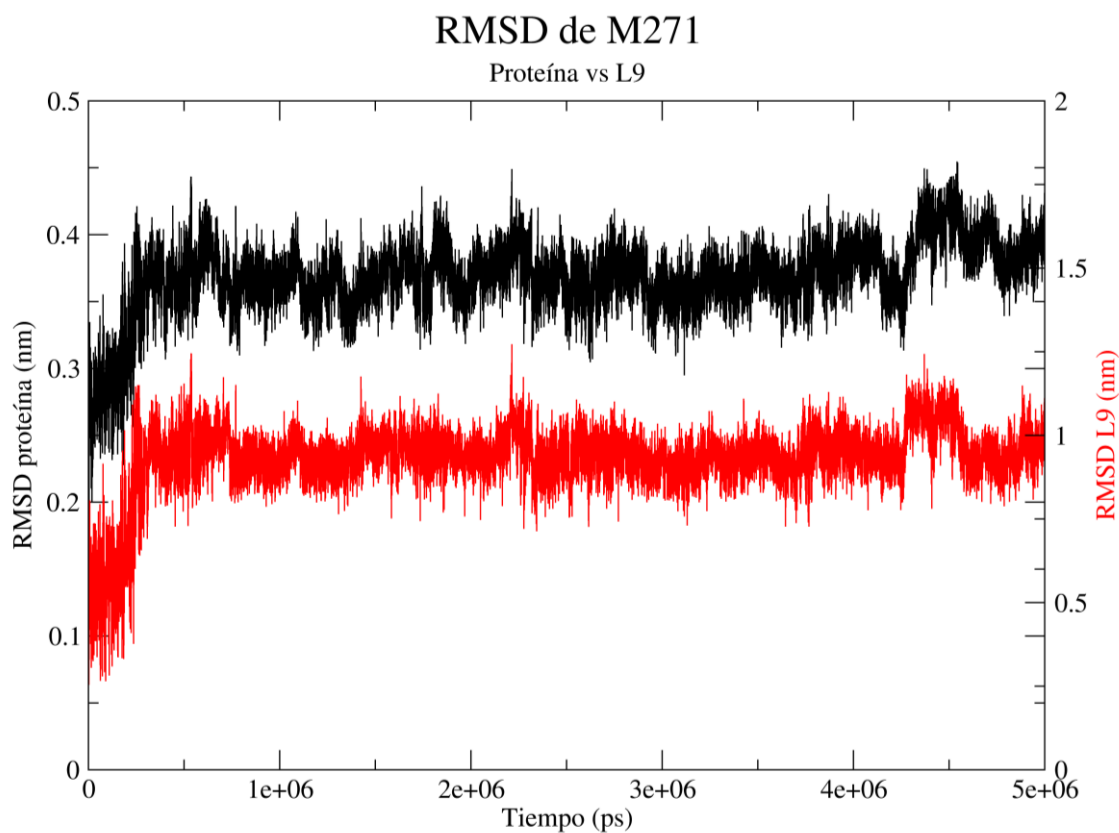


Figura 23: RMSD de M271 y el asa L9 de la misma. En negro M271 y en rojo el asa L9.

En la figura 24 se muestra la comparación entre el RMSD general de la proteína (eje “Y” izquierda) contra el RMSD del asa L9 (eje “Y” derecha) de PDI. Al igual que en la comparación de M271, aquí se muestra que el comportamiento del RMSD general de la proteína está ligado al comportamiento del RMSD del asa L9,

sin embargo, una vez más no podemos atribuir todas estas fluctuaciones únicamente al asa L9.

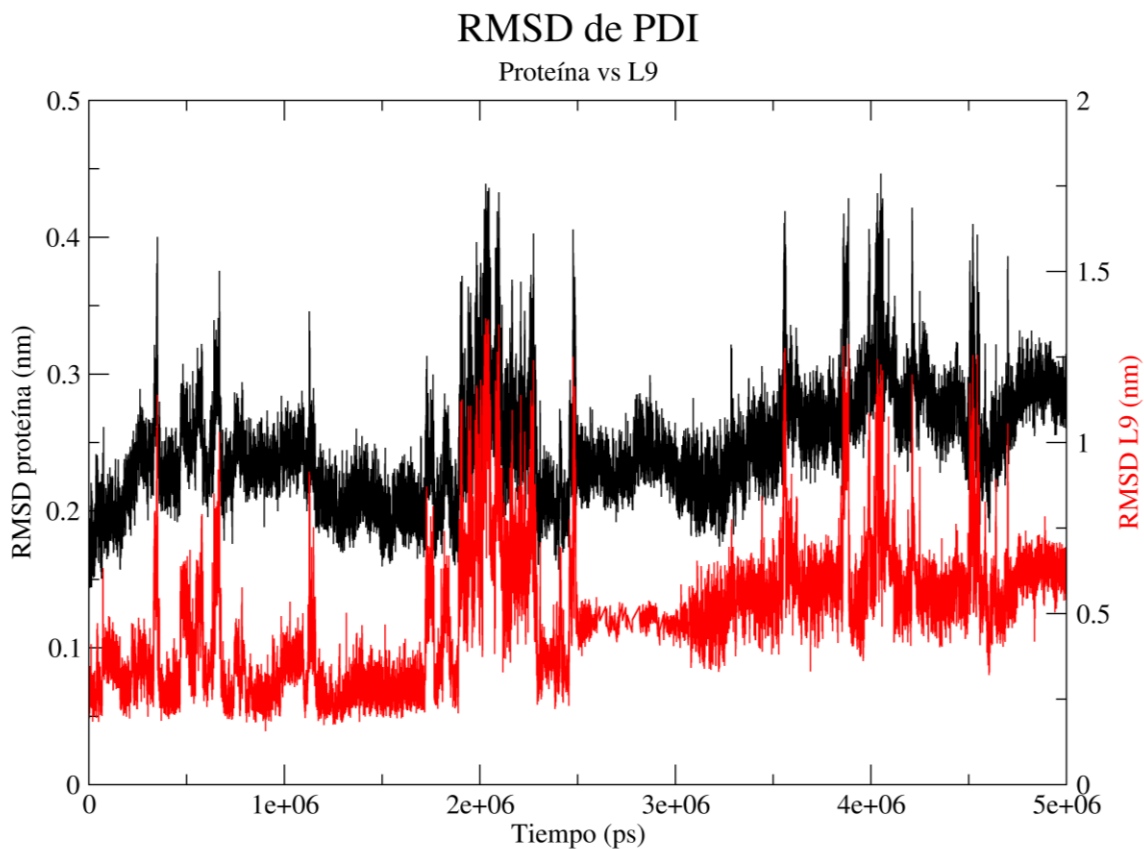


Figura 24: RMSD de PDI y el asa L9 de la misma. En negro PDI y en rojo el asa L9.

En la figura 25 se muestra la comparación entre el RMSD general de la proteína (eje "Y" izquierda) contra el RMSD del asa L9 (eje "Y" derecha) de E3Ad. En esta comparación, a diferencia de las anteriores, el comportamiento del RMSD general de la proteína y del asa L9 no tienen un comportamiento similar. Por lo cual las variaciones que se presentan en E3Ad, no se deben principalmente al comportamiento del asa L9.

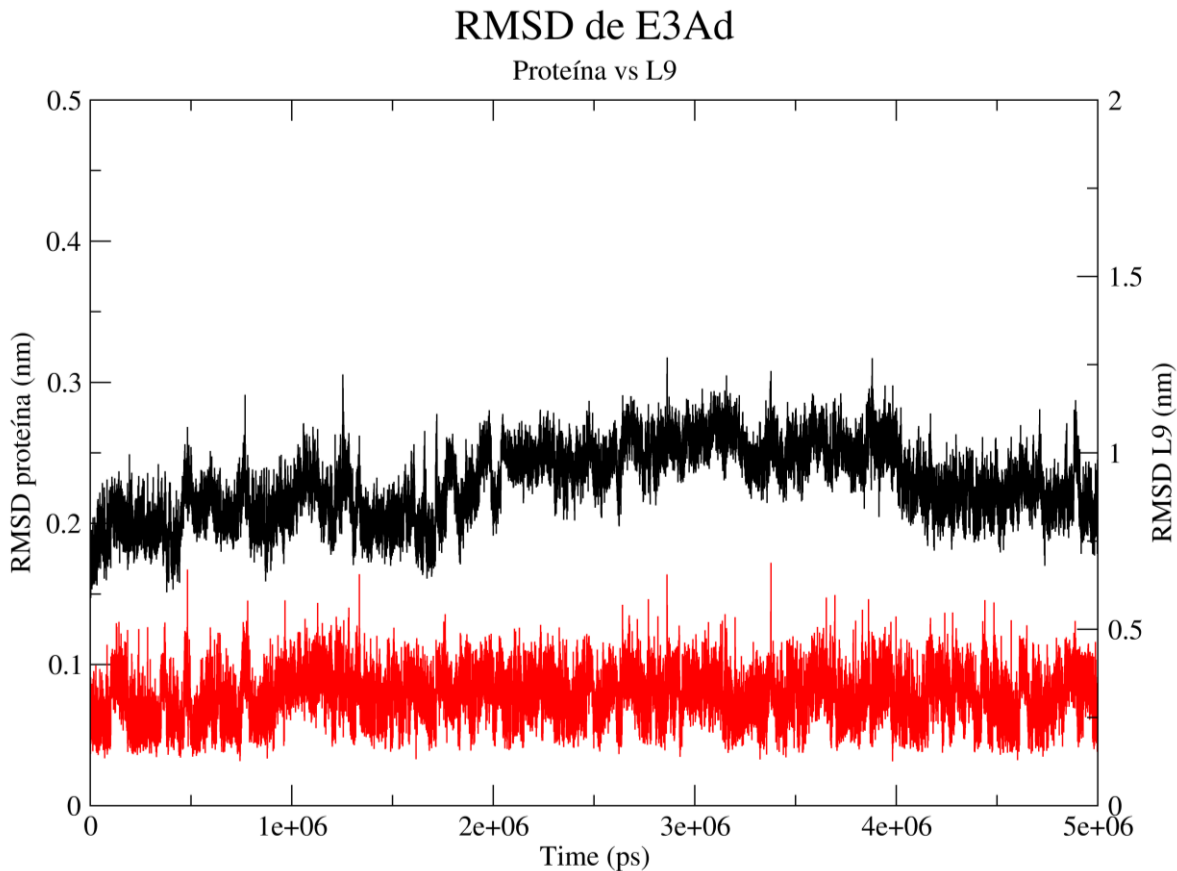
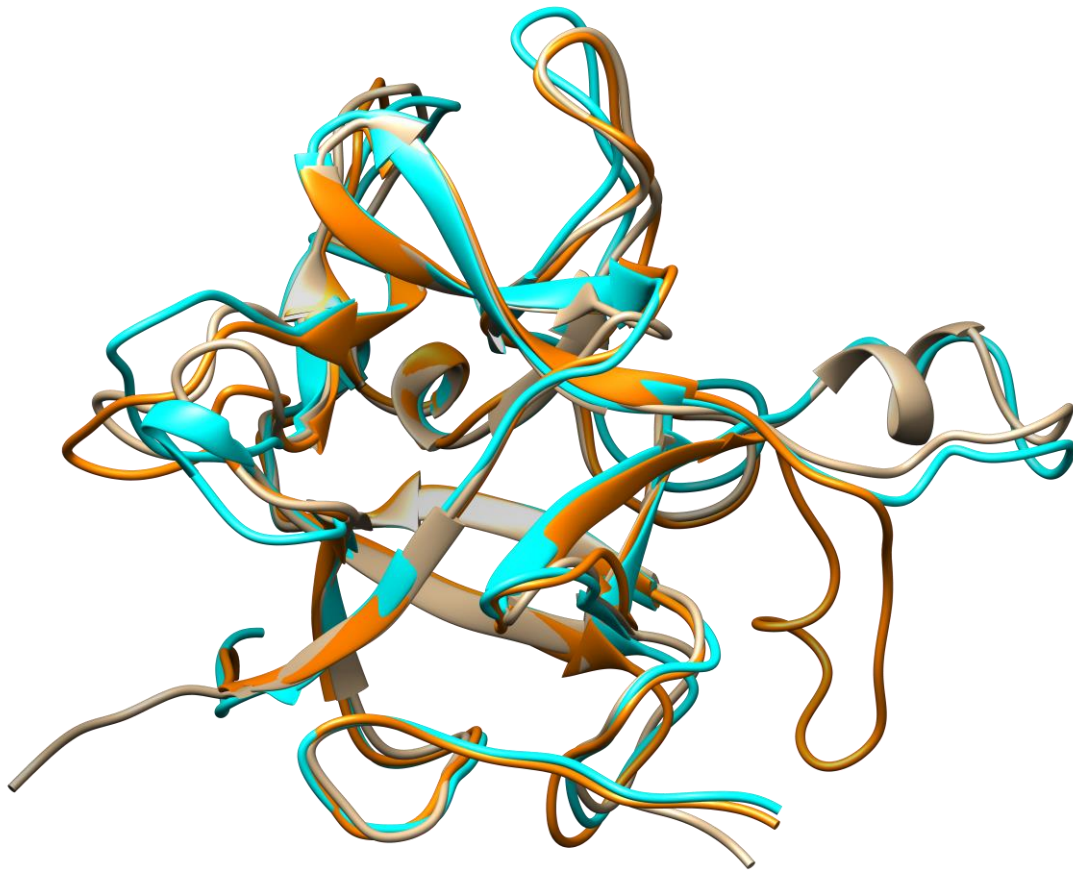


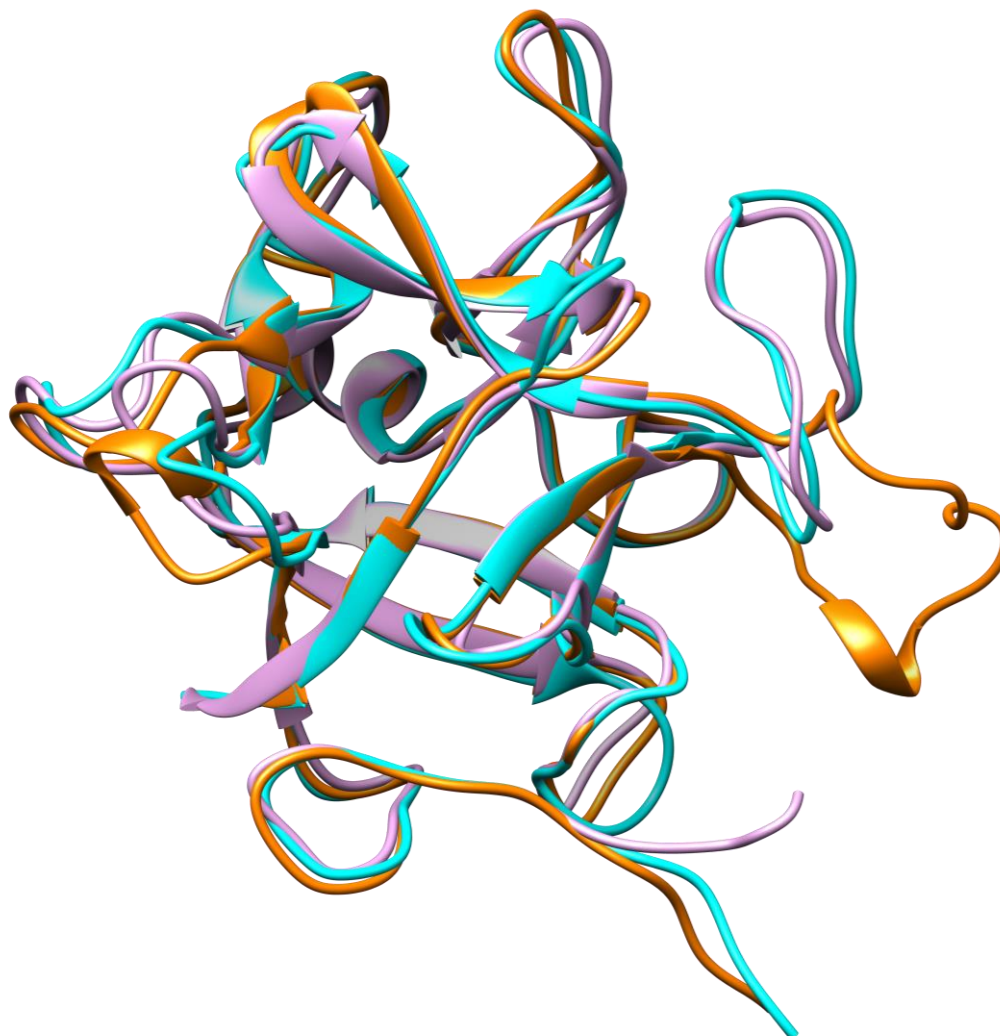
Figura 25: RMSD de E3Ad y el asa L9 de la misma. En negro E3Ad y en rojo el asa L9.

Durante la simulación se observó que M271, tiende a un comportamiento en el cual el asa L9 se contrae hacia el núcleo de la proteína y se mantiene en un equilibrio en el cual visita diferentes conformaciones que podrían llevarle a extender de nuevo el asa hacia el solvente. Esto se puede observar en los datos de RMSD del asa L9 de M271 los cuales son 0.2377 nm y 1.2704 nm para el valor mínimo y máximo respectivamente. (Figura 26)



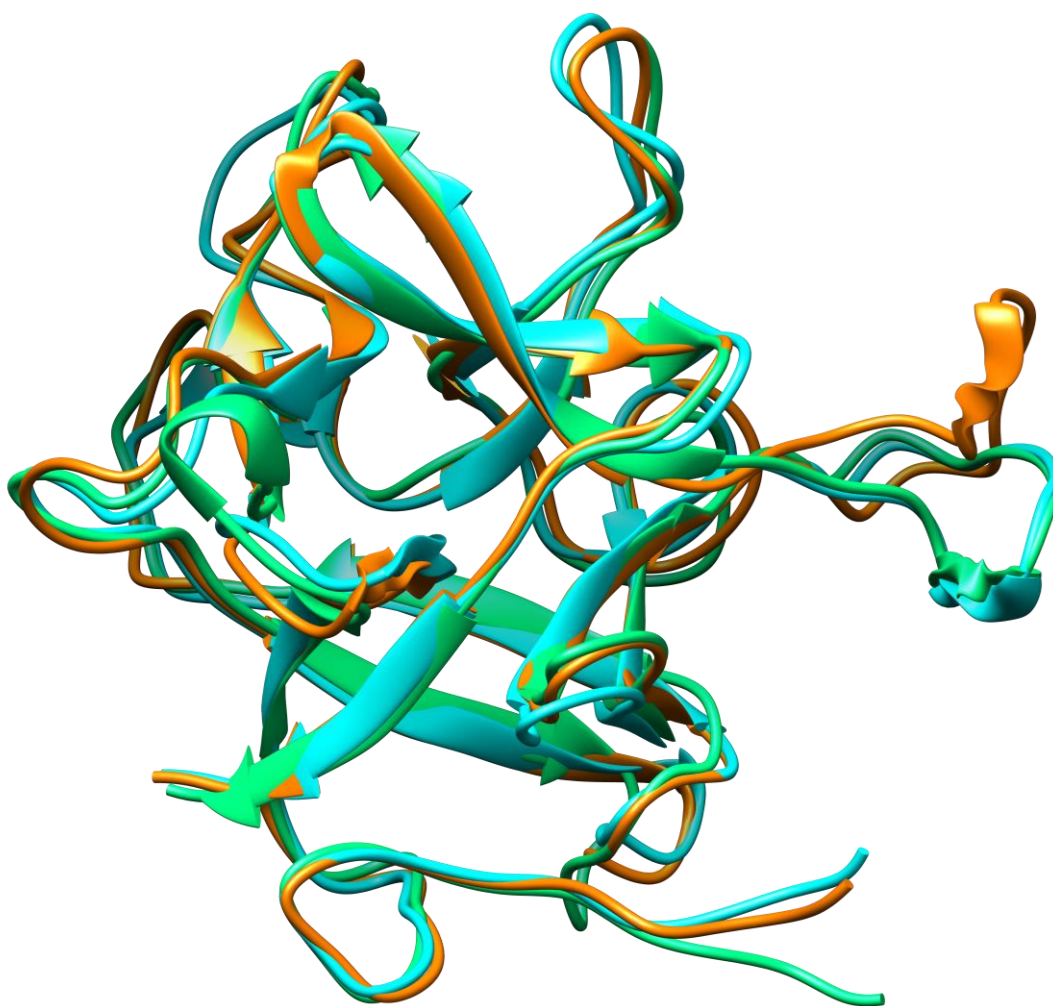
*Figura 26:* Comparación de los valores mínimos y máximos de RMSD de la simulación de dinámica molecular de M271 con la estructural inicial posterior a la minimización de energía. En dorado M271 después de la minimización de energía, en azul valor mínimo de RMSD y en naranja valor máximo de RMSD.

Por otro lado, PDI extiende el asa L9 hacia el solvente y se mantiene en un equilibrio de conformaciones entre el estado extendido y el estado contraído. Los datos de RMSD del asa L9 de PDI son 0.1548 nm y 1.3607 nm para el valor mínimo y máximo respectivamente. (Figura 27)



*Figura 27:* Comparación de los valores mínimos y máximos de RMSD de la simulación de dinámica molecular de PDI con la estructural inicial posterior a la minimización de energía. En morado PDI después de la minimización de energía, en azul valor mínimo de RMSD y en naranja valor máximo de RMSD.

Por último, E3Ad presenta cambios mínimos durante toda la dinámica en comparación con M271 y PDI, presentando un estado en el cual visita diferentes conformaciones alrededor de la conformación inicial. Los datos de RMSD del asa L9 de E3Ad son 0.1212 nm y 0.6866 nm para el valor mínimo y máximo respectivamente. (Figura 28)



*Figura 28:* Comparación de los valores mínimos y máximos de RMSD de la simulación de dinámica molecular de E3Ad con la estructural inicial posterior a la minimización de energía. En verde E3Ad después de la minimización de energía, en azul valor mínimo de RMSD y en naranja valor máximo de RMSD.

Adicionalmente se realizaron cálculos de RMSF (Root-Mean-Square Fluctuation). En la figura 29 se observa el comportamiento de M271. Se pueden observar 11 picos en la gráfica a lo largo de la secuencia de aminoácidos de la estructura (eje x) correspondientes a cada asa del inhibidor, además de 2 picos en los extremos correspondientes al amino inicial en la izquierda del gráfico y el carboxilo terminal en la derecha del gráfico. Se observa que en el residuo 143, correspondiente a una cisteína, el inicio del asa L9 en M271 se presenta un pico hasta el residuo 161, correspondiente a una cisteína, el cual es el final del asa L9. Es importante mencionar que no solamente L9 presenta fluctuaciones durante la simulación, sino que es un comportamiento habitual en las 11 asas que conforman el inhibidor. Además, es importante mencionar que se puede observar un comportamiento de mucha variabilidad en la región del residuo 26 al 36,

correspondiente al asa L1. Al igual que el asa L3, L6 y L8 presenta fluctuaciones semejantes al asa L9.

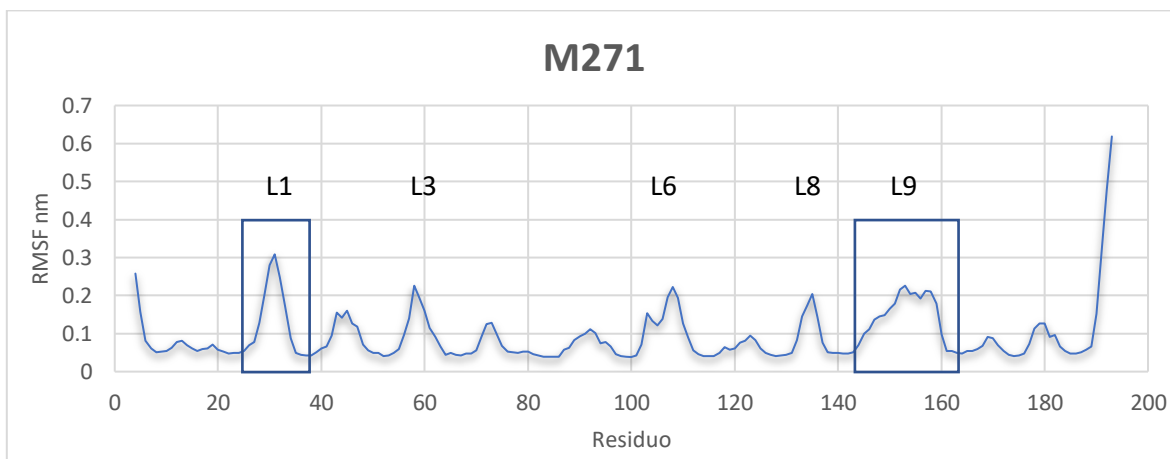


Figura 29: Cálculo de RMSF de M271. Se muestra el número de residuos de la proteína en el eje X y la variación de las fluctuaciones en el eje Y.

En la figura 30 se observa el comportamiento de PDI. En esta gráfica se observan mayores fluctuaciones tanto en el asa L9, la cual inicia en la cisteína 142 y termina en la cisteína 160, como en el asa L1, la cual inicia en el residuo 25 y termina en el residuo 35. Estas fluctuaciones junto con el comportamiento del asa L9 en el cálculo de RMSD, nos muestran una vez más que PDI presenta una mayor flexibilidad en la región del asa L9 y además también en la región del asa L1.

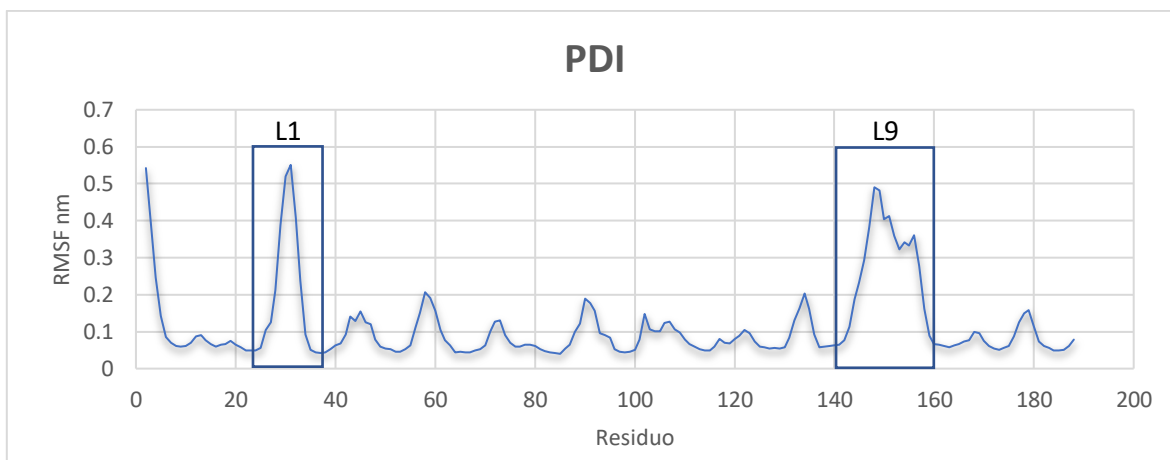


Figura 30: Cálculo de RMSF de PDI. Se muestra el número de residuos de la proteína en el eje X y la variación de las fluctuaciones en el eje Y.

En la figura 31 se observa el comportamiento de E3Ad. En esta gráfica se observa un comportamiento para el asa L9, la cual inicia en la cisteína 142 y termina en la cisteína 159, que reafirma lo observado en el cálculo de RMSD, observando que las fluctuaciones del asa L9 son menores en comparación con PDI, pero



similares en comparación con M271, sin embargo, también se observa que el asa 2, 3 y 4, tienen la misma magnitud en las fluctuaciones. Al igual que con M271 y PDI se observa que el asa L1, la cual inicia en el residuo 25 y termina en el residuo 36, presenta altas fluctuaciones.

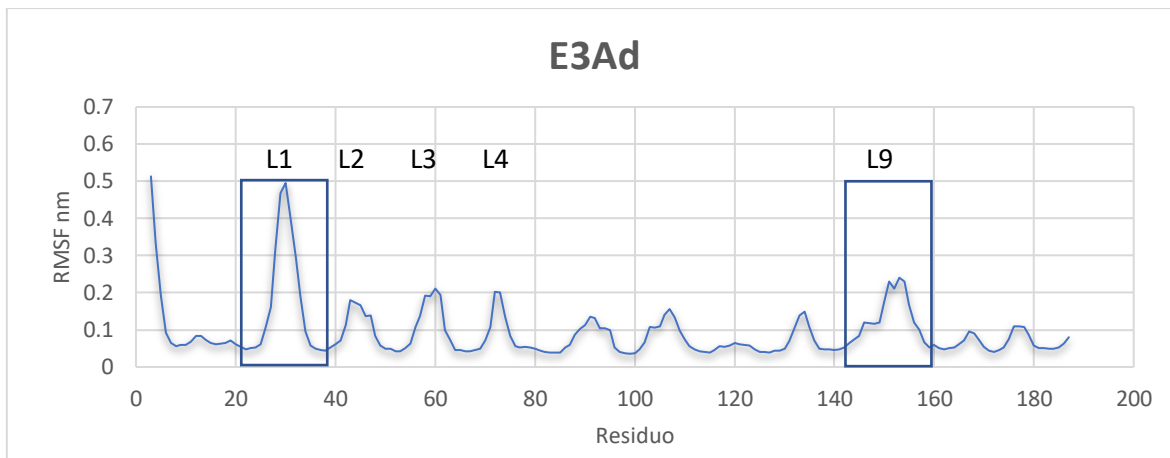


Figura 31: Cálculo de RMSF de E3Ad. Se muestra el número de residuos de la proteína en el eje X y la variación de las fluctuaciones en el eje Y.

Adicionalmente se realizó un análisis de estructura secundaria a los datos obtenidos de la simulación. Se calculó la probabilidad de la proteína a presentar un tipo de estructura secundaria particular durante toda la simulación.

En la figura 32 se observa la probabilidad de estructura secundaria de M271 a lo largo de la simulación de dinámica molecular, en la cual se observa que la región del asa L9 (aumentada, a la derecha de la imagen) presenta principalmente la estructura de asa (representada con una "T"), pero también se observa que en los residuos 148, 149 y 150 y 157, 158 y 159 se presenta una estructura de hélice 3-10 (representada con una "G"), aunque la probabilidad de encontrar la hélice es menor que la del asa, esto nos indica que la hélice está presente, con su subsecuente aporte estructural al asa L9. Además, esto también muestra que, a diferencia del modelo obtenido por cristalografía, donde solo se observa una hélice, durante la simulación el inhibidor puede presentar dos hélices en la región del L9. Igualmente se observan las 12 láminas beta que conforman al inhibidor (representadas con una "E").

# M271

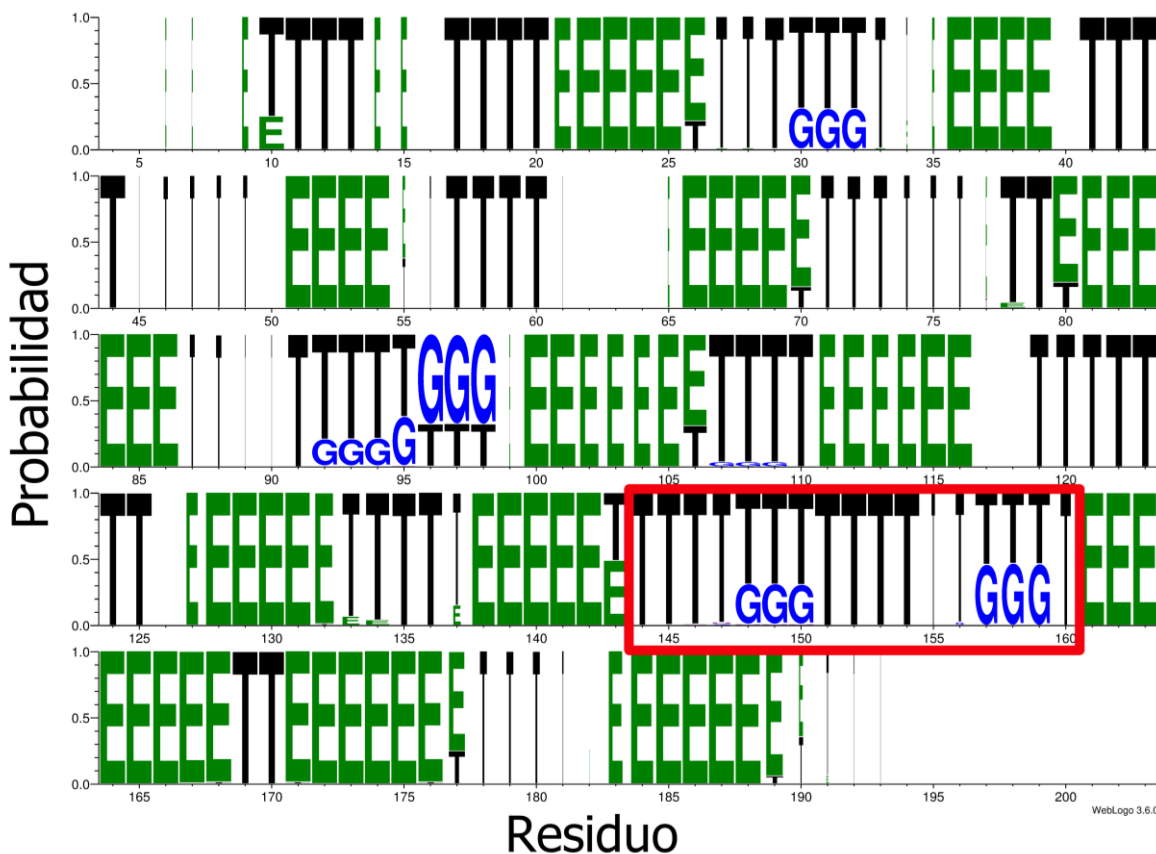


Figura 32: Cálculo de estructura secundaria de M271. En el eje X se muestran los residuos y en el eje Y la probabilidad de cada estructura secundaria. T=asa, E=lámina beta, G=hélice 3-10, H=hélice.

En la figura 33 se observa la probabilidad de estructura secundaria de PDI a lo largo de la simulación de dinámica molecular, en la cual se observa que la región del asa L9 (aumentada, a la derecha de la imagen) presenta principalmente la estructura de asa (representada con una "T"), pero también se observa que en los residuos 147, 148 y 149 se presenta una estructura de hélice 3-10 (representada con una "G"), esto es interesante dado que a diferencia del modelo por cristalografía, donde no se observa ninguna hélice (lo cual se atribuye principalmente al orden de las moléculas dentro del cristal) durante la simulación el inhibidor presenta una hélice 3-10, la cual le otorgaría estabilidad al asa. Igualmente se observan las 12 láminas beta que conforman al inhibidor (representadas con una "E").

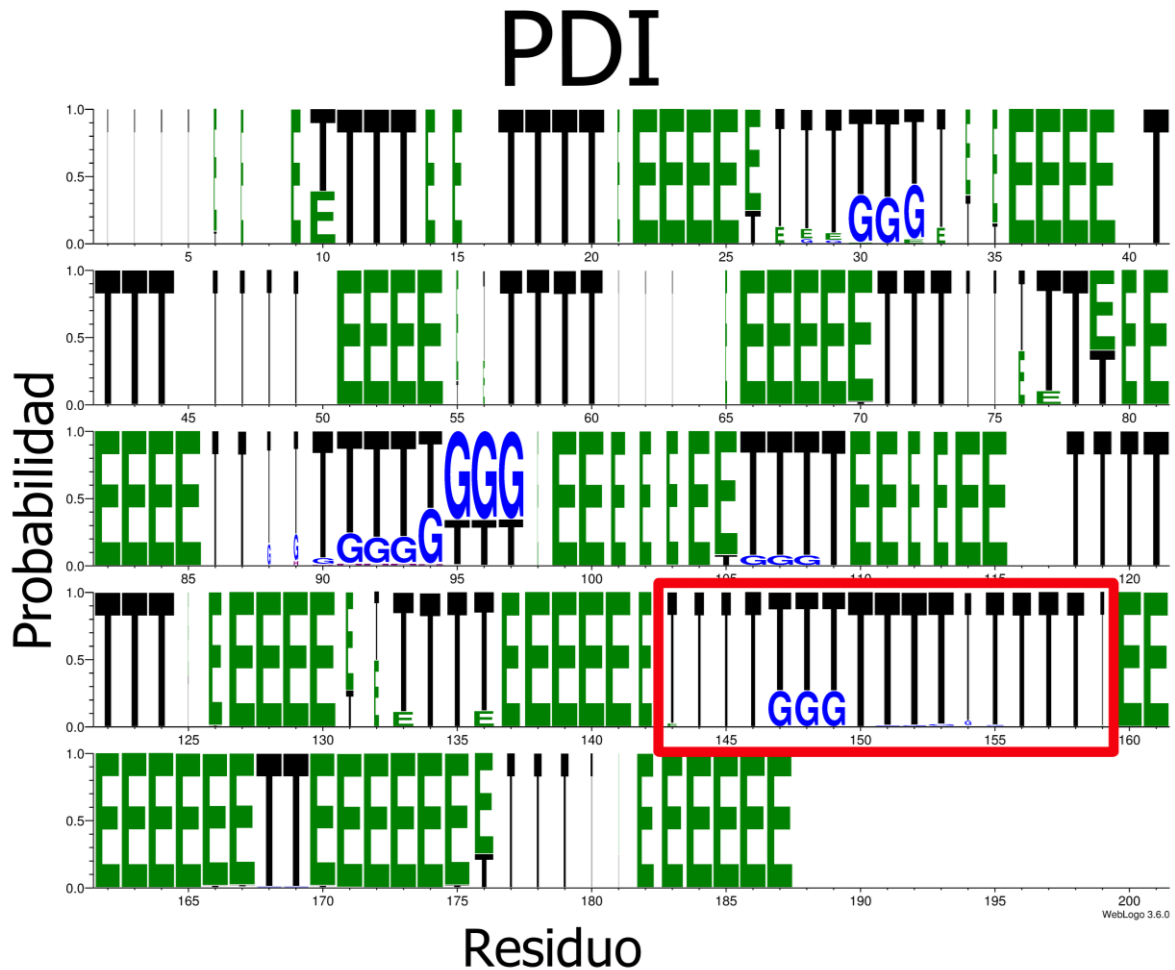


Figura 33: Cálculo de estructura secundaria de PDI. En el eje X se muestran los residuos y en el eje Y la probabilidad de cada estructura secundaria. T=asa, E=lámina beta, G=hélice 3-10, H=hélice.

En la figura 34 se observa la probabilidad de estructura secundaria de E3Ad a lo largo de la simulación de dinámica molecular, en la cual se observa que la región del asa L9 (aumentada, a la derecha de la imagen) presenta la estructura de asa (representada con una “T”), pero también se observa que en los residuos 153, 154 y 155 se presenta mayoritariamente una estructura de hélice 3-10 (representada con una “G”), a diferencia de los otros dos inhibidores que tienen una probabilidad de presentar la hélice 3-10, E3Ad tiene una alta probabilidad de presentarla, lo cual podría estar ligado a la estabilidad mayor del inhibidor durante la simulación y la poca flexibilidad presentada por el asa L9 del mismo. Igualmente se observan las 12 láminas beta que conforman al inhibidor (representadas con una “E”)

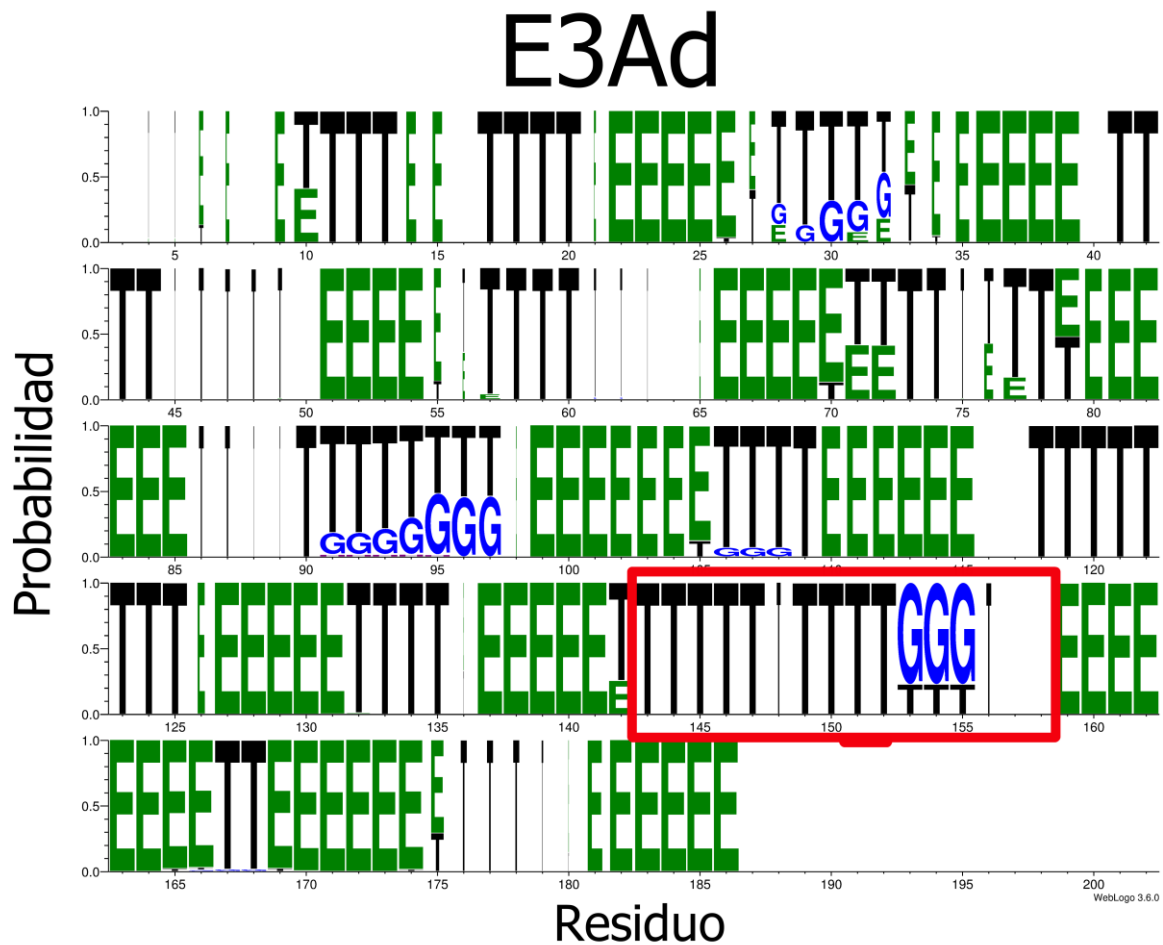


Figura 34: Cálculo de estructura secundaria de E3Ad. En el eje X se muestran los residuos y en el eje Y la probabilidad de cada estructura secundaria. T=asa, E=lámina beta, G=hélice 3-10, H=hélice.

En términos generales se observó que tanto M271 como PDI poseen una flexibilidad mayor, esto también podría estar relacionado con la mayor similitud entre los dos inhibidores. Por otro lado, E3Ad presenta una flexibilidad menor adoptando cambios al inicio de la simulación y permaneciendo en ellos por el resto de esta, siendo el inhibidor con menor identidad entre los 3, aun cuando esta diferencia es de solo el 5% de la secuencia.

## 8 CONCLUSIONES.

- El inhibidor M271 presenta el plegamiento típico de la familia Kunitz-STI.
- La capacidad inhibitoria de M271, es semejante a la de E3Ad.
- El asa L9 es una zona, pero no la única, de alta flexibilidad estructural en M271 y PDI, pero no en E3Ad.
- El comportamiento estructural en la zona del asa L9 de M271 y PDI es semejante, mientras que el comportamiento del asa L9 de E3Ad es diferente

al de M271 y PDI. Este cambio en el comportamiento puede estar ligado a la diferencia en secuencia en el asa L9 entre los 3 inhibidores.

- El asa L1, es la zona que presenta mayores fluctuaciones en los tres inhibidores estudiados.

## 9 PERSPECTIVAS.

1. Obtención del complejo proteasa inhibidor, lo que nos permitiría observar las regiones de interacción del complejo.
2. Realizar un mayor número de estudios estructurales con el inhibidor M271 y Plm II. Estudios de NMR y estudios de cryomicroscopía. Con lo cual estudiar la interacción responsable de la inhibición.
3. Obtención de mutantes de M271. Escindiendo primero la totalidad del asa L9, observando con esto si la inhibición se mantiene o se pierde.
4. Obtención de mutantes de M271. Mutaciones puntuales de diferentes residuos a lo largo del asa L9, buscando responder cuales son los residuos responsables de la formación del complejo proteasa inhibidor.
5. Realizar un mayor número de simulaciones de dinámica molecular. Usando principalmente el método de pHREMD (pH Replica Exchange Molecular Dynamics), en el cual se realizan cambios en el estado de protonación de la proteína. Simulando de esta forma cambios de pH durante la simulación.
6. Realizar simulaciones de dinámica molecular con el método de T-REM (Temperature Replica Exchange Molecular Dynamics). Simulando cambios de temperatura durante la simulación.

## 10 BIBLIOGRAFÍA.

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindah, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1–2, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001>
- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., ... Zwart, P. H. (2010). PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, 66(2), 213–221. <https://doi.org/10.1107/S0907444909052925>
- Bao, R., Zhou, C.-Z., Jiang, C., Lin, S.-X., Chi, C.-W., & Chen, Y. (2009). The Ternary Structure of the Double-headed Arrowhead Protease Inhibitor API-A Complexed with Two Trypsins Reveals a Novel Reactive Site Conformation \*. <https://doi.org/10.1074/jbc.M109.022095>
- Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8), 3684–3690. <https://doi.org/10.1063/1.448118>

- Botella, M. a, Xu, Y., Prabha, T. N., Zhao, Y., Narasimhan, M. L., Wilson, K. a, ... Hasegawa, P. M. (1996). Differential expression of soybean cysteine proteinase inhibitor genes during development and in response to wounding and methyl jasmonate. *Plant Physiology*, 112(1996), 1201–1210. <https://doi.org/10.1104/pp.112.3.1201>
- Cater, S. A., Lees, W. E., Hill, J., Brzin, J., Kay, J., & Phylip, L. H. (2002). Aspartic proteinase inhibitors from tomato and potato are more potent against yeast proteinase A than cathepsin D. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, 1596(1), 76–82. [https://doi.org/10.1016/S0167-4838\(02\)00206-6](https://doi.org/10.1016/S0167-4838(02)00206-6)
- Choi, B.-K., Bobrowicz, P., Davidson, R. C., Hamilton, S. R., Kung, D. H., Li, H., ... Gerngross, T. U. (2003). Editorial Expression of Concern: Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichiapastoris*. *Proceedings of the National Academy of Sciences*, 100(15), 5022–5027. <https://doi.org/10.1073/pnas.1003237107>
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *Journal of Chemical Physics*, 98, 10089–10092. <https://doi.org/10.1063/1.464397>
- Dunaevsky, Y. E., Gladysheva, I. P., Pavlukova, E. B., Beliakova, G. A., Gladyshev, D. P., Papisova, A. I., ... Belozersky, M. A. (1997). The anionic protease inhibitor BWI-1 from buckwheat seeds. Kinetic properties and possible biological role. *Physiologia Plantarum*, 101(3), 483–488. <https://doi.org/10.1034/j.1399-3054.1997.1010305.x>
- Emsley, P., & Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*, 60(12 I), 2126–2132. <https://doi.org/10.1107/S0907444904019158>
- Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallographica Section D: Biological Crystallography*, 62(1), 72–82. <https://doi.org/10.1107/S0907444905036693>
- González-Rábade, N., Badillo-Corona, J. A., Aranda-Barradas, J. S., & Oliver-Salvador, M. del C. (2011). Production of plant protease in vivo and in vitro - A Review. *Biotechnology Advances*, 29(5), 983–996.
- Guerra, Y., Valiente, P. A., Berry, C., & Pons, T. (2012). Predicting functional residues of the *Solanum lycopersicum* aspartic protease inhibitor (SLAPI) by combining sequence and structural analysis with molecular docking. *Journal of Molecular Modeling*, 18(6), 2673–2687. <https://doi.org/10.1007/s00894-011-1290-2>
- Guerra, Y., Valiente, P. A., Pons, T., Berry, C., & Rudiño-piñera, E. (2016). Structures of a bi-functional Kunitz-type STI family inhibitor of serine and aspartic proteases : Could the aspartic protease inhibition have evolved from a canonical serine protease-binding loop ? *Journal of Structural Biology*, 195(2),

259–271. <https://doi.org/10.1016/j.jsb.2016.06.014>

- Guo, J., Erskine, P. T., Coker, A. R., Wood, S. P., & Cooper, J. B. (2015). Structure of a Kunitz-type potato cathepsin D inhibitor. <https://doi.org/10.1016/j.jsb.2015.10.020>
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins*, *65*, 712–725. <https://doi.org/10.1002/prot>
- Huang, Y. J., To, K. Y., Yap, M. N., Chiang, W. J., Suen, D. F., & Chen, S. G. (2001). Cloning and characterization of leaf senescence up-regulated genes in sweet potato. *Physiologia Plantarum*, *113*(3), 384–391. <https://doi.org/10.1034/j.1399-3054.2001.1130312.x>
- Jorgense, W. L., Chandrasekhar, J., & Madura, J. D. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, *79*, 926–935. <https://doi.org/10.1063/1.445869>
- Joshi, B. N., Sainani, M. N., Bastawade, K. B., Gupta, V. S., & Ranjekar, P. K. (1998). Cysteine protease inhibitor from pearl millet: A new class of antifungal protein. *Biochemical and Biophysical Research Communications*, *246*(2), 382–387. <https://doi.org/10.1006/bbrc.1998.8625>
- Kabsch, W. (2010). XDS. *Acta Crystallographica Section D Biological Crystallography*, *66*(2), 125–132. <https://doi.org/10.1107/S0907444909047337>
- Koiwa, H., Bressan, R. A., & Hasegawa, P. M. (1997). Regulation of protease inhibitors and plant defense. *Trends in Plant Science*, *2*(10), 379–384. [https://doi.org/10.1016/S1360-1385\(97\)01107-2](https://doi.org/10.1016/S1360-1385(97)01107-2)
- Koiwa, H., Shade, R. E., Zhu-Salzman, K., Subramanian, L., Murdock, L. L., Nielsen, S., ... Hasegawa, P. M. (1998). Phage display selection can differentiate insecticidal activity of soybean cystatins. *Plant Journal*, *14*(3), 371–379. <https://doi.org/10.1046/j.1365-313X.1998.00119.x>
- Krattiger, A. F. (1996). Insect Resistance in Crops: A Case Study of *Bacillus thuringiensis* (Bt) and its Transfer to Developing Countries. *Director*, (2), 1–42.
- Lawrence, P. K., & Koundal, K. R. (2002). Plant protease inhibitors in control of phytophagous insects. *Electronic Journal of Biotechnology*, *5*(1). <https://doi.org/10.2225/vol5-issue1-fulltext-3>
- Ledoigt, G., Griffaut, B., Debiton, E., Vian, C., Mustel, A., Evray, G., ... Madelmont, J. C. (2006). Analysis of secreted protease inhibitors after water stress in potato tubers. *International Journal of Biological Macromolecules*, *38*(3–5), 268–271. <https://doi.org/10.1016/j.ijbiomac.2006.03.016>
- Leslie, A. G. W., & Powell, H. R. (2007). Processing diffraction data with mosflm, 41–51. [https://doi.org/10.1007/978-1-4020-6316-9\\_4](https://doi.org/10.1007/978-1-4020-6316-9_4)
- Li, H., Robertson, A. D., & Jensen, J. H. (2005). Very fast empirical prediction and

- rationalization of protein pK<sub>a</sub> values. *Proteins: Structure, Function and Genetics*, 61(4), 704–721. <https://doi.org/10.1002/prot.20660>
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*, 40(4), 658–674. <https://doi.org/10.1107/S0021889807021206>
- Merritt, E. A. (1999). Expanding the model : anisotropic displacement parameters in protein structure refinement. *Acta Crystallographica Section D*, 1109–1117. <https://doi.org/10.1107/S0907444999003789>
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., ... Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, 67(4), 355–367. <https://doi.org/10.1107/S0907444911001314>
- Painter, J., & Merritt, E. A. (2006a). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallographica Section D: Biological Crystallography*, 62(4), 439–450. <https://doi.org/10.1107/S0907444906005270>
- Painter, J., & Merritt, E. A. (2006b). TLSMD web server for the generation of multi-group TLS models. *Journal of Applied Crystallography*, 39(1), 109–111. <https://doi.org/10.1107/S0021889805038987>
- Pannetier, C., Giband, M., Couzi, P., Le Tan, V., Mazier, M., Tourneur, J., & Hau, B. (1997). Introduction of new traits into cotton through genetic engineering: Insect resistance as example. *Euphytica*, 96(1), 163–166. <https://doi.org/10.1023/A:1002906308304>
- Park, Y., Bo, H. C., Kwak, J. S., Kang, C. W., Lim, H. T., Cheong, H. S., & Hahm, K. S. (2005). Kunitz-type serine protease inhibitor from potato (*Solanum tuberosum* L. cv. Jopung). *Journal of Agricultural and Food Chemistry*, 53(16), 6491–6496. <https://doi.org/10.1021/jf0505123>
- Parrinello, M., & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12), 7182–7190. <https://doi.org/10.1063/1.328693>
- Pernas, M., Sánchez-Monge, R., & Salcedo, G. (2000). Biotic and abiotic stress can induce cystatin expression in chestnut. *FEBS Letters*, 467(2–3), 206–210. [https://doi.org/10.1016/S0014-5793\(00\)01157-1](https://doi.org/10.1016/S0014-5793(00)01157-1)
- Pouvreau, L., Gruppen, H., Van Koningsveld, G. A., Van Den Broek, L. A. M., & Voragen, A. G. J. (2003). The most abundant protease inhibitor in potato tuber (cv. Elkana) is a serine protease inhibitor from the Kunitz family. *Journal of Agricultural and Food Chemistry*, 51(17), 5001–5005. <https://doi.org/10.1021/jf0212486>
- Rawlings, N. D., & Barrett, A. J. (2013). Introduction: Aspartic and Glutamic Peptidases and Their Clans. *Handbook of Proteolytic Enzymes*, 1, 3–19.



<https://doi.org/10.1016/B978-0-12-382219-2.00001-6>

- Rawlings, N. D., Tolle, D. P., & Barrett, A. J. (2004). Evolutionary families of peptidase inhibitors. *Biochem. J*, 378, 705–716.
- Renko, M., Sabotič, J., & Turk, D. (2012). ??-Trefoil inhibitors - From the work of Kunitz onward. *Biological Chemistry*. <https://doi.org/10.1515/hsz-2012-0159>
- Ritonja, a, Krizaj, I., Mesko, P., Kopitar, M., Lucovnik, P., Strukelj, B., ... Turk, V. (1990). The amino acid sequence of a novel inhibitor of cathepsin D from potato. *FEBS Letters*, 267(1), 13–15.
- Ryan, C. A. (1990). Protease Inhibitors in Plants: Genes for Improving Defenses Against Insects and Pathogens. *Annual Review of Phytopathology*, 28(1), 425–449. <https://doi.org/10.1146/annurev.py.28.090190.002233>
- Schechter, I., & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications*, 27(2), 157–162. [https://doi.org/10.1016/0006-291X\(72\)90076-9](https://doi.org/10.1016/0006-291X(72)90076-9)
- Tang, J., & Wong, R. N. S. (1987). Evolution in the structure and function of aspartic proteases. *Journal of Cellular Biochemistry*, 33(1), 53–63. <https://doi.org/10.1002/jcb.240330106>
- Theron, L. W., & Divol, B. (2014). Microbial aspartic proteases: current and potential applications in industry. *Applied Microbiology and Biotechnology*, 98(21), 8853–8868. <https://doi.org/10.1007/s00253-014-6035-6>
- Valueva, T. A., & Mosolov, V. V. (2004). Role of inhibitors of proteolytic enzymes in plant defense against phytopathogenic microorganisms. *Biochemistry (Moscow)*, 69(11), 1305–1309. <https://doi.org/10.1007/s10541-005-0015-5>
- Valueva, T. A., Revina, T. A., Gvozdeva, E. L., Gerasimova, N. G., Il'inskaia, L. I., & Ozeretskova, O. L. (2001). [Effect of elicitors on accumulation of protease inhibitors in injured potato tubers]. *Prikl Biokhim Mikrobiol*, 37(5), 601–606.
- Van der Vyver, C., Schneidereit, J., Driscoll, S., Turner, J., Kunert, K., & Foyer, C. H. (2003). Oryzacystatin I expression in transformed tobacco produces a conditional growth phenotype and enhances chilling tolerance. *Plant Biotechnology Journal*, 1(2), 101–112. [https://doi.org/PBI010 \[pii\]r10.1046/j.1467-7652.2003.00010.x](https://doi.org/PBI010[pii]r10.1046/j.1467-7652.2003.00010.x)
- Williamson, V. M., & Hussey, R. S. (1996). Nematode Pathogenesis and Resistance in Plants. *The Plant Cell*, 8(10), 1735. <https://doi.org/10.2307/3870226>
- Zhao, Y., Botella, M. A., Subramanian, L., Niu, X., Nielsen, S. S., Bressan, R. A., & Hasegawa, P. M. (1996). Two wound-inducible soybean cysteine proteinase inhibitors have greater insect digestive proteinase inhibitory activities than a constitutive homolog. *Plant Physiology*, 111(4), 1299–1306. <https://doi.org/10.1104/pp.111.4.1299>