



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

**Búsqueda de inhibidores de la ribonucleasa P bacteriana que
interaccionen con su subunidad proteica**

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

M. en C. Ezequiel Alejandro Madrigal Carrillo

Asesor principal: Dr. Alfredo Torres Larios
[Instituto de Fisiología Celular](#)

Comité tutor: Dr. Arturo Rojo Domínguez
[UAM Cuajimalpa](#)

Comité tutor: Dr. Enrique García Hernández
[Instituto de Química](#)

Ciudad Universitaria, CD. MX. Marzo de 2021



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.

Contexto del Trabajo

Este trabajo fue realizado bajo la tutoría del Dr. Alfredo Torres Larios en el laboratorio de Biofísica de Biomacromoléculas, en el laboratorio No. 205, Edificio Sur, el cual forma parte del Departamento de Bioquímica y Biología Estructural de la División de Ciencias Básicas del Instituto de Fisiología Celular en la Universidad Nacional Autónoma de México, Ciudad Universitaria en la Ciudad de México. Exceptuando el uso del sincrotrón para la difracción de cristales de la proteína P (en el Advanced Photon Source del Laboratorio Nacional Argonne, Illinois, parte del Departamento de Energía de USA), así como los experimentos con compuestos líder mediante termoforesis a microescala (llevados a cabo en la empresa Nanotemper en San Francisco, California, USA), los experimentos descritos en el presente estudio fueron llevados a cabo en nuestro laboratorio o en el Laboratorio Nacional de Estructura de Macromoléculas (LANEM) del Instituto de Química de la UNAM.

Agradecimientos

Es imposible dar crédito a la gran cantidad de personas que hicieron posible este trabajo y que me pusieron donde estoy. Uno se debe a la sociedad y de manera aislada no hubiera podido lograr todo este trabajo. Espero no pecar de omisión en mis agradecimientos.

En primer lugar, quiero agradecer especialmente al Dr. Alfredo Torres Larios por haberme aceptado en su equipo de investigación hace ya 8 años desde mis estudios de maestría, en donde aprendí la mayor parte de mis habilidades experimentales, de quien aprendí disciplina y carácter al enfrentar la investigación científica, me enseñó resiliencia y sobre todo le agradezco infinitamente por haberme dado la oportunidad de participar en este proyecto tan único, retador y apasionante.

Agradezco a Nallely Cabrera González y al Dr. Ruy Pérez Montfort por su apoyo en protocolos de purificación, por permitirme hacer uso de su equipo de FPLC, por facilitarme reactivos como proteasa TEV, ADN polimerasas, cepas bacterianas, facilitarme el termociclador cada vez que lo necesitaba y guiarme en mis primeros experimentos de mutagénesis dirigida.

Gracias a todos mis colegas del laboratorio: Lalo Campos, Brian López, Uriel Dimas, Vicky Godínez, Hugo Santamaría, Marilú Torres Martínez y Roberto Bahena por todo su apoyo tanto dentro como fuera del laboratorio, así como por los tantos momentos amenos que pasamos.

Agradezco a la Dra. Annia Rodríguez por sus comentarios objetivos de mis experimentos y su ayuda técnica en experimentos que involucraron el manejo de ARN.

Agradezco bastante a la Dra. Adela Rodríguez Romero del Instituto de Química por su gran apoyo durante la difracción de nuestras muestras de proteína P en el difractor del laboratorio nacional de estructuras de macromoléculas (LANEM) del Instituto de Química de la UNAM, así como a la M. en C. Georgina E. Espinosa Pérez. Asimismo, la Dra. Rodríguez Romero colaboró con nosotros ayudándonos en identificar y confirmar nuestros ligandos principales mediante espectrometría de masas con la cual confirmamos la identidad química del “hit” cristalográfico 2-MBX, el Compuesto C-10 y la pupurina.

Al Laboratorio Nacional del Departamento de Energía de los Estados Unidos de Norteamérica, Argonne National Laboratory, en Argonne, Illinois por permitirnos hacer uso de las instalaciones del sincrotrón APS LS-CAT, donde llevamos a cabo los experimentos de difracción de rayos X de cristales de la proteína P en complejo con moléculas pequeñas.

Al comité de mi examen de candidatura por su disposición, tiempo y atinadas observaciones sobre mi trabajo: Dr. Federico del Río, Dra. Gloria Saab, Dr. Enrique García Hernández y al Dr. Alejandro Sosa.

A los administrativos de la Secretaría Académica del Posgrado en Ciencias Bioquímicas de la UNAM, Lety García, Julio Ignacio Palacios Ordoñez y Adelina González Pérez por toda su ayuda y paciencia durante el proceso de titulación.

Agradezco especialmente al Dr. Marco Igor Valencia por sus invaluable observaciones acerca de mis experimentos, hipótesis, tutorales, tesis y demás pláticas filosóficas sobre innumerables temas.

A la valiosa colaboración del M. en C. Carlos Díaz Tufinio en este proyecto con el desarrollo del ensayo de actividad de la RNasa P por fluorescencia.

Al Dr. Pedro Jiménez y al Dr. Luis Brieba del CINVESTAV Irapuato LANGEBIO por invitarme a colaborar en mi primer artículo internacional relacionado con mis estudios de posgrado.

Al Dr. Chi Hao Luan, Director del High-Throughput Laboratory de la Northwestern University en Chicago, USA, por su apoyo al facilitarnos una muestra de la colección de compuestos Spectrum y por haberme enseñado la técnica del Thermal Shift Assay.

Al Dr. Aaron Hernández Cid, ex compañero de mi laboratorio por apoyarme en aspectos técnicos y numerosos tips al usar el equipo FPLC para purificación de proteínas.

Al Dr. Marcelino Arciniega por su valiosa colaboración en nuestro proyecto y su apoyo durante mi proceso de titulación.

Me encuentro bastante agradecido con los miembros de mi comité tutorial, conformado por el Dr. Enrique García Hernández del Instituto de Química y al Dr. Arturo Rojo Domínguez de la UAM-Cuajimalpa por sus atinadas y profundas aportaciones durante los exámenes tutorales que siempre guiaron de manera segura el desarrollo del proyecto.

Asimismo, quiero agradecer a la Dra. Alejandra Hernández Santoyo del LANEM en el Instituto de Química por su tiempo y sus atinadas observaciones acerca del manuscrito de esta tesis.

Gracias a la que fue por muchos años mi segunda casa: el Instituto de Fisiología Celular.

Gracias a todo el personal de mantenimiento, cómputo, almacén, compras, biblioteca y limpieza del instituto. Sin su vital labor no es posible llevar a cabo la investigación científica.

Especiales gracias al Dr. Daniel Maturana y a la Dra. Dinorha Leyva, científicos de aplicación y de campo de la compañía Nanotemper, por haberme enseñado a usar la instrumentación para experimentos de termoforesis a microescala en San Francisco Bay area, USA.

Agradezco a la Dra. Martha L. Harber, en ese entonces adscrita a la compañía FortéBio, Inc., por haberme enseñado a usar el equipo para experimentos de interferometría de biocapas y cómo analizar e interpretar los resultados provenientes de los sensorgramas.

En especial quiero agradecer la ayuda y el apoyo que siempre me han proporcionado mis padres: el Profesor Ezequiel Alejandro Madrigal Arellano y la Profesora Ma. del Rosario Estela Carrillo García, así como a mi hermana Andrea Madrigal Carrillo, apoyo sin el cual nada de este trabajo hubiera sido posible. Todo mi amor y todo mi agradecimiento infinitamente.

Esta es una especial dedicatoria a mi querido sobrino Andrés Martínez, quien ojalá cuando lea esto le sirva de inspiración para convertirse en científico: que sigas así de curioso, tocando todo, aprendiendo y riendo.

Erica, por tus atinadas observaciones al revisar el manuscrito de esta tesis, así como por tu valiosa e incondicional amistad, gracias.

Aunque su existencia física haya cesado, quiero agradecer por todo su apoyo a mi abuela Ma. del Socorro Arellano quien sigue viviendo en nuestra memoria. A mi tía Esperanza Madrigal Arellano y familia por todo su apoyo a través de todos estos años.

A mis amigos Johan Olguín, Héctor Muñoz, Alan Torres, Walter Buchanan, Mayra Cruz, Iván Díaz, Pedro Hernández, Mireya Gutiérrez e Ilse Rivas. Gracias por su apoyo.

A la Universidad Nacional Autónoma de México, la casa de estudios de la nación por haberme permitido crecer profesionalmente en un espacio académico enriquecedor y laico donde pude alcanzar el máximo nivel de estudios que la sociedad ofrece.

También se reconoce el apoyo de la Fundación Marcos Moshinsky, del Consejo Nacional de Ciencia y Tecnología (CONACYT) proyectos CB-253323, PDCPN2014-47543, INFR2014-225216, así como también al programa: Apoyo a Proyectos de Investigación e Innovación Tecnológico (PAPIIT-UNAM), IN201213 e IN204820. Asimismo, se agradece al CONACYT por la beca de Doctorado con la cuenta No. 417328 dentro del programa “Programa de Maestría y Doctorado en Ciencias Bioquímicas” de la UNAM.

Lista de Contenidos

Abreviaturas y Acrónimos

Resumen

Abstract

1. Introducción

2. Antecedentes

3. Planteamiento del problema

4. Hipótesis

5. Objetivos

6. Materiales y Métodos

7. Resultados

8. Discusión

9. Conclusiones

10. Perspectivas

11. Bibliografía

12. Apéndices

Apéndice A. Materiales suplementarios

Apéndice B. Métodos suplementarios

Apéndice C. Resultados suplementarios

Apéndice D. Recomendaciones para experimentos futuros

Apéndice E. Glosario de términos

13. Anexos

Anexo A. Artículo Nucleic Acids Research 2019 (primer autor)

Anexo B. Artículo Proteins 2018 (co-autor)

Abreviaturas y Acrónimos

Å	Ångström (véase Apéndice E. Glosario de términos)
ADN	Ácido desoxiribonucleico
APS	Ammonium persulfate; Persulfato de amonio
AmOAc	Acetato de amonio
ARN	Ácido ribonucleico
ARN P	Subunidad de ARN de la ribonucleasa P bacteriana
A.u.	Arbitrary Fluorescent Units; Unidades Arbitrarias de Fluorescencia
BHQ1	Apagador de fluorescencia 1; Black Hole Quencher 1
BLI	Biolayer interferometry; Interferometría de Biocapas
BsP	<i>Bacillus subtilis</i> P protein; Proteína P de <i>Bacillus subtilis</i>
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
°C	Grado Celsius, unidad de temperatura
CMBX	5-Cloro-2-Mercaptobenzoxazol
CR	Conserved Region; Región Conservada
C5	Subunidad proteica de la ribonucleasa P de <i>E. coli</i>
C-7	Compuesto optimizado No. 7; 5-DTS; 5-(2,3-Dihidro-2-Tioxo-3-Benzoxazolilmetilamino)-ácido salicílico
C-10	Compuesto optimizado No. 10; 4-DTS; 4-(2,3-Dihidro-2-Tioxo-3-Benzoxazolilmetilamino)-ácido salicílico
Da	Daltones, unidad de masa atómica unificada
DEPC	Diethyl pyrocarbonate; Dietilpirocarbonato
DMSO	Dimethyl sulfoxide; Dimetilsulfóxido
DTT	Ditiotreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EcP	<i>Escherichia coli</i> P protein; Proteína P de <i>Escherichia coli</i>
EtBr	Ethidium bromide; Bromuro de etidio
EtOH	Etanol
FBDD	Fragment-Based Drug Design; Diseño de fármacos basado en fragmentos
FPLC	Fast Protein Liquid Chromatography; Cromatografía líquida rápida de proteínas
FTU	1-4-Fluorofenil-2-tiourea, fragmento molecular, ligando de la proteína P
g	Gramos, unidad de masa
GST	Glutación S-Transferasa
h	Hora, unidad de tiempo
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; ácido 4-(2-hidroxietil)-1-1-piperazinaetanosulfónico
HTS	High Throughput Screening; Búsqueda de compuestos en formato de alta eficiencia
His-tag	Etiqueta de fusión de 6 histidinas
IC₅₀	The half maximal inhibitory concentration; Mitad de la máxima concentración inhibitoria
IPTG	Isopropyl-β-D-1-thiogalactopyranoside; Isopropil-β-D-1-tiogalactopiranosido
Ir6Ac	Iriginol Hexaacetato
K	Grado Kelvin, unidad de temperatura
K₂SO₄	Sulfato de potasio
K_D	Constante de disociación en el equilibrio
kDa	Kilodalton, unidad de masa atómica unificada
K_{off}	Constante de velocidad de disociación
K_{on}	Constante de velocidad de asociación
LB	Medio de cultivo bacteriano Luria-Bertani
M	Molar, unidad de concentración

M	Peso molecular o masa molecular
mM	Milimolar, unidades de concentración
Mh	RNA Minihelix artificial substrate; Sustrato artificial Minihélice de ARN
2-MBX	2-Mercaptobenzoxazol
MD	Molecular Dynamics; Dinámica Molecular
μM	Micromolar, unidad de concentración
mL	Mililitro, unidad de volumen
mol	Moles
min	Minutos, unidades de tiempo
MIC₅₀	Minimum Inhibitory Concentration required to inhibit the growth of 50% of microorganisms; Concentración mínima requerida para inhibir el crecimiento del 50% de los microorganismos
mg	Miligramo, unidad de masa
MgCl₂	Cloruro de Magnesio
MST	Microscale Thermophoresis; Termoforesis a Microescala
NaOAc	Acetato de sodio
NaCl	Cloruro de Sodio
Ni-NTA	Nickel-Nitrilotriacetic acid; Ácido nitrilotriacético-Níquel
nM	Nanomolar, unidad de concentración
nm	Nanómetro, unidad de distancia
NMR	Nuclear magnetic resonance; Resonancia Magnética Nuclear
O.D.	Optical density; Densidad óptica
PanDDA	Pan-Data set Density Analysis; Análisis densidad de todos los grupos de datos
pre-tRNA	Precursor of transfer RNA; Precursor del ARN de transferencia
Proteína P	Subunidad proteica de la ribonucleasa P bacteriana
PyMOL	Molecular graphics system in Python; Sistema de gráficos moleculares en Python
PDB	Protein Data Bank; Banco de Datos de Proteínas
pH	Potencial de hidrógeno
PCR	Polymerase chain reaction; Reacción en cadena de la polimerasa
PEG	Polietilenglicol
RNR	Motivo RNR
RNasa P	Ribonucleasa P
rpm	Revoluciones por minuto
s	Segundo, unidad de tiempo
SA	Streptavidin biosensor; Biosensor de estreptavidina
SaP	<i>Staphylococcus aureus</i> P protein; Proteína P de <i>Staphylococcus aureus</i>
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis; Electroforesis en gel de poliacrilamida-Dodecilsulfato de sodio
SSA	SuperStreptavidin biosensor; Biosensor de Superestreptavidina
T	Temperatura
t	Tiempo
TCEP	Tris(2-carboxyethyl)phosphine; Tris(2-carboxietil)fosfina
TEV	Tobacco etch virus protease; Proteasa del virus del grabado del tabaco
TEMED	Tetrametiletildiamina
<i>T. maritima</i>	<i>Thermotoga maritima</i>
TmP	<i>Thermotoga maritima</i> P protein; Proteína P de <i>Thermotoga maritima</i>
Tris	Trisaminomethane; Trisaminometano
V	Volt, unidad de potencial eléctrico
Wt	Wild-type strain; Cepa tipo silvestre
XRC	X-Ray Crystallography; Cristalografía de Rayos X

Búsqueda de inhibidores de la ribonucleasa P bacteriana que interaccionen con su subunidad proteica

Resumen

La ribonucleasa P (RNasa P) es un complejo ribonucleoproteico esencial que participa en la biosíntesis de proteínas en todos los organismos vivos vía la maduración de los precursores del ARN de transferencia (pre-tRNAs) mediante la escisión del enlace fosfodiéster de su extremo 5'-líder. En este trabajo, consideramos a la subunidad proteica de la RNasa P (proteína P) como blanco promisorio para el descubrimiento de nuevos antibióticos dado el incremento de patógenos resistentes a múltiples fármacos a escala global. Esta consideración, se debe a que existen diferencias de los componentes proteicos en eucaria y archaea con respecto a la RNasa P bacteriana. La subunidad de ARN (ARN P) contiene regiones universalmente conservadas, incluyendo el sitio catalítico, lo cual compromete la especificidad de un inhibidor desarrollado contra esta subunidad. En el presente estudio desarrollamos una plataforma de búsqueda en librerías de compuestos a través del análisis actividad-unión-estructura, para lo cual se montó un ensayo de unión de moléculas pequeñas para los diferentes componentes de la RNasa P mediante interferometría de biocapas (BLI). Se optimizó un ensayo de actividad en tiempo real de alto rendimiento mediante espectroscopía de fluorescencia y se montó una metodología para obtener cristales y resolver la estructura terciaria de la proteína P en complejo con diferentes moléculas pequeñas. A partir de esta última estrategia, se determinaron las primeras tres estructuras cristalográficas de la proteína P en complejo con moléculas pequeñas. Dos de estas estructuras confirmaron la existencia de un sitio en la proteína P para la unión de moléculas pequeñas. Una de estas estructuras se encuentra en complejo con el fragmento 2-mercaptobenzoxazol (2-MBX), el cual pudo ser optimizado para obtener dos compuestos (C-7 y C-10) que muestran una mayor afinidad por la proteína P y actividad antibacteriana. La tercera estructura resuelta corresponde a la proteína P en complejo con el inhibidor purpurina, que actúa mediante la unión a la subunidad proteica. Este inhibidor se detectó a través del ensayo de actividad de la RNasa P mediante fluorescencia. Posteriormente, se recurrió a la técnica de BLI y a ensayos de actividad adicionales para caracterizar el mecanismo de inhibición, a partir de lo cual se infirió que la purpurina se une a la proteína P y a la holoenzima, pero no a la subunidad de ARN. La estructura cristalográfica de la proteína P en complejo con la purpurina, confirmó que el sitio de unión del inhibidor es importante para la función biológica de la holoenzima.

Palabras clave: diseño de fármacos basado en fragmentos, ensayo de actividad por fluorescencia, interferometría de biocapas, ribonucleasa P, antibióticos, cristalografía de rayos X, búsqueda de inhibidores.

Bacterial ribonuclease P inhibitor screening through binding to its protein subunit

Abstract

Ribonuclease P (RNase P) is an essential ribonucleoprotein complex involved in the biosynthesis of proteins in all organisms *via* the maturation of all transfer RNA precursors (pre-tRNAs) through the phosphodiester bond cleavage of its 5'-leader sequence. In this work, we consider the RNase P protein subunit (P protein) as a promising target for the discovery of antibiotics due to the emergence of multiple resistant-drugs pathogens at a global scale. This consideration is derived from the fact that there are key differences between the RNase P protein subunits from eukarya and archaea compared to that of bacterial RNase P. The RNA subunit (P RNA) contains universally conserved regions, including the catalytic site, compromising the specificity of any inhibitor developed against this subunit. Here, we developed an activity-binding-structure platform for compound library screening, for which an assay was set to search for the binding of small molecules to the RNase P subunits and substrates by biolayer interferometry (BLI); a high-throughput activity assay by fluorescence spectroscopy was optimized for detection of RNase P inhibitors and an experimental methodology was set for growing crystals and solving crystallographic structures of the P protein in complex with small molecules. From the latter strategy, it was possible to obtain the first three crystal structures of the P protein in complex with small molecules. Two of these structures showed that a binding site for small molecules on the P protein exists. One of these structures is in complex with the fragment 2-mercaptobenzoazole (2-MBX), which was optimized to compounds C-7 and C-10, which displayed higher affinity for the P protein and antibacterial activity. The third crystal structure solved is that of the P protein in complex with the inhibitor purpurin, which acts through binding to the P protein subunit. This inhibitor was detected through the RNase P activity assay by fluorescence spectroscopy. We applied the BLI technique and further activity assays to characterize the inhibition mechanism, and from this approach it was inferred that purpurin binds to the P protein and to the holoenzyme, but not to P RNA subunit. The crystallographic structure of the P protein in complex with purpurin confirmed that the binding site of the inhibitor is important for the biological function of the holoenzyme.

Keywords: fragment-based drug design, fluorescence activity assay, biolayer interferometry, ribonuclease P, antibiotics, X-ray crystallography, inhibitor screening.

1. Introducción

Han pasado 40 años desde que las últimas clases de entidades químicas usadas como antibióticos fueron descubiertas: en 1962 las fluoroquinolonas y 40 años después la oxazolidinona (linezolid) en 2000 (1, 2, 3). Desde entonces, ninguna clase de antibiótico clínicamente útil y comercial ha sido desarrollado; aunque en enero de 2015, la última novedosa entidad química antibacteriana, la teixobactina, desarrollada por NovoBiotic Pharmaceuticals, fue reportada. Sin embargo, aún no se ha producido industrialmente y no ha sido aprobada por la FDA (4). Estos pasados 40 años en los que no se han descubierto nuevas clases de antibióticos marcan la necesidad de nuevas clases de antibióticos para combatir bacterias resistentes a múltiples fármacos que están emergiendo a nivel global (5, 6). Otra preocupación actual es que varias grandes compañías farmacéuticas están abandonando las áreas de terapias antibacterianas debido al tiempo prolongado de investigación, altos costos en inversión y bajo éxito en los proyectos (2, 7). Expertos del Wellcome Trust Institute advierten que, si no se ponen en marcha esfuerzos globales inmediatamente, se estima que se estarán registrando anualmente casi 10 millones de muertes para el año 2050 debido a enfermedades relacionadas con resistencia a antibióticos (8). Este fenómeno creciente de resistencia bacteriana a múltiples fármacos hace un llamado tanto a la comunidad científica como a los gobiernos para tomar acciones globales coordinadas.

El ribosoma bacteriano, un complejo ARN-proteína, es uno de los modelos para el descubrimiento de antibióticos, ya que es el blanco para la mitad de ellos (15). No obstante, muchas mutaciones que confieren resistencia a diversos antibióticos han sido reportadas durante la última década (16-20). Por lo tanto, es deseable estudiar como blancos terapéuticos a otros actores relacionados con el ribosoma y esenciales para la vida, como lo es la ribonucleasa P bacteriana (RNasa P). La RNasa P es una de las dos ribozimas naturales conocidas de recambio múltiple (11, 21), siendo la otra el ribosoma mismo (22). La RNasa P está presente y conservada en todos los organismos vivos. Exceptuando a la archaea *Nanoarchaeum equitans*, que es un simbiote obligado de la archaea *Ignicoccus*, todos los demás organismos requieren de la RNasa P (23). La RNasa P está involucrada en el dogma central de la biología molecular, ya que juega un papel crucial en la biosíntesis de proteínas en todos los organismos vivos (24). Durante el proceso de biogénesis de los ARN de transferencia, cuando éstos son sintetizados, se encuentran formando parte de transcritos primarios más largos (25), por lo que antes de poder ser aminoacilados, deben ser editados para generar un tRNA maduro aminoacilable (26).

La RNasa P es una endonucleasa cuya función principal es hidrolizar el enlace fosfodiéster de los precursores del tRNA, generando un ARN de transferencia maduro y un extremo 5' fosforilado (29). Su catálisis, una hidrólisis que es absolutamente dependiente de cationes de metales divalentes (30), es llevada a cabo por la subunidad de ARN, lo que la hace una verdadera ribozima (11). *In vitro*, la RNasa P es capaz de llevar a cabo catálisis en ausencia de la subunidad proteica (31), pero en un ambiente de alta fuerza iónica (11).

La RNasa P bacteriana es un complejo ribonucleoproteico

El descubrimiento de que el componente de ARN de la RNasa P era el agente catalítico fue publicado en 1983 (11). Sin embargo, no fue sino hasta casi 30 años después, en el año 2010, que la estructura cristalográfica a 3.8 Å de resolución de la RNasa P bacteriana en complejo con el tRNA y el producto 5'-líder fue publicada (35). Esta información permitió tener conocimiento, antes sólo inferido por estudios de bioquímica clásica, de las interacciones entre los componentes de la holoenzima. En bacteria, la RNasa P es un complejo ribonucleoproteico (Figura 1) formado por una subunidad grande de ARN, denominada "ARN P", así como por una subunidad pequeña de proteína, llamada proteína P.

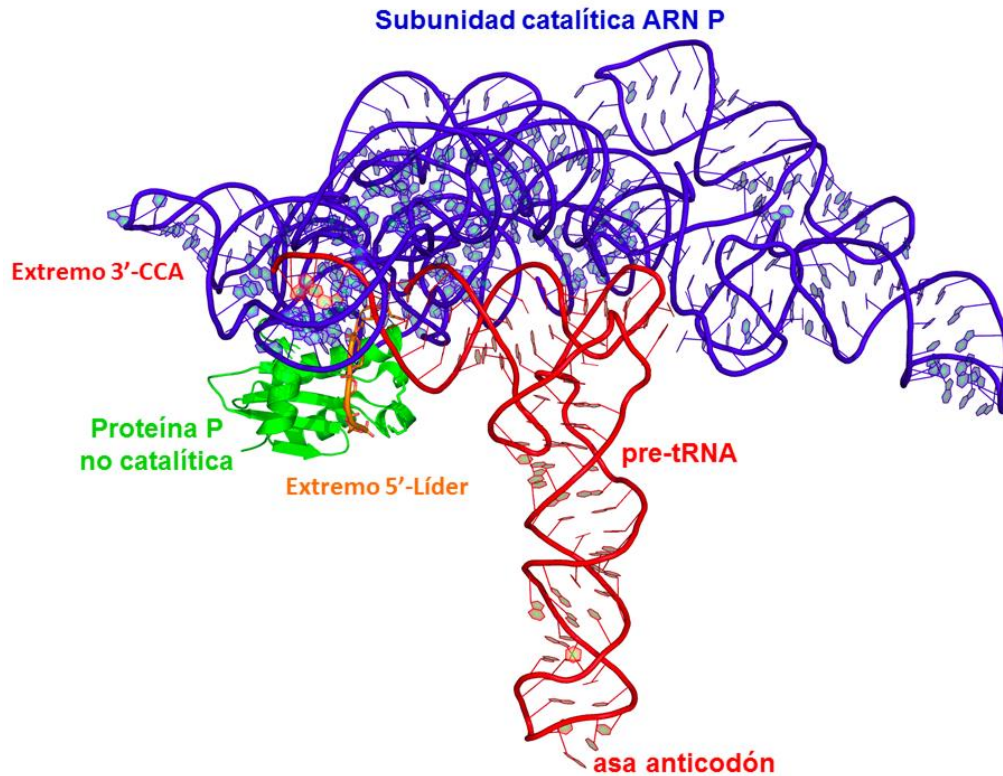


Figura 1. Estructura tridimensional de la holoenzima de la RNasa P de *T. maritima* en complejo con tRNA y 5'-líder. En color azul se muestra la subunidad catalítica de ARN de la RNasa P de 110 kDa y una secuencia lineal de 338 bases. En color rojo se muestra el tRNA de 25 kDa y 76 bases que interacciona con la holoenzima mediante el extremo 3'-CCA y el asa TΨC. En color naranja se muestra el subproducto 5'-líder escindido de 5 bases que interacciona con un surco superficial de la proteína P. En color verde se muestra la subunidad proteica de la RNasa P (proteína P) de aproximadamente 14.3 kDa y ~120 aa. Modificado de Reiter y col. (35).

La subunidad de ARN se compone de 338 nucleótidos y tiene un peso aproximado de 110 kDa. Estructuralmente, se divide en dos regiones: el dominio S y el dominio C. El dominio S o dominio de especificidad, es la región de la subunidad de ARN que se encuentra involucrada en el reconocimiento del pre-tRNA e interacciona tanto con el asa D como con el asa TΨC del pre-tRNA. El dominio C, o dominio catalítico, es donde se encuentra el sitio activo y presenta interacciones con la proteína P y con el extremo 3'-CCA del pre-tRNA (35).

La subunidad proteica de la RNasa P, o proteína P, se compone de 117 aminoácidos y tiene un peso molecular aproximado de 14.3 kDa (35). Estructuralmente, el pre-tRNA interacciona con la proteína P a través de contactos con el extremo 5'-líder de cadena sencilla (35). Así, la proteína P no tiene función catalítica. Su actividad radica en aumentar la afinidad de los cationes divalentes (36), estabilizar al extremo 5'-líder y ayudar a la liberación del producto (37). También aumenta la afinidad de la holoenzima por el pre-tRNA (38) y ayuda al plegamiento correcto de la subunidad de ARN (39, 40).

La estructura cristalográfica de la proteína P de la RNasa P bacteriana de *Thermotoga maritima* muestra una proteína globular que pertenece a la clase estructural α/β (Figura 3). Dicho inusual motivo estructural entrecruzado, así como un gran surco central, es compartido por la proteína ribosomal S5 y el

factor ribosomal de elongación translocasa, observaciones que sugieren que la proteína P evolucionó a partir de un ancestro común con la maquinaria ribosomal (41).

La subunidad de ARN no es un buen blanco farmacéutico

La subunidad de ARN de la RNasa P bacteriana ha sido el blanco predilecto para el descubrimiento de nuevos antibióticos debido a su actividad catalítica; de manera indirecta se ha estudiado el funcionamiento de esta ribozima a través de la inhibición de su actividad en un contexto de una gama amplia de mutantes y condiciones (42, 43).

Los componentes de la subunidad de ARN de la RNasa P en bacteria, archaea y eucaria se encuentran estructuralmente relacionados (44, 45, 46). La subunidad de ARN posee cinco regiones conservadas (CR-I a CR-V) las cuales son comunes a todos los organismos (47). La región conservada CR-IV es donde se localiza el sitio catalítico (35). En esta región se localiza la hélice P4, encargada de la coordinación de los metales divalentes de los cuales depende la reacción de la ribozima (48, 49). Con base en estas observaciones, la especificidad/selectividad por la RNasa P bacteriana se vería comprometida si un inhibidor es desarrollado para interactuar con la subunidad ARN P.

La proteína P podría ser blanco para el descubrimiento de inhibidores

La proteína P tiene una alta conservación en estructura, secuencia y función en todo el dominio bacteria, y no cuenta con homólogos en estructura ni en secuencia en otros dominios de la vida (34, 35, 50, 51) (Figura 2). La proteína P (Figura 3) posee una región que representa la mayor identidad de secuencia entre proteínas P bacterianas, conocida como el motivo RNR por la presencia de residuos arginina-asparagina (41); dicho motivo interactúa con el dominio C de la subunidad de ARN. También están conservados ciertos residuos que conforman el surco de unión al extremo 5'-líder (35, 52, 53) (Figura 4).

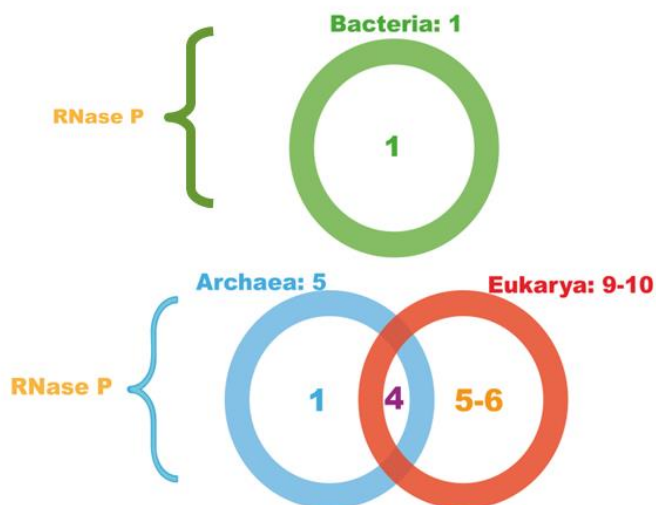


Figura 2. La proteína P no tiene homólogos en otros dominios de la vida. Diagrama de Venn Euler de la divergencia de la proteína P. **Superior.** En bacteria, solo existe una subunidad de proteína para la RNasa P. **Inferior.** En los otros dominios de la vida, la RNasa P se compone en arquea de cinco subunidades y en eucarya se compone de hasta 10 subunidades proteicas. Arquea y eucaria comparten cuatro subunidades proteicas.

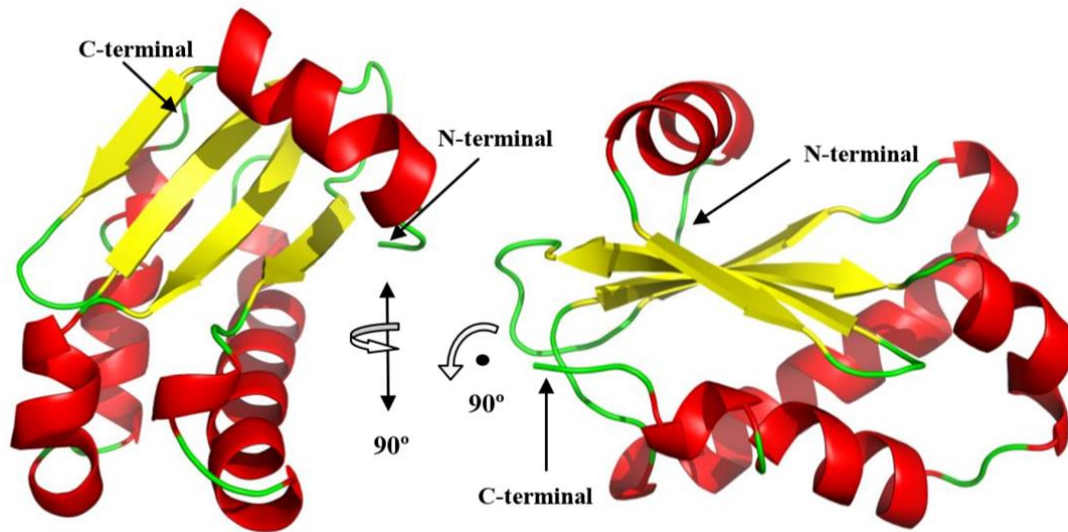


Figura 3. Estructura de la proteína P de *Thermotoga maritima*. La subunidad proteica de la RNasa P tiene 120 aminoácidos y un peso molecular aproximado de 14.3 kDa. Es una proteína globular α/β , con un surco central donde interacciona el extremo 5'-líder. La proteína P está compuesta por tres hélices alfa y cuatro hebras beta anti-paralelas y es homóloga a la proteína ribosomal S5, por lo que se hipotetiza que tienen un ancestro en común (41) (código PDB: 1NZ0; Kazantsev y col. (54)).

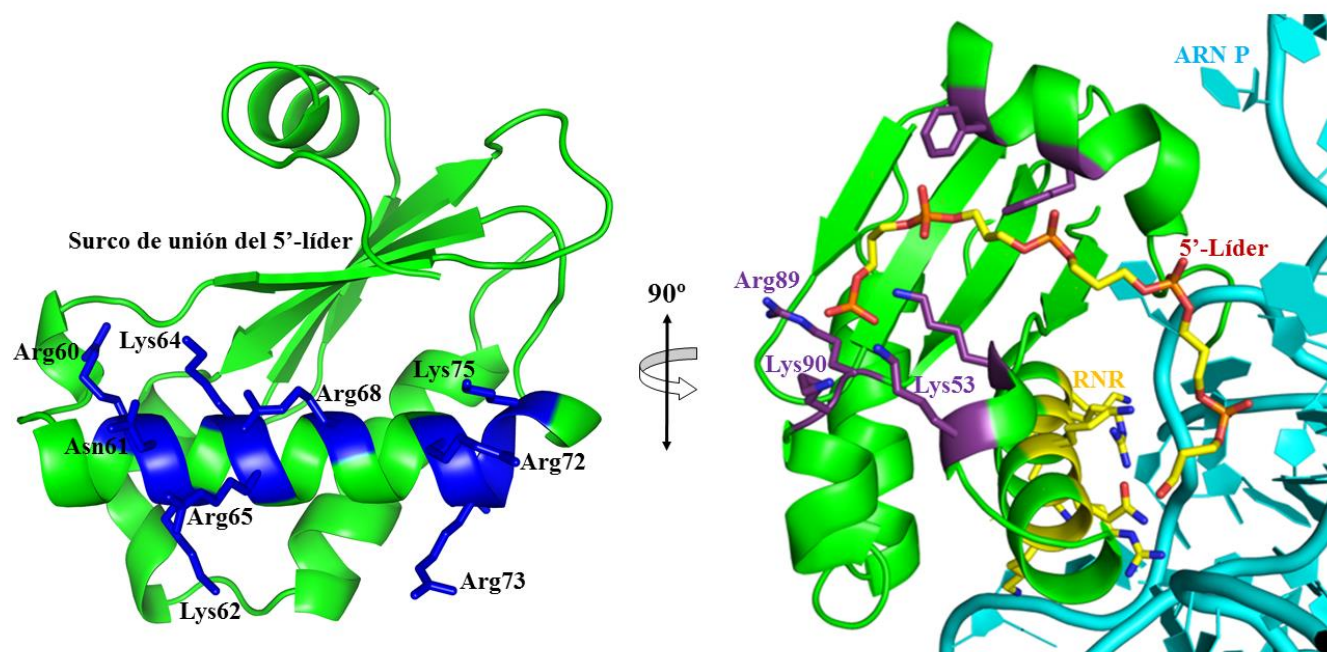


Figura 4. Regiones conservadas en la proteína P que interaccionan con el extremo 5'-líder y la subunidad ARN P. **Izquierda.** En verde, se muestra la estructura de la proteína P. En azul, se muestra el motivo RNR que abarca los residuos Arg60-Lys75 que interaccionan con la subunidad de ARN. El nombre del motivo RNR se debe a los residuos arginina y asparagina que lo componen. **Derecha.** Acercamiento de la interacción de la proteína P-ARN P de la holoenzima de la RNasa P bacteriana. En azul claro, se muestra la subunidad de ARN P. En morado, se muestran los residuos conservados que interaccionan con el 5'-líder (cadena de carbonos amarillos y fosfatos rojos). En amarillo, se muestra el motivo RNR con los residuos que interaccionan con la subunidad ARN P (código PDB: 3Q1R).

En la actualidad se cuenta con la estructura tridimensional apo de la proteína P de varias bacterias modelo (Tabla 1). Sin embargo, antes de este trabajo, no existían estructuras de la proteína P en complejo con moléculas pequeñas tipo fármaco.

Tabla 1. Estructuras tridimensionales apo disponibles de la proteína P.

Organismo	Código PDB	Año	Técnica	Resolución (Å)
<i>B. subtilis</i>	1AF	1998	X-ray	2.6
<i>S. aureus</i>	1D6T	2000	NMR	-
<i>T. maritima</i>	1NZ0	2003	X-ray	1.2
<i>E. coli</i>	2LJP	2011	NMR	-
<i>B. subtilis</i>	4JG4	2013	X-ray	2.3
<i>S. aureus</i>	6DIR	2018	X-ray	2.0

Con el fin de que una molécula biológica sea considerada como un buen blanco para el desarrollo de fármacos, ésta tiene que ser susceptible de someterse a ensayos reproducibles y que permitan la búsqueda de colecciones de compuestos. Para ello se debe obtener pura y activa (55). Además, debe existir algún efecto o relación evidente del blanco con alguna enfermedad específica (56). En el caso de la subunidad proteica de la RNasa P bacteriana, si se inhibe su función, la célula bacteriana cesa la biosíntesis de proteínas y por tanto el metabolismo celular. Así, la proteína P se ha establecido como un blanco atractivo para el descubrimiento y desarrollo de nuevos inhibidores (42).

La proteína P presenta una baja tasa de mutación ya que la expresión de muchos operones policistrónicos es controlada por la RNasa P y se ha demostrado que mutaciones puntuales de residuos en la proteína P afectan la correcta expresión de genes río abajo de operones (32). Se cuenta también con evidencia experimental de residuos en la proteína P que al ser mutados ocasionan una disminución importante en la eficiencia catalítica de la holoenzima (58).

Dada la función de la proteína P, si algún compuesto químico logra unirse en la zona de interacción con el 5'-líder, se podría inhibir la actividad de la holoenzima. Sin embargo, el gran reto de descubrir y desarrollar inhibidores contra la proteína P, es que es una proteína globular sin cavidades y sólo presenta el surco donde se une el 5'-líder (59).

Descubrimiento de inhibidores utilizando fragmentos moleculares

La metodología llamada diseño de fármacos basado en fragmentos (FBDD), permite obtener ligandos más eficientes (véase *“Eficiencia de ligando” en Apéndice E. Glosario de términos*) y explora un espacio químico (véase *Apéndice E. Glosario de términos*) más amplio que las metodologías de alta eficiencia que usan compuestos más grandes y complejos (63). El principio es que una biblioteca de compuestos orgánicos de bajo peso molecular (<300 Da) es probada para determinar unión a un blanco terapéutico de interés y, un compuesto que se haya unido o “hit” (véase *Apéndice E. Glosario de términos*), puede ser detectado a través de varias técnicas biofísicas, después de lo cual los “hits” son susceptibles de optimización y crecimiento por diseño basado en estructura en un proceso iterativo que incrementa su afinidad al blanco (Figura 5).

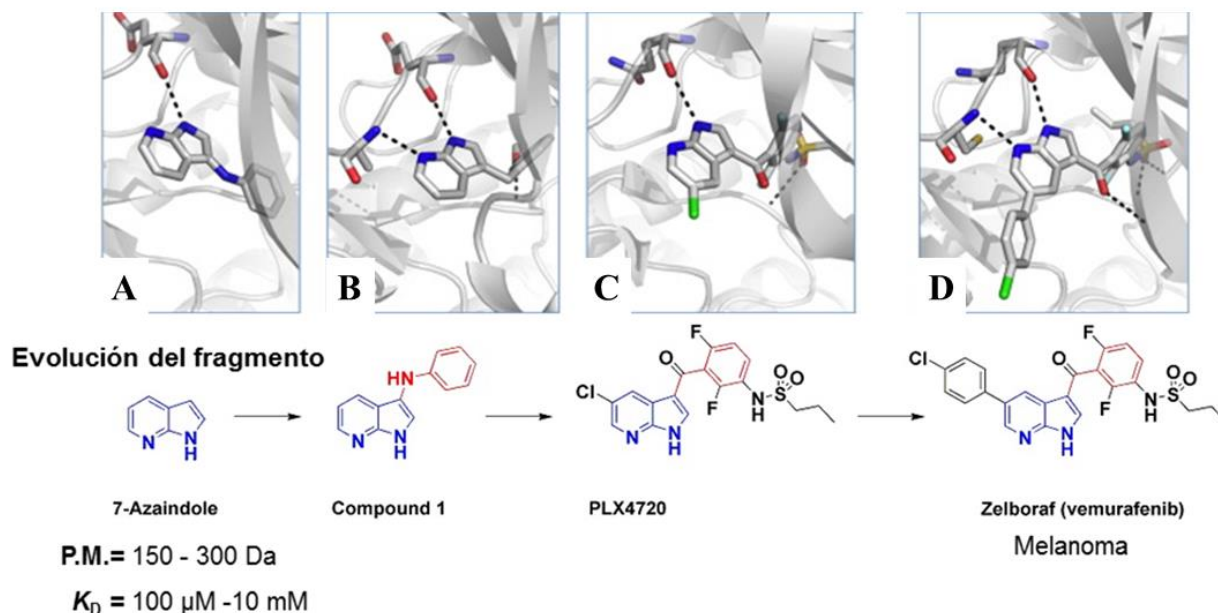


Figura 5. Diseño de fármacos basado en estructura. **A.** Compuesto 7-azaindol detectado mediante cristalografía de rayos X. **B.** Compuesto 1 optimizado a partir del 7-azaindol. **C.** Inhibidor líder PLX4720. **D.** Resultado final con Vemurafenib, un fármaco comercial para tratar melanoma aprobado en 2011 por la FDA con afinidad y características optimizadas a partir del fragmento mostrado en A. Las características típicas iniciales de los fragmentos comprenden un intervalo de peso molecular de entre 150-300 Da y una afinidad de unión por el blanco en el intervalo de K_D de 100 μM – 10 mM. Tomado de Bolag y col. (64).

Una vez que un evento de unión es detectado mediante algún método biofísico, se modifica el compuesto o “hit” para disminuir su toxicidad, aumentar el número de interacciones con el blanco, mejorar las características de solubilidad y biodisponibilidad.

Existen tres metodologías para el diseño de fármacos basado en fragmentos (63):

- 1) Evolución del fragmento.** Es el método más común. Una vez que es detectado un evento de unión, se hace un diseño basado en la estructura tridimensional para lograr más contactos específicos con el blanco añadiendo grupos químicos para generar un inhibidor más potente.
- 2) Unión de fragmentos.** Se realiza cuando se detectan dos eventos de unión en sitios diferentes pero cercanos, en cuyo caso se pueden unir covalentemente mediante un *linker* (véase Apéndice E. *Glosario de términos*), generando un compuesto de peso molecular mayor y más potente.
- 3) Química “click” del fragmento.** Se da en el sitio activo, donde la misma enzima cataliza la generación de su propio inhibidor al hacer reaccionar uno o dos fragmentos.

- Las técnicas biofísicas para la detección de fragmentos, utilizadas en este estudio, son:

A) Interferometría de biocapas

La interferometría de biocapas (BLI) es una tecnología que mide interacciones biomoleculares. En esta técnica, se mide el patrón de interferencia de la luz blanca reflejada desde dos superficies: una capa de proteína inmovilizada y una capa interna de referencia. Al analizar los patrones interferométricos de las ondas resultantes, se pueden medir cambios en el grosor de la punta del biosensor. Un evento de unión entre el ligando inmovilizado sobre la punta del biosensor y un analito en solución produce un incremento en grosor óptico en el biosensor que resulta en un cambio en el patrón de interferencia (65-67).

B) Termoforesis a microescala

La termoforesis a microescala (MST) mide el movimiento de partículas en un gradiente microscópico de temperatura. Cualquier cambio en la estructura o conformación debido a la interacción con un ligando, produce un cambio en la velocidad de difusión molecular a través del gradiente de temperatura, el cual es usado para determinar parámetros cinéticos de unión (68). El movimiento es detectado y cuantificado usando fluoróforos intrínsecos (triptófano, fenilalanina o tirosina) o unidos covalentemente (68-70). MST es una tecnología libre de inmovilización y de micro-fluídica.

C) Cristalografía de rayos X de biomacromoléculas

La cristalografía de rayos X es una técnica que permite obtener la estructura tridimensional de las macromoléculas biológicas a nivel atómico (71, 72), contando así con evidencia experimental del sitio y modo de unión de moléculas pequeñas. El análisis de este sitio puede ayudar a optimizar los compuestos detectados en un proceso de descubrimiento de inhibidores (73). Existen bibliotecas de fragmentos moleculares diseñadas para aplicaciones de cristalografía para llevar a cabo rápidamente la detección de la forma del “hit” de acuerdo a la forma de su mapa de densidad electrónica (74, 75, 76) (Figura 6).

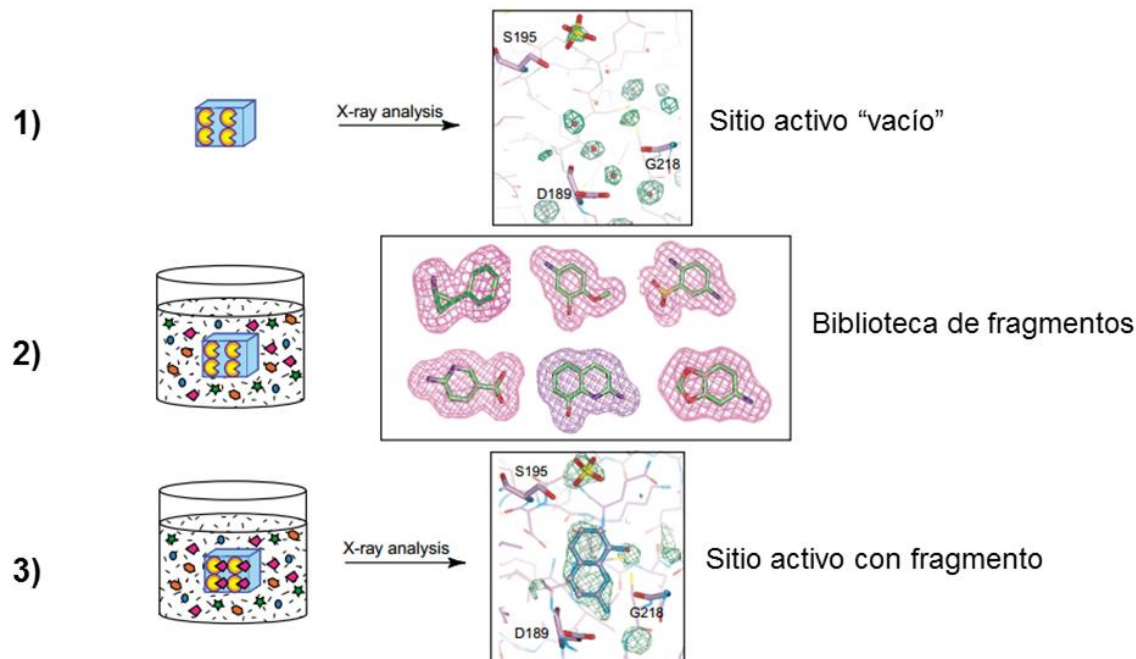


Figura 6. Proceso de identificación de un fragmento o "hit" mediante cristalografía de rayos X. 1) Se resuelve la estructura de la proteína y se analiza el sitio activo, que en esta etapa del proceso se encuentra "vacío". 2) Se lleva a cabo un remojo de cristales de la proteína blanco en cocteles de fragmentos (2-10 fragmentos por coctel). 3) Se resuelven las estructuras de los cristales que fueron remojados y se vuelve a analizar el sitio activo para determinar si existe un evento de unión. Los cocteles de fragmentos se diseñan de tal manera que los compuestos en ese coctel difieran de forma unos de otros, de tal manera que sean fácilmente discernibles a través del análisis de sus densidades electrónicas. Modificado de Nienaber y col. (76).

Evidencia de la utilidad de la técnica es que en 2011 fue aprobado por la FDA el primer fármaco comercial para tratar melanoma, Vemurafenib (77, 78), el cual comenzó su desarrollo a partir de un compuesto de bajo peso molecular detectado por cristalografía aplicando el FBDD (Figura 5).

Estructuras de complejos ligando-proteína han podido ser obtenidas utilizando estrategias como: intercambio de ligando, remojo, co-cristalización, expresión en presencia del ligando, la adición de ligandos durante la purificación de biomoléculas blanco y concentración de solución de una mezcla blanco-ligando (79, 80, 81).

D) Ensayos de actividad por fluorescencia

El método convencional para la medición de la actividad de RNasa P emplea pre-tRNA marcado con ^{32}P , el cual, después de incubarlo con RNasa P, se carga en geles urea-acrilamida, seguido de un revelado mediante auto-radiografía (82), por lo que este formato no es práctico para ensayar de cientos a miles de compuestos en un intervalo de tiempo razonable.

En cambio, las técnicas por fluorescencia permiten obtener resultados con mayor rapidez y una buena reproducibilidad (82, 83), de tal manera que es posible ensayar volúmenes por pocillo de 10 a 20 μL en microplacas de 96 o 384 pocillos que pueden ser utilizadas en lectores de placa automatizados. Dicha

configuración experimental puede ser adaptada a ensayos en tiempo real, permitiendo la determinación de parámetros cinéticos de actividad, así como el monitoreo de cientos a miles de compuestos en poco tiempo (83).

Una de las ventajas de emplear sondas sintéticas fluorescentes es que es posible diseñar sustratos que mimeticen al pre-tRNA y que a la vez tengan menos afinidad por la holoenzima de la RNasa P (83, 84), por lo que el ensayo podría ser sensible a la detección de moléculas pequeñas que presentaran constantes de disociación bajas. Dichas moléculas pequeñas con interacciones débiles son la semilla para el desarrollo subsecuente de inhibidores (63, 64).

Inhibidores de molécula pequeña de la RNasa P reportados a la fecha

A) Aminoglucósidos

Los aminoglucósidos son moléculas pequeñas, policatiónicas, con un sistema de anillos conectados que consisten en uno o muchos azúcares aminados unidos mediante enlaces glicosídicos a un aminocitol (85). Se sabe que varios antibióticos aminoglucósidos interaccionan con blancos bacterianos de ARN y complejos ARN-proteína, como la paramomicina y tobramicina, los cuales obstruyen el proceso de traducción en el sitio de decodificación del ARN de transferencia-aminoacilado (sitio A) de la subunidad de ARN ribosomal 16S (86).

Los aminoglucósidos son promiscuos, ya que se unen y perturban la actividad de otros blancos diferentes al ARN, como fosfolípidos ácidos de la membrana epitelial humana, varias fosfolipasas, la ATPasa Na^+/K^+ (85, 86); de manera que estas perturbaciones generan toxicidad en humanos. Sin embargo, los aminoglucósidos pueden ser considerados como compuestos líderes para estudiar interacciones ARN-molécula pequeña y para el desarrollo de potenciales antibióticos no tóxicos.

Para el caso de la RNasa P, usando como modelo a la ribozima de *E. coli* y *M. hyopneumoniae*, se midió la actividad de RNasa P en presencia y ausencia de la proteína P (también llamada C5) con varios aminoglucósidos, los resultados sugieren que la neomicina B inhibe la actividad de la RNasa P uniéndose al ARN e impidiendo la unión de metales catiónicos divalentes (Mg^{2+}) (87).

B) Compuestos conjugados derivados de aminoglucósidos

Ya que las proteínas de unión al ARN reconocen a sus blancos mediante zonas cargadas altamente electropositivas y regiones ricas en arginina (88, 89), los aminoglucósidos han sido modificados con sustituciones de argininas para diseñar compuestos peptidomiméticos con mayor afinidad por blancos bacterianos de ARN que los aminoglucósidos simples.

Debido al hecho de que los aminoglucósidos se unen a varios ARNs celulares no relacionados entre sí, el grupo de Gopalan llevó a cabo estudios con derivados conjugados de aminoglucósidos con el fin de aumentar la especificidad hacia la RNasa P bacteriana. Por ejemplo, el derivado de neomicina B, NeoR6 y un derivado de gentamicina tri-arginina (R3G) fueron usados para observar los efectos sobre la inhibición de la RNasa P de *E. coli* como modelo experimental y en las bacterias patogénicas *Neisseria gonorrhoeae*, *Porphyrromonas gingivalis* y *Streptococcus pneumoniae* (90). La actividad de la RNasa P de *E. coli* es inhibida por neomicina B y NeoR6 con valores IC_{50} de 60 μM y 125 nM respectivamente. La RNasa P de bacterias patogénicas fue completamente inhibida con 500 nM de NeoR6 o 1500 nM de

R3G. La actividad de la RNasa P humana no fue inhibida a concentraciones de NeoR y R3G que son 10 veces más altas que los valores de IC_{50} para la RNasa P de *E. coli*. Sin embargo, sí se presentó inhibición completa a 7.5 μ M de NeoR y 20 μ M de R3G (90).

En el segundo estudio, el grupo de Gopalan sintetizó tres compuestos derivados de Neomicina B: hexaguanidina y los conjugados penta-arginil y hexa-lisil (91). Los derivados de neomicina B inhiben la RNasa P de *E. coli* con valores de IC_{50} fueron ~400, 6, 3 y 0.5 μ M para NeoB, NeoG6, NeoK6 y NeoR5, respectivamente, indicando que el conjugado neomicina-penta-arginina, fue el inhibidor más potente. El mecanismo de inhibición de estos compuestos, al igual que el de los aminoglicósidos, es mediante la interferencia de unión de iones metálicos divalentes Mg^{2+} .

C) Fentotiazinas y derivados

Wu y colaboradores, llevaron a cabo el primer estudio que demuestra que ARNs funcionales no codificantes son inhibidos por diferentes derivados de fentotiazinas usando a la subunidad de ARN de la RNasa P de *E. coli*, *M. hyopneumoniae* y *M. tuberculosis* para estudiar los efectos de varios compuestos que comparten la misma estructura núcleo de fentotiazina. La toluidina azul-O mostró ser el inhibidor más potente de todos los derivados de fentotiazinas, que podrían estar involucradas en interacciones de apilamiento (véase Apéndice E. Glosario de términos) poco específicas con las bases de ARN (92).

D) Benzimidazoles

Hori y colaboradores, llevaron a cabo el primer reporte de inhibidores del ARN de la RNasa P con ligandos sintéticos orgánicos cuyas estructuras no están basadas en productos naturales. Los benzimidazoles se unen reversiblemente al pre-tRNA con constantes de disociación aparente en el intervalo micromolar, de esta manera evitando la unión del pre-tRNA e inhibiendo la actividad de RNasa P (93). A diferencia de los aminoglicósidos, el efecto inhibitorio de los benzimidazoles parece no ser resultado de un desplazamiento de iones metálicos, sino que el mecanismo de inhibición podría ser debido a un cambio conformacional inducido en uno o más componentes de la holoenzima.

E) Irginol hexaacetato

El primer ensayo de actividad de RNasa P en tiempo real para búsqueda de alta eficiencia mediante polarización/anisotropía fluorescente (FP/FA) usando un sustrato pre-tRNA con una etiqueta de fluoresceína en el extremo 5', fue aplicado para identificar inhibidores de la RNasa P a partir de una biblioteca de 2880 compuestos. El resultado principal de esta búsqueda fue la identificación de un derivado natural, el irginol hexaacetato (Ir6Ac), como nuevo inhibidor de la RNasa P de *Bacillus subtilis* (94). La IC_{50} de Ir6Ac para la inhibición de la RNasa P de *B. subtilis* fue de 820 nM y es un inhibidor reversible con más alta especificidad que aquellos inhibidores reportados como la neomicina B. La adición de Ir6Ac reduce los valores k_{cat} y K_M , indicando un modelo de inhibición mixto (el inhibidor se une tanto a la enzima, como al complejo ES con diferente afinidad). No hay información estructural del sitio y modo de unión a la RNasa P, por lo que no es posible progresar al diseño basado en estructura.

2. Antecedentes

Durante el proyecto de Maestría, a través de una búsqueda en una biblioteca de 352 fragmentos moleculares (150-300 Da) provenientes de la biblioteca ZenobiaTM, se detectó mediante cristalografía de rayos X la unión de un compuesto de bajo peso molecular ($M=151$ Da) en la zona de unión de la proteína P para el extremo 5'-líder del pre-tRNA (Figura 7, panel izquierdo). Dicho fragmento es el 2-mercaptobenzoxazol (llamado en este estudio 2-MBX) (Figura 7, panel derecho).

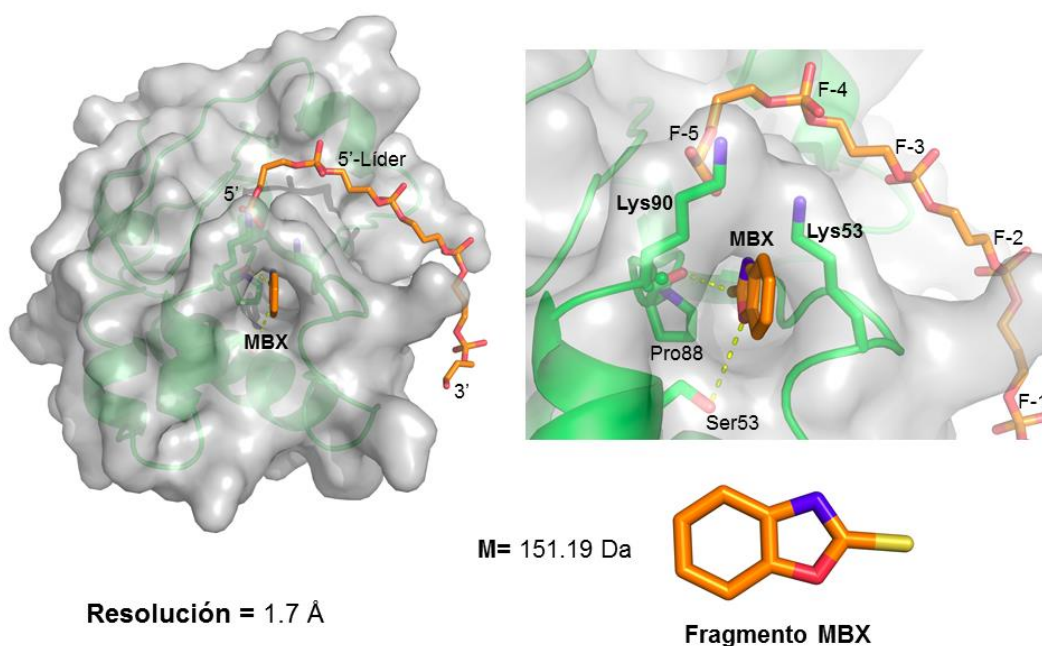


Figura 7. El fragmento 2-MBX se une en un sitio de la proteína P importante para la función de la holoenzima. Estructura cristalográfica de la proteína P en complejo con el fragmento 2-mercaptobenzoxazol (2-MBX) a 1.7 Å de resolución. Existen dos interacciones de apilamiento con los residuos Lys53 y Lys90, y un puente de hidrógeno con la Ser53. En un estudio independiente (58), la mutación de los residuos Lys53 y Lys90 por alanina disminuye cinco veces la actividad de la holoenzima. El extremo 5'-líder se muestra en naranja como cadena fosfatada. F-1 a F-5 indica la posición de los grupos fosfato en el líder. En la parte inferior se muestra la estructura del fragmento 2-MBX.

El descubrimiento del fragmento 2-MBX generó la posibilidad de descubrir inhibidores que actúen por medio de la unión a la proteína P, siempre y cuando se cuente también con metodologías que permitan la medición de la unión a los diferentes componentes de la holoenzima y el sustrato, así como la actividad de la misma, todo ello, en un formato de medio o incluso de alto rendimiento.

Detectar la unión de moléculas pequeñas a la proteína P por cristalografía de rayos X no ha sido sencillo, ya que los fragmentos generalmente presentan baja ocupación y son poco afines al blanco, lo cual es parte de la naturaleza del diseño basado en fragmentos. Otro de los retos a superar es que los cristales de proteína P son sensibles a bajas concentraciones de DMSO (1%), disolvente en el cual son solubles los fragmentos, por lo que se tuvo que buscar otra condición en la que fueran solubles y

compatibles con las condiciones de cristalización de la proteína P. Las ventajas que ha aportado nuestra metodología de cristalogénesis, son que el licor madre donde crecen los cristales se compone del agente precipitante PEG-1000 al 12%, mientras que la solución crioprotectora se compone de PEG-1000 30-35% y los compuestos son solubles en PEG-200 50% y PEG-400 50%, de tal manera que el remojo de cristales realizado con compuestos disueltos en PEG reduce el daño al cristal de proteína. Dicho daño se manifiesta en el incremento de la mosaicidad (véase Apéndice E. *Glosario de términos*), en la disolución del cristal y la disminución en la resolución. Otras ventajas de nuestra metodología son: que la proteína P cristaliza en tres días y que es una proteína globular pequeña (14.3 kDa), en cuya superficie es fácil y rápido visualizar un evento de unión. Con base en estos avances metodológicos, resulta interesante aplicarlos en la búsqueda y caracterización de la unión de moléculas pequeñas a la proteína P con el fin de obtener información estructural valiosa de esta interacción y poder llevar a cabo una optimización guiada por estructura de estos compuestos.

Criterio de selección del tipo de RNasa P bacteriana

Se eligió para trabajar en este proyecto a la ribonucleasa P de *Thermotoga maritima* por las siguientes razones: 1) la RNasa P de *T. maritima* representa al tipo más común de ribozima “tipo A”, el mismo que se encuentra presente en *Escherichia coli* (95). 2) Se conoce la estructura cristalográfica de la holoenzima de la RNasa P de *T. maritima* en complejo con tRNA y el producto 5'-líder (11) así como la estructura apo de la subunidad de ARN (35), facilitando los estudios de estructura función. 3) El protocolo de plegamiento de la holoenzima de la RNasa P de *T. maritima* (véase 6. *Materiales y métodos*) provee cristales que presentan un complejo molecular homogéneo y funcional (11). 4) Aunque la RNasa P de *T. maritima* presenta una actividad óptima a 50°C (96), la holoenzima es activa en el intervalo de temperatura de 20-55°C.

3. Planteamiento del problema

Todos los inhibidores de la RNasa P bacteriana reportados a la fecha actúan uniéndose a la subunidad de ARN, la cual posee regiones universalmente conservadas, incluyendo el sitio catalítico, por lo que la posibilidad de generar inhibidores inespecíficos es muy alta. En ese sentido, la proteína P ofrece un blanco alternativo atractivo no explorado aún. Además, no existe información estructural disponible de un inhibidor en complejo con alguna subunidad de la RNasa P bacteriana.

4. Hipótesis

Es posible detectar moléculas orgánicas de bajo peso molecular que se unan a la subunidad proteica de la RNasa P bacteriana y desarrollar inhibidores de la holoenzima de la RNasa P a partir de las mismas.

5. Objetivos

Objetivo general

Descubrir inhibidores de bajo peso molecular de la RNasa P bacteriana que actúen mediante la interacción con la subunidad proteica a través de la implementación de una metodología que involucre ensayos de actividad, ensayos de unión y determinación de la estructura tridimensional de complejos inhibidor-proteína P.

Objetivos particulares

1. Desarrollar un ensayo para analizar la unión de moléculas pequeñas a la proteína P, al ARN P y a la holoenzima de la RNasa P bacteriana.
2. Optimizar un ensayo para analizar la actividad de RNasa P bacteriana en un formato de medio o alto rendimiento.
3. Detectar inhibidores de la holoenzima de la RNasa P bacteriana.
4. Determinar el modo de unión de inhibidores en la proteína P mediante cristalografía de rayos X.

6. Materiales y Métodos

6a) Materiales

Reactivos generales. Todos los reactivos fueron adquiridos del Sigma-Aldrich a menos que se especifique algún otro proveedor.

Reactivos para cristalografía. Todos los reactivos y materiales utilizados en el presente estudio para experimentos de cristalogénesis y remojo fueron adquiridos de la compañía Hampton Research.

Las listas de cepas, vectores y oligonucleótidos utilizados se encuentran en el Apéndice A. Materiales suplementarios, Tabla S1-S3.

Extremo ARN 5'-líder del pre-tRNA. Para ensayos mediante interferometría de biocapas se hizo uso del oligonucleótido de cadena sencilla de ARN del extremo 5'-líder del pre-tRNA. Esta secuencia (5'-GGAAAAAGAU-3', adquirida de la compañía Integrated DNA Technologies (IDT) es la misma secuencia que se usó en la cristalización de la estructura de la holoenzima de la RNasa P (35), y es una secuencia consenso de todos los extremos 5'-líder de los pre-tRNA de *T. maritima*.

Extremo 5'-líder de ADN. La secuencia de 10 bases 5'-GGAAAAAGAT-3', es la misma que la secuencia del extremo 5'-líder de ARN, pero en bases de ADN. Fue adquirida de la compañía IDT y se utilizó como control para comparar la preferencia de la proteína P por líderes de ARN o ADN durante experimentos mediante interferometría de biocapas.

Oligonucleótido biotinilado complementario a la subunidad ARN P. Para llevar a cabo ensayos mediante interferometría de biocapas, se utilizó un oligo de ARN complementario a la subunidad de ARN P con el fin de inmovilizarla sobre biosensores de Super Streptavidina (SSA), con la secuencia 5'-CAUCCAGACGAU-3' adquirida de IDT. Las modificaciones en el ARN P consistieron en una secuencia con parte de la secuencia del tallo P1 complementario al oligonucleótido biotinilado (AUCGUCUGGA). El tallo P1 se localiza lejos del sitio activo o de otras regiones involucradas en interacciones específicas (35). La actividad de esta construcción fue indistinguible de la holoenzima tipo silvestre.

Sustrato artificial minihélice (Mh). El sustrato minihélice (Mh), que tiene la secuencia de 24 bases de ARN de cadena sencilla: 5'-GAUCUGAAUGCGGAAACGCGCCAC-3', fue modificado a partir de un sustrato previamente caracterizado de la RNasa P, pMini3bpUG (31, 33) y acoplado a un apagador llamado Black Hole Quencher 1 (BHQ1) o Black Hole Quencher 2 (BHQ2) en el término 5' y un fluoróforo acoplado al término 3' llamado Alexa Fluor 488. Esta sonda fue adquirida de Sigma-Aldrich o de IDT. Una cantidad adquirida de 0.05 mmoles tuvo como rendimiento entre 6.3 y 20 nmol (57.5 a 181 µg, suficiente para 1300 a 4000 reacciones dependiendo de la concentración experimental de sustrato). Antes de su uso, la minihélice fue diluida a partir de una solución de 100 mM (disuelta en agua) a 3.6 mM en amortiguador de actividad H (Tris-HCl pH 7.4 33 mM, HEPES 66 mM, 400 mM AmOAc y 100 mM MgCl₂).

Sustrato pre-tRNA-mitades. El sustrato pre-tRNA-mitades fue diseñado para facilitar la síntesis química y concebido a partir del hecho de que el asa anticodon no es requerida para el reconocimiento por la RNasa P (35). Así como el sustrato Mh, el pre-tRNA-mitades fluorescente fue acoplado al apagador Black Hole Quencher 1 (BHQ1) al término 5' del oligo 5' y el fluoróforo Alexa Fluor 488 al término del 3' del oligo 3'. Las sondas fueron adquiridas de IDT. Una escala solicitada de 250 nmol tuvo como rendimiento 27.4 nmol (0.36 mg) para el oligo 5' y 3.8 nmol (50 µg) para el oligo fluorescente 3'. Antes de su uso, las sondas fueron diluidas a partir de una solución de 100 mM (disuelta en agua) y ensambladas mediante la mezcla y calentamiento de 3.6 µM en amortiguador de actividad L (Tris-HCl pH 7.4, 33 mM, HEPES 66 mM, 100 mM acetato de amonio (AmOAc) y 10 mM MgCl₂).

Biblioteca Zenobia Library-1TM. La biblioteca de 352 fragmentos de 100-300 Da Zenobia-Library-1, fue adquirida de la compañía Zenobia Therapeutics, Inc. Se conforma por 352 compuestos liofilizados para ser disueltos en 50 µL de DMSO 100% y dar una concentración final de 200 µM por compuesto y está diseñada para búsqueda de compuestos mediante cristalografía de rayos X para que “hits” unidos al blanco puedan ser identificados por su forma a través de la inspección de su densidad electrónica.

Microbiblioteca de compuestos derivados del 2-MBX. Los 11 compuestos que forman la microbiblioteca de compuestos derivados del fragmento 2-MBX fueron adquiridos en polvo de manera individual de la compañía Sigma-Aldrich.

Biblioteca TimTec FL-500TM. La biblioteca de TimTec FL-500 fue adquirida de la compañía TimTec LLC y se compone de 500 moléculas de origen natural entre las cuales se encuentran flavonoides, isoflavonoides, flavonas y sus derivados.

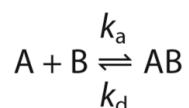
Biblioteca The Spectrum CollectionTM. Una muestra de la biblioteca The Spectrum Collection (MicroSource Discovery Systems, Inc), fue adquirida a través del Dr. Chi-Hao Luan, Director del High-Throughput Laboratory de la Northwestern University, Chicago, Illinois, USA. Esta colección de moléculas pequeñas se conforma por 2,560 compuestos dentro de los que se encuentran fármacos patentados, insecticidas, herbicidas y demás compuestos bioactivos.

6b) Métodos

La metodología relacionada con clonación y purificación de biomacromoléculas (como el ARN y proteínas utilizadas) se encuentra en el Apéndice B. Métodos suplementarios.

Reconstitución de la holoenzima de la ribonucleasa P bacteriana. La reconstitución de la holoenzima de la RNasa P de *T. maritima* fue llevada a cabo mezclando el ARN P y la proteína P en una proporción de 1:1.1 en amortiguador de actividad TH 1X (Tris-HCl 33 mM pH 7.4, HEPES 66 mM). La concentración final a la cual se mezcló la subunidad de ARN fue de 400 nM, mientras que la concentración final de la subunidad de proteína fue de 440 nM. A esta mezcla, se añadieron 100-400 mM de acetato de amonio (400 mM en ausencia de la proteína para aumentar actividad catalítica). Esta solución fue calentada a 94°C durante 2 min en un thermoblock e inmediatamente incubada en hielo durante 2 min, al término de este tiempo se añadió MgCl₂ hasta una concentración final de 100 mM. La concentración de la holoenzima reconstituida fue cuantificada utilizando un espectrofotómetro NanoDrop ND-100 y se almacenó a -20°C hasta su uso.

Procesamiento y ajuste matemático de los sensorgramas para experimentos de unión por interferometría de biocapas. Los experimentos fueron controlados y monitoreados con ayuda del programa Data Acquisition 8.2, mientras que los parámetros de la cinética de unión fueron calculados usando el programa Data Analysis 8.2 (FortéBio, Inc.). El análisis cinético es usado para determinar la afinidad de una interacción y medir la velocidad de asociación y de disociación para unión reversible, no covalente. En la técnica de interferometría de biocapas, un interactuante es inmovilizado sobre la superficie del biosensor (denominado ligando) y otro interactuante que permanece en solución (denominado analito). En un experimento típico de cinética de unión usando un biosensor en la plataforma Octet, un ensayo comienza con una línea base inicial o paso de equilibrio utilizando un amortiguador. Después, un ligando (tal como la proteína P) es inmovilizado sobre la superficie del biosensor (carga o *loading*). Luego de la inmovilización del ligando, los biosensores son sumergidos dentro de la solución amortiguadora para generar una línea base con el fin de evaluar y determinar el nivel de carga del ligando. Después del paso de la línea base, los biosensores son sumergidos en la solución que contiene al analito (paso de asociación). En este paso, es medida la interacción de unión del analito con el ligando inmovilizado. Posterior a la asociación del analito, el biosensor es sumergido dentro de la solución amortiguadora sin el analito, y el analito unido se disocia del ligando (paso de disociación). Varias concentraciones de analito pueden ser corridas en paralelo (hasta ocho mediciones simultáneas en el sistema OctetRED96). La serie de pasos del ensayo es repetida en biosensores nuevos para cada analito siendo probado. Cada respuesta de unión es medida y reportada en tiempo real en una gráfica de respuesta vs tiempo llamada sensorgrama. La afinidad mide que tan fuerte un es complejo, o cuánto complejo es formado cuando la unión se acerca al equilibrio. El modelo más simple usado para describir esta interacción entre dos biomoléculas es representado por la ecuación siguiente:



Donde, A representa a la molécula ligando inmovilizada sobre la superficie del biosensor, y B es el analito en solución. Este modelo de unión asume una interacción simple 1:1, donde una molécula de ligando interactúa con una molécula de analito, y la unión es independiente y de igual fuerza para todos los sitios de unión. La formación del complejo en este caso sigue una cinética de pseudo primer orden. La velocidad de la formación del complejo AB es dictado por una velocidad constante de asociación y las concentraciones de ligando y analito sin unirse. La ecuación usada para ajustar la etapa de asociación es una integración de diferentes ecuaciones que muestran que la velocidad de asociación es una función de la disminución de concentración de moléculas de ligando no unidas mientras ocurre la unión de analito:

$$Y = Y_0 + A(1 - e^{-k_{obs} * t})$$

Donde, Y= nivel de unión, Y_0 = unión al comienzo de la asociación, A es una asíntota y t = tiempo. k_{obs} es la constante de velocidad observada. Al mismo tiempo que el complejo AB se está formando, también se separa de nuevo en sus componentes A y B. El término k_{obs} refleja la velocidad general de la

asociación y disociación combinada de las dos moléculas interactuantes. Cuando los biosensores son sumergidos en solución amortiguadora sin analito, el complejo comienza a disociarse a una velocidad que es una función de la constante de la velocidad de disociación y de la concentración del complejo ligando-analito. La ecuación usada para ajustar la disociación es:

$$Y = Y_0 + Ae^{-k_d \cdot t}$$

Donde, Y_0 es la unión al comienzo de la disociación, y k_d es la constante de la velocidad de disociación. Esta ecuación refleja el decremento en la velocidad de disociación con respecto al tiempo mientras la concentración de complejos unidos sobre la superficie disminuye. El término k_d mide la estabilidad del complejo, o la fracción de complejos que se disocian por segundo, y es expresado en unidades de s^{-1} . Al resolver las ecuaciones de arriba para k_{obs} y k_d , la constante de la velocidad de asociación puede entonces ser calculada con la ecuación siguiente:

$$k_a = \frac{k_{obs} - k_d}{[Analyte]}$$

La constante de la velocidad de asociación, k_a , representa el número de complejos AB formados por segundo en una solución 1 molar de A y B. El término k_a es expresado en $m^{-1}s^{-1}$. La K_D , expresada en unidades de molaridad (m), es la constante de afinidad, o la constante de disociación en el equilibrio, la cual mide que tan fuerte se une el ligando con el analito; representa la relación de la velocidad de asociación con la velocidad de disociación y puede ser calculada usando k_a y k_d :

$$K_D = \frac{[A] \cdot [B]}{[AB]} = \frac{k_d}{k_a}$$

En una interacción bimolecular 1:1, las etapas de asociación y disociación muestran una señal que está descrita por una sola función exponencial. Las moléculas de analito se unen a la misma velocidad a cada sitio de unión del ligando. La curva de asociación sigue un perfil de unión hiperbólico característico, con un incremento exponencial en la señal seguido por una meseta o *plateau* cuando la unión alcanza el equilibrio. La curva de disociación sigue un decremento exponencial con la señal eventualmente retornando a la línea base. La solución del ajuste general para una unión 1:1 es como sigue:

Etapas de asociación:

$$y = R_{max} \frac{1}{1 + \frac{k_d}{k_a \cdot [Analyte]}} (1 - e^{-(k_a \cdot [Analyte] + k_d) \cdot x})$$

Fase de disociación:

$$y = y_0 e^{-k_d(x-x_0)}$$

$$y_0 = R_{\max} \frac{1}{1 + \frac{k_d}{k_a * [\text{Analyte}]}} (1 - e^{-(k_a * [\text{Analyte}] + k_d)x_0})$$

Ensayo de interferometría de biocapas con la proteína P. La inmovilización de la proteína P con etiqueta de histidinas de *T. maritima* a 0.05 mg/mL (3.7 μ M), fue llevada a cabo en biosensores Ni-NTA. Para las pruebas iniciales de unión de la proteína P, una concentración de compuesto de 1 mM fue usada. Los periodos de tiempo para cada paso del ensayo fueron como sigue: inmovilización (400 s), lavado (100 s), establecimiento de la línea base (100 s), asociación (250 s) y disociación (250 s). Para cada curva de titulación, seis puntos fueron empleados correspondientes a concentraciones de compuesto de 125 μ M, 62.5 μ M, 31.25 μ M, 15.6 μ M, y 7.8 μ M. Biosensores control, sin proteína P inmovilizada, fueron usados para medir la unión inespecífica de compuestos a la superficie del biosensor y llevar a cabo la substracción o resta de esta señal al calcular la señal neta de unión del compuesto al blanco. Todos los pasos del ensayo (inmovilización, lavado, línea base, asociación y disociación) fueron llevados a cabo en un amortiguador compuesto por 20 mM HEPES pH 7.4, 100 mM acetato de amonio, 10 mM MgCl₂ y 5% DMSO. Los ensayos fueron llevados a cabo en microplacas de 96 pocillos de fondo antireflejante negro (Greiner Bio-One Cat. No. 655209), en un volumen total de 200 μ L a 30°C y en un régimen de agitación orbital a 1000 rpm. Los experimentos fueron controlados y monitoreados con ayuda del programa Data Acquisition 8.2 (FortéBio, Inc.).

Ensayos de interferometría de biocapas para ARN P y holoenzima. Ensayos de unión en tiempo real con ARN P, holoenzima, proteína P, sustrato minihélice y 5'-líder con compuestos individuales, fueron llevados a cabo mediante interferometría de biocapas usando el sistema Octet RED96 (FortéBio, Inc.). Para experimentos involucrando a la holoenzima o ARN P, un oligonucleótido de ARN, biotinilado en el término 5' con la secuencia 5'-biotina-CAUCCAGACGAU, fue inmovilizado en biosensores de Super Streptavidin (SSA, FortéBio, Inc.). Después, un ARN P modificado u holoenzima fueron inmovilizados a esta superficie. Las modificaciones en el ARN P consistieron en una secuencia con parte de la secuencia del tallo P1 complementario al oligonucleótido biotinilado (AUCGUCUGGA). El tallo P1 se localiza lejos del sitio activo y de otras regiones involucradas en interacciones específicas (35). La actividad de esta construcción fue indistinguible de la holoenzima tipo silvestre. Alternativamente, una vez reconstituida la holoenzima, ésta fue inmovilizada a través de la etiqueta de histidinas de la proteína P de *T. maritima* sobre biosensores Ni-NTA a 0.05 mg/mL (3.7 μ M). Biosensores control fueron usados para medir la unión inespecífica de compuestos a la superficie del biosensor y llevar a cabo la substracción de esta señal al calcular la señal neta de unión del compuesto al blanco. Todos los pasos del ensayo (inmovilización, lavado, línea base, asociación y disociación) fueron llevados a cabo en un amortiguador compuesto por 20 mM HEPES pH 7.4, 100 mM acetato de amonio,

10 mM MgCl₂ y DMSO al 5%. Los ensayos fueron llevados a cabo en microplacas de 96 pocillos de fondo antirreflejante negro (Greiner Bio-One Cat. No. 655209), en un volumen total de 200 µL a 30°C y en un régimen de agitación orbital a 1000 rpm. Los experimentos fueron controlados y monitoreados con ayuda del programa Data Acquisition 8.2 (FortéBio, Inc.). Los parámetros de la cinética de unión fueron calculados usando el programa Data Analysis 8.2 (FortéBio, Inc.). Después de la substracción de la señal de los biosensores control, las curvas de unión fueron alineadas al eje X y Y, así como a la curva inter-paso de asociación-disociación para obtener una línea base común para las fases de asociación y disociación.

Unión del ARN 5'-líder a la proteína P y bloqueo de unión del ARN 5'-líder mediado por compuestos. Un oligonucleótido de cadena sencilla con una secuencia de ARN 5'-líder de 10 bases (5'-GGAAAAGAU-3') fue adquirido de la compañía IDT y usado a concentraciones de 5 µM. Esta secuencia es la misma que fue usada para la obtención de la estructura del cristal de la holoenzima de la RNasa P (35). La proteína P con etiqueta de histidinas fue inmovilizada sobre biosensores Ni-NTA a 0.05 mg/mL (3.7 µM). El ARN 5'-líder no se une inespecíficamente a la superficie de los biosensores Ni-NTA, por lo que no fue necesario llevar a cabo experimentos con biosensores control para abstraer o restar unión inespecífica. Los pasos llevados a cabo durante el ensayo fueron como se describe a continuación: inmovilización de la proteína P (400 s, 1000 rpm), lavado (100 s, 1000 rpm), línea base (100 s, 1000 rpm), asociación de compuesto (800 s, 500 rpm), disociación de compuesto (200 s, 500 rpm), línea base (100 s, 1000 rpm), asociación del ARN 5'-líder (200 s, 1000 rpm). Todos los pasos del ensayo fueron llevados a cabo a 30°C. Los compuestos fueron usados a una concentración de 100 µM en todos los experimentos. Todos los pasos del ensayo fueron llevados a cabo en 20 mM HEPES pH 7.4, 100 mM acetato de amonio, 10 mM MgCl₂ y 5% DMSO

Ensayo de interferometría de biocapas con el sustrato artificial minihélice. El sustrato artificial minihélice, biotinilado en el término 3', fue adquirido de la compañía Integrated DNA Technologies (IDT) (secuencia de 24 bases de ARN, 5'-GAUCUGAAUGCGGAAACGCGCCAC-3'-Biotina). La minihélice biotinilada fue inmovilizada sobre biosensores Super Streptavidin (SSA) a una concentración de 10 µM. Una búsqueda inicial de compuestos a 1 mM resultó en una señal de unión a la minihélice únicamente para dos compuestos: el violeta de genciana y el azul de metileno. Para los experimentos de titulación, 10 µM de minihélice 3'-biotinilada fue inmovilizada durante 300 s sobre biosensores Super Streptavidin (SSA), seguido de un paso de lavado de 100 s, así como pasos de asociación y disociación de 200 s en presencia de seis concentraciones diferentes de compuestos abarcando desde 500 µM a 15.6 µM. Los biosensores SSA sin minihélice biotinilada fueron usados para medir unión inespecífica de los compuestos al biosensor. Todos los demás parámetros (amortiguador, temperatura, control del instrumento y análisis) fueron como se describió anteriormente.

Optimización del fragmento 2-MBX guiado por estructura. Con base en el modelo de la estructura cristalográfica de la proteína P de *T. maritima* en complejo con el fragmento 2-mercaptobenzoxazol, se llevó a cabo un estudio de optimización del "hit" basado en estructura con el fin de obtener un ligando más afín y específico a la proteína P. Para ello, se llevó a cabo un análisis con el módulo de programa QikProp (Maestro programa, Versión. 9.3. (2012). New York, NY: Schrödinger, LLC), el cual predice una amplia variedad de propiedades farmacológicas relevantes del "hit" como el coeficiente de partición octanol/agua y agua/gas, actividad sobre el sistema nervioso central y enzimas humanas, grado de

permeabilidad a través de membranas celulares, IC_{50} teórico para bloqueo de HERG K de tal manera que decisiones adecuadas puedan llevarse a cabo al analizar un “hit” unido a un blanco. Posteriormente, se utilizó la suite de programas MAESTRO (Schrödinger) la cual está enfocada en diseño, optimización y evaluación de fármacos basado en fragmentos o moléculas pequeñas. Para ello, se utilizó el programa SiteMap (Schödinger) el cual localiza sitios de unión para moléculas pequeñas, califica y jerarquiza sitios de unión de acuerdo a la tendencia de un sitio para unir ligandos y elimina aquellos sitios que no son relevantes farmacológicamente, es decir, predice la “drogabilidad” de un blanco (véase *Apéndice E. Glosario de términos*) de esos sitios y ayuda en el proceso de optimización de un “hit” al proveer información acerca de interacciones ligando-receptor y sugerir estrategias para modificar compuestos líder para aumentar y mejorar la complementariedad del ligando al receptor. Por último, se llevó a cabo un análisis con ayuda del módulo Phase (Schödinger), el cual es un programa que ayuda a crear hipótesis para la optimización a partir de complejos proteína-ligando aplicando metodología de farmacóforos o a partir de interacciones ligando-proteína observadas en una estructura cristalográfica.

Construcción de la micro-biblioteca de derivados del fragmento 2-MBX. Se buscaron compuestos basados en la estructura del fragmento 2-MBX de acuerdo con el buscador por estructura en el servidor de la compañía Sigma-Aldrich que correspondiera con el 70% de identidad estructural y que tuvieran grupos químicos funcionales a partir del Nitrógeno 5 del anillo oxazol. A partir de esta búsqueda, se obtuvieron 11 compuestos que cumplían con las características deseadas, los cuales fueron adquiridos de la misma compañía.

Ensayos de compuestos optimizados derivados del fragmento 2-MBX mediante BLI. Los 11 compuestos elegidos se probaron a una concentración de 200 μ M y la proteína TmPwt-His se usó a 0.05 $\text{mg}\cdot\text{mL}^{-1}$ (3.7 μ M). Biosensores Ni-NTA (FortéBio, Inc.) fueron hidratados por 1 h en amortiguador PBS 1X, 5% DMSO previamente a los experimentos. Asimismo, biosensores control en los que no se inmovilizó la proteína P fueron usados como medio para substrar la unión no específica de los compuestos hacia la superficie del biosensor. Los ensayos de unión en tiempo real se llevaron a cabo en un sistema OCTET RED96 (FortéBio, Inc.) para interferometría de biocapas y las condiciones de ensayo fueron 1000 rpm de agitación orbital, $T= 30^{\circ}\text{C}$ y un volumen de pocillo de 200 μ L en placas de 96 pocillos de fondo negro (Greiner bio-one). Los tiempos para cada paso fueron: inmovilización de ligando 300 s, lavado 100 s, línea base 50 s, asociación de analito 50 s y disociación de analito 50 s y los parámetros cinéticos de unión fueron calculados con ayuda del programa Data Analysis 8.2. Como criterio de elección de “hits” en el laboratorio se estableció que la respuesta BLI de unión fuera mayor a 1 nm, tuviera k_{on} (o k_a) alta y una mínima disociación k_{off} (o k_d). Las gráficas de los sensorgramas fueron generadas y editadas con ayuda del programa GraphPad Prism 6.

Ensayos de unión de la proteína P silvestre y mutantes con compuestos optimizados mediante termoforesis a microescala. Con el fin de comparar y confirmar la robustez y confiabilidad de los parámetros cinéticos de unión obtenidos mediante interferometría de biocapas de la proteína P con los compuestos optimizados C-7 y C-10, así como con el ARN 5'-líder, se procedió a llevar a cabo experimentos de titulación de la proteína P con estos ligandos mediante termoforesis a microescala.

A) *Búsqueda de capilares óptimos para la proteína P.* Después de haber preparado las soluciones seriales de los compuestos optimizados C-7 y C-10 y mezclarlos con la proteína P etiquetada con el colorante fluorescente NT-647, las muestras fueron cargadas en capilares y colocados en la bandeja de

capilares. La primera medición que el instrumento lleva a cabo se conoce como “Capillary Scan” o *Cap Scan*. El instrumento escanea la fluorescencia a través de la longitud de la bandeja de capilares. Por lo tanto, los capilares que contienen moléculas etiquetadas con fluoróforo resultan en picos de fluorescencia. Este procedimiento asegura una determinación exacta de la posición de cada capilar previamente a que la medición mediante MST inicie. A través de la selección automática de los picos de fluorescencia, la posición exacta para cada capilar individual es anotada automáticamente después de 60 s del *Cap Scan*. Con el fin de determinar la concentración adecuada de la proteína P etiquetada con colorante fluorescente NT-647, y para identificar el tipo de capilar de vidrio adecuado para esta muestra en la cual no sea observada absorción inespecífica a la pared del capilar, una prueba previa fue llevada a cabo. Para ello, la proteína P etiquetada fue diluida en el amortiguador de ensayo para MST (en proporción 1:2) y cargada en tres diferentes tipos de capilares para posteriormente llevar a cabo un *Cap Scan*. Asimismo, la curva MST fue registrada para evaluar la presencia de agregados o heterogeneidad.

B) Termoforesis a microescala. Para un experimento típico por MST, se requiere un compañero fluorescente que interaccione (el blanco) y otro compañero no fluorescente que interaccione (el ligando). La fluorescencia puede ser proveída por una etiqueta extrínseca (p.ej. colorante NT-647) covalentemente unida a la molécula de interés, una proteína de fusión (GFP) o a través de fluorescencia intrínseca del triptófano. La concentración de la molécula fluorescente es mantenida constante y baja (típicamente 50 pM, 10 nM y 10 nM para el NT.115^{Pico}, NT.115 y el NT.LabelFree, respectivamente). La molécula no fluorescente (ligando) es entonces variada en su concentración y 16 muestras fueron preparadas a diferentes concentraciones. Después de mezclar los agentes interactuantes, las muestras se incubaron brevemente durante 5 min a T ambiente para alcanzar el equilibrio. Posteriormente, se cargaron las muestras en los capilares mediante acción de capilaridad y se cargaron en la bandeja de capilares del equipo. En el experimento típico por MST con etiquetas, una dilución serial de hasta 16 diluciones fue preparada. Aquí, la concentración de la molécula fluorescente etiquetada fue mantenida constante entre 50-300 nM y la concentración de ligando se varió. Por lo tanto, 2-20 μ M de la proteína P tipo silvestre o mutante fue etiquetada usando el colorante fluorescente NT-647 con ayuda del kit Monolith NT Protein Labeling (reactivo a aminas). El procedimiento de etiquetado y la remoción subsecuente de colorante libre sin reaccionar mediante columna de exclusión de tamaño molecular fue llevado a cabo dentro de 45 min. El compuesto optimizado C-4 fue usado como control negativo en estos experimentos ya que en experimentos mediante BLI no se observó alguna evidencia de unión a la proteína P. Una dilución serial de los compuestos C-4, C-7 y C-10 fue preparada usando el amortiguador PBS 1X pH 7.4 adicionado con 500 mM NaCl y 1 mg/mL BSA. Para la dilución serial, la concentración más alta fue usada al menos 20 veces más alta que la K_D esperada. 10 μ L de la dilución serial de la molécula no etiquetada fue mezclada con 10 μ L de la molécula etiquetada con fluoróforo. Las muestras mezcladas fueron cargadas en capilares de vidrio y el análisis de MST fue llevado a cabo usando el instrumento Monolith NT.115 (Nanotemper Technologies).

Para llevar a cabo experimentos mediante termoforesis a microescala sin etiqueta fluorescente alguna, se utilizó el instrumento Monolith NT.LabelFree el cual mide la fluorescencia intrínseca de los aminoácidos aromáticos de la proteína P como el triptófano. Una dilución serial de 16 diluciones fue preparada donde la concentración de la molécula con triptófano (proteína P), fue mantenida constante a 50-500 nM y la concentración del ligando fue variado. Una solución inicial de la proteína P que contiene triptófano fue preparada y se eligió una concentración a la cual la señal de fluorescencia fuera

significativamente más alta que aquella del amortiguador usado (ruido de fondo). Una dilución serial del ligando no fluorescente fue preparada usando el amortiguador optimizado para MST (50 mM Tris pH 7.6, 150 mM NaCl, 10 mM MgCl₂ adicionado con 0.1% del surfactante no iónico hidrofílico Pluronic F-127). La concentración más alta de esta dilución elegida fue 20 veces más alta que la K_D esperada. 10 μ L de la dilución serial de los compuestos optimizados fueron mezclados con 10 μ L de la proteína P intrínsecamente fluorescente. La mezcla de las muestras fueron cargadas en capilares de vidrio y el análisis por MST fue llevado a cabo en un equipo para termoforesis a microescala Monolith NT.LabelFree (Nanotemper Technologies). Para ensayos mediante MST para medir interacción entre la proteína P marcada con etiqueta fluorescente NT-647 y el ARN 5'-líder, se utilizó el instrumento Monolith NT.115^{pico} (Nanotemper Technologies).

Ensayo de inhibición de crecimiento por difusión en placa. Para estos ensayos, se utilizó kanamicina como control positivo a una concentración final de 10 mM, mientras que la concentración final utilizada para los compuestos optimizados C-7 y C-10 fue de 25 mM. Los compuestos optimizados a probarse, así como los antibióticos usados como controles positivos, fueron absorbidos en discos de celulosa de 0.5 cm de diámetro y colocados de manera equidistante dentro de una placa Petri con medio de cultivo LB suplementado con glucosa o arabinosa con el fin de inducir ya sea la RNasa P bacteriana o la PRORP de *A. thaliana*. Este procedimiento se llevó a cabo una vez que 50-200 μ L de cepas *E. coli* BW (97, 98), fueran plaqueadas en la caja Petri. Las condiciones de incubación fueron 37°C durante 16 h, y al final de este periodo se analizó la caja Petri para determinar la presencia de halos de inhibición alrededor de los discos de celulosa adicionados con inhibidores o antibiótico. Un disco de DMSO 100% fue aplicado a la caja con cultivo como control para evaluar el efecto del DMSO en el crecimiento de las bacterias.

Búsqueda de compuestos con la biblioteca The Spectrum Collection™ por interferometría de biocapas. Para la búsqueda de moléculas pequeñas que interaccionaran con la proteína P, se utilizó la biblioteca comercial de compuestos The Spectrum Collection™ (MicroSource Discovery Systems Inc.) que comprende 2560 compuestos patentados y bioactivos. La biblioteca Spectrum se preparó en un formato de cocteles de ocho compuestos a una concentración final de 200 μ M por compuesto en 16% DMSO final. Biosensores Ni-NTA (FortéBio, Inc.) fueron hidratados por 1 h en amortiguador PBS 1X pH 7.4, 5% DMSO previamente a los experimentos. Los ensayos de unión en tiempo real se llevaron a cabo en un sistema OCTET RED96 (FortéBio, Inc.) para interferometría de biocapas y las condiciones de ensayo fueron 1000 rpm de agitación orbital, T= 30°C y un volumen de pocillo de 200 μ L en placas de 96 pocillos de fondo negro (Greiner bio-one). Los tiempos para cada paso fueron: inmovilización de ligando 300 s, lavado 100 s, línea base 50 s, asociación de analito 50 s y disociación de analito 50 s. Los datos experimentales fueron monitoreados con ayuda del programa Data Acquisition 8.2 mientras que los parámetros cinéticos de unión fueron calculados con ayuda del programa Data Analysis 8.2. La proteína P con etiqueta de histidinas se inmovilizó en los biosensores Ni-NTA a 0.1 mg/ml (7.4 μ M), con tiempos para los pasos como se describe: inmovilización (300s), lavado (100s), línea base (50s) y asociación-disociación (250s).

Para llevar a cabo la deconvolución e identificación de “hits” derivados de la campaña de búsqueda en la biblioteca Spectrum, los tiempos de cada etapa se configuraron como sigue: inmovilización (300 s), lavado (100 s), línea base (50 s), asociación (50 s) y disociación (50 s); mientras que los parámetros del experimento fueron; T= 30°C, agitación orbital 1000 rpm, el volumen de cada pocillo fue de 200 μ L y la

concentración de inmovilización de la proteína P fue de 0.1 mg/mL (7.4 μ M). Aquellos cocteles de ocho compuestos en los cuales se identificó una señal significativa de unión, fueron divididos posteriormente en cocteles de cuatro compuestos con una concentración individual final de 50 μ M, los cuales, a su vez, una vez identificada una señal de unión correspondiente, fueron divididos en cocteles de dos compuestos hasta identificar el compuesto que mostró señal de unión a la proteína P. Los criterios para identificar un “hit” fueron los siguientes: 1) que tuviera un límite inferior mínimo de 3 nm de respuesta ya que la mayoría de los compuestos en la biblioteca presentaron una señal de 1 nm o menos; 2) fueron descartadas aquellas señales con velocidades de disociación (k_{off}) demasiado rápidas al observar los sensorgramas; 3) se consideraron únicamente velocidades de asociación (k_{on}) muy rápidas que fueran acompañadas de velocidades de disociación lentas; y 4), se evaluó la forma de las curvas del sensorgrama y se descartaron aquellas que no presentaron geometrías canónicas hiperbólicas Michaelianas. Al final de la campaña de búsqueda de compuestos, se identificaron 30 “hits” de los 2560 compuestos totales de la biblioteca Spectrum, que equivalen al 1.17% de éxito, y los cuales se distribuyeron en un intervalo amplio de afinidades desde altas (llamados en el presente trabajo como “hits” primarios) a bajas (llamados en este trabajo como “hits” secundarios).

Búsqueda de compuestos con la biblioteca TimTec FL-500™ por interferometría de biocapas. La biblioteca de compuestos TimTec FL-500™ (TimTec LLC Drug Discovery), la cual comprende 500 compuestos entre los que se encuentran flavonoides, flavonas e isoflavonoides y sus derivados, fue usada para la búsqueda de ligandos específicos de la proteína P mediante interferometría de biocapas.

La proteína P con etiqueta de histidinas se inmovilizó en biosensores Ni-NTA a 0.1 mg/mL (7.4 μ M) en amortiguador PBS 1X pH 7.4 adicionado con 5% DMSO en un volumen de 200 μ L, mientras que los compuestos fueron ensayados en una búsqueda inicial en cocteles de ocho compuestos a una concentración final por compuesto de 100 μ M a partir de una solución inicial por compuesto de 8 mM. La concentración final de DMSO en el ensayo fue de 10% máximo y los parámetros de experimentación fueron T= 30°C, agitación orbital 1000 rpm y los tiempos de cada paso fueron como sigue: inmovilización (300 s), lavado (100 s), línea base (50 s), asociación (50 s) y disociación (50 s).

Para llevar a cabo la deconvolución e identificación de “hits” derivados de la campaña de búsqueda en la biblioteca TimTec, los tiempos de cada etapa se configuraron como sigue: inmovilización (300 s), lavado (100 s), línea base (50 s), asociación (50 s) y disociación (50 s); mientras que los parámetros del experimento fueron: T= 30°C, agitación orbital 1000 rpm, el volumen de cada pocillo fue de 200 μ L y la concentración de inmovilización de la proteína P fue de 0.1 mg/mL (7.4 μ M). Aquellos cocteles de ocho compuestos en los cuales se identificó una señal significativa de unión, fueron divididos posteriormente en cocteles de cuatro compuestos, los cuales, a su vez, una vez identificada una señal de unión correspondiente, fueron divididos en cocteles de dos compuestos hasta identificar el compuesto que mostró señal de unión a la proteína P.

Ensayo de actividad RNasa P basado en gel. Para llevar a cabo en ensayo de actividad en gel, se mezcló holoenzima a una concentración final de 300 nM con fragmentos moleculares o inhibidores a ensayar. Los fragmentos usualmente se mezclan en concentraciones molares de entre 10,000 a 50,000 veces más altas que la holoenzima y se lleva la mezcla a 10 μ L de volumen total con amortiguador TH 1X. Posteriormente, se añade el sustrato minihélice fluorescente a una concentración final de 4 μ M inmediatamente antes de comenzar el tiempo de incubación para la reacción. Para reacciones con pre-

tRNA canónico, una concentración final de 5 μM debe ser añadida; estas concentraciones, en un volumen de 10 μL , son las cantidades mínimas requeridas que pueden ser detectadas en geles bajo luz UV. Las condiciones estándar de la reacción fueron 55°C durante 10 min. Las condiciones para el sustrato minihélice fluorescente fueron 37°C durante 10 min, mientras que las condiciones al usar el sustrato canónico pre-tRNA fueron 37°C durante 10 min y para la subunidad catalítica de ARN fueron 50°C durante 25 min. La reacción se detuvo añadiendo a la mezcla de reacción 5 μL de amortiguador de carga para ARN (urea 10 M, Tris 89 mM, ácido bórico 89 mM, EDTA 20 mM y bromophenol blue 1% (P/V)). Para la visualización del gel se usó un gel de poli(acrilamida) desnaturante al 18% (acrilamida 18%, urea 7 M, Tris 89 mM, ácido bórico 89 mM, EDTA 20 mM, APS 10% y 40 μL de TEMED). Se corrió a 250-300 V durante 30-60 min. Las bandas se visualizaron bajo luz UV sobre un fondo oscuro para visualizar el sustrato minihélice fluorescente o ARN con bromuro de etidio a 10 mg/mL.

Ensayo de actividad RNasa P bacteriana mediante espectroscopía de fluorescencia. Las reacciones fueron llevadas a cabo en un volumen total de 20 μL . El pipeteo fue llevado a cabo a mano usando pipetas multicanal. Primeramente, una mezcla de reacción de 18 μL que contenía 16.5 μL de amortiguador de actividad (amortiguador H para sustrato Mh o amortiguador L adicionado con 10% PEG-200 para el pre-tRNA-mitades) y 270 nM (Mh) o 135 nM (pre-tRNA-mitades) (1.5 o 0.7 μL de 3.6 μM de la solución inicial) de la sonda sustrato fluorescente fue incubada previamente en hielo durante 5 min en una microplaca de poliestireno de 384 pocillos con un fondo transparente (Corning ref 3540). La holoenzima fue añadida a una concentración final de 33 nM (2 μL de una solución inicial de 333 nM) en amortiguador TH 1X y mezclado. La cinética de actividad fue monitoreada en un detector para microplacas Synergy MX Multimode Detector (BioTek) a 37°C en régimen de agitación rápida. Al menos 20 lecturas fueron tomadas por pozo cada 15 s. La sonda fue excitada a 485 nm, y su emisión fue monitoreada a 535 nm con un tiempo de integración de 1 s. A partir de las gráficas de Unidades Relativas de Fluorescencia (RFU) vs tiempo, la regresión lineal fue usada para determinar las pendientes en la porción lineal de las curvas, usando al menos 12 puntos de datos para obtener las velocidades iniciales. El ajuste no lineal para obtener los parámetros Michaelis-Menten fue llevado a cabo usando el software GraphPad Prism® 6.01 programa.

Ensayo de inhibición de la RNasa P *in vivo*. Con el fin de llevar a cabo experimentos para analizar inhibición diferencial de la RNasa P bacteriana *in vivo* por efecto de inhibidores durante curvas de muerte vs tiempo, se utilizó la cepa *E. coli* BW, la cual expresa o no expresa la RNasa P dependiendo de la condición de crecimiento. Específicamente, la expresión del gen *rnpB* (el cual codifica para la subunidad de ARN catalítica de la RNasa P) en la cepa *E. coli* BW, depende estrictamente de la presencia de arabinosa en el medio de cultivo como inductor (97), mientras que, en presencia de glucosa, en lugar de la arabinosa como fuente de carbono, la expresión cromosomal del gen *rnpB* es apagada en la cepa *E. coli* BW, y en ese momento, el crecimiento celular requiere la expresión (basada en vector de expresión) de otro gen que provea la función de RNasa P. La enzima con actividad de RNasa P basada únicamente en proteína llamada PRORP, de *Arabidopsis thaliana*, tiene la habilidad de reemplazar funcionalmente *in vivo* a la actividad de la RNasa P de *E. coli*, y, por tanto, permite el rescate del crecimiento de la cepa la *E. coli* BW a 37°C bajo condiciones no permisivas en placas o medio LB adicionadas con glucosa (98).

Escaneo de la biblioteca The Spectrum Collection™ mediante actividad por fluorescencia. Se trabajó con la biblioteca The Spectrum Collection™ (MicroSource Discovery Systems, Inc), la cual comprende 2560 compuestos, cada uno disuelto en DMSO al 100% a 10 mM. Para la búsqueda inicial, los compuestos de la biblioteca fueron probados en cocteles de ocho compuestos en un formato de alta eficiencia en microplacas de 384 pocillos. La concentración final de cada compuesto en el coctel fue de 1.25 mM. Para el ensayo de actividad, las reacciones fueron configuradas para un volumen total de 20 μL (13.5 μL de amortiguador de actividad, 1.5 μL de sustrato Mh fluorescente a 3.6 μM , 3 μL de DMSO al 100% o coctel). Este procedimiento resultó en una concentración final de of 187.5 μM por cada compuesto individual. Los pozos control incluidos para cada serie de experimentos fueron: 1) la reacción con el doble de volumen de la Mh fluorescente (3 μL) para evitar la señal de saturación para el resto de los pozos; 2) dos reacciones con 10% de DMSO (sin coctel de compuestos); 3) sustrato artificial fluorescente Mh únicamente sin ribozima, y 4) reacción con 100 mM CaCl_2 . La reacción se inició mediante la adición de 2 μL de la holoenzima a 333 nM y monitoreados como se describe en la sección previa. Cada ronda de experimentos contuvo 20 diferentes cocteles correspondientes en una fila de la microplaca de 384 pocillos. Los cocteles que mostraron inhibición (considerados como aquellos con un valor de pendiente más bajo que cualquiera de los controles positivos en le misma placa) fueron verificados y deconvolucionados para determinar los fragmentos individuales responsables por la inhibición observada. En esta etapa, cada compuesto proveniente del coctel que mostró inhibición fue examinado a una concentración final de 1 mM (2 μL de la solución inicial de 10 mM). La correlación entre el ensayo basado en fluorescencia y el ensayo basado en gel para los nueve compuestos identificados en esta etapa fue llevada a cabo a 250 y 500 μM , respectivamente.

Cristalización de la proteína P apo. Para llevar a cabo la cristalización de la proteína P, se realizó una búsqueda de las condiciones idóneas mediante el método de difusión de vapor en modalidad de gota sedente. Las soluciones, concentraciones, así como los porcentajes y pH, fueron:

- Acetato de Sodio (NaOAc), 100 mM, pH = 4.8 - 5.2
- Polietilenglicol (PEG) 1000 = 12-18%
- Sulfato de Potasio (K_2SO_4) = 200 mM
- Proteína P = 2-14 mg/mL

La búsqueda de las condiciones ideales de cristalización dentro del estos intervalos se realizó en placas de cristalización Swissci™ de 96 pocillos. En estos experimentos se colocó una gota con 1 μL de proteína a 3 mg/mL y 1 μL de solución madre añadiendo 40 μL de la misma en el reservorio, y fue sellado con una película plástica (HR3-609, Crystal Clear Sealing Film). En esta configuración, los experimentos se llevaron a cabo a 23°C. El licor madre detectado en el que crecieron los mejores cristales de proteína P apo tiene la siguiente formulación:

- Acetato de Sodio (NaOAc), 100 mM, pH = 4.8
- Polietilenglicol (PEG) 1000 = 12%
- Sulfato de Potasio (K_2SO_4) = 200 mM
- Proteína P = 3 mg/mL

En estas condiciones, los cristales crecieron dentro de una semana hasta alcanzar dimensiones aproximadas de 300 μm x 200 μm x 90 μm . Previo al sobre enfriamiento en nitrógeno líquido, los cristales se sumergieron durante 1 min en una solución crioprotectora que tuvo la siguiente formulación:

- Polietilenglicol-1000, 35%
- Acetato de Sodio trihidratado (NaOAc), pH 5.2, 100 mM
- Sulfato de Potasio (K_2SO_4), 20 mM
- Ditiotreitól (DTT), 1 mM

Experimentos de remojo de cristales de la proteína P con purpurina. Los cristales se obtuvieron a partir de la proteína P con etiqueta de hexa-histidinas en las mismas condiciones que la proteína P apo sin etiqueta de histidinas. Se preparó una solución de purpurina inicial de 100 mM con PEG-400, 50%; se agitó en un vortex en presencia de perlas de vidrio y se centrifugó a 14000 rpm durante 10 min. El sobrenadante fue usado para diluciones subsecuentes. La composición de la solución crio-protectora para equilibrar el cristal fue: PEG-1000 30%, 20 mM K_2SO_4 , 100 mM NaOAc trihidratado pH 4.8 y 1 mM DTT; mientras que para experimentos de remojo fue la misma solución, pero adicionada con purpurina 20 mM. El método de equilibrio consistió en retirar todo el licor madre en el que creció el cristal y sustituirlo por la solución crio-protectora sin purpurina durante 24 h. Se dejó estabilizar o equilibrar por 24 h y después esta solución crio-protectora fue sustituida 3 veces (1 vez por día) con solución crio-protectora adicionada con 20 mM de purpurina para incrementar gradualmente y gentilmente la concentración neta de ligando dentro del cristal apo. A esta solución crio-protectora, 20 mM de purpurina fueron añadidos para llevar a cabo experimentos de remojo, de esta forma disminuyendo la concentración final de todos los demás componentes en la solución crio-protectora. Para experimentos de remojo con purpurina, después de 24 h de equilibrar con la solución crioprotectora sin ligando, 2 μL de esta solución fueron retirados de la gota del cristal y sustituidos por 2 μL de solución crio-protectora adicionada con purpurina 20 mM. En total, se llevaron a cabo tres sustituciones de la solución crio-protectora, y al tercer día, la totalidad de la solución crio-protectora con 20 mM de purpurina fue retirada y sustituida o repuesta con solución fresca de solución crio-protectora con purpurina. El tiempo total que se dejó el cristal apo en remojo a partir de la tercera y última sustitución de licor madre por solución crio-protectora adicionada con purpurina fue de 2 meses y 23 días. La temperatura durante los experimentos de remojo fue de 25°C durante la noche y de 30°C durante el día.

Colecta de datos mediante difracción de rayos X y determinación de la estructura tridimensional.

Posterior a la cristalización de la proteína P apo, o después de experimentos de remojo con compuestos, los cristales de la proteína P fueron remojados con 8 μL de solución crio-protectora con fragmentos individuales durante 4 min, e inmediatamente después, los cristales fueron ultra-enfriados y almacenados en nitrógeno líquido hasta su análisis. Los datos de difracción fueron colectados a partir de monocristales de proteína P a una $T=100$ K y a una $\lambda=0.9785$ Å en la línea 21-ID-G del Life Sciences Collaborative Access Team (LS-CAT) en el Advanced Photon Source (Argonne National Laboratory, Argonne, Illinois, USA). Los datos fueron indexados y procesados con el programa XDS (99) y la reducción de datos fue llevada a cabo con ayuda del programa Aimless (100). Como principio de prueba, y dado que en la estrategia de búsqueda de fragmentos es esperado que se detecten ligandos unidos con baja ocupación, se aplicó el método Pan-Dataset Density Analysis method (PanDDA) (101). Esta metodología lleva a cabo un análisis estadístico de múltiples conjuntos de datos cristalográficos, lo

cual permite contrastar la densidad electrónica de diferentes estados de fondo, y localiza regiones que presentan un evento de unión en el mapa parcial de diferencia (p. ej. un potencial ligando unido). En total, se analizaron conjuntos de datos de un cristal de referencia (sin remojo de compuesto) y 24 cristales (cada uno remojado en diferentes compuestos individuales de la biblioteca Zenobia Fragment Library-1TM), los cuales difractaron a 1.5 Å. La estructura de referencia fue resuelta mediante reemplazo molecular usando el programa Phaser (102), con un modelo de búsqueda basado en la proteína P de *Thermotoga maritima* (código PDB: 1NZ0). Se observaron cuatro monómeros en la unidad asimétrica, con un contenido de disolvente en el cristal del 41%. El resto de los 24 conjuntos de datos fueron afinados contra esta solución, la cual tuvo como resultado valores iniciales de R_{free} de entre 0.20-0.22.

Posteriormente, se corrió el programa PanDDA de acuerdo a como está implementado en CCP4 (103) y se detectaron ocho conjuntos de datos con un potencial evento de unión. Tres conjuntos de datos presentaron eventos de unión cerca de los residuos Lys90 y Lys53, y se eligieron conjuntos de datos con la densidad electrónica más limpia de los ocho conjuntos de datos con eventos de unión potenciales y se identificó el correspondiente fragmento, que fue remojado, unido a una molécula (cadena D) de las cuatro presentes en la unidad asimétrica. El fragmento detectado es el compuesto 1-4-Fluorofenil-2-tiourea (nombrado en adelante como FTU) con coordenadas dentro de la biblioteca Zenobia fragment Library-1TM, Placa-4, Fila-D, Columna-8. De acuerdo con el fabricante, el número de identificación del compuesto corresponde al ZT0773, tiene una fórmula lineal $\text{FC}_6\text{H}_4\text{NHCSNH}_2$ y un peso molecular de 170.21 Da. El ligando FTU fue ajustado manualmente y los límites y afinamiento fue generado por medio del programa eLBOW de la suite PHENIX (104), usando el código SMILES C1=CC(=CC=C1NC(=S)N)F. El afinamiento fue alternado con ciclos manuales de construcción y afinamiento en COOT (105), PHENIX y el servidor PDB_REDO (106). El factor B del ligando es alto (70.5 Å² vs 55.9 Å² y 60.9 Å² para los residuos circundantes Lys53 y Lys90, respectivamente) y la densidad del ligando es inequívoca y clara en el mapa 2mFo-DFc a bajo nivel de contorno de 0.5 σ . La presencia del FTU fue validada mediante un mapa de omisión (mapa polder) (107) como se encuentra implementado en la suite PHENIX. El coeficiente de correlación (CC) obtenido fue de 0.78, así como una ocupación de 1.0, valores calculados con ayuda del programa MOLPROBITY como se encuentra implementado en la Suite PHENIX. El modelo resultante no presenta residuos en zonas prohibidas del mapa de Ramachandran. Las coordenadas atómicas y los factores de estructura de la proteína P de *T. maritima*, en complejo con el fragmento 1-4-fluorofenil-tiourea, han sido depositados en el Protein Data Bank (PDB) con el código de acceso 6CQC.

Con respecto a la estructura cristalográfica de la proteína P de *T. maritima* en complejo con el inhibidor purpurina, se observa un monómero en la unidad asimétrica, con un contenido de disolvente en el cristal del 40%. La presencia de la purpurina fue confirmada inicialmente mediante mapas de diferencia Fourier, que fueron calculados usando la estructura de la proteína P apo (Código de acceso PDB: 1NZ0). Los ligandos fueron manualmente ajustados y las restricciones para el afinamiento fueron generados mediante el programa eLBOW de la suite PHENIX (104) usando el código SMILES C1=CC=C2C(=C1)C(=O)C3=C(C2=O)C(=C(C=C3O)O)O para la purpurina. El afinamiento fue alternado con ciclos de construcción y afinamiento manual en COOT (105), PHENIX y el servidor PDB_REDO (106). Para la purpurina, el factor B del ligando es de 42.5 Å² vs. 16.0 Å², 25.4 Å² y 21.1 Å² para los residuos circundantes Val33, Gln28 y Arg89, respectivamente, y la densidad electrónica del ligando es evidente en el mapa 2mFo-DFc a un nivel de contorno de 1.5 σ . Mapas de omisión 2Fo-Fc de

recocido simulado σA -weighted fueron usados para validar la presencia del ligando. La ocupación de la purpurina es de 1.0 y su coeficiente de correlación (CC) es de 0.87, valores calculados con ayuda del programa MOLPROBITY de acuerdo a como se encuentra implementado en la Suite PHENIX. Las figuras fueron preparadas con Pymol (The PyMOL Molecular Graphics System, Versión 1.7.4, Schrödinger, LLC). Las coordenadas atómicas y los factores de estructura de la subunidad proteica de la RNasa P de *T. maritima* en complejo con la purpurina han sido depositadas en el Protein Data Bank (PDB) con el código de acceso 6MAX.

7. Resultados

Con el fin de lograr un mayor enfoque y claridad en esta sección sobre los resultados principales, se ha incluido en el “Apéndice C. Resultados suplementarios”, los resultados sobre: 1) la expresión y purificación de biomacromoléculas, 2) el establecimiento del ensayo de actividad medido por fluorescencia por nuestro grupo de investigación y su optimización, 3) el establecimiento de los ensayos de unión, y 4) la cristalización de la proteína P.

Los resultados se dividen en las aproximaciones estructurales, de medición de actividad, medición de unión y de cómo se lograron integrar estas estrategias en la obtención de un inhibidor que tiene un sitio de unión en la proteína P.

7a) Enfoque sobre cristalografía de rayos X

La re-evaluación de la búsqueda de eventos de unión por cristalografía de rayos X de la biblioteca Zenobia Fragment Library-1TM resulta en la detección de un nuevo ligando

Dado que logramos tener un éxito inicial en la búsqueda de fragmentos a través de la cristalografía de rayos X, realizamos varios intentos de obtener complejos utilizando compuestos optimizados como el C-7 o C-10, sin lograrlo. (véase *Resultados. Enfoque sobre ensayos de unión*). Una estrategia a partir de la cual logramos detectar otro fragmento fue al retomar la búsqueda por cristalografía de la biblioteca Zenobia partiendo de los juegos de datos colectados previamente, en donde los cristales fueron remojados en cocteles de 8 compuestos. Esta etapa del proyecto, coincidió con el surgimiento de una nueva técnica en cristalografía de rayos X para detectar eventos débiles de unión o con baja señal, adecuada para fragmentos llamada PaNDDA (véase *Métodos. Colecta de datos mediante difracción de rayos X y determinación de la estructura tridimensional*). De esta manera, se logró detectar el fragmento 1-(4-fluorofenil)-2-tiourea (FTU), el cual se localiza exactamente en el mismo sitio de unión del fragmento 2-MBX (Figuras 8, 9). El fragmento FTU interacciona con la proteína P mediante un puente de hidrógeno con la cadena principal de la Pro88, mientras que el anillo del fragmento se encuentra flanqueado por Lys53 y Lys90 (Figuras 8, 9), de igual manera que el “hit” 2-MBX. Los datos de colecta y afinamiento de datos se muestran en la Tabla 2.

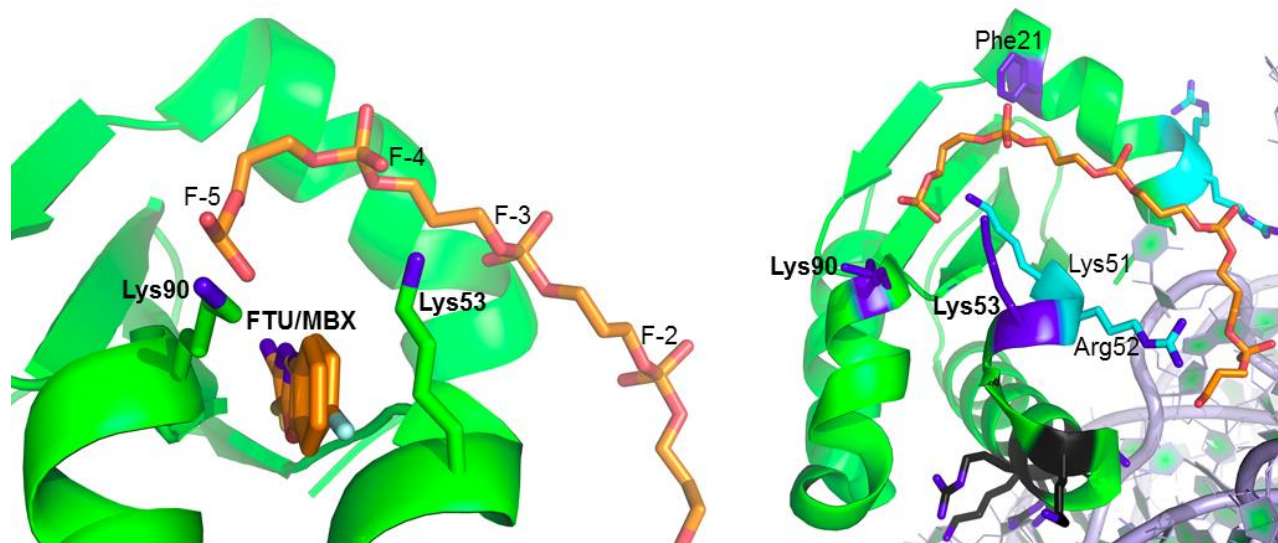


Figura 8. Dos fragmentos se unen en un sitio de la proteína P importante para la función de la holoenzima. Izquierda. Modelo de la estructura cristalográfica de la proteína P en complejo con la superposición de los fragmentos 2-mercaptobenzoxazol (2-MBX) y 4-fluorofenil-tiourea (FTU). Tanto el 2-MBX como el FTU se unen en la misma cavidad, muy cerca de la zona de interacción con la base -5 del extremo 5'-líder que se muestra en naranja como cadena fosfatada. Ambos fragmentos comparten interacciones de apilamiento con la proteína P mediante los residuos Lys53 y Lys90 que se encuentran flanqueando a los fragmentos. **Derecha.** Residuos mutados en la proteína en un estudio independiente (58). La mutación de los residuos Lys53 y Lys90 por Ala disminuye cinco veces la actividad de la holoenzima.

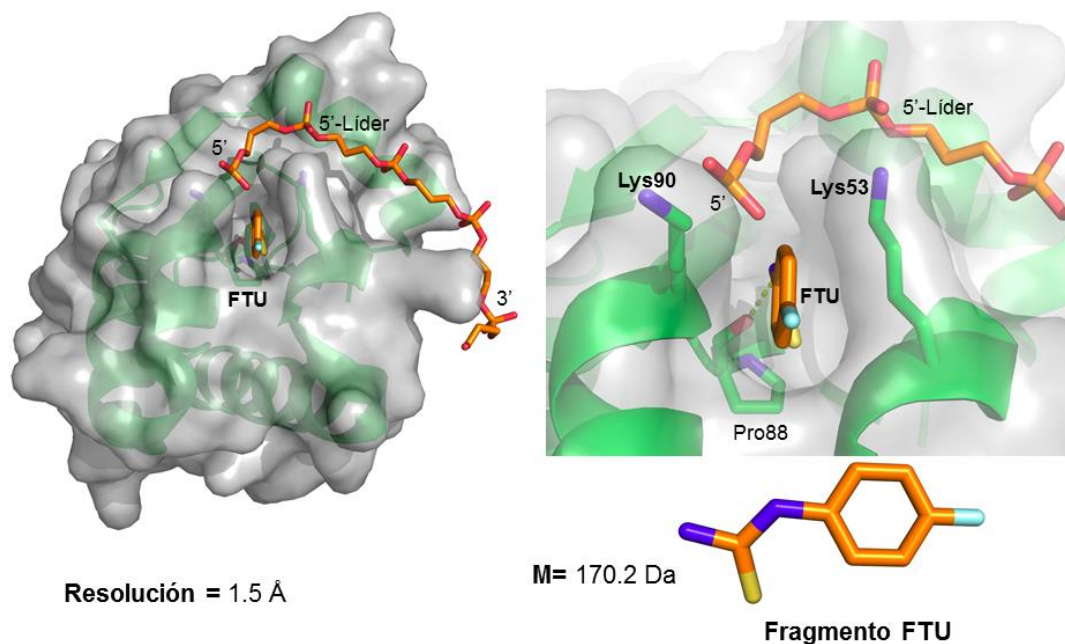


Figura 9. El fragmento FTU se une en una zona importante para la proteína P. Izquierda. Sitio de unión del fragmento FTU en el contexto de la proteína P a 1.5 Å de resolución. **Derecha.** Modo de unión del fragmento FTU donde se muestran las interacciones del FTU con la proteína P. Se observan interacciones del FTU con los residuos Lys90 y Lys53, así como un puente de hidrógeno con el grupo carboxilo de la cadena principal de la Pro88. En la parte inferior derecha se muestra la estructura del fragmento FTU.

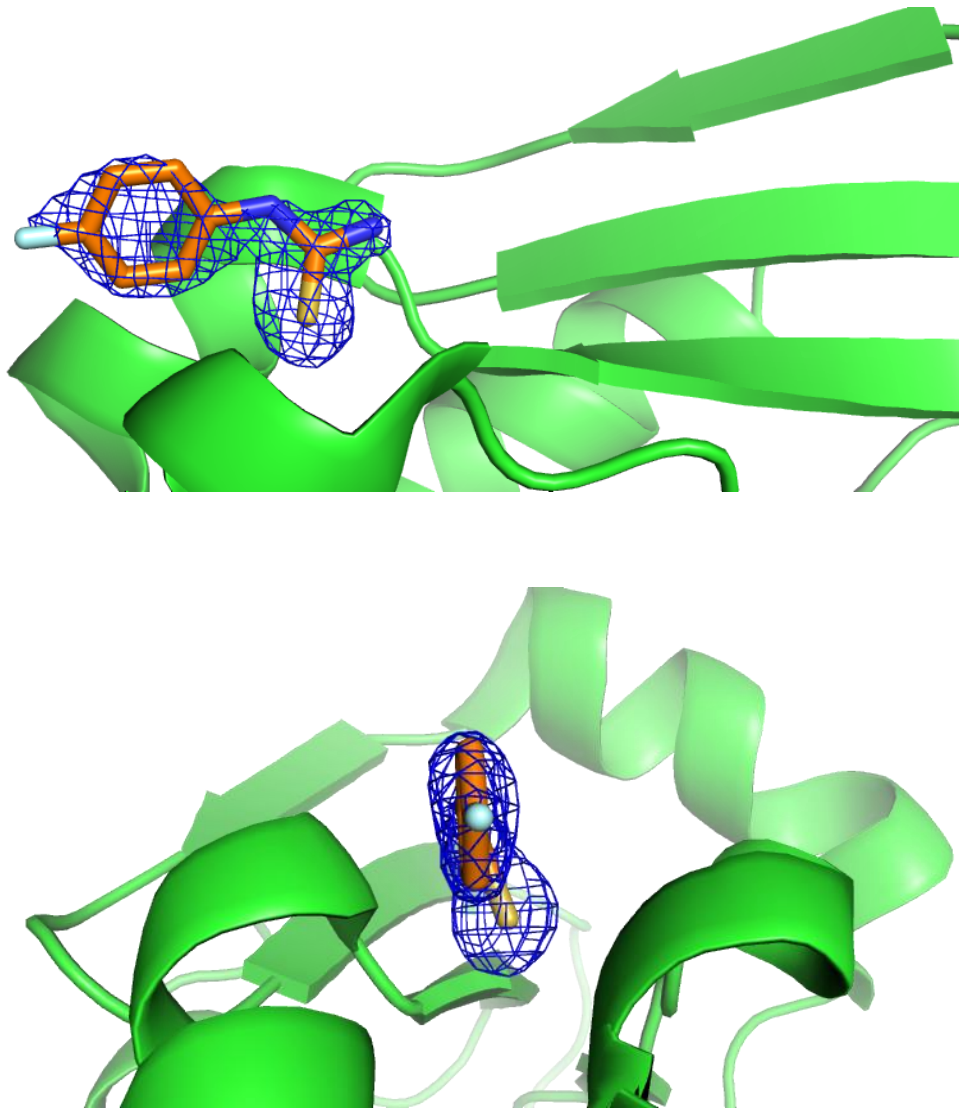


Figura 10. Densidad electrónica del fragmento FTU. Dos diferentes perspectivas del mapa de densidad electrónica 2mFo-DFc (en color malla azul), contorneado a 1σ , alrededor del fragmento FTU en complejo con la proteína P (color verde).

Tabla 2. Estadísticos de colección de datos y afinamiento de la estructura cristalográfica de la proteína P en complejo con el fragmento FTU.

TmP-(4-fluorofenil)-tiourea Código PDB 6CQC	
<u>Colección de datos</u>	
Grupo espacial	P2 ₁
Dimensiones de celda	
<i>a, b, c</i> (Å)	56.1, 64.2, 68.1
<i>a, b, c</i> (°)	90.0, 101.6, 90.0
Resolución (Å)	46.29-1.54 (1.57-1.54) *
<i>R</i> _{sym} o <i>R</i> _{merge}	0.037 (0.205)
<i>I</i> /σ	19.5 (3.7)
Compleción (%)	96.9 (78.5)
Redundancia	3.2 (2.9)
Mean (<i>I</i>) half-set correlation CC _(1/2)	0.993 (0.915)
Factor B de Wilson (Å ²)	18.0
<u>Afinamiento</u>	
Resolución (Å)	46.28-1.54 (1.56-1.54)
No. de reflexiones	67950(2083)
<i>R</i> _{work} / <i>R</i> _{free}	0.193/0.213 (0.241/0.271)
No. átomos	
Proteína	7824
Ligando/ion	18 (Lig), 75 (sulfato)
Agua	512
Factores B	
Proteína (por cadena)	26.7, 26.2, 23.2, 33.4
Ligando/ion	70.5 (Lig), 43.2 (sulfato)
Agua	34.4
Desviaciones RMS	
Longitudes de enlace (Å)	0.005
Ángulos de enlace (°)	0.685
Un cristal fue usado.	
*La capa de más alta resolución se muestra en paréntesis.	

Lamentablemente, la unión del fragmento FTU no pudo detectarse por interferometría de biocapas y tampoco inhibe la actividad ni el crecimiento bacteriano, por lo que perseguimos otras estrategias alternas iniciales. Sin embargo, el FTU sirve como punto de partida para optimizarlo.

7b) Enfoque sobre ensayos de unión

En otra aproximación, consideramos que, dado que la función de la proteína P es la de reconocer al 5'-líder, el monitoreo de la unión de compuestos podría ser una buena estrategia para iniciar el descubrimiento de nuevos inhibidores. A continuación, se describen las estrategias que se siguieron en este sentido.

1) Optimización del fragmento inicial encontrado por cristalografía de rayos X

El fragmento 2-MBX se une con baja afinidad a la proteína P

El “hit” cristalográfico 2-MBX (véase sección 2. Antecedentes), presentó una constante de disociación de 5.3 mM (Figura 11) a partir de ensayos de titulación mediante interferometría de biocapas. Un elemento importante en el ensayo es que fue reproducible y que se pudo detectar un comportamiento dosis-respuesta del fragmento 2-MBX.

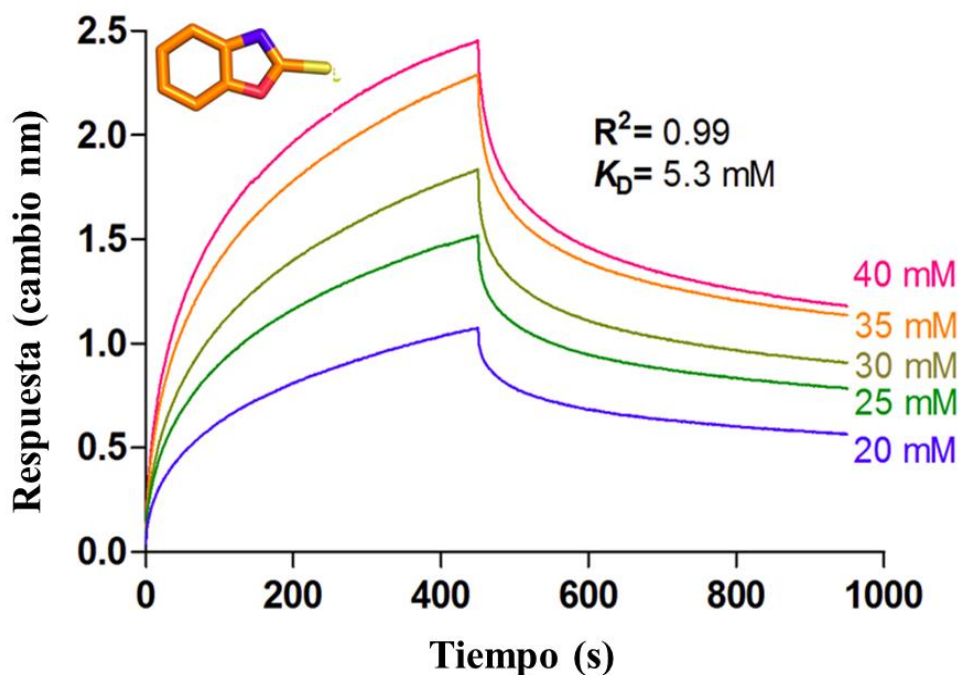


Figura 11. Sensorgrama de la titulación del “hit” 2-MBX. La proteína P fue inmovilizada a 0.05 mg/mL (3.7 μ M) en biosensores Ni-NTA y la concentración del 2-MBX (M=151 Da) se varió de 20-40 mM. La constante de disociación en el equilibrio es de 5.3 mM. En la parte superior izquierda se aprecia la estructura del 2-MBX.

El fragmento 2-MBX se optimizó mediante diseño basado en estructura

Con base en el análisis de la estructura cristalográfica de la proteína P en complejo con el 2-MBX, se observó que existe espacio para añadir grupos funcionales a partir del nitrógeno 5 del anillo oxazol del compuesto (Figura 12). Dicho espacio abarca el residuo Arg89, el cual se encuentra conservado y es importante para la función de la holoenzima según resultados de mutagénesis dirigida obtenidos por Reiter y col. (58) (Figura 9, derecha). El espacio disponible para crecer el fragmento 2-MBX se comparte con el sitio de unión del extremo 5'-líder, (Figura 12, derecha), por lo que, de crecerse el fragmento hacia este espacio, podría generarse un inhibidor que actúa bloqueando estéricamente la unión del sustrato.

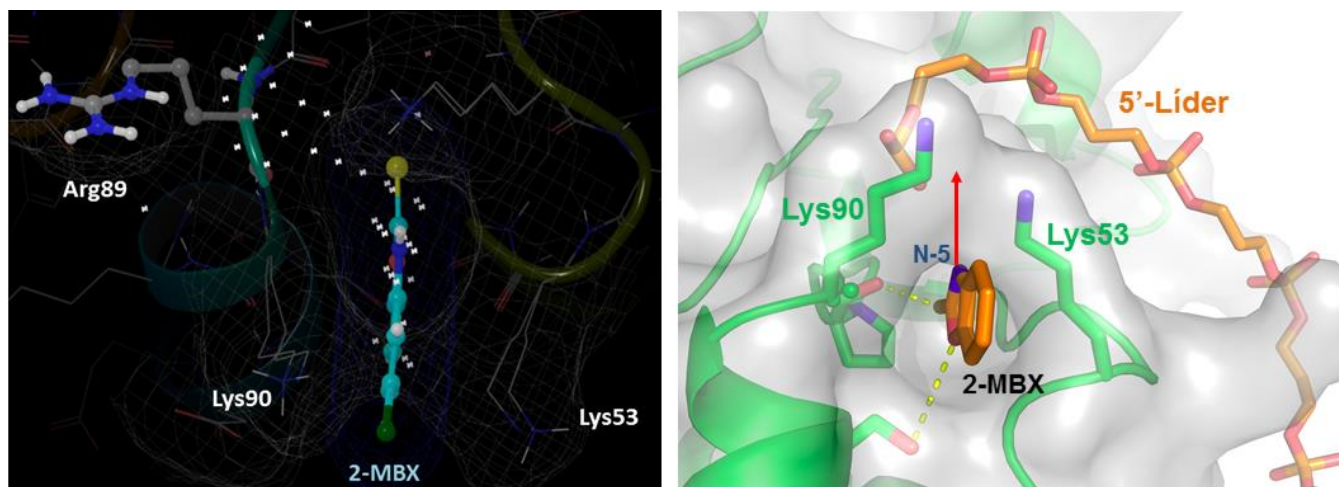


Figura 12. Espacio disponible a partir del N-5 del anillo oxazol del 2-MBX en el sitio de unión. Izquierda. Resultado del análisis mediante MAESTRO (Schrödinger, Inc). Los puntos blancos indican el espacio disponible para crecer el fragmento 2-MBX hacia el residuo Arg89 que se muestra en la zona superior izquierda. En color cian se muestra el fragmento 2-MBX flanqueado por los residuos Lys90 y Lys53. **Derecha.** Perspectiva alterna del resultado de MAESTRO. En naranja se muestra el “hit” 2-MBX con el nitrógeno 5 del anillo oxazol en azul y una flecha denota el espacio disponible hacia donde crecer el fragmento flanqueado por Lys53 y Lys90. El extremo 5'-líder se muestra como cadena fosfatada en color naranja.

Posteriormente, añadimos grupos funcionales *in silico* a partir del nitrógeno 5 del anillo oxazol (Figura 12, izquierda) y obtuvimos un compuesto de mayor peso molecular que tiene interacciones putativas con la Arg89 (Figura 13, derecha).

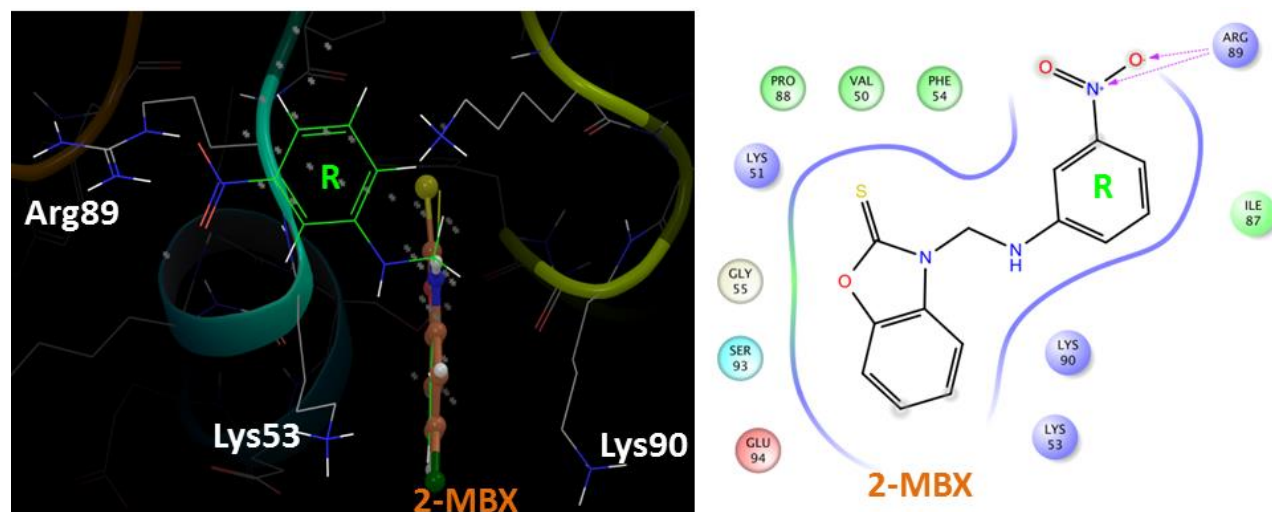


Figura 13. Crecimiento del fragmento 2-MBX basado en estructura. Izquierda. Los puntos grises denotan el espacio disponible cerca del sitio de unión para crecer el 2-MBX (en color naranja opaco). En color verde se muestra el grupo químico nitrobenzeno añadido *in silico*. **Derecha.** Representación en 2D de las interacciones del nitrobenzeno hipotético con los residuos circundantes. Las flechas en color púrpura indican la interacción hidrofílica hipotética del motivo nitrato del nitrobenzeno añadido *in silico*, con el residuo Arg89.

A partir de este análisis, diseñamos una pequeña biblioteca de 11 compuestos derivados del 2-MBX (Figura 14), los cuales servirían para estudios mediante BLI, actividad y cristalografía de rayos X. Estos compuestos se adquirieron de la compañía Sigma-Aldrich, y se eligieron con base en una búsqueda de compuestos semejantes en estructura utilizando el servidor de Sigma-Aldrich, localizando subestructuras con 70% de similaridad y que tuvieran grupos funcionales añadidos a partir del nitrógeno 5 del anillo oxazol.

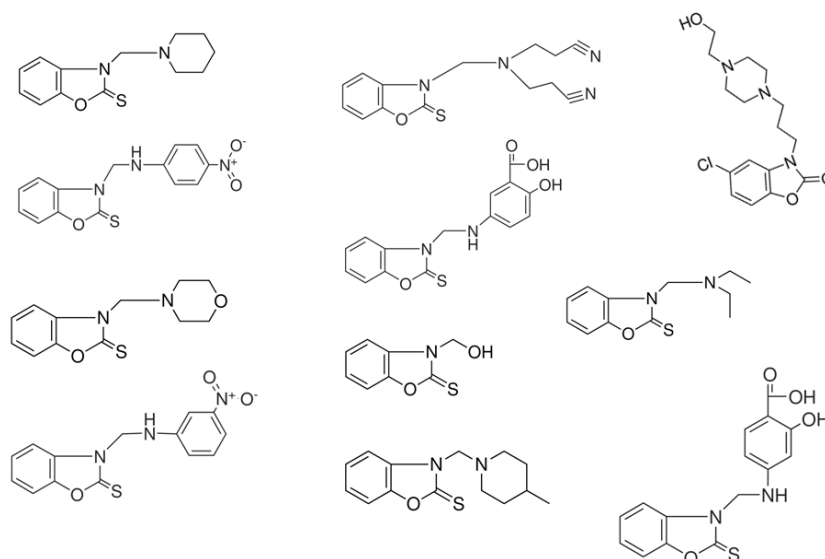


Figura 14. Micro-biblioteca de compuestos derivados del 2-MBX. Colección de 11 compuestos que comparten el núcleo 2-MBX y tienen grupos funcionales añadidos a partir del Nitrógeno 5 del anillo oxazol. Se ensayaron estos compuestos mediante interferometría de biocapas para medir su unión a la proteína P.

La búsqueda mediante BLI de derivados del 2-MBX resultó en la detección de dos ligandos de la proteína P

Una vez construida la biblioteca de derivados del 2-MBX, éstos fueron ensayados contra la proteína P mediante BLI con el fin de detectar algún evento de unión. Los resultados de este experimento permitieron la detección de tres compuestos que mostraron señal (Figura 15).

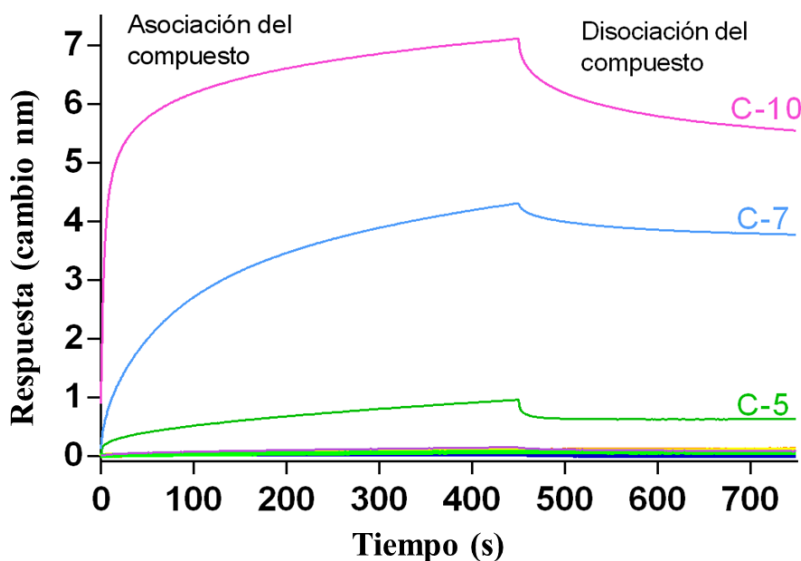


Figura 15. Sensorgrama obtenido a partir de la búsqueda con la biblioteca de derivados del 2-MBX. Dos señales de unión sobresalieron, correspondientes a los compuestos optimizados C-7 y C-10. La respuesta del sensorgrama de color verde se debe al compuesto falso positivo C-5, que no presentó un comportamiento dosis-respuesta durante ensayos posteriores. El experimento se realizó usando 200 μM de cada compuesto ensayado, 0.05 mg/mL (3.7 μM) de proteína P silvestre con etiqueta de histidinas inmovilizada en biosensor Ni-NTA, en amortiguador PBS 1X pH 7.4 y 200 μL de volumen total por pocillo.

Los compuestos C-7 y C-10 se unen a la proteína P de manera dosis-dependiente

El comportamiento de los compuestos C-7 y C-10 hacia la proteína P fue dosis-dependiente (Figuras 16, 17), a diferencia del compuesto C-5. Los compuestos optimizados presentan una afinidad mayor a la proteína P ($K_D = 2.2 \mu\text{M}$ y $0.7 \mu\text{M}$) en comparación con el “hit” inicial 2-MBX ($K_D = 5.3 \text{ mM}$).

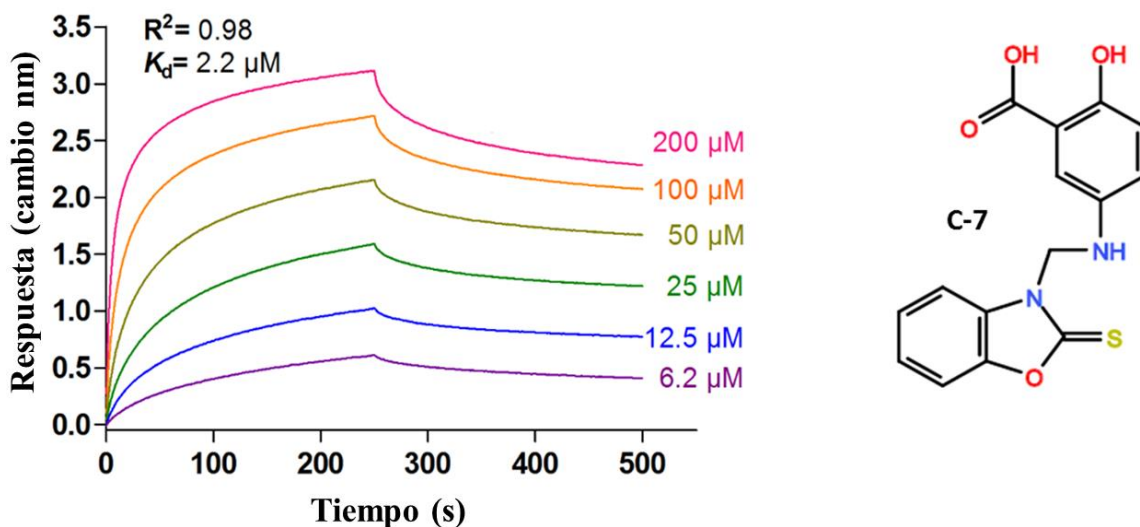


Figura 16. Unión del compuesto C-7 (5-DTS) a la proteína P. Izquierda. Sensorgramas de las curvas de titulación del compuesto C-7 con un intervalo de seis puntos de concentración de $6.2 \mu\text{M}$ – $200 \mu\text{M}$. La concentración de la proteína P con etiqueta de histidinas se mantuvo constante a 0.05 mg/mL ($3.7 \mu\text{M}$). Una K_D de $2.2 \mu\text{M}$ fue determinada para esta interacción. **Derecha.** Estructura del compuesto C-7.

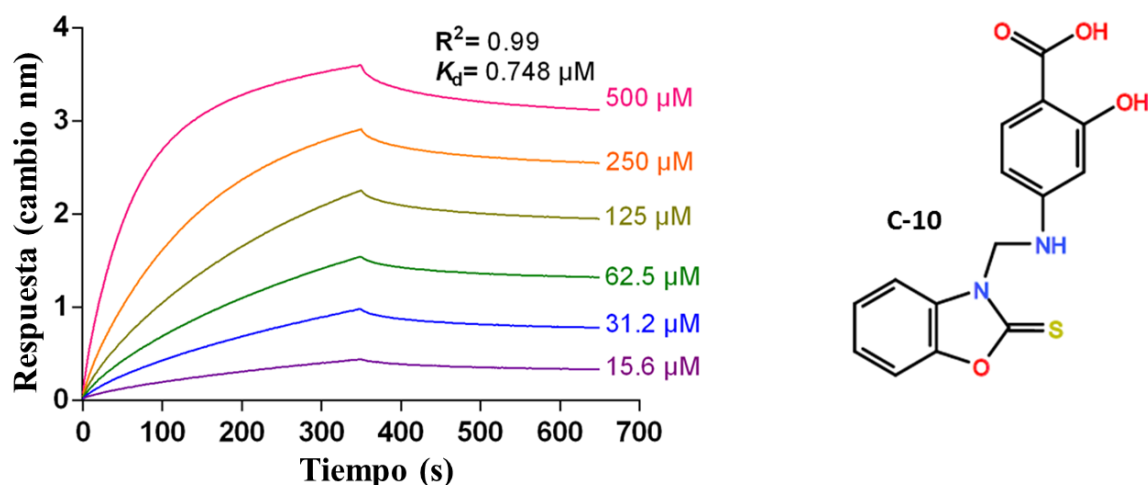


Figura 17. Unión del compuesto C-10 (4-DTS) a la proteína P. Izquierda. Sensorgramas de las curvas de titulación del compuesto C-10 con un intervalo de seis puntos de concentración de $15.6 \mu\text{M}$ – $500 \mu\text{M}$. La concentración de la proteína P con etiqueta de histidinas se mantuvo constante a 0.05 mg/mL ($3.7 \mu\text{M}$). Una K_D de $0.7 \mu\text{M}$ fue determinada para esta interacción. **Derecha.** Estructura del compuesto C-10.

Los compuestos C-7 y C-10 se unen a la proteína P mediante MST

Utilizando una técnica ortogonal para la medición de afinidad (MST), el compuesto C-7 mostró una afinidad por la proteína P en el intervalo micromolar ($K_D = 25.4 \mu\text{M}$), valor que se encuentra en relativo acuerdo con los resultados de BLI ($K_D = 2.2 \mu\text{M}$) (Figura 18).

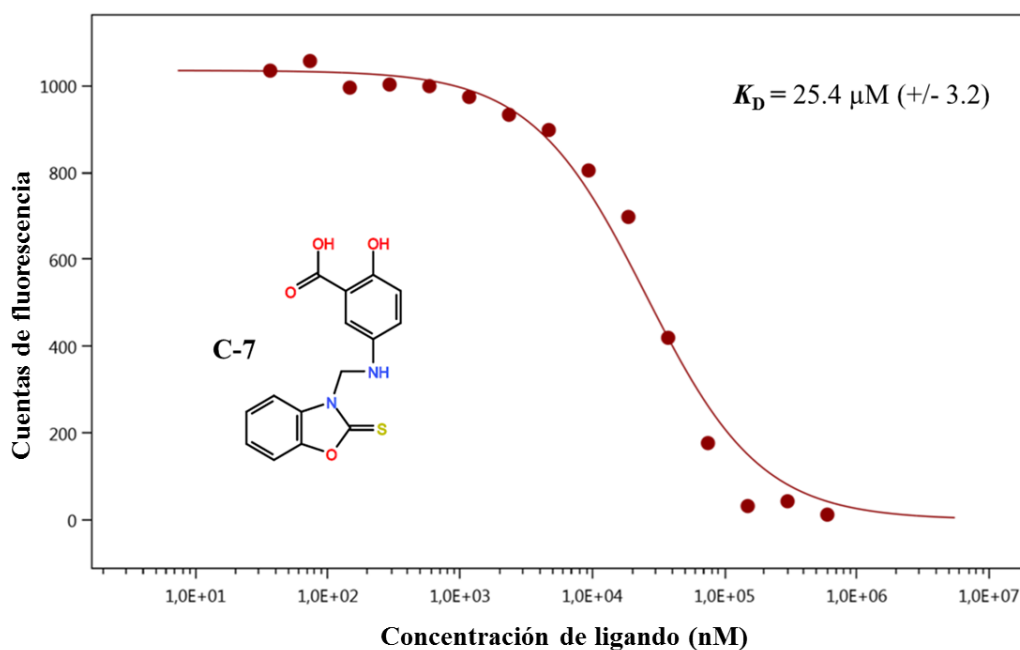


Figura 18. El compuesto C-7 interacciona con la proteína P medido por MST. La etiqueta fluorescente NT-647 fue covalentemente unida a la proteína P vía acoplamiento de grupos NHS. Se mantuvo constante a la proteína P, mientras que el C-7 fue titulado en 15 puntos de 3 nM – 600 μM . El ensayo se llevó a cabo en amortiguador de fosfatos, pH 7.4, adicionado con 500 mM NaCl y 1 mg/mL BSA. Las muestras fueron cargadas dentro de capilares *premium glass* y la medición realizada usando el equipo para termoforesis a microescala Monolith NT.115. Las concentraciones en el eje X están graficadas en escala logarítmica en nM. La K_D fue de $25.4 \mu\text{M} +/- 3.2 \mu\text{M}$.

El compuesto C-10 mostró la misma afinidad medido por MST que la calculada mediante BLI (747 nM vs 748 nM) (Figura 19).

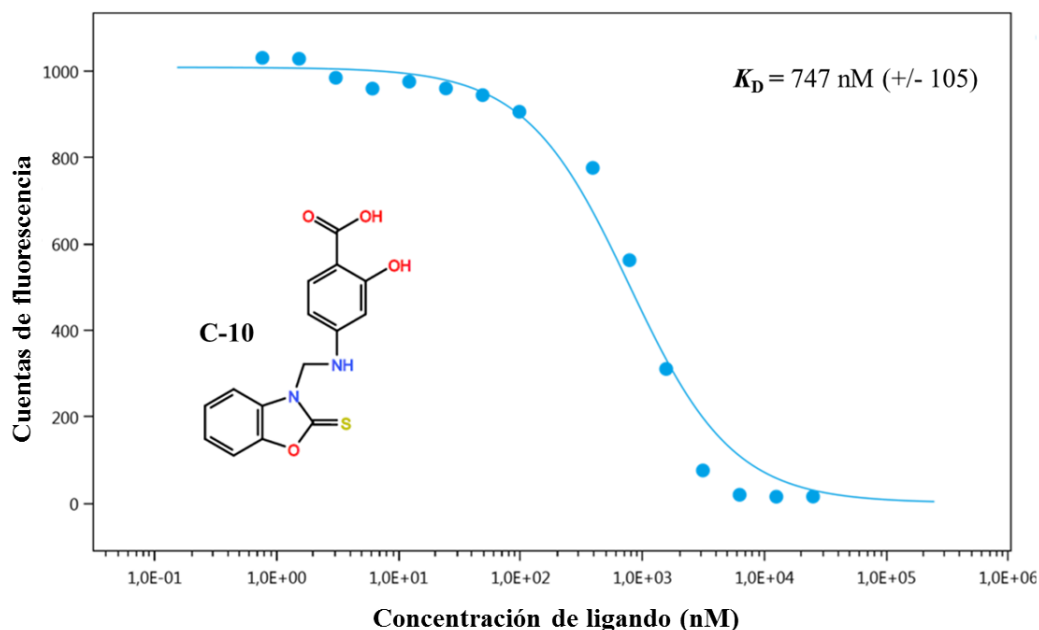


Figura 19. El compuesto C-10 interacciona con la proteína P medido por MST. La etiqueta fluorescente NT-647 fue covalentemente unida a la proteína P vía acoplamiento de grupos NHS. Se mantuvo constante la proteína P etiquetada, mientras que el C-10 fue titulado en 15 puntos de 3 nM – 600 μ M. El ensayo se llevó a cabo en amortiguador de fosfatos, pH 7.4, adicionado con 500 mM NaCl y 1 mg/mL BSA. Las muestras fueron cargadas dentro de capilares *premium glass* y la medición fue llevada a cabo usando el equipo Monolith NT.115. Las concentraciones en el eje X están graficadas en escala logarítmica en nM. La K_D fue de 747 nM +/- 105 nM.

Los resultados de MST y BLI validan la unión a la proteína P. Un resumen de la información obtenida se muestra en la Tabla 3.

Tabla 3. Resumen de la corroboración de valores de afinidades mediante MST.

Proteína marcada con fluoróforo	Fuente de la fluorescencia	Ligando	kD obtenida por MST	kD esperada	Corroboración con respecto a BLI
TmP wt	NT-647	Líder de RNA	1.1 nM	1.65 nM	✓
TmP wt	NT-647	Comp-4	No hay interacción	No interacción	✓
TmP wt	NT-647	Comp-7	25.4 μ M	3.3 μ M	✓
TmP wt	NT-647	Comp-10	747 nM	748 nM	✓

Los compuestos C-7 y C-10 presentan actividad antibacteriana

En ensayos en placa *in vitro* con la cepa *E. coli* BW para expresión diferencial (véase *Métodos. Ensayo de inhibición de la RNasa P in vivo*), los compuestos optimizados no inhiben el crecimiento de bacterias que expresan la enzima PRORP (“*Proteinaceus RNase P*” de *A. thaliana*), pero sí inhiben bacterias que expresan RNasa P bacteriana mediante ensayos de difusión en disco con ambas cepas (Figura 20).

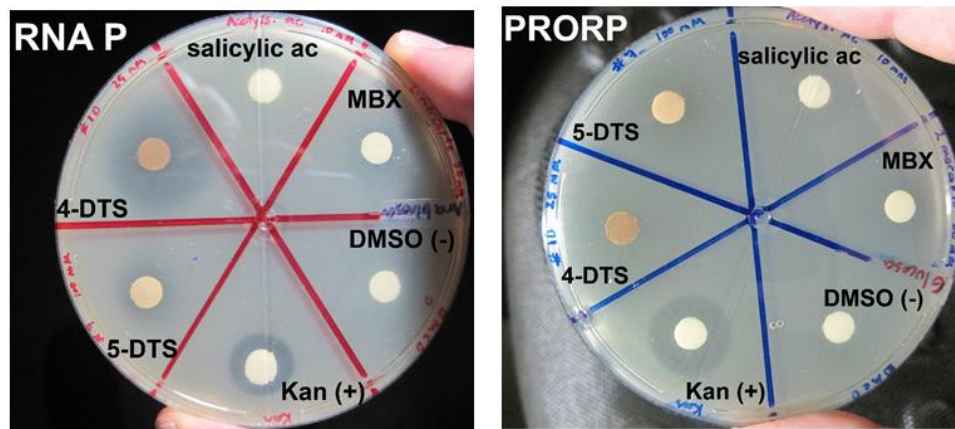


Figura 20. Los compuestos C-7 (5-DTS) y C-10 (4-DTS) inhiben el crecimiento de *E. coli*. Ensayos de difusión de disco en donde las placas se dividieron en seis cuadrantes con papel filtro embebido en 10 mL de los siguientes compuestos: 4-DTS (25 mM), 5-DTS (100 mM), MBX (50 mM), ácido acetilsalicílico (10 mM), kanamicina (control positivo, 10 mM) y DMSO (control negativo). El MBX y el ácido acetilsalicílico se probaron por ser precursores de los compuestos optimizados. **Izquierda.** Crecimiento en presencia de 10 mM de arabinosa (expresión de ARN P). **Derecha.** Crecimiento en condiciones no permisivas en presencia de 10 mM glucosa (expresión de enzima PRORP).

El compuesto C-10 presenta parámetros cinéticos de unión a las proteínas P de *S. aureus* y *B. subtilis* idénticos a aquellos de la proteína P de *T. maritima*

Dado que el compuesto C-10 tiene un mayor efecto inhibitorio (Figura 19-20), se utilizaron las proteínas P de *S. aureus* y *B. subtilis* para corroborar y comparar los resultados obtenidos mediante BLI con la proteína P de *T. maritima*. Los valores de afinidad se encuentran en buen acuerdo con la K_D obtenida para *T. maritima*. La K_D calculada para la BsP fue de 1.4 μM (Figura 21, panel superior) y el valor para SaP fue de 1.2 μM (Figura 21, panel inferior). La K_D calculada para la interacción con TmP es de 0.7 μM .

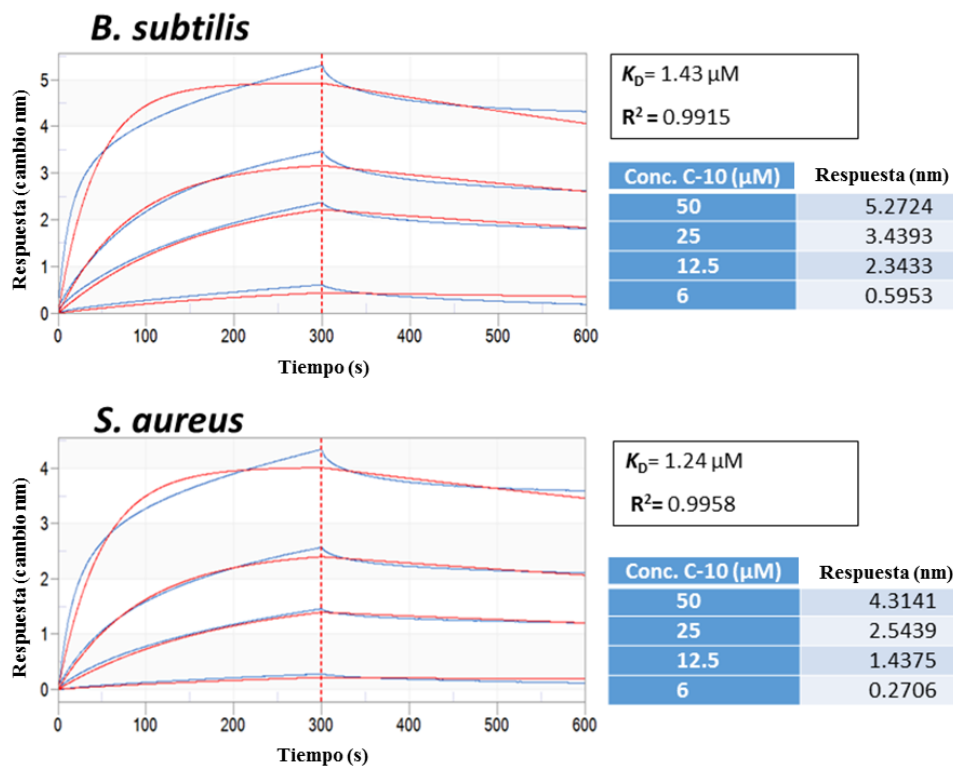


Figura 21. El compuesto C-10 interacciona con las proteínas P de *B. subtilis* y *S. aureus*. **Panel superior.** Titulación del compuesto C-10 contra la proteína P de *B. subtilis* ($K_D=1.43 \mu\text{M}$). **Panel inferior.** Titulación del compuesto C-10 contra la proteína P de *S. aureus* ($K_D=1.24 \mu\text{M}$) fue calculada para esta interacción. A la derecha de cada gráfica se muestra la curva de concentraciones en micromolar de C-10 usadas para cada titulación, así como su respuesta BLI en nanómetros. Las proteínas P fueron inmovilizadas en biosensores Ni-NTA a 0.05 mg/mL ($3.7 \mu\text{M}$).

Al final de la optimización del “hit” cristalográfico 2-MBX, obtuvimos dos compuestos líder optimizados, el C-7 y el C-10, los cuales mostraron inhibir el crecimiento bacteriano y unirse con mayor afinidad a la proteína P que el 2-MBX (Figura 22). Sin embargo, no logramos obtener las estructuras cristalográficas de los complejos de la proteína P con estos compuestos optimizados, muy probablemente debido a que no existe espacio disponible en los canales intersticiales del cristal, o bien, que el grupo espacial cristalino de la proteína P oculta los sitios de unión bajo estudio con sus vecinos cristalinos en el empaquetamiento tridimensional del cristal. Se intentaron una serie de estrategias de cristalogénesis (datos no mostrados) para inducir la formación de grupos espaciales diferentes que permitieran la unión de los ligandos a la proteína P, sin éxito. Debido a este obstáculo técnico, nos enfocamos en intentar otras alternativas de estrategias experimentales, que se describen más adelante, con el fin de poder detectar otros inhibidores de bajo peso molecular.

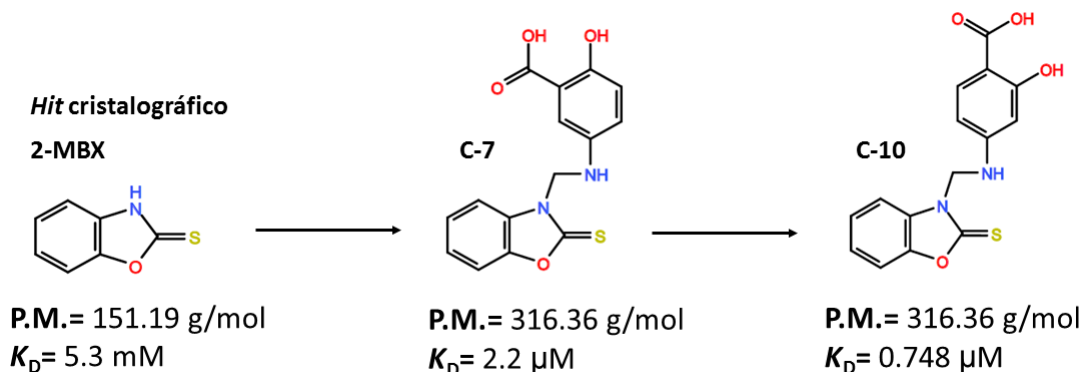


Figura 22. Evolución del fragmento cristalográfico 2-MBX en un compuesto con mayor afinidad por la proteína P. El “hit” cristalográfico 2-MBX (panel izquierdo), que presentó una K_D por la proteína P de 5.3 mM y $M=151$ Da, pudo ser optimizado a los compuestos C-7 y C-10 que presentan un peso molecular de 316 Da y una afinidad por la proteína P de 2.2 μM y 0.7 μM , respectivamente (panel central y derecho).

2) Búsqueda de unión de análogos de un inhibidor ya descrito

En 2014, Liu y col. (94) llevaron a cabo una búsqueda de inhibidores mediante un ensayo *in vitro* de actividad de fluorescencia anisotrópica con la RNasa P de *B. subtilis*, resultando en un inhibidor llamado iriginol hexaacetato, el cual inhibe a la RNasa P con una $K_i= 80$ nM. Uno de los problemas con este inhibidor es que no es factible que funcione *in vivo*, ya que puede ser rápidamente desacetilado y perder su actividad.

El iriginol hexaacetato se une a la proteína P

Determinamos si el iriginol hexaacetato interacciona con la proteína P mediante experimentos de titulación por BLI en condiciones asociadas a búsqueda de compuestos por BLI y a ensayos de actividad por fluorescencia. Este compuesto tiene una alta afinidad por la proteína P y se une de manera dosis-respuesta a nuestro blanco (Figura 23).

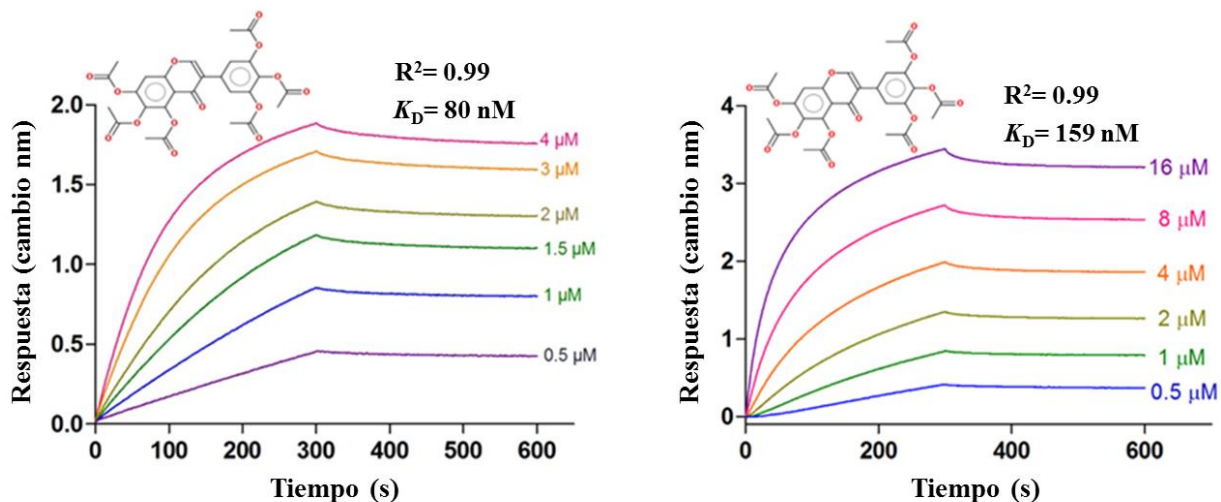


Figura 23. El iriginol hexaacetato se une a la proteína P de manera dosis-respuesta. Panel izquierdo. El iriginol hexaacetato se tituló con seis puntos en un intervalo de concentraciones de 0.5 μM – 4 μM en amortiguador PBS (PBS 1X, 5% DMSO). Una K_D de 80 nM fue determinada para esta interacción. **Panel derecho.** El iriginol hexaacetato se tituló con seis puntos en un intervalo de concentraciones de 0.5 μM – 16 μM usando el amortiguador HAAMD (HEPES 20 mM, acetato de amonio 100 mM, MgCl_2 10 mM adicionado con 5% DMSO). Una K_D de 159 nM fue determinada para esta interacción. En ambos experimentos la proteína P (0.05 mg/mL) se inmovilizó en biosensores Ni-NTA.

Ya que el iriginol hexaacetato es un isoflavonoide, llevamos a cabo una búsqueda de ligandos para la proteína P a través de interferometría de biocapas usando una colección de flavonoides, flavonas, isoflavonas, isoflavonoides y sus derivados.

“Hits” detectados a partir de la búsqueda en la biblioteca de derivados de flavonoides FL-500

Se llevó a cabo una búsqueda de ligandos de la proteína P usando la biblioteca FL-500 (TimTec, Inc), la cual es una colección de 500 fitocompuestos. Se detectaron un total de nueve compuestos que se unieron de manera dosis-dependiente a la proteína P.

Estos “hits” pueden ser clasificados en dos grupos según sus afinidades calculadas por la proteína P: “hits” principales y secundarios. Los “hits” principales abarcaron valores de K_D desde 0.4 μM hasta 6.7 μM (Figura 24).

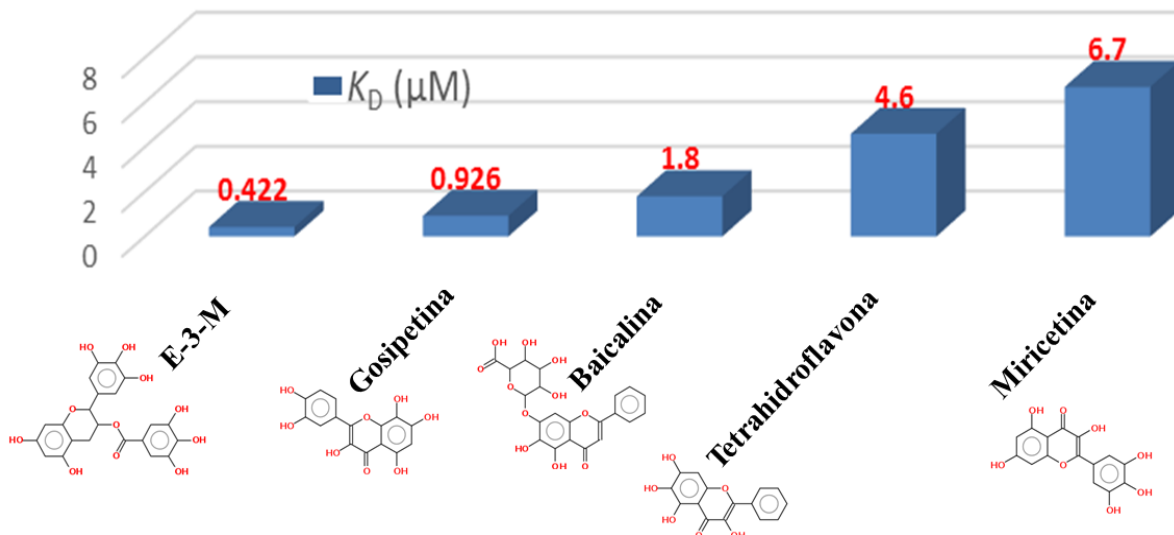


Figura 24. “Hits” principales de la biblioteca FL-500. Se muestran los “hits” principales, los números rojos indican la constante de afinidad para la interacción de cada compuesto con la proteína P calculada a partir de experimentos de titulación de los mismos. De estos “hits”, la epigallocatequina-3-monogallato (E-3-M) fue el compuesto que mostró una mayor afinidad.

En cuanto al grupo de los “hits” secundarios, éstos presentaron valores de K_D desde 15.8 μM hasta 77.7 μM (Figura 25).

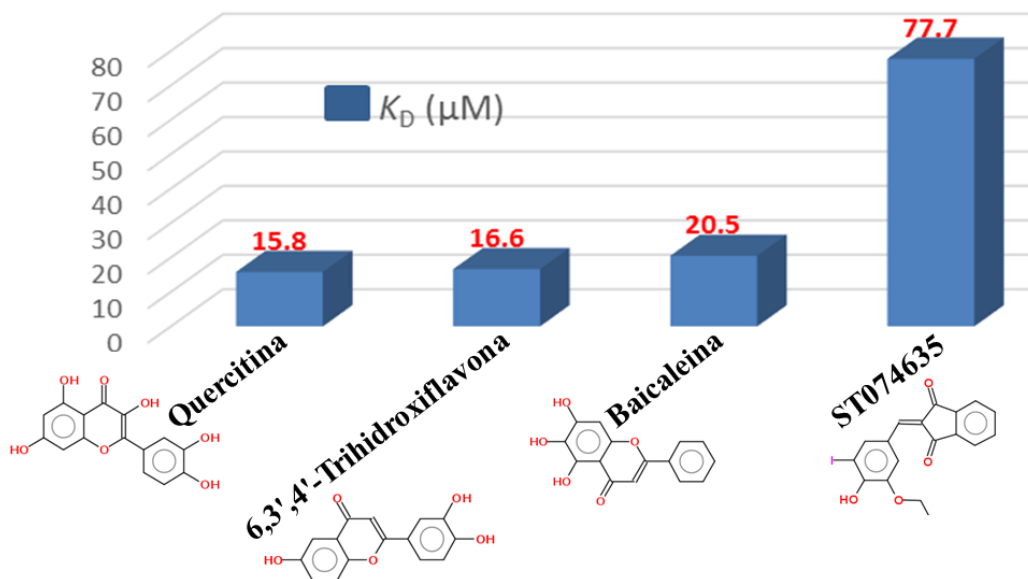


Figura 25. “Hits” secundarios de la biblioteca FL-500. Se muestran los cuatro “hits” secundarios. Los números rojos indican la constante de afinidad para la interacción de cada compuesto con la proteína P calculada a partir de experimentos de titulación de los mismos.

En este punto, descubrimos que la única manera de observar inhibición de la actividad de la RNasa P de *T. maritima* en presencia de iriginol hexaacetato era mediante una preincubación de al menos 30 minutos, en donde el compuesto cambiaba de color de amarillo a café, por lo que pensamos que podría existir un proceso de oxidación en el mismo y suspendimos la caracterización de los demás compuestos relacionados descritos en esta sección.

7c) Enfoque sobre los ensayos de actividad

Dado el éxito limitado obtenido usando inicialmente los enfoques estructurales y de unión, consideramos entonces optimizar el ensayo de actividad desarrollado con anterioridad con objeto de tomarlo como punto de partida para la búsqueda de inhibidores. La optimización implicó reducir el volumen del ensayo a 20 microlitros, la transferencia a placas de 384 pozos, la optimización de la temperatura y el desarrollo del ensayo utilizando el sustrato pre-tRNA-mitades.

Búsqueda de inhibidores de la RNasa P de *T. maritima* mediante ensayo de actividad monitoreado por espectroscopía de fluorescencia

Se llevó a cabo una búsqueda de inhibidores mediante el ensayo de actividad de RNasa P por fluorescencia en formato de placas de 384 pocillos utilizando la biblioteca Spectrum con 2850 compuestos. El resultado final de esta búsqueda, una vez deconvolucionados los cocteles iniciales, fueron cuatro compuestos de bajo peso molecular que inhiben a la RNasa P en presencia del sustrato Mh y del pre-tRNA. Estos “hits” son el violeta de genciana, la juglona, la purpurina y la hemateína, los cuales mostraron inhibir la actividad de la RNasa P en el intervalo micromolar (Figura 26).

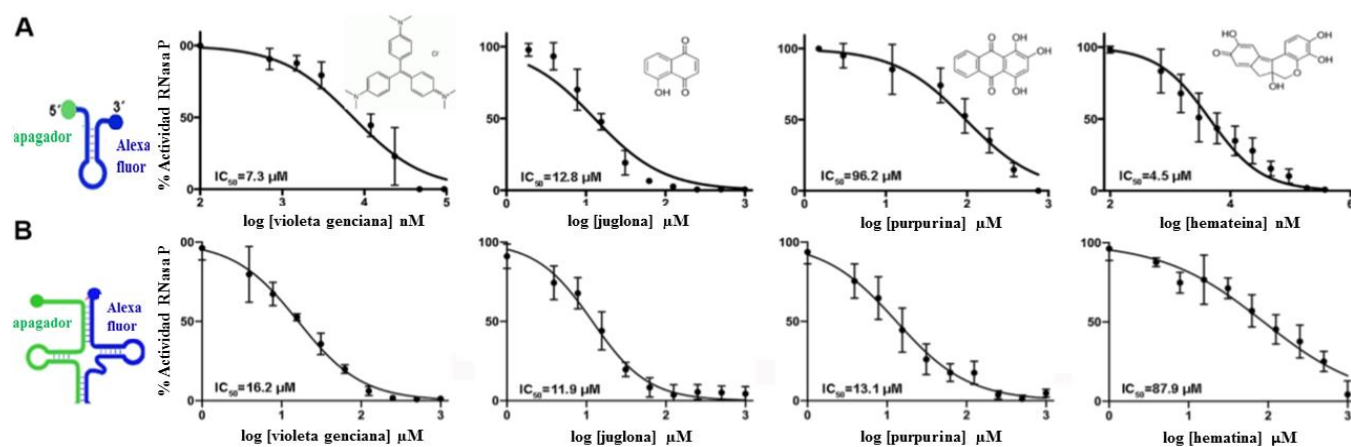


Figura 26. Se detectaron cuatro compuestos que inhiben a la RNasa P en el intervalo micromolar. Gráficas de porcentaje de actividad de RNasa P vs concentración. **Superior.** Ensayos de actividad empleando el sustrato fluorescente Mh en presencia de los cuatro inhibidores detectados. **Inferior.** Ensayo de actividad usando el sustrato fluorescente pre-tRNA-mitades en presencia de los cuatro inhibidores detectados. De izquierda a derecha se muestran las estructuras del violeta de genciana, juglona, purpurina y hemateína, respectivamente.

-- Integración de la plataforma actividad-unión-estructura

Ya que el objetivo principal de este proyecto fue detectar e identificar compuestos de bajo peso molecular que inhiban a la RNasa P vía la unión a la proteína P, se procedió a determinar si estos compuestos se unían a la proteína P de manera selectiva. De ellos, el violeta de genciana mostró unirse de manera dosis-respuesta a la subunidad ARN P y no a la proteína P (Figura 34), por lo que ya no se persiguió el estudio de este compuesto. Los restantes tres inhibidores, se unieron a la proteína P de manera dosis-respuesta durante experimentos de titulación mediante BLI (Figura 27).

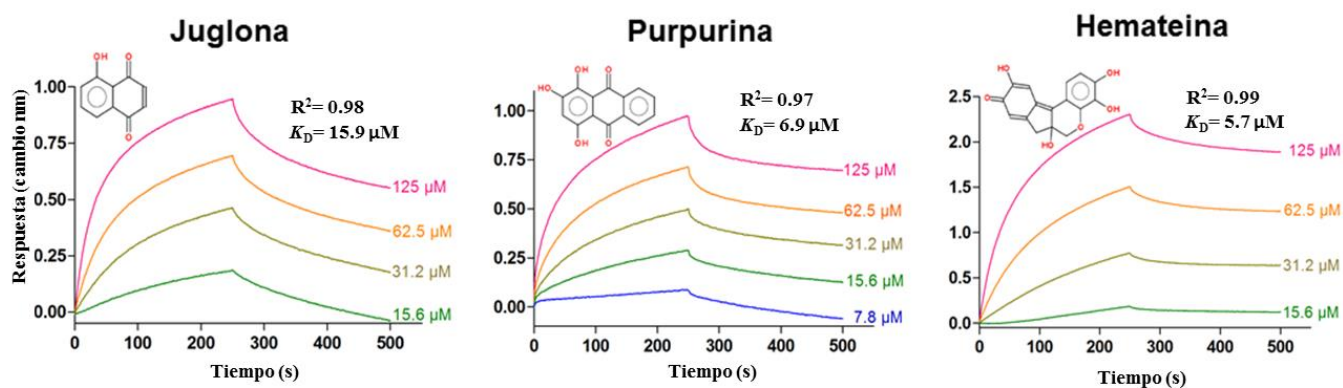


Figura 27. Tres inhibidores se unen de manera dosis-respuesta a la proteína P. Sensorgramas de la titulación de los tres inhibidores descubiertos. Curvas de titulación de la juglona (**panel izquierdo**), de la purpurina (**panel central**) y de la hemateína (**panel derecho**). La proteína P fue inmovilizada en biosensores Ni-NTA a una concentración final de 0.05 mg/mL (3.7 μM). Las tres titulaciones se llevaron a cabo en un intervalo de concentraciones desde 7.8 hasta 125 μM del respectivo inhibidor.

Mediante estudios de dinámica molecular, realizados por el Dr. Marcelino Arciniega Castro del IFC-UNAM, se observó que los compuestos hemateína y purpurina permanecen en la superficie de la proteína P en comparación con los otros inhibidores (Figura 28).

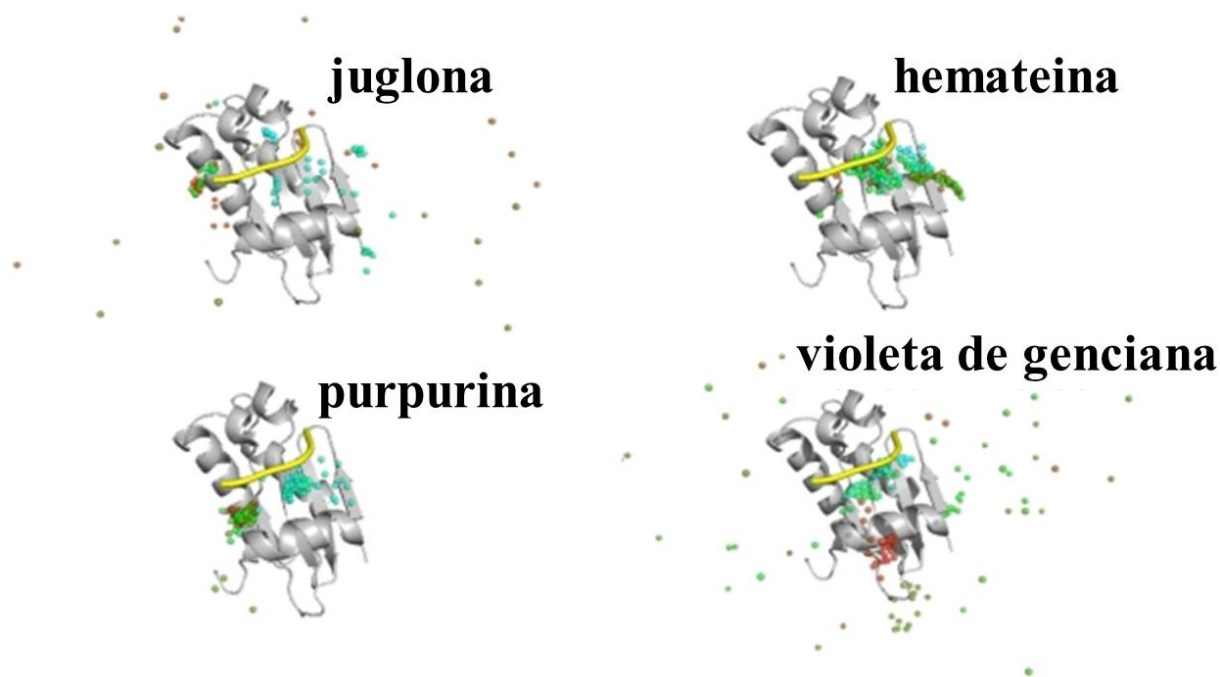


Figura 28. Simulaciones de dinámica molecular de la proteína P en complejo con inhibidores. En cada panel, esferas semitransparentes representan el centro de geometría, a lo largo de una trayectoria de 200 ns, de los correspondientes compuestos probados. El 5'-líder (en amarillo) se muestra como referencia, pero no fue incluido en la simulación. El código de color de esferas indica las posiciones iniciales y finales en azul y rojo, respectivamente. La conformación del complejo inicial para el violeta de genciana, la juglona y la hemateína corresponden a la mejor pose sugerida por acoplamiento molecular. La conformación inicial de la purpurina corresponde a aquella observada en la estructura del cristal en complejo con la proteína P (código PDB: 6MAX). La hemateína y la purpurina permanecen unidas a la proteína P a lo largo del protocolo de la dinámica.

Validación de la purpurina como un inhibidor de la RNasa P a través de su unión a la proteína P

Cristalización de la proteína P de *T. maritima* en complejo con purpurina

Después de intentar una serie de protocolos que involucraron remojo, co-cristalización y co-purificación de la proteína P con purpurina, la técnica que resultó en la obtención del complejo con purpurina fue un remojo de casi 3 meses de duración en presencia de 20 mM de purpurina (Figura 29).

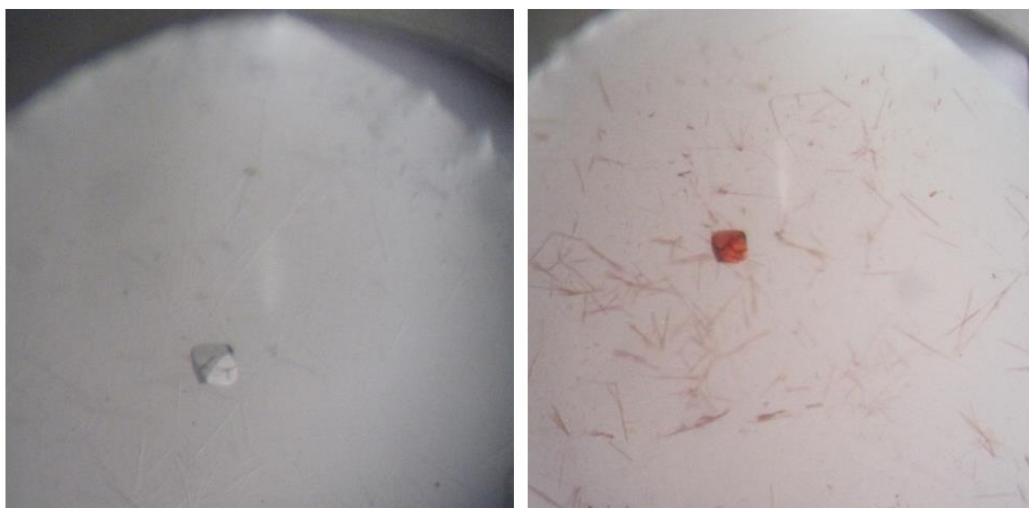


Figura 29. Cristalización y experimento de remojo con purpurina. Izquierda. Cristal apo de la proteína P de *T. maritima*. Derecha. Cristal remojado con 20 mM de purpurina durante 2 meses y 23 días.

Estructura cristalográfica de la proteína P de *T. maritima* en complejo con purpurina

El sitio de unión de la purpurina (1,2,4-trihidroxiantraquinona) se localiza en un extremo final del surco que interacciona con el extremo 5'-líder donde se encontraría la base -5 y -4 del líder (Figura 30).

La ocupación del inhibidor purpurina fue de 1.0, y su densidad electrónica se visualiza claramente a través del mapa de diferencia contorneado a 1σ (Figura 31).

Se observa una zona hidrofóbica con una red de interacciones que involucra a tres residuos: Val33, Leu35 e Ile87; mientras que, rodeando a esta zona hidrofóbica, también se observa una zona hidrofílica con interacciones que involucran a tres residuos: Gln28, Lys51 y Arg89 (Figura 32).

El modo de unión de la purpurina se puede describir a través de sus tres motivos estructurales: el motivo benceno, el motivo benzoquinona y el motivo catecol. El motivo benceno interacciona con la proteína P mediante interacción *Pi-alkyl* con la Leu35 e Ile87. Asimismo, el motivo benzoquinona interacciona con la proteína mediante interacción *Pi-alkyl* con la Val33 y el benceno del motivo benzoquinona, así como un puente de hidrógeno con la Gln28. El motivo catecol de la purpurina interacciona con la proteína P mediante puentes de hidrógeno con la Lys51 y la Arg89 (Figura 32).

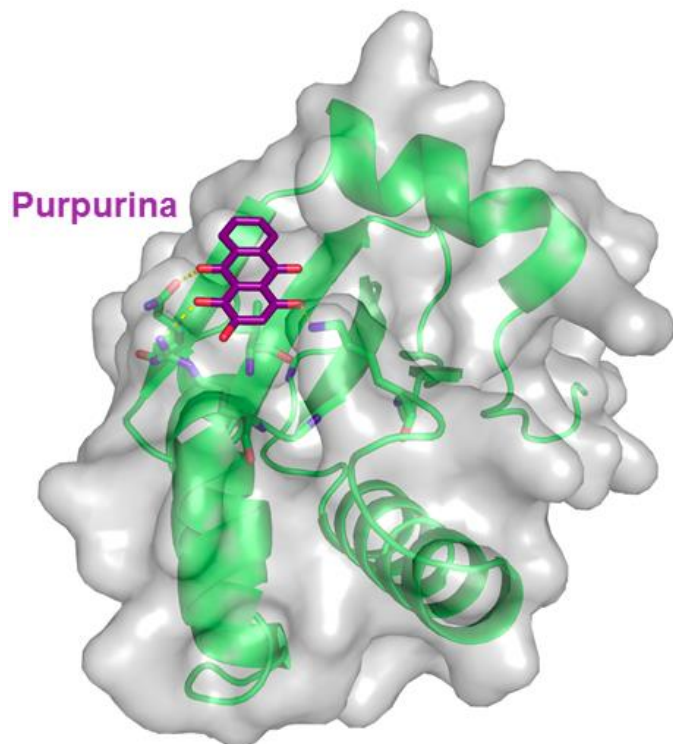


Figura 30. Estructura cristalográfica de la proteína P en complejo con purpurina. Sitio de unión de la purpurina que coincide con el sitio de unión del extremo 5'-líder.

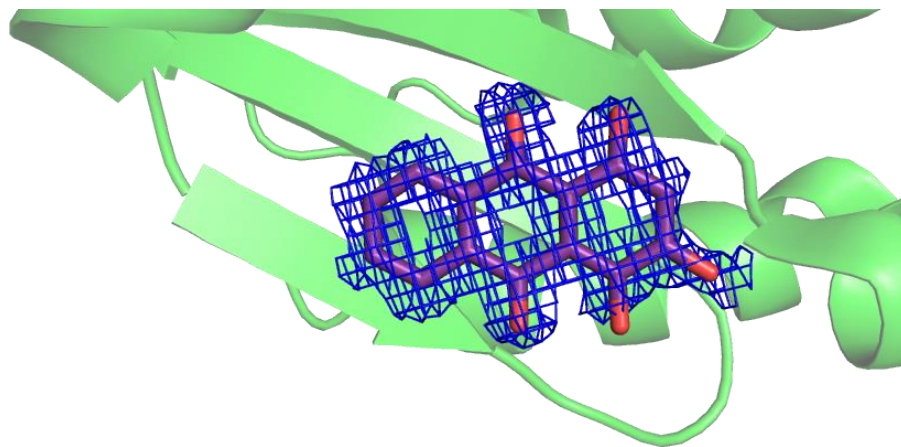


Figura 31. Densidad electrónica del inhibidor purpurina. Mapa de densidad electrónica 2mFo-DFc (en malla azul), contorneado a 1σ alrededor de la purpurina en complejo con la proteína P (color verde). La purpurina se disolvió en 50% de PEG-400 y el cristal de la proteína P apo fue remojado en purpurina 20 mM durante 3 meses.

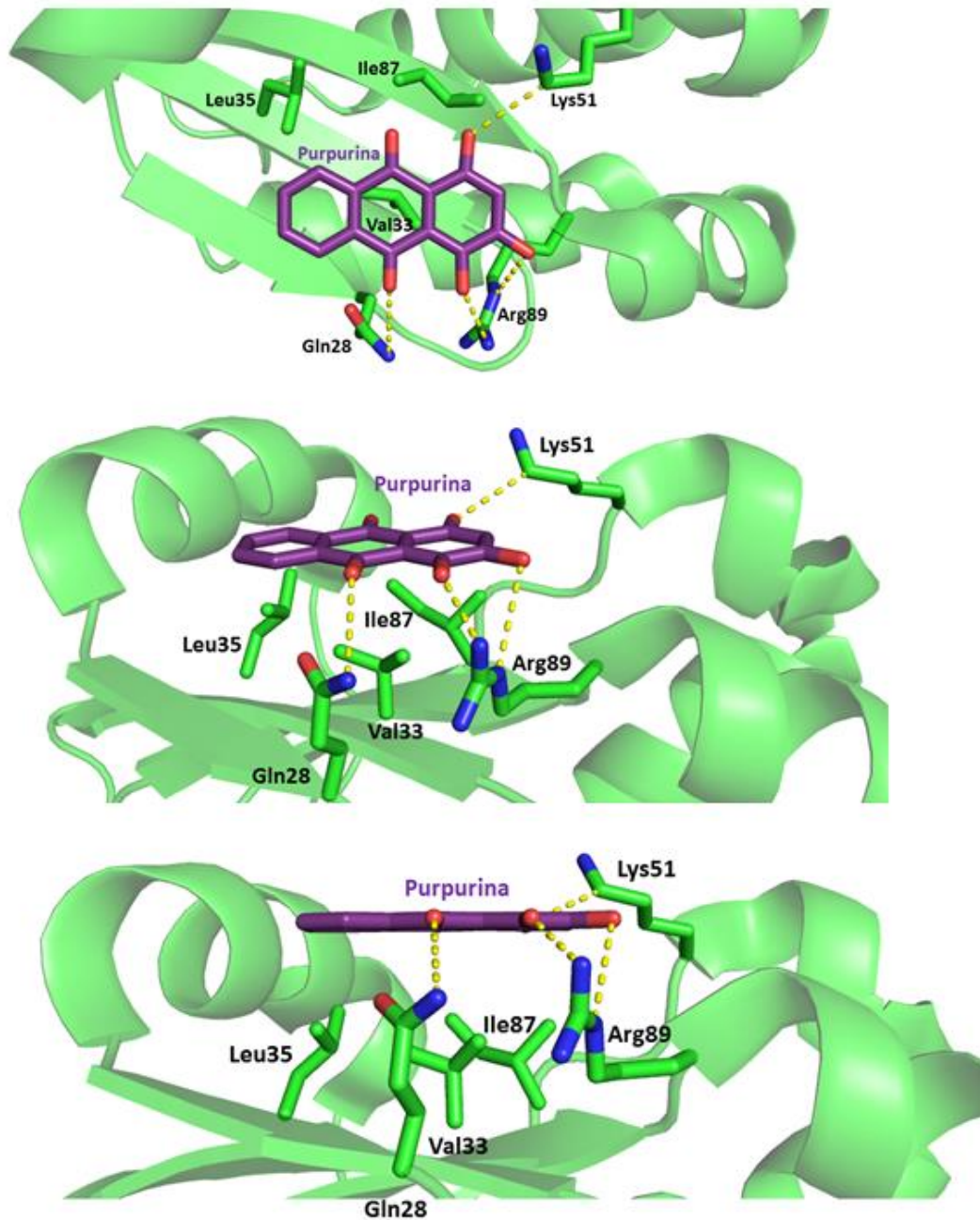


Figura 32. Diferentes vistas del modo de unión de la purpurina. A través de interacciones hidrofóbicas con los residuos Val33, Leu35 e Ile87, así como con interacciones hidrofílicas con los residuos Gln28, Lys51 y Arg89. La estructura del complejo se resolvió a una resolución de 1.4 Å.

La estructura del complejo se resolvió a una resolución de 1.4 Å (Tabla 4) mediante reemplazo molecular, usando como templado a la estructura cristalográfica de la proteína P de *T. maritima* código PDB 1NZ0 (54), aunque para estudios posteriores se utilizó la estructura apo de la proteína P resuelta en nuestro laboratorio. Se utilizó sólo un cristal para la difracción por rayos X y su grupo espacial fue I2₁. En cuanto al ligando observado, el valor del factor B de la purpurina fue de 42 Å², mientras que los residuos interactuantes mostraron un factor B de 26.5 Å². Los estadísticos de afinamiento indican que no existen residuos que se encuentren fuera de los límites en el gráfico de Ramachandran (Tabla 4).

La identidad del compuesto también se confirmó mediante espectrometría de masas (datos no mostrados) por la Dra. Adela Rodríguez Romero del Instituto de Química de la UNAM.

Tabla 4. Estadísticos de colecta de datos y de afinamiento de la estructura cristalográfica de la proteína P en complejo con purpurina.

TmP-Purpurina	
Código PDB 6MAX	
<u>Colección de datos</u>	
Grupo espacial	I2 ₁
Dimensiones de celda	
<i>a, b, c</i> (Å)	60.4, 32.7, 64.1
<i>a, b, c</i> (°)	90.0, 113.6, 90.0
Resolución (Å)	51.91-1.42 (1.44-1.42) *
R_{sym} o R_{merge}	0.055 (0.394)
I/σ	11.2 (1.7)
Compleción (%)	99.7 (97.2)
Redundancia	4.7 (4.1)
Mean (I) half-set correlation CC _(1/2)	0.998 (0.886)
Factor B de Wilson (Å ²)	17.2
<u>Afinamiento</u>	
Resolución (Å)	51.91-1.42 (1.46-1.42)
No. de reflexiones	21010(1544)
$R_{\text{work}}/R_{\text{free}}$	0.143/0.181 (0.249/0.308)
No. átomos	
Proteína	952
Ligando/ion	19 (Lig), 10 (sulfato)
Agua	102
Factores B	
Proteína (por cadena)	26.4
Ligando/ion	42.5 (Lig), 32.5 (sulfato)
Agua	31.1
Desviaciones RMS	
Longitudes de enlace (Å)	0.008
Ángulos de enlace (°)	1.098
Un cristal fue usado.	
*La capa de más alta resolución se muestra en paréntesis.	

La purpurina no inhibe la actividad de la subunidad ARN P

Ya que durante experimentos de actividad por fluorescencia se detectaron cuatro compuestos que inhiben la actividad de la holoenzima, probamos si estos “hits” actúan inhibiendo la subunidad catalítica ARN P o la holoenzima. Realizamos también experimentos en gel que corroboraran lo obtenido mediante fluorescencia (Figura 33). Si bien los cuatro compuestos inhiben la actividad de la holoenzima, la actividad del ARN P no se ve afectada por la adición de hemateína o purpurina (Figura 33).

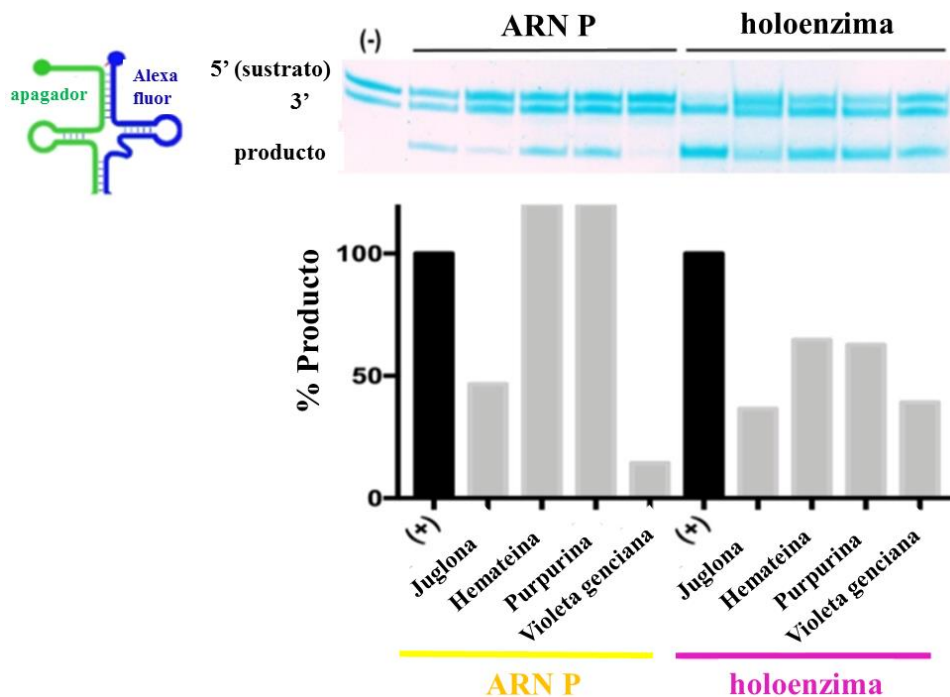


Figura 33. La purpurina inhibe selectivamente a la holoenzima de la RNasa P. Densitometría del ensayo de actividad de punto final en gel. **Panel superior.** Ensayo de actividad en gel en presencia de diferentes inhibidores con la subunidad ARN P u holoenzima. **Panel inferior.** Densitometría del porcentaje de producto escindido en la reacción por efecto de diferente inhibidor en presencia de la subunidad ARN P u holoenzima con base en el gel del panel superior.

La purpurina se une a la holoenzima de la RNasa P

Para corroborar los resultados obtenidos a partir de los ensayos de actividad, llevamos a cabo ensayos de unión por BLI para analizar la interacción con el ARN P o la holoenzima, inmovilizando estos blancos en biosensores SSA a través de un oligo de ARN complementario biotinilado. De estos cuatro inhibidores, el cristal violeta fue el único que mostró más señal de unión por el ARN P que por la holoenzima, aunque se unió a ambos blancos (Figura 34, panel superior izquierdo). La juglona y la hemateína se unen tanto a la holoenzima como al ARN P (Figura 34, panel superior derecho e inferior izquierdo). La purpurina fue el único compuesto que mostró unirse a la holoenzima y no a la subunidad ARN P (Figura 34, panel inferior derecho). Este resultado fue indicativo de que el modo de acción de la purpurina podría ser a través de la interacción con la proteína P.

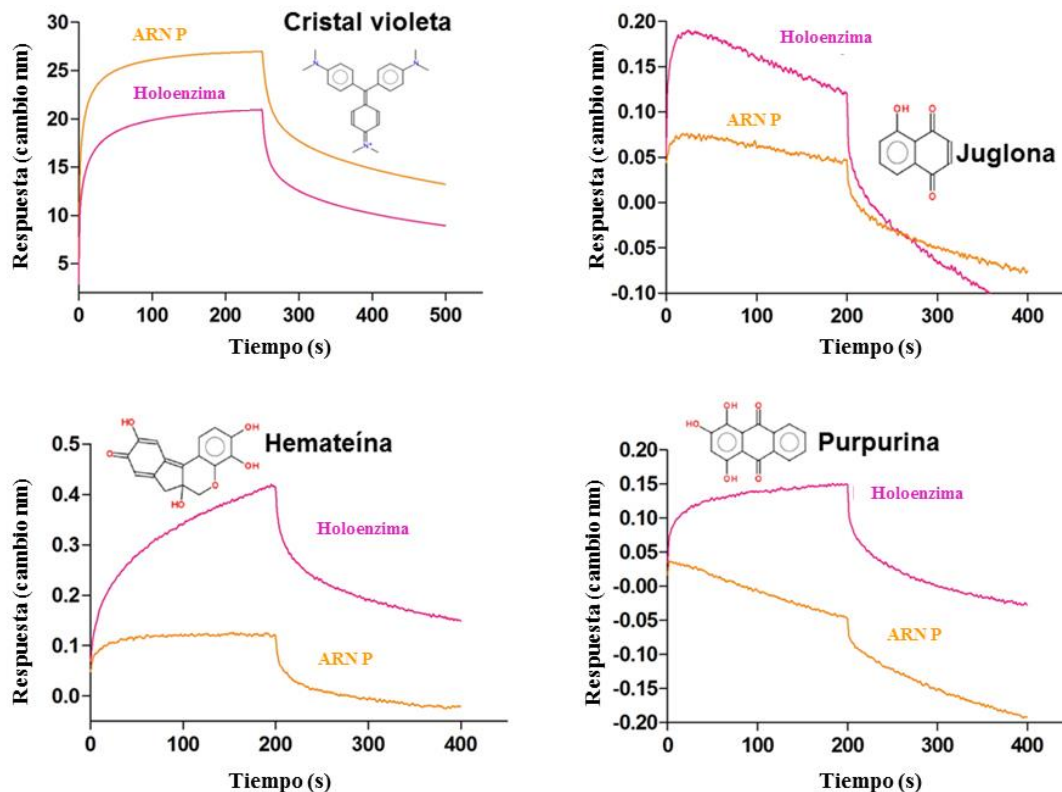


Figura 34. La purpurina se une selectivamente a la holoenzima de la RNasa P bacteriana. Afinidades relativas de los inhibidores del ARN P obtenidos mediante BLI. En color rosa se muestran los sensorgramas correspondientes para la interacción de la holoenzima con cada inhibidor. En color amarillo se muestra el sensorgrama para la interacción de la subunidad ARN P con cada inhibidor. El ARN P o la holoenzima de la RNasa P fue inmovilizada sobre biosensores SSA (Superstreptavidin) a través de un oligonucleótido biotinilado de ARN complementario a una región en el tallo P1 del ARN P. Sorprendentemente, la purpurina no se une al ARN P de manera efectiva. En contraste, el violeta de genciana, se une al ARN P.

Es interesante notar que, en el caso de la juglona y la purpurina cuando se unen al ARN P, se observa un decremento o pérdida en la señal (Figura 34, panel inferior y superior derecho). La curva de unión negativa puede deberse a la ocurrencia de cambios conformacionales del blanco, la cual se compacta a partir de la unión con el compuesto, cambiando su tamaño lo cual provoca cambios negativos de unión. También puede significar que los compuestos se unen transitoria y muy débilmente al blanco inmovilizado sin presentar alguna interacción efectiva y duradera dentro de la etapa de unión mediante BLI. En comparación, un ligando verdadero y eficiente (véase Apéndice E. Glosario de términos) mantiene su unión demostrando una interacción efectiva y duradera con el blanco incluso durante la etapa de disociación, tal como el caso de la purpurina cuando se une a la holoenzima (Figura 34, panel inferior derecho).

Además, a través de experimentos de titulación, se determinó que la purpurina se une de manera dosis-dependiente tanto a la proteína P como a la holoenzima de la RNasa P con un valor de K_D muy similar (Figura 35). El valor de K_D calculado para la interacción de la purpurina con la proteína P es de $6.8 \mu\text{M}$ (Figura 35, panel derecho), mientras que para la holoenzima fue de $13.2 \mu\text{M}$ (Figura 35, panel izquierdo).

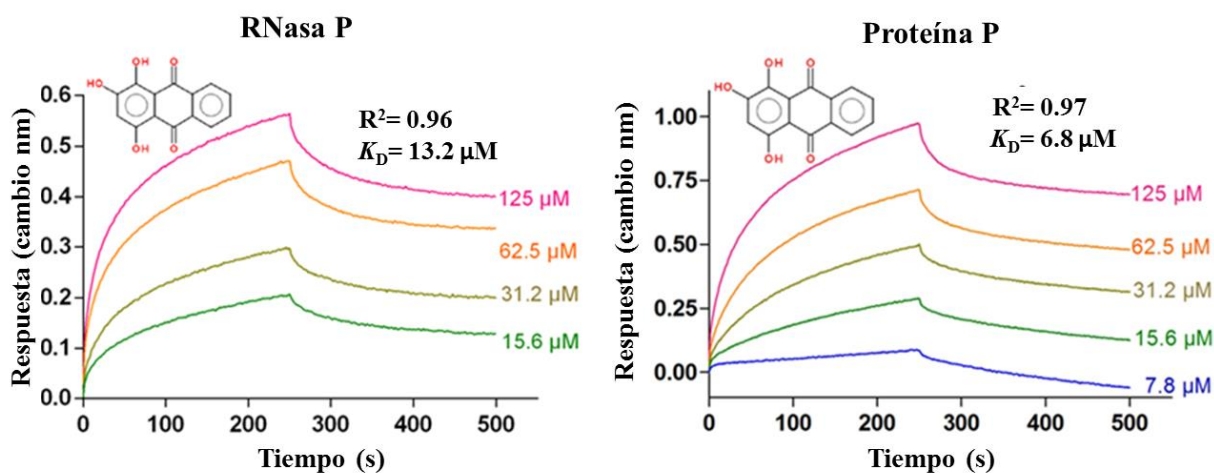


Figura 35. La purpurina interactúa en el mismo intervalo de afinidad tanto con la holoenzima como con la proteína P. Senogramas de la titulación de la purpurina obtenidas mediante BLI empleando el mismo intervalo de concentraciones de purpurina. **Izquierda.** La holoenzima de la RNasa P fue inmovilizada a una concentración de 800 nM sobre biosensores Ni-NTA vía la etiqueta de histidinas de la proteína P en amortiguador HAAMD. La K_D calculada para esta interacción fue de $13.2 \mu\text{M}$. **Derecha.** La proteína P fue inmovilizada a biosensores Ni-NTA a una concentración de 0.05 mg/mL. La K_D calculada para esta interacción fue de $6.8 \mu\text{M}$.

La purpurina inhibe a la RNasa P en el intervalo micromolar

Se llevaron a cabo ensayos de actividad de punto final a diferentes concentraciones de purpurina en presencia del sustrato Mh o del sustrato pre-tRNA-mitades, con el fin de saber más acerca de su mecanismo de inhibición y determinar la constante de inhibición.

El valor de IC_{50} calculado para la inhibición de la RNasa P en presencia de sustrato Mh fue de $96.2 \mu\text{M}$ (Figura 36, panel izquierdo), mientras que en presencia de sustrato pre-tRNA-mitades fue de $125 \mu\text{M}$ (Figura 36, panel derecho).

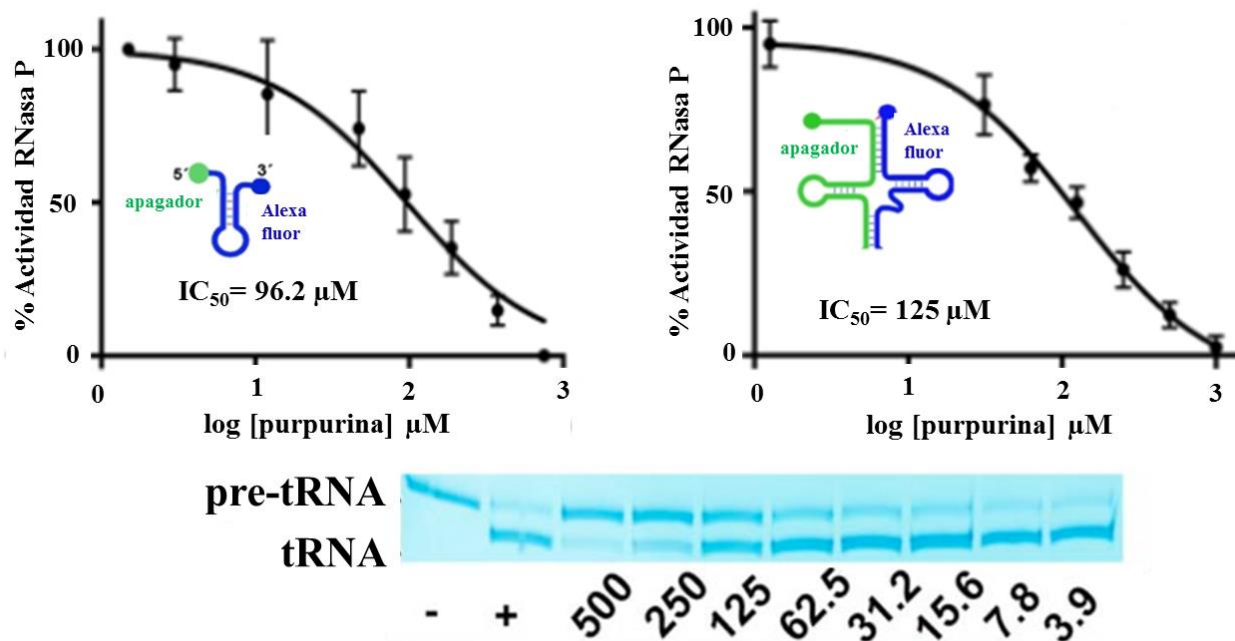


Figura 36. Inhibición de la RNasa P en presencia de sustratos. Izquierda. Cálculo de la IC_{50} empleando el sustrato minihélice. **Derecha.** Cálculo de la IC_{50} empleando el sustrato pre-tRNA-mitades. **Panel inferior.** Ensayo de actividad en gel.

También se caracterizó cinéticamente la inhibición de la actividad de la RNasa P por efecto de la purpurina, lo que resultó en una $K_i=1.9 \mu\text{M}$ de manera dosis-dependiente (Figura 37). Este valor de K_i se encuentra en buen acuerdo con la K_D con la holoenzima (13.2 μM) y la K_D con la proteína P (6.8 μM).

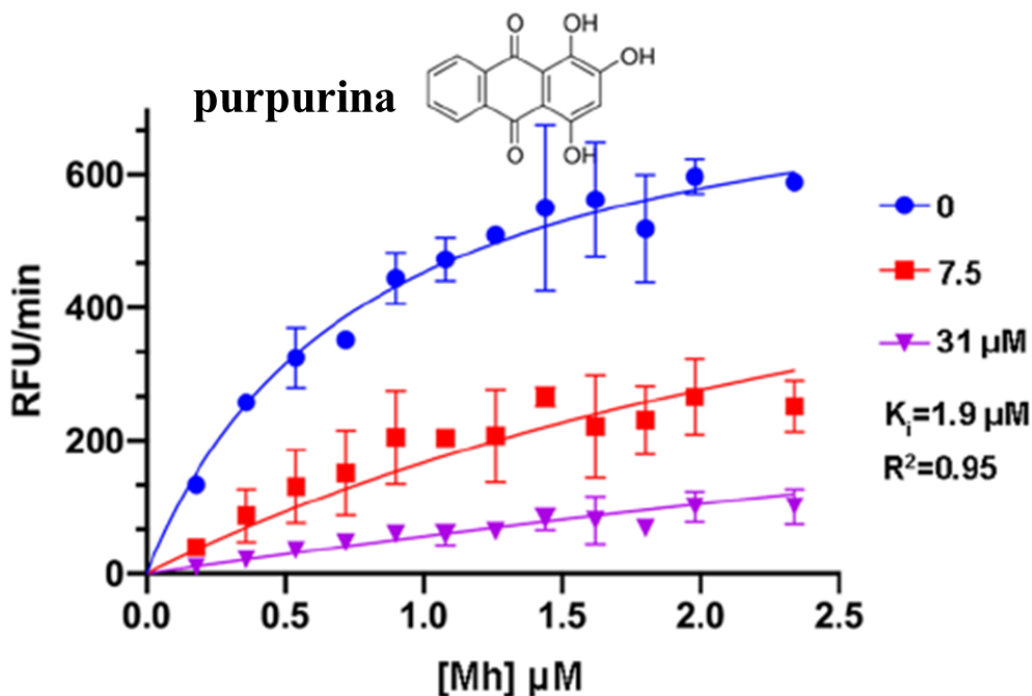


Figura 37. Caracterización cinética de la inhibición de la RNasa P por purpurina. Una constante de inhibición (K_i) de 1.9 μM fue calculada para esta reacción obtenida mediante ensayos de actividad por fluorescencia en presencia de sustrato fluorescente Mh. Un modelo competitivo es el que mejor se ajusta a los datos experimentales. El intervalo de K_i obtenida es semejante a la K_D obtenida para la interacción RNasa P-purpurina medida mediante BLI de 13 μM (Figura 35).

La purpurina se une en una región de la proteína P importante para la función de la holoenzima de la RNasa P

A través del análisis del sitio y modo de unión del inhibidor purpurina en la proteína P, se observó que la purpurina interacciona con residuos importantes en la función de la holoenzima, pues de acuerdo a un estudio independiente por Reiter y col. (58), si estos residuos son mutados por alanina, la eficiencia catalítica de la holoenzima de la RNasa P disminuye (Figura 38, panel inferior). Estos residuos son la Arg89 y la Lys51 que se encuentran interaccionando con la purpurina a través de puentes de hidrógeno con el motivo catecol (Figura 38, panel izquierdo). La Lys51 también interacciona con los fragmentos 2-MBX y FTU (Figuras 7 y 8).

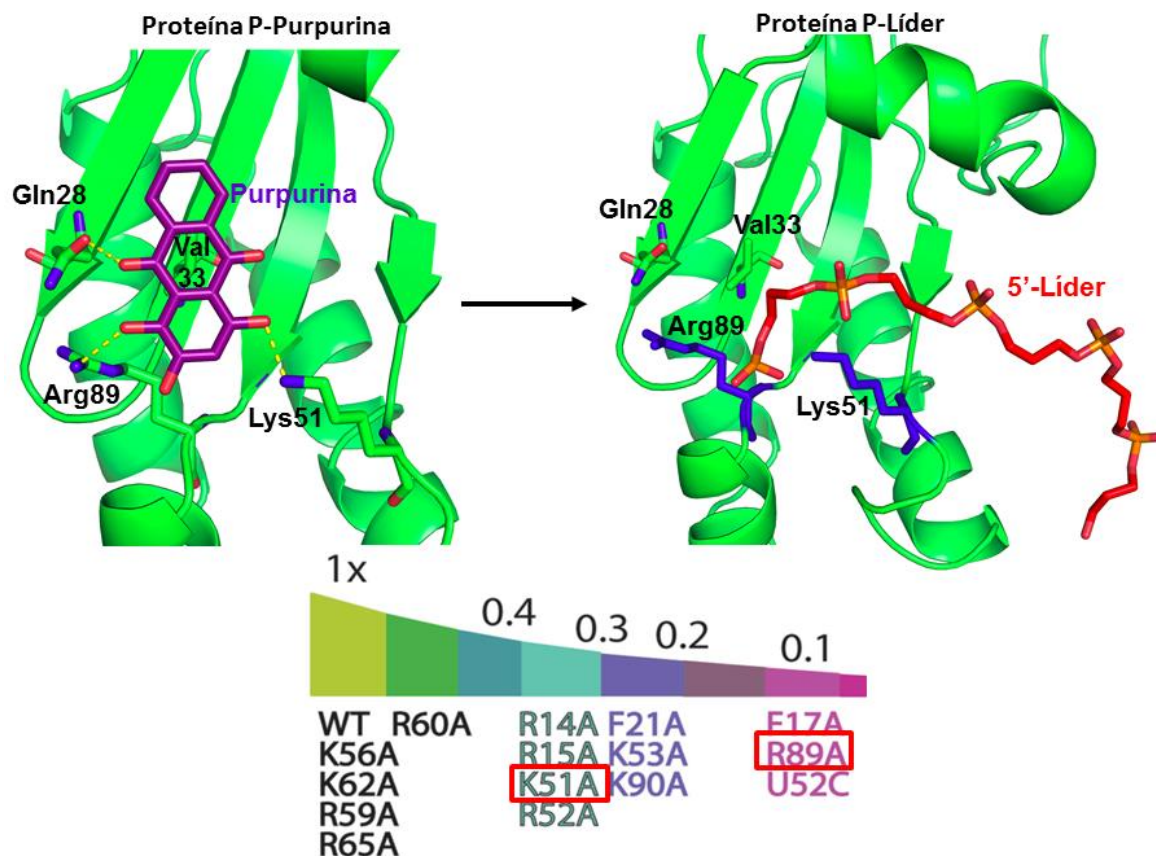


Figura 38. Los residuos con los que interactúa la purpurina son importantes para la actividad de la holoenzima. **Izquierda.** Estructura de la proteína P en complejo con purpurina donde se observan sus interacciones con residuos Lys51 y Arg89. **Derecha.** Estructura de la proteína P en complejo con 5'-líder donde se observa que ocupa el mismo sitio que la purpurina en la región del Arg89 y Lys51. **Inferior.** Fracción de la disminución en la actividad de la holoenzima por efecto de las mutaciones por alanina señaladas donde los residuos Lys51 y Arg89 interactúan con la purpurina (58).

Para obtener otro indicativo de que la purpurina bloquea la interacción del extremo 5'-líder con la proteína P, el blanco de interés (proteína P o RNasa P) fue inmovilizado en biosensores de BLI; el blanco se sometió a interacción con la purpurina y el complejo blanco-purpurina se sometió a interacción con algún sustrato (5'-líder, Mh, pre-tRNA) (Figura 39).

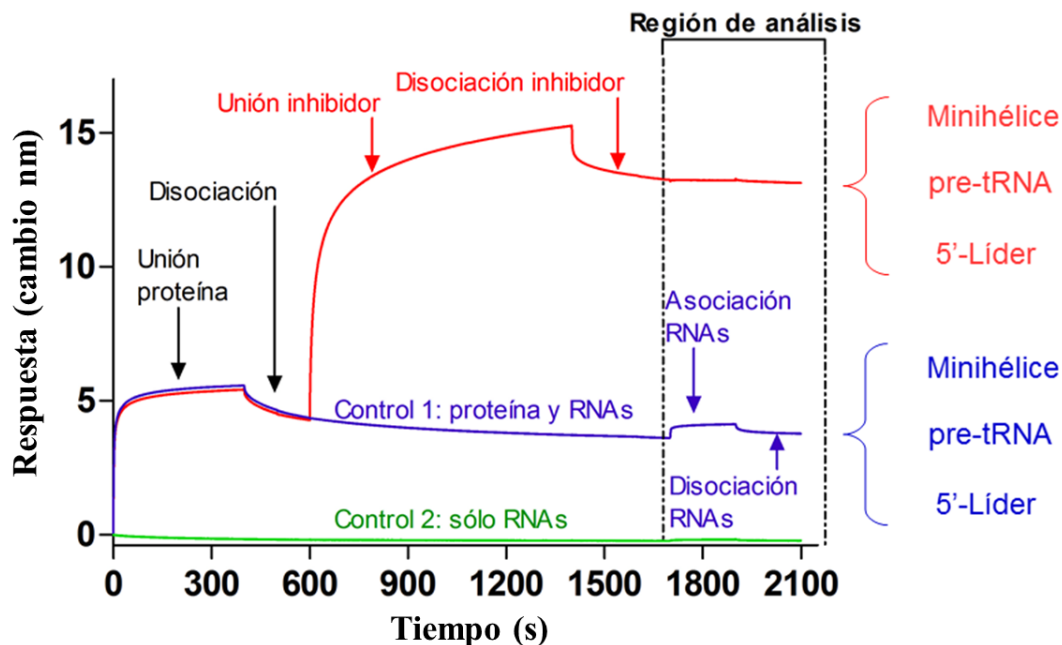


Figura 39. Diseño del ensayo de interferencia de unión de sustratos por la purpurina. Sensorgrama de un experimento típico para detectar interferencia de unión de diferentes sustratos. Primeramente, se une el blanco a evaluar (proteína P, ARN P u holoenzima); posteriormente se hace un lavado, una línea base y luego se asocia el inhibidor bajo estudio para después ensayar la unión de diversos sustratos. Los controles incluyen biosensor sin inmovilizar algún blanco, y blanco sin unir purpurina. En los tres tipos de experimento permanece constante la concentración de sustratos. La región de análisis del sensorgrama, denotada en líneas semicontinuas, es la que se presentará en los sensorgramas de las Figuras 40 y 41.

Tanto el sustrato minihélice como el extremo 5'-líder ya no se unen a la proteína P cuando previamente se ha unido la purpurina a la proteína (Figura 40, izquierda y centro). El pre-tRNA se unió parcialmente a la holoenzima en presencia de purpurina (Figura 40, panel derecho). Debe resaltarse que la subunidad ARN P reconoce al pre-tRNA en dos zonas del brazo aceptor que no contactan a la proteína P.

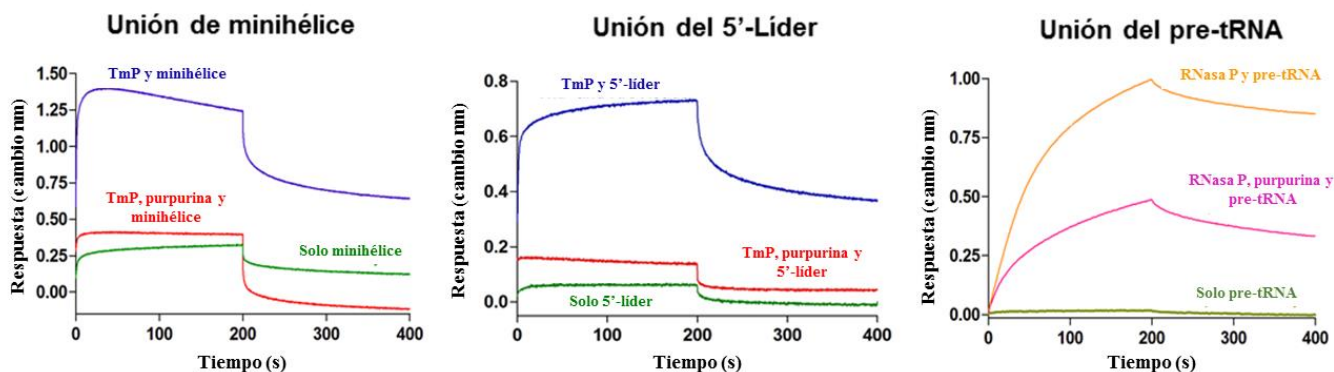


Figura 40. Inhibición de la unión de diferentes sustratos por efecto de la purpurina. Región de análisis de sensorgramas obtenidos a partir de experimentos de interferencia de unión. **Panel izquierdo:** la purpurina impide la unión del sustrato Mh no fluorescente a la proteína P. **Panel central:** la purpurina impide la unión del 5'-líder a la proteína P. **Panel derecho:** la purpurina impide la unión del sustrato canónico pre-tRNA a la RNasa P.

También se llevaron a cabo experimentos por BLI para determinar la interacción de la proteína P con el pre-tRNA y tRNA en presencia de purpurina (Figura 41). En presencia de purpurina, la proteína P no puede unirse al pre-tRNA (Figura 41).

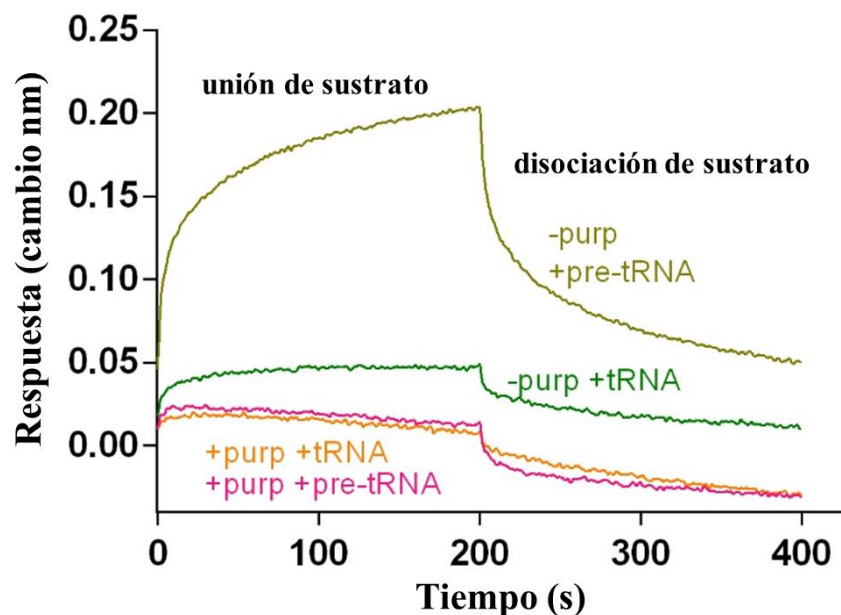


Figura 41. La purpurina impide la unión del sustrato pre-tRNA a la proteína P. Sensorgramas de impedimento de unión de sustrato a la proteína P. El sensorgrama verde oscuro con la señal más alta corresponde a la unión del sustrato pre-tRNA sin purpurina, mientras que, el sensorgrama verde fuerte es el experimento donde primeramente se unió la purpurina a la proteína P y después se midió la unión del tRNA. El sensorgrama rosa indica la unión de la purpurina y posteriormente la unión del pre-tRNA.

Otro indicador del efecto de la purpurina en el impedimento de la unión del sustrato pre-tRNA a la RNasa P se muestra en la figura 42. En este experimento se inmovilizó la holoenzima a biosensores Ni-NTA a través de la etiqueta de histidinas de la proteína P para que pudiera interaccionar con diferentes concentraciones de purpurina y después unirse a una concentración constante de sustrato. Cuanto más se incrementa la concentración de purpurina, menos proporción de sustrato pre-tRNA se une a la holoenzima (Figura 42).

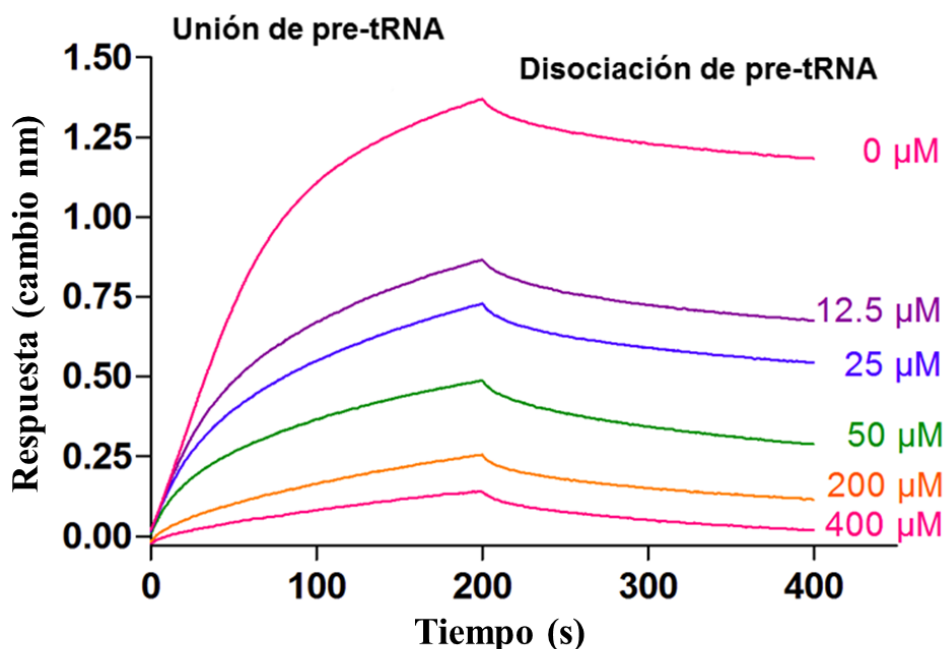


Figura 42. La purpurina inhibe la unión del pre-tRNA a la RNasa P de manera dosis-respuesta. Región de análisis del sensorgrama de la asociación del pre-tRNA a la holoenzima. El impedimento de unión del pre-tRNA ocurre de manera dosis respuesta en un intervalo de concentraciones crecientes de 0-400 µM de purpurina, mientras que permaneció constante la concentración de pre-tRNA. La RNasa P fue inmovilizada en biosensores Ni-NTA a 400 nM en amortiguador HAAMD.

Con base en los parámetros cinéticos de unión de la purpurina con la proteína P, la holoenzima y el ARN P, así como a sus propiedades moleculares y fisicoquímicas, la purpurina es un excelente compuesto líder (véase Apéndice E. Glosario de términos) para perseguir su optimización desde un punto de vista del diseño de fármacos basado en fragmentos.

Propuesta para el desarrollo de un inhibidor más potente mediante *fragment-linking*

Una de las estrategias para desarrollar un inhibidor más potente es el *fragment-linking* o la fusión de fragmentos, la cual es aplicable en casos en donde se localizan dos compuestos en sitios de unión cercanos dentro del mismo blanco, de tal manera que puedan ser unidos covalentemente por medio de un *linker* adecuado. Algunos de los *linkers* de estructura general que se utilizan para conectar fragmentos son conectores flexibles y con grupos funcionales como las oxiaminas (108) o aminoxi o conectores de polietilenglicol que contengan ya sea grupos aminoxi, oxiamina o sus combinaciones. La presencia de grupos amino puede además facilitar la permeabilidad de los compuestos en la membrana de bacterias Gram-negativas. Aplicamos estos principios para sugerir un nuevo inhibidor (compuesto B) que se muestra en la figura 43.

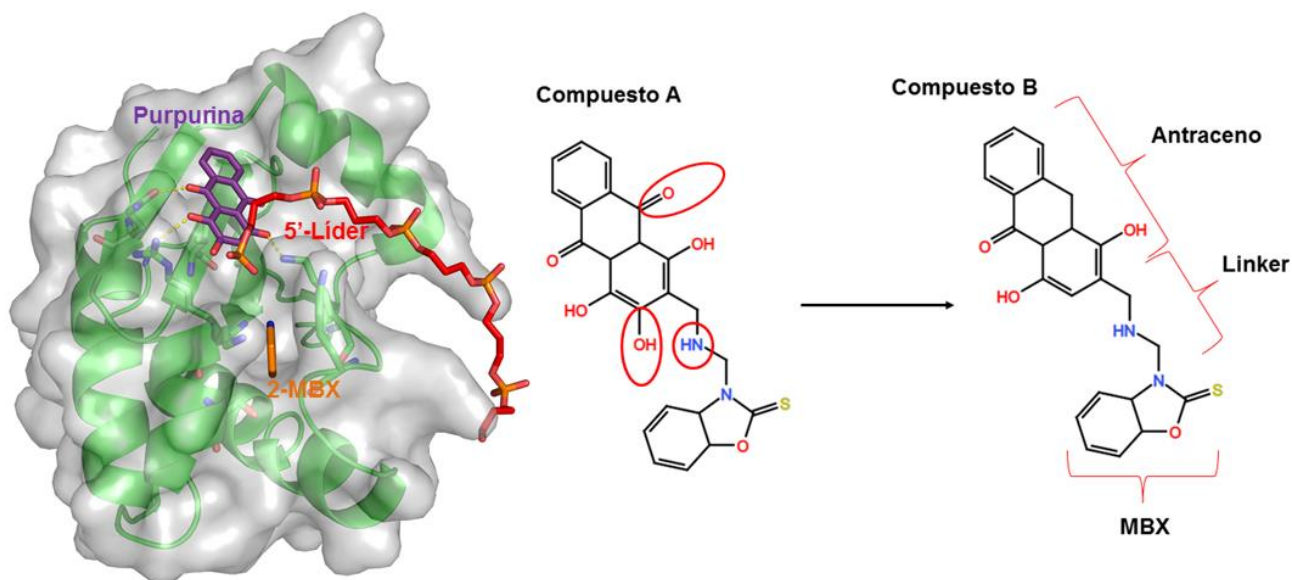


Figura 43. Propuesta de optimización de inhibidor mediante fusión de fragmentos. Panel izquierdo. Estructura cristalográfica de la proteína P en complejo con purpurina (morado) con la sobreposición del fragmento 2-MBX (naranja) y del extremo 5'-líder (rojo y naranja como cadena fosfatada) como se encuentra en la estructura cristalográfica de la holoenzima de la RNasa P en complejo con el líder. **Panel derecho.** El Compuesto A se obtendría mediante síntesis química al fusionar la purpurina y el fragmento 2-MBX mediante un *linker* adecuado, mientras que el compuesto B ya no tendría el grupo carbonilo (C=O) ni el grupo hidroxilo (R-OH) ya que de acuerdo a la estructura cristalográfica no están interactuando con algún residuo; el grupo amino añadido en el *linker* ayudaría a aumentar su permeabilidad en membrana.

Utilizando el servidor SwissSimilarity (111), el cual llevó a cabo búsquedas para encontrar ligandos similares al Compuesto B propuesto, se verificó que éste no se encuentra en ninguna de las bases de datos de fase investigacional o comercial de acceso público. Tampoco se encontró algún ligando similar en la base de datos del PDB (utilizando el servidor <http://ligand-expo.rcsb.org/>). De acuerdo con el servidor SwissADME (112), al llevar a cabo un análisis *in silico* de la dificultad de síntesis del Compuesto B, el servidor calcula un valor de 5.3, (siendo el número 10 la dificultad más alta y menos factible).

8. Discusión

A continuación, se destacan los aspectos más sobresalientes de este trabajo que permitieron conjuntar técnicas de medición de actividad, unión y determinación de estructuras tridimensionales para la búsqueda y descubrimiento de nuevos inhibidores de la ribonucleasa P bacteriana.

1) Unión

1a) Desarrollo de ensayos mediante interferometría de biocapas de la RNasa P bacteriana

Existen varias técnicas biofísicas para la detección de unión de ligandos de molécula pequeña a proteínas. De todas ellas, una tecnología que sobresale por sus características ventajosas es la interferometría de biocapas, la cual carece de microfluídica, requiere poco mantenimiento y es más barata y fácil de usar si se le compara con NMR o SPR. Además, esta técnica es apta para la búsqueda de ligandos de bajo peso molecular en formato de media y alta eficiencia. Otra característica sobresaliente, y que además es *ad hoc* a nuestras necesidades experimentales, es que es posible el uso de hasta un 10% de DMSO sin que se presente ruido en la señal, lo cual es conveniente para disolver compuestos poco solubles, como lo son los fragmentos de la biblioteca Zenobia Fragment Library-1, la biblioteca de flavonoides FL-500 y la biblioteca The Spectrum Collection. Dentro de este contexto, usando la interferometría de biocapas, se logró montar una plataforma específica para la proteína P, se validaron los biosensores Ni-NTA para medición y monitoreo de parámetros cinéticos de unión en interacciones proteína-molécula pequeña y se corroboraron los valores obtenidos mediante termoforesis a microescala.

Además, con esta técnica no solamente fue posible la medición y detección de unión de moléculas pequeñas a la proteína P, sino que se inmovilizó a la holoenzima y al ARN P catalítico a diferentes biosensores (p. ej. Biosensores Ni-NTA, SuperStreptavidin y Streptavidin), con lo cual también se midió y se corroboró de manera reproducible la interacción del extremo 5'-líder con la proteína P, la interacción del pre-tRNA con la holoenzima, la interacción de la proteína P con la subunidad ARN P y la interacción del sustrato minihélice, tanto a la proteína P como a la holoenzima (en concordancia como lo reportado en la literatura véase *Apéndice C. Resultados suplementarios. Validación de biosensores Ni-NTA mediante BLI de la interacción de proteína P y RNasa P con sus sustratos canónicos*). Finalmente, el uso de esta técnica fue crucial para inferir el mecanismo de impedimento de unión del sustrato a la holoenzima por efecto de la purpurina.

Al inicio de este proyecto, se pensó que muchos de los compuestos que se unieran a la proteína P presentarían también inhibición. Empleamos diversas bibliotecas y encontramos una serie de compuestos que presentan afinidad desde 18.4 nM hasta centenas de micromolar. Los compuestos se probaron en el ensayo de actividad por fluorescencia desarrollado y la gran mayoría de ellos no tuvieron un efecto. Así, aunque en principio sería posible utilizar a estos compuestos como guías para desarrollar inhibidores, la estrategia de iniciar la búsqueda monitoreando la unión de compuestos no es tan exitosa como lo es la búsqueda inicial de inhibición de actividad (véase "*Actividad*" más abajo).

2) Estructura

2a) Desarrollo de una metodología para la obtención de cristales de la proteína P en complejo con diferentes moléculas pequeñas

Antes de este trabajo, no se había reportado la obtención de alguna estructura de la subunidad proteica de la RNasa P bacteriana en complejo con moléculas pequeñas. Por ello, la identificación de condiciones en las que la proteína P pudiera ser cristalizada en complejo con ligandos se presentó como un reto crucial a superar.

Puntos clave a resolver durante el proyecto para la obtención de dichos cristales fueron: la optimización de condiciones de la solución madre y de los componentes de la solución crioprotectora, el tiempo de remojo, la composición de la solución madre, la identificación de componentes que inhiben la formación del complejo proteína-ligando, la identificación de disolventes adecuados para compuestos poco solubles y diferentes construcciones de la proteína P para generar diferentes empaquetamientos cristalográficos (*véase sección 10. Perspectivas*).

Un aspecto muy importante en el proceso de obtención de complejos con moléculas pequeñas fue el tiempo necesario de remojo de cristales apo de la proteína P. Se observó que entre más grande el compuesto, mayor será el tiempo necesario para la obtención del complejo proteína-ligando. Por ejemplo, para el caso del 2-MBX fue un remojo relativamente rápido de 4 min, mientras que, para la purpurina, remojos de 1 semana no fueron suficientes ya que el complejo proteína P-purpurina se formó hasta los dos meses y medio de remojo.

2b) Detección por cristalografía de rayos X de dos sitios en la proteína P para unión de molécula pequeña

A través de una búsqueda mediante cristalografía de rayos X, se detectaron dos fragmentos usando la biblioteca de compuestos The Zenobia Fragment Library-1 (Zenobia Therapeutics), los cuales se unieron en el mismo sitio en la proteína P. A diferencia del fragmento 2-MBX, no fue posible detectar la unión del fragmento FTU a la proteína P a través de BLI, por lo que la única manera en que pudo ser detectada su unión fue mediante la metodología combinada de cristalografía de rayos X en conjunto con el método PanNNDA. A nuestro conocimiento, la estructura de la proteína P en complejo con la purpurina (código PDB: 6MAX), así como el complejo con el fragmento FTU (código PDB: 6CQC) y el fragmento 2-MBX (estructura no depositada aún en el PDB), son las primeras tres estructuras cristalográficas de la proteína P en complejo con moléculas orgánicas de bajo peso molecular tipo fármaco.

Los fragmentos FTU y 2-MBX se unen en una zona de la proteína P que es importante para la función de la holoenzima, ya que ambos compuestos son flanqueados por las cadenas alifáticas de los residuos Ly53 y Lys90, que son residuos clave relacionados con la eficiencia catalítica de la holoenzima de la RNasa P (58). Aunque en la estructura cristalográfica de la holoenzima el líder sólo es observado a través de su cadena principal fosfatada, de acuerdo a la secuencia de 10 bases obtenida a partir de la secuencia consenso de todos los precursores 5'-líder en bacteria (35), es posible predecir que en la zona de los residuos Lys53 y Lys90 está presente una guanina, aunque la estructura de alta resolución de la proteína P en complejo con el líder es necesaria para entender su interacción con las bases del líder.

Por otro lado, el sitio de unión de la purpurina se localiza dentro del surco de unión del extremo 5'-líder del pre-tRNA y se encuentra a una distancia de 24 Å del sitio activo de la holoenzima. El modo de unión de la purpurina es a través de una serie de interacciones hidrofílicas e hidrofóbicas. La purpurina interactúa con una zona hidrofóbica en el surco de unión del extremo 5'-líder formado por los residuos Val33, Leu35 e Ile87, mientras que las interacciones hidrofílicas están mediadas por los residuos Gln28, Lys51 y Arg89.

La purpurina, con un peso molecular de 256 Da, se encuentra en el intervalo de peso de un fragmento (150-300 Da) y es susceptible de optimización hacia inhibidores más potentes, eficientes y específicos a través de la funcionalización en sitios externos de la molécula.

2c) Integración de la plataforma estructura-unión

La optimización de un compuesto se refiere a la adición de grupos funcionales o modificaciones químicas, de tal manera que se obtenga un ligando con características mejoradas con respecto al “hit” inicial en cuanto a una mayor afinidad. Además, es necesario identificar si el sitio de unión de un ligando en la estructura cristalográfica es importante para la función del blanco para perseguir su estudio y descartar que sea un artefacto cristalográfico, por ejemplo, un evento de unión fortuito debido al empaquetamiento del cristal sin relevancia biológica en la función de una enzima. El sitio de unión del 2-MBX es biológicamente relevante, pues de acuerdo a alineamientos de secuencias de regiones 5'-líder de pre-tRNAs en *T. maritima*, es muy probable que ese sitio lo ocupe una guanina, en cuyo caso el 2-MBX mimetiza la base que se aloja en esa posición.

Gracias al ensayo de unión por BLI, conjuntamos la información estructural y obtuvimos compuestos con una mayor afinidad a la proteína P, obteniendo los compuestos optimizados C-7 y C-10, que, si bien presentan actividad antibacteriana, no fue posible la obtención de cristales en complejo con la proteína P, por lo que el proceso de optimización hacia compuestos tipo fármaco se vio comprometido utilizando esta estrategia. Este hecho es probablemente atribuible a que no hay espacio dentro del cristal por el grupo espacial y se obstruye el acceso de compuestos al sitio de unión de interés en la proteína P que es el surco de unión la extremo 5'-líder. Con base en esta hipótesis, es recomendable inducir nuevos grupos espaciales de la proteína P (véase sección 9. Conclusiones y sección 10. Perspectivas), que puedan tener un arreglo tridimensional tal que permita la unión de ligandos en este sitio.

3) Actividad

3a) Optimización de un ensayo mediante espectroscopía de fluorescencia

Tradicionalmente, se han llevado a cabo ensayos de actividad en geles empleando radioisótopos, lo cual hace al ensayo tardado, costoso y potencialmente tóxico (82); además, esta configuración experimental en geles hace impráctica su aplicación para un formato de búsqueda de inhibidores de alta eficiencia con bibliotecas de cientos a miles de compuestos.

El ensayo desarrollado en este trabajo hace uso de un sustrato fluorescente artificial que presenta una menor afinidad por la holoenzima que el sustrato canónico pre-tRNA (84), característica que lo hace apto para la detección de moléculas pequeñas que presentan una afinidad débil por el blanco (63, 74). Además, este ensayo es apto para un formato de alta eficiencia, la caracterización cinética puede ser llevada de manera rápida y es posible ensayar de cientos a miles de compuestos en pocos días, permitiendo el uso de cantidades pequeñas de muestra, lo cual disminuye costos comparado con ensayos que emplean radioisótopos.

Otra de las ventajas importantes del ensayo es que soporta hasta 30% de DMSO, característica que le confiere una gran ventaja sobre otras técnicas que requieren de la dilución de compuestos orgánicos poco solubles en soluciones biológicas. Además, el uso del fluoróforo AlexaFluor y del apagador Black Hole Quencher no interfieren con la actividad de la holoenzima y el ensayo es fácilmente adaptable para llevar a cabo experimentos con placas de 96 o 384 pocillos en lectores de placas.

Inicialmente, este ensayo puede ser usado con el sustrato artificial Mh para llevar a cabo búsquedas de cientos o miles de fragmentos y posteriormente corroborar estos fragmentos o probar los compuestos optimizados mediante estructura con el sustrato fluorescente pre-tRNA-mitades.

3b) Integración de la plataforma actividad-uniión-estructura

A diferencia de la plataforma de unión, iniciar una búsqueda a través del ensayo de actividad resultó en la obtención de compuestos más promisorios. Después de una búsqueda con la biblioteca de compuestos The Spectrum Collection, fue posible detectar cuatro nuevos inhibidores de la RNasa P bacteriana: la juglona, la hemateína, la purpurina y el violeta de genciana, los cuales mostraron inhibir la actividad de la RNasa P en el intervalo micromolar bajo.

Una vez determinada la inhibición de la holoenzima por algún compuesto en presencia de sustrato Mh (Figura 26, panel superior), se corroboró la inhibición empleando el sustrato fluorescente pre-tRNA-mitades (Figura 26, panel inferior).

Por primera vez, se ha desarrollado un ensayo de actividad para la RNasa P bacteriana diseñado específicamente para la detección de fragmentos que inicialmente presentan una afinidad baja por el blanco, por lo que su aplicación es muy apropiada para la metodología del diseño de fármacos basado en fragmentos.

4) Mecanismo de inhibición de la RNasa P bacteriana por la purpurina

A la fecha, todos los inhibidores reportados de la RNasa P bacteriana tienen como blanco a la subunidad catalítica de ARN y a nuestro conocimiento, no existe algún inhibidor reportado que actúe mediante la interacción con la subunidad proteica. Con base en los resultados, el modelo de inhibición que mejor se ajusta a los datos experimentales es un modelo de inhibición competitiva.

La purpurina puede ser considerada un inhibidor competitivo clásico si se considera a la proteína P no como una subunidad, sino como parte de una unidad con la holoenzima, y la proteína puede ser vista como parte del sitio activo que se conformaría por el sitio catalítico del ARN P y el surco de unión del extremo 5'-líder en la proteína P.

Como un ejemplo análogo al mecanismo de inhibición descrito en este trabajo, Wang y col. (108), reportan la estructura del dominio catalítico de la proteasa USP14 unida al inhibidor IU1 y sus análogos. A diferencia del clásico modo de unión en el cual compuestos de molécula pequeña interactúan directamente con la cisteína del centro catalítico, el inhibidor IU1 y sus tres análogos se unen a un sitio en el surco en forma de palma-pulgar que se encuentra a 12 Å del centro catalítico. Dicho evento de unión no promueve un cambio conformacional en USP14 que lleve a la inactividad de la enzima; en cambio, la unión del inhibidor bloquea el acceso de la región C-terminal de la ubiquitina hacia el centro catalítico de la enzima. A pesar de este complejo y peculiar modo de inhibición reportado, los autores indican que el inhibidor se comporta de manera competitiva. Este mecanismo es idéntico en principio al observado en la RNasa P por efecto de la purpurina, la cual se localiza a 24 Å de distancia del sitio catalítico, por lo que también puede ser considerada como un inhibidor alostérico que actúa de manera competitiva. Los moduladores alostéricos ofrecen ventajas que los hacen apropiados como candidatos a fármacos (109).

9. Conclusiones

Con base en el análisis de los resultados obtenidos en este trabajo, se llegó a las siguientes conclusiones:

1. Una de las conclusiones más relevantes del presente estudio es que, con base en los resultados obtenidos, fue posible confirmar nuestra hipótesis que fue la detección e identificación de compuestos orgánicos de bajo peso molecular que inhiban a la holoenzima de la RNasa P bacteriana vía la unión a la proteína P.
2. En cuanto el objetivo general, se descubrieron nuevos inhibidores de bajo peso molecular de la RNasa P bacteriana que actúan mediante la interacción con la subunidad proteica a través de la implementación de una metodología que involucró ensayos de actividad, ensayos de unión y determinación de la estructura tridimensional de complejos inhibidor-proteína P.
3. Con respecto a los objetivos particulares, 1) se desarrolló un ensayo para analizar la unión de moléculas pequeñas a la proteína P, al ARN P y a la holoenzima de la RNasa P bacteriana; 2) se desarrolló un ensayo para analizar la actividad de RNasa P bacteriana en un formato de medio rendimiento, 3) se identificaron nuevos inhibidores de la holoenzima de la RNasa P bacteriana y 4) se determinó el modo de unión de inhibidores en la proteína P mediante cristalografía de rayos X.
4. Uno de los resultados más sobresalientes del presente trabajo fue la determinación del orden secuencial óptimo de técnicas en la plataforma desarrollada para el descubrimiento de nuevos inhibidores de la RNasa P bacteriana el cual fue: actividad-unión-estructura. Esto lo determinamos así, ya que, como parte de nuestro proceso de discriminación de “hits”, en primer lugar lo más importante es asegurar una inhibición de la holoenzima por efecto de algún compuesto, que el compuesto no inhiba la actividad de la subunidad aislada ARN P, posteriormente asegurar que ese compuesto se une a la proteína P de manera dosis-dependiente y descartar aquellos que se unieron la subunidad ARN P y, finalmente, es crucial conocer el sitio y modo de unión en la proteína P del compuesto que pasó los filtros mencionados con el fin generar hipótesis de un mecanismo de inhibición de la holoenzima.
5. Se desarrolló un ensayo de actividad de RNasa P bacteriana mediante espectroscopía de fluorescencia en un lector de placas en formato de 384 pocillos, donde el tiempo de ensayo y el volumen de muestra se optimizó.
6. Se montó y se optimizó un ensayo de unión en formato de media eficiencia en placas de 96 pocillos mediante interferometría de biocapas para la proteína P con el fin de llevar a cabo la búsqueda de ligandos de bajo peso molecular.
7. Se logró desarrollar una metodología para obtener cristales de la proteína P en complejo con moléculas orgánicas de bajo peso molecular, empleando el agente precipitante PEG-1000 a

diferentes porcentajes tanto para la cristalogénesis de los complejos en el licor madre como para la formulación de la solución crioprotectora, así como PEG-200 y PEG-400 para disolver dichos compuestos en experimentos de remojo.

8. A partir de la detección del fragmento 2-MBX por cristalografía de rayos X, éste pudo ser optimizado mediante diseño guiado por estructura hacia el compuesto optimizado C-7 y el compuesto optimizado C-10, los cuales se unen con una afinidad mejorada a la proteína P de manera dosis-dependiente, inhiben la actividad de la holoenzima de la RNasa P e inhiben el crecimiento bacteriano.
9. Se logró obtener la estructura cristalográfica de la proteína P en complejo con el fragmento FTU el cual se une en el mismo sitio que el fragmento 2-MBX. El hallazgo de este segundo fragmento, en el mismo sitio que el 2-MBX, ayudó al descubrimiento de un *hot spot* o sitio privilegiado para la unión de fragmentos en la proteína P.
10. Se determinó que los residuos Lys53 y Lys90, los cuales se encuentran flanqueando tanto al fragmento 2-MBX como al fragmento FTU, son de vital importancia para la actividad de la holoenzima de la RNasa P bacteriana.
11. Lamentablemente, no fue posible obtener cristales en complejo con los compuestos optimizados C-7 y C-10, posiblemente debido al empaquetamiento de la proteína P dentro del cristal donde los canales intersticiales no son suficientemente grandes para que estos compuestos difundan hasta el sitio de unión del 2-MBX. Sin embargo, esta observación es importante ya que nos marca la pauta para planear una estrategia con la cual inducir nuevos grupos espaciales de la proteína P con los que sea posible albergar compuestos optimizados de un peso molecular incrementado en los sitios de interés.
12. El iriginol hexaacetato y el irigenol son candidatos interesantes de perseguir su optimización ya que se sabe que el iriginol hexaacetato es un inhibidor de la RNasa P de *B. subtilis* y de *T. maritima* y que el irigenol es un ligando dosis-dependiente de la proteína P.
13. A partir de la búsqueda mediante BLI en la biblioteca Spectrum, se detectaron 27 “hits” en total los cuales se unen de manera dosis-dependiente a la proteína P. De ellos, 18 presentan afinidad alta en el intervalo nanomolar (18.4 – 692 nM), mientras que 9 de ellos tienen afinidad en el intervalo micromolar (1-913 μ M).
14. En la biblioteca de flavonoides FL-500 se detectaron nueve “hits” que interaccionan de manera dosis-respuesta con la proteína P. De ellos, cinco tienen afinidad en el intervalo de fracciones y unidades de micromolar (0.42 – 6.7 μ M); los cuatro restantes presentan afinidades en el intervalo de decenas de micromolar (15.8-77.7 μ M).

15. Con el objetivo de validar a la purpurina como un verdadero inhibidor de la RNasa P bacteriana, fue necesaria la corroboración de los resultados empleando técnicas ortogonales como BLI, MST, cristalografía de rayos X y actividad por espectroscopía de fluorescencia.
16. Se logró obtener la estructura de la proteína P en complejo con el inhibidor purpurina cuyo sitio de unión se localiza en el sitio de unión para la parte distal del 5'-líder. La estructura del complejo se resolvió a una resolución de 1.4 Å y la ocupación de la purpurina fue total.
17. Por primera vez, a nuestro conocimiento, se logró determinar la estructura cristalográfica de la proteína P en complejo con un inhibidor validado de la RNasa P.
18. El modo de unión de la purpurina a través de sus diferentes motivos estructurales es la siguiente: se observa un parche hidrofóbico en el sitio de unión con una red de interacciones hidrofóbicas que involucran a la Leu35, Val33 e Ile87. El motivo benceno de la purpurina interacciona con la Leu35, mientras que el motivo benzoquinona interacciona con la Val33 y con la Gln28. El motivo catecol de la purpurina interacciona con la proteína P mediante puentes de hidrógeno con la Lys51 y la Arg89.
19. Se determinó que los residuos de la proteína P, Arg89 y Lys51, que interaccionan con la purpurina, son de vital importancia para la actividad de la holoenzima de la RNasa P bacteriana, ya que, tal como lo reporta un grupo de investigación independiente (54), mutaciones puntuales de estos residuos son perjudiciales para la eficiencia catalítica de la ribozima. Existe evidencia no estructural de que estos mismos residuos son importantes para la interacción de la proteína P con el extremo 5'-líder (50-53, 58). Esto es indicativo de que el inhibidor purpurina interacciona en una zona de la proteína P que es importante para la actividad de la ribozima, por lo que ayuda a explicar como es que la purpurina es capaz de modular la actividad general de la holoenzima aunque éste no se encuentre interaccionando con el sitio catalítico del ARN P.
20. Con base en la evidencia obtenida, la hipótesis del mecanismo de inhibición de la purpurina es que el compuesto actúa vía la unión a la proteína P de manera específica, lo que ocasiona un impedimento estérico en la unión del sustrato específicamente en su extremo 5'-líder, provocando una escisión fallida del líder, de tal manera inhibiendo efectivamente la actividad de la holoenzima.
21. La plataforma desarrollada en este estudio de investigación permitió el descubrimiento y validación de un verdadero inhibidor de bajo peso molecular de la holoenzima de la RNasa P bacteriana que actúa por una vía novedosa al inhibir a la RNasa P interaccionando con la subunidad proteica.
22. Determinamos que la subunidad proteica de la RNasa P representa un muy buen blanco farmacéutico para el desarrollo de inhibidores. Nosotros logramos demostrar que la proteína P es un blanco terapéutico que puede ser modulado por moléculas orgánicas de bajo peso molecular. Esto, visto a través de evidencia como lo es la inhibición de la actividad en presencia de la

purpurina, la unión de la purpurina de manera dosis-dependiente a la proteína P y la estructura cristalográfica de la proteína P en complejo con la purpurina.

- 23.** El Compuesto B, hipotetizado en este estudio como posible estrategia de optimización del inhibidor purpurina, es una metodología interesante que podría resultar en el desarrollo de un inhibidor más potente y específico de la RNasa P bacteriana.

10. Perspectivas

1. Optimizar el inhibidor purpurina mediante diseño basado en estructura para obtener un compuesto líder con buenas características de biodisponibilidad, afinidad y potencia inhibitoria incrementada hacia el blanco.
2. Optimizar el ligando FTU hacia compuestos líder con una afinidad por la proteína P que sea semejante a los compuestos optimizados C-7 y C-10 o con el inhibidor purpurina (en el intervalo de unidades de micromolar).
3. Generar un inhibidor más potente al conectar la purpurina con el fragmento 2-MBX o bien, con el fragmento FTU aplicando la estrategia de unión de fragmentos.
4. Realizar una nueva estrategia de cristalización apta para generar complejos proteína P-compuestos optimizados, por medio del diseño de una red cristalina en la que exista espacio suficiente para la difusión de los compuestos optimizados C-7 y C-10 entre los canales intersticiales del cristal. Una manera de llevar a cabo este enfoque, es empleando proteínas de fusión como las repeticiones de anquirinas, las cuales han sido utilizadas en otros sistemas biológicos (113).

11. Bibliografía

1. Walsh C. (2003) Where will new antibiotics come from? *Nat Rev Microbiol.* 1:65-70.
2. National Research Council (2006) Treating infectious diseases in a microbial world: report of two workshops on novel antimicrobial therapeutics. National Academies Press, Washington, DC, USA. pp 8-11.
3. Coates ARM, Halls, G, Hu Y (2011) Novel classes of antibiotics or more of the same? *Br J Pharmacol.* 163:184-194.
4. Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, Mueller A, Schäberle TF, Hughes DE, Epstein S, Jones M, Lazarides L, Steadman VA, Cohen DR, Felix CR, Fetterman KA, Millett WP, Nitti AG, Zullo AM, Chen C, Lewis K (2015) A new antibiotic kills pathogen without detectable resistance. *Nature.* 517:455-459.
5. Falagas ME, Karageorgopoulos DE (2008) Pandrug resistance (PDR), extensive drug resistance (XDR), and multidrug resistance (MDR) among gram-negative bacilli: need for international harmonization in terminology. *Clin Infect Dis.* 46:1121-2.
6. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J, Struelens MJ, Vatopoulos A, Weber JT, Monnet DL (2012) Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect.* 18:268-281.
7. Roca I, Akova M, Baquero F, Carlet J, Cavaleri M, Coenen S, Cohen J, Findlay D, Gyssens I, Heuer OE, Kahlmeter G, Kruse H, Laxminarayan R, Liébana E, López-Cerero L, MacGowan A, Martins M, Rodríguez-Baño J, Rolain JM, Segovia C, Sigauque B, Tacconelli E, Wellington E, Vila J (2015) The global threat of antimicrobial resistance: science for intervention. *New Microbes New Infect.* 6:22-29.
8. Antimicrobial resistance: Tackling a crisis for the health and wealth of nations. The Review on antimicrobial resistance. Chaired by Jim O'Neill. HM Government & Wellcome Trust. 2014.
9. Wainwright M, Kristiansen JE (2011) On the 75th anniversary of prontosil. *Dyes Pigm.* 88:231-234.
10. Colebrook L, Kenny M (1936) Treatment of puerperal infections and experimental infection in mice with prontosil. *Lancet.* 227:1279-1281.
11. Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell.* 35:649-857.
12. Knowling S, Morris KV (2011) Non-coding RNA and antisense RNA. Nature's trash or treasure? *Biochimie.* 93:1922-1927.
13. Sherwood AV (2016) Riboswitch-mediated gene regulation: novel RNA architectures dictate gene expression responses. *Annu Rev Microbiol.* 70:361-374.
14. Deigan KE, Ferre-D'amaré AR (2011) Riboswitches: discovery of drugs that target bacterial gene-regulatory RNAs. *Acc Chem Res.* 44(12):1329-1338.
15. Franceschi F (2007) Back to the future: the ribosome as an antibiotic target. *Future Microbiol.* 2(6):571-574.
16. Canu A, Malbruny B, Coquemont M, Davies TA, Appelbaum PC, Leclercq R (2002) Diversity of ribosomal mutations conferring resistance to macrolides, clindamycin, streptogramin, and telithromycin in *Streptococcus pneumoniae*. *Antimicrob Agents Chemother.* 46(1):125-131.
17. Cocozaki AI, *et al* (2016) Resistance mutations generate divergent antibiotic susceptibility profiles against translation inhibitors. *Proc Natl Acad Sci USA.* 113(29):8188-8193.

18. Wilcox SK, Cavey GS, Pearson JD (2001) Single ribosomal protein mutations in antibiotic-resistant bacteria analyzed by mass spectrometry. *Antimicrob Agents Chemother.* 45(11):3046-3055.
19. Springer B, Kidan YG, Prammananan T, Ellrott K, Böttger EC, Sander P (2001) Mechanisms of streptomycin resistance: selection of mutations in the 16s rRNA gene conferring resistance. *Antimicrob Agents Chemother.* 45(10):2877-2884.
20. Tait-Kamradt A, Davies T, Cronan M, Jacobs MR, Appelbaum PC, Sutcliffe J (2000) Mutations in 23S rRNA and ribosomal protein L4 account for resistance in pneumococcal strains selected *in vitro* by macrolide passage. *Antimicrob Agents Chemother.* 44(8):2118-2125.
21. Masquida B, Westhof E (2011) RNase P: At last, the key finds its lock. *RNA.* 17:1615-1618.
22. Cech TR (2000) The ribosome is a ribozyme. *Science.* 289:878-879.
23. Randau L, Schröder I, Söll D (2008) Life without RNase P. *Nature.* 453:120-124.
24. Altman S (2017) The Origin of Life, “the World of RNA,” and the Problem of Antibiotic Resistance. *Herald of the Russian Academy of Sciences.* 87(6):474-479. Pleiades Publishing, Ltd.
25. De la Sierra-Gallay IL, Pellegrini O, Condon C (2005) Structural basis for substrate binding, cleavage and allostery in the tRNA maturase RNase Z. *Nature.* 433(7026):657-661.
26. Hartmann RK, Gößringer M, Späth B, Fischer S, Marchfelder A (2009) Chapter 8: The making of tRNAs and more – RNase P and tRNase Z. *Molecular Biology of RNA Processing and Decay in Prokaryotes.* 319-368.
27. De la Sierra-Gallay IL, Mathy N, Pellegrini O, Condon C (2006) Structure of the ubiquitous 3' processing enzyme RNase Z bound to transfer RNA. *Nat Struct Mol Biol.* 13(4):376-377.
28. Kunzmann A, Brennicke A, Marchfelder A (1998) 5' end maturation and RNA editing have to precede tRNA 3' processing in plant mitochondria. *Proc Natl Acad Sci USA.* 95:108-113.
29. Kazantsev AV, Pace NR (2006) Bacterial RNase P: a new view of an ancient enzyme. *Nat Rev Microbiol.* 4(10):729-740.
30. Smith D, Pace NR (1993) Multiple magnesium ions in the ribonuclease P reaction mechanism. *Biochem.* 32(20):5273-5281.
31. Kikovska E, Staffan GS, Kirsebom LA (2007) Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proc Natl Acad Sci USA.* 104(7):2062-2067.
32. Li Y, Altman S (2003) A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs. *Proc Natl Acad Sci USA.* 100(23):13213-13218.
33. Altman S, Wesolowski D, Guerrier-Takada C, Yong L (2005) RNase P cleaves transient structures in some riboswitches. *Proc Natl Acad Sci USA.* 102(32):11284-11289.
34. Hernandez-Cid A, Aguirre-Sampieri S, Diaz-Vilchis A, Torres-Larios A (2012) Ribonucleases P/MRP and the expanding ribonucleoprotein world. *IUBMB Life.* 64:521-528.
35. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragon A (2010) Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature.* 468:784-789.
36. Sun L, Harris ME (2007) Evidence that binding of C5 protein to P RNA enhances ribozyme catalysis by influencing active site metal ion affinity. *RNA.* 13:1505-1515.
37. Reich C, Olsen GJ, Pace B, Pace NR (1988) Role of the protein moiety of ribonuclease P, a ribonucleoprotein enzyme. *Science.* 239:178-181.
38. Kurz JC, Niranjanakumari S, Fierke CA (1998) Protein component of *Bacillus subtilis* RNase P specifically enhances the affinity for precursor-tRNA^{Asp}. *Biochem.* 37:2393-2400.
39. Buck AH, Dalby AB, Poole AW, Kazantsev AV, Pace NR (2005) Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J.* 24:3360-3368.

40. Son A, Choi S, Han G, Seong BL (2015) M1 RNA is important for the in-cell solubility of its cognate C5 protein: implications for RNA-mediated protein folding. *RNA Biol.* 12(11):1198-1208.
41. Stams T, Niranjanakumari S, Fierke CA, Christianson DW (1998) Ribonuclease P protein structure: evolutionary origins in the translational apparatus. *Science.* 280(5364):752-755.
42. Willkomm, D.K., Willkomm DK, Gruegelsiepe H, Goudinakis O, Kretschmer-Kazemi FR, Bald R, Erdmann VA, Hartmann RK (2003) Evaluation of bacterial RNase P RNA as a drug target. *ChemBioChem.* 4:1041-1048.
43. Drainas D (2016) Antibiotics and RNase P. *Antibiotics.* 5:15.
44. Ellis JC, Brown JW (2009) The RNase P family. *RNA Biol.* 6:362-9.
45. Walker SC, Engelke DR (2006) Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol.* 41(2):77-102.
46. Marquez SM, Chen JL, Evans D, Pace NR (2006) Structure and function of eukaryotic Ribonuclease P RNA. *Mol Cell.* 24(3):445-456.
47. Chen JL, Pace NR (1997) Identification of a universally conserved core of ribonuclease P RNA. *RNA.* 3:557-560.
48. Christian EL, Kaye NM, Harris ME (2000) Helix P4 is a divalent metal ion binding site in the conserved core of the ribonuclease P ribozyme. *RNA.* 6:511-519.
49. Christian EL, Smith KMJ, Perera N, Harris ME (2006) The P4 metal binding site in RNase P RNA affects active site metal affinity through substrate positioning. *RNA.* 12:1463-1467.
50. Kirsebom LA, Vioque A (1996) RNase P from bacteria. Substrate recognition and function of the protein subunit. *Mol Biol Rep.* 22(2-3):99-109.
51. Mondragón A (2010) En Protein Reviews Volumen 10: Ribonuclease P. Editores: Liu, F, Altman S Ch.4:63-78 Springer.
52. Smith JK, Hsieh J, Fierke CA (2007) Importance of RNA-protein interactions in bacterial ribonuclease P structure and catalysis. *Biopol.* 87:329-338.
53. Koutmou KS, Zahler, NH, Kurz JC, Campbell FE, Harris ME, Fierke CA (2010) Protein-precursor tRNA contact leads to sequence-specific recognition of 5' leaders by bacterial ribonuclease P. *J Mol Biol.* 396(1):195-208.
54. Kazantsev AV, Krivenko AA, Harrington DJ, Carter RJ, Holbrook SR, Adams PD, Pace NR (2003) High-resolution structure of RNase P protein from *Thermotoga maritima*. *Proc Natl Acad Sci USA.* 100(13):7497-7502.
55. Chen XP, Du GH (2007) Target validation: A door to drug discovery. *Drug Discov Ther.* 1(1):23-29.
56. Gashaw I, Ellinghaus P, Sommer A, Asadullah K (2011) What makes a good drug target? *Drug Discov Today.* 16(23/24):1037-1043.
57. Prinz F, Schlange T, Asadullah K (2011) Believe it or not: how much can we rely on published data on potential drug targets? *Nat Rev Drug Discov.* 10(9):712-712.
58. Reiter NJ, Osterman, AK, Mondragón A (2012) The bacterial ribonuclease P holoenzyme requires specific, conserved residues for efficient catalysis and substrate positioning. *Nucl Ac Res.* 40(20):10384-10393.
59. Wang J, Li S, Li H (2019). Drug design of “undruggable” targets. *Chin J Chem.* 37(5):501-512.
60. Yang NJ, Hinner MJ (2015) Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Methods Mol Biol.* 1266:29-53.
61. Dworkina JP, Lazcano A, Miller SL (2003) The roads to and from the RNA world. *J Theoret Biol.* 222:127-134.
62. Nelson JW, Breaker RR (2017) The lost language of the RNA World. *Sci Signal.* 10(483).

63. Rees DC, Congreve M, Murray CW, Carr R (2004) Fragment-based lead discovery. *Nat Rev Drug Discov.* 3:660-672.
64. Bollag G. (2010) Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature.* 30:467(7315):596-9.
65. Concepcion J, Witte K, *et al* (2009) Label-free detection of biomolecular interactions using biolayer interferometry for kinetic characterization. *Comb Chem High Throughput Screen.* 12:791-800.
66. Wartchow CA, Podlaski F, Li S, Rowan K, Zhang X, Mark D, Huang KS (2011) Biosensor-based small molecule fragment screening with biolayer interferometry. *J Comput Aided Mol Des.* 25:669-676.
67. Roh C, Kim SE, Jo SK (2011) Label free inhibitor screening of hepatitis C virus (HCV) NS5B viral protein using RNA oligonucleotide. *Sensors.* 11:6685-6696.
68. Jerabek-Willemsen M, André T, Wanner R, Roth HM, Duhr S, Baaske P, Breitsprecher D (2014) MicroScale Thermophoresis: Interaction analysis and beyond. *J Mol Struct.* 1077:101-113.
69. Gaffarogullari EE, Krause A, Balbo J, Herten DP, Jäschke A (2013) Microscale thermophoresis provides insights into mechanism and thermodynamics of ribozyme catalysis AlphaScreen assay. *RNA Biol.* 10(12):1815-1821.
70. Entzian C, Schubert T (2016) Studying small molecule–aptamer interactions using MicroScale Thermophoresis (MST). *Methods.* 97:27-34.
71. Smyth MS, Martin JHJ (2000) X Ray crystallography. *J Clin Pathol: Mol Pathol.* 53:8-14.
72. Jaeger J (2004) Macromolecular structure determination by X-ray crystallography. *Encyclopedia of Life Sciences, John Wiley & Sons.* doi: 10.1038/npg.els.0002723.
73. Zheng H, Hou J, Zimmerman MD, Wlodawer A, Minor W (2014) The future of crystallography in drug discovery. *Expert Opin Drug Discov* 9(2):125-137.
74. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW (1996) Discovering high-affinity ligands for proteins: SAR by NMR. *Science.* 274:1531-1534.
75. Spurlino JC (2011) Fragment screening purely with protein crystallography. *Methods in Enzymology, Ch.13.* 493:321-356.
76. Nienaber VL, Richardson PL, Klighofer V, Bouska JJ, Giranda VL, Greer J (2000) Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat Biotechnol.* 18:1105-1108.
77. US Food and Drug Administration. FDA labeling information — Zelboraf. FDA website [online], http://www.accessdata.fda.gov/drugsatfda_docs/label/2011/202429s000lbl.pdf (2011).
78. Tsai J, Lee JT, Wang W, Zhang J, Cho H, Mamo S, Bremer R, Gillette S, Kong J, Haass NK, Sproesser K, Li L, Smalley KS, Fong D, Zhu YL, Marimuthu A, Nguyen H, Lam B, Liu J, Cheung I, Rice J, Suzuki Y, Luu C, Settachatgul C, Shellooe R, Cantwell J, Kim SH, Schlessinger J, Zhang KY, West BL, Powell B, Habets G, Zhang C, Ibrahim PN, Hirth P, Artis DR, Herlyn M, Bollag G. (2008) Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proc Natl Acad Sci USA.* 105:3041-3046.
79. Hassell AM, An G *et al* (2007) Crystallization of protein–ligand complexes. *Acta Cryst.* D63:72-79.
80. Peselis A, Gao A, Serganov A (2016) Preparation and crystallization of riboswitches. *Methods Mol Biol.* 1320:21-36
81. Ke A, Doudna JA (2004) Crystallization of RNA and RNA–protein complexes. *Methods.* 34:408-414.
82. Ying BW, Fourmy D, Yoshizawa S (2007) Substitution of the use of radioactivity by fluorescence for biochemical studies of RNA. *RNA.* 13:2042-2050.
83. Madrigal-Carrillo EA, Díaz-Tufinio CA, Santamaría-Suárez HA, Arciniega M, Torres-Larios A (2019) A screening platform to monitor RNA processing and protein-RNA interactions in ribonuclease P uncovers a small molecule inhibitor. *Nucl Ac Res.* doi.org/10.1093/nar/gkz285.

84. Brännvall M, Kikovska E, Wu S, Kirsebom LA (2007) Evidence for induced fit in bacterial RNase P RNA-mediated cleavage. *J Mol Biol.* 372(5):1149-1164.
85. Mingeot-Leclercq MP, Glupczynski Y, Tulkens PM (1999) Aminoglycosides: activity and resistance. *Antimicrob Agents Chemother.* 43(4):727-737.
86. Vicens Q, Westhof E (2003) RNA as a drug target: the case of aminoglycosides. *ChemBioChem* 4:1018-1023.
87. Mikkelsen NE, Brännvall M, Virtanen A, Kirsebom LA (1999) Inhibition of RNase P RNA cleavage by aminoglycosides. *Proc Natl Acad Sci USA.* 96:6155-6160.
88. Tan R, Frankel AD (1995) Biochemistry structural variety of arginine-rich RNA-binding peptides. *Proc Natl Acad Sci USA.* 92:5282-5286.
89. Bayer TS, Booth LN, Knudsen SM, Ellington AD (2005) Arginine-rich motifs present multiple interfaces for specific binding by RNA. *RNA.* 11:1848-1857.
90. Eubank TD, Biswas R, Jovanovic M, Litovchick A, Lapidot A, Gopalan V (2002) Inhibition of bacterial RNase P by aminoglycoside-arginine-conjugates. *FEBS Lett.* 511:107-112.
91. Kawamoto SA, Sudhakar CG, Hatfield CL, Sun J, Behrman EJ, Gopalan V (2008) Studies on the mechanism of inhibition of bacterial ribonuclease P by aminoglycoside derivatives. *Nucl Ac Res.* 36:697-704.
92. Wu S, Mao G, Kirsebom LA (2016) Inhibition of bacterial RNase P RNA by phenothiazine derivatives. *Biomol.* 6:38.
93. Hori Y, Bichenkova EV, Wilton AN, El-Attug MN, Sadat-Ebrahimi S, Tanaka T, Kikuchi Y, Araki M, Sugiura Y, Douglas KT (2001) Synthetic inhibitors of the processing of pretransfer RNA by the ribonuclease P ribozyme: Enzyme inhibitors which act by binding to substrate. *Biochem.* 40:603-608.
94. Liu X, Chen Y, Fierke CA (2014) A real-time fluorescence polarization activity assay to screen for inhibitors of bacterial ribonuclease P. *Nucl Ac Res.* 42(20):e159.
95. Masquida B, Westhof E (2011) RNase P: at last, the key finds its lock. *RNA.* 17:1615-1618.
96. Paul R, Lazarev D, Altman S (2001) Characterization of RNase P from *Thermotoga maritima*. *Nucleic Acids Res.* 29:880-885.
97. Wegscheid B, Hartmann RK (2006) The precursor tRNA 3'-CCA interaction with *Escherichia coli* RNase P RNA is essential for catalysis by RNase P *in vivo*. *RNA.* 12:2135-2148.
98. Gößringer M, Kretschmer-Kazemi Far R, Hartmann RK (2006) Analysis of RNase P protein (rnpA) expression in *Bacillus subtilis* utilizing strains with suppressible rnpA expression. *J Bacteriol.* 188(19):6816-6823.
99. Kabsch, W. (2010) Xds. *Acta Crystallogr D Biol Crystallogr.* 66:125-132.
100. Evans, P.R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr.* 67:282-292.
101. Pearce NM, Krojer T, Bradley AR, Collins P, Nowak RP, Talon R, Marsden BD, Kelm S, Shi J, Deane CM, von Delft F. (2017) A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat Commun.* 8:15123.
102. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic program. *J Appl Crystallogr.* 40:658-674.
103. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* 67:235-242.
104. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, *et al.* (2010) PHENIX: a comprehensive python based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 66:213-221.

- 105.** Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 60:2126-2132.
- 106.** Joosten RP, Long F, Murshudov GN, Perrakis A (2014) The PDB_REDO server for macromolecular structure model optimization. *IUCrJ.* 1:213-220.
- 107.** Liebschner D, Afonine PV, Moriarty NW, Poon BK, Sobolev OV, Terwilliger TC, Adams PD (2017) Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr D Struct Biol.* 73:148-157.
- 108.** Wang Y, *et al* (2018) Small molecule inhibitors reveal allosteric regulation of USP14 via steric blockade. *Cell Res.* 28:1186-1194.
- 109.** De Vivo M, Masetti M, Bottegoni G, Cavalli A (2016) Role of molecular dynamics and related methods in drug discovery. *J Med Chem.* 59(9):4035-61.
- 110.** Munneke S, Prevost JRC, Painter GF, Stocker BL, Timmer MSM (2015) The rapid and facile synthesis of oxyamine linkers for the preparation of hydrolytically stable glycoconjugates. *Org Lett.* 173:624-627.
- 111.** Zoete V, Daina A, Bovigny C, Michielin O (2016) SwissSimilarity: A web tool for low to ultra high throughput ligand-based virtual screening. *J Chem Inf Model.* 56(8):1399-1404.
- 112.** Daina A, Michielin O, Zoete V (2017) SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep.* 7:42717. doi: 10.1038/srep42717.
- 113.** Batyuk A, Wu Y, Honegger A, Heberling MM, Plückthun A (2016) DARPIn-based crystallization chaperones exploit molecular geometry as a screening dimension in protein crystallography. *J Mol Biol.* 428(8):1574-1588.

12. Apéndices

Apéndice A. Materiales suplementarios

Cepas bacterianas. *Escherichia coli* BL21(DE3)plysS, *Escherichia coli* XL-GOLD, *Escherichia coli* BL21(DE3) y *Escherichia coli* BW (1). *Staphylococcus aureus* str. 29213 ATCC. *Bacillus subtilis* str. 6051 ATCC.

Tabla S1. Descripción de las cepas bacterianas usadas en este trabajo.

Cepa	Descripción	Referencia/Origen
<i>Escherichia coli</i> BL21(DE3)plysS	<i>E. coli</i> str. B F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB ⁺] _{K-12} (λ ^S) pLysS[T7p20 ori _{p15A}](Cm ^R)	Studier y col. (2)
<i>Escherichia coli</i> BL21(DE3)	<i>E. coli</i> str. B F ⁻ ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB ⁺] _{K-12} (λ ^S)	Studier y col. (2)
<i>Escherichia coli</i> XL-GOLD	endA1 glnV44 recA1 thi-1 gyrA96 relA1 lac Hte Δ(mcrA)183Δ(mcrCB-hsdSMR-mrr)173 tet ^R F[proAB lac ^R ZΔM15 Tn10(Tet ^R Amy Cm ^R)]	Stratagene
<i>Escherichia coli</i> BW	La expresión del gen <i>rnpB</i> (codifica la subunidad de ARN de la RNasa P bacteriana) en cepas <i>E. coli</i> BW depende estrictamente de la adición de arabinosa al medio. En ausencia de arabinosa, y en presencia de glucosa, la expresión cromosomal de <i>rnpB</i> en la cepa BW está reprimida y la supervivencia de las células depende completamente de la actividad de RNasa P codificada en plásmido.	Wegscheid y col. (3)
<i>Staphylococcus aureus</i> str. 29213	Las aplicaciones de esta cepa incluyen control de calidad, pruebas de susceptibilidad en disco, pruebas de susceptibilidad en general, descubrimiento de antibióticos, investigación entérica. Aislada de herida, nivel de bioseguridad 2.	ATCC
<i>Bacillus subtilis</i> str. 6051	Las aplicaciones de esta cepa incluyen aseguración de esterilidad, huésped para bacteriófagos, producción de isopreno, pruebas de resistencia bacteriana. Nivel de bioseguridad 1.	ATCC

Genomas. *Staphylococcus aureus* str. 29213 ATCC. *Bacillus subtilis* str. 6051 ATCC.

Vectores de expresión. Los plásmidos fueron visualizados con ayuda del software SnapGene viewer 3.2.1 <https://www.snapgene.com> y analizadas sus características generales, longitud en pares de bases y sitios de restricción con ayuda del repositorio global de plásmidos de addgene: <https://www.addgene.org/tools>. Algunos de los vectores fueron adquiridos de ATUM o DNA 2.0. La información completa de los plásmidos usados en este estudio puede ser consultada en la Tabla S2.

Oligonucleótidos. Todos los oligonucleótidos fueron sintetizados por alguna de las siguientes compañías: DNA 2.0, ATUM, IDT o Sigma-Aldrich. Fueron diseñados y analizados para determinar su T_m, secuencia, %GC, tendencia a dimerizar, tendencia a heterodimerizar, tendencia de formación de asas, tendencia a formar estructuras secundarias, peso molecular, con ayuda de alguna de las siguientes herramientas en línea, OligoEvaluatorTM de Sigma-Aldrich: <http://www.oligoevaluator.com/LoginServlet> u OligoAnalyzer Tool de IDT: <https://www.idtdna.com/pages/tools/oligoanalyzer>. Los oligonucleótidos usados en el presente estudio se describen en la Tabla S3.

Tabla S2. Descripción de las características de los vectores usados en este estudio.

Plásmido	Características relevantes	Aplicación e inserto
pET23-TEV-amp	amp ^R , promotor T7 inducible on IPTG, etiqueta de hexa-histidinas	Sobreexpresión de la proteasa TEV
pGEX-4ta-TmP	LacI, amp ^R , sitio de corte para proteasa trombina, proteína de fusión GST	Sobreexpresión de la proteína P de <i>T. maritima</i> tipo silvestre para purificación mediante cromatografía por intercambio catiónico
pD454-GST-TmP	Promotor T7, inducible con IPTG, RBS fuerte, amp ^R , pUC Ori, proteína de fusión GST	Sobreexpresión de la proteína P de <i>T. maritima</i> tipo silvestre
pUC19-MIRNAP	Inducible con IPTG, amp ^R	Producción de la subunidad de ARN P de <i>T. maritima</i>
pUC19-RNAPolT7	Inducible con IPTG, amp ^R	Producción de la ARN polimerasa T7 para transcripción de la subunidad de ARN de <i>T. maritima</i>
pMCSG10	amp ^R , sitio para LIC, sitio de corte para proteasa TEV, proteína de fusión GST, etiqueta de hexa-histidinas	Plásmido parental para llevar a cabo clonación independiente de ligación
pMCSG10-BsP	amp ^R , sitio para LIC, sitio de corte para proteasa TEV, proteína de fusión GST, etiqueta de hexa-histidinas	Vector de expresión para producción de la proteína P de <i>B. subtilis</i>
pMCSG10-SaP	amp ^R , sitio para LIC, sitio de corte para proteasa TEV, proteína de fusión GST, etiqueta de hexa-histidinas	Vector de expresión para producción de la proteína P de <i>S. aureus</i>
pJ204-TmR-Amp-IPTG	amp ^R , Inducible con IPTG	Producción de la subunidad de ARN de la RNasa P de <i>T. maritima</i>
pJ204-pretRNA-ChI-IPTG	Resistencia a cloranfenicol y amp ^R , inducible con IPTG	Producción del sustrato canónico de la RNasa P de <i>T. maritima</i> el precursor del ARN de transferencia

Tabla S3. Secuencia, descripción y blanco de los oligonucleótidos usados en este estudio.

Oligo	Secuencia (5'-3')	Blanco
T7 prom	TAATACGACTCACTATAGGG	Secuencia del promotor de la ARN polimerasa T7 del bacteriófago T7 para secuenciación del vector entero
T7 term	GCTAGTTATTGCTCAGCGG	Secuencia del terminador de la transcripción de la ARN polimerasa T7 del bacteriófago T7 para secuenciación del vector entero
pGEX-3'	CCGGGAGCTGCATGTGTCAGAGG	Oligo reverso del término 3' del sitio de clonación múltiple (MCS) del plásmido pGEX para secuenciación de todo el plásmido
pGEX-5'	GGGCTGGCAAGCCACGTTTGGTG	Oligo delantero del término 5' del sitio de clonación múltiple (MCS) del plásmido pGEX para secuenciación de todo el plásmido
SaP3	CTTACTTAATCTTTTTATTA AAAACTTT	Oligo reverso para amplificación de gen proteína P de <i>S. aureus</i> a partir de ADN genómico
SaP5	ATGTTATTGGAAAAAGCTTACCGA	Oligo delantero para amplificación de gen proteína P de <i>S. aureus</i> a partir de ADN genómico
BsP3	TGGTGATGCCTACTTTGACGAA	Oligo reverso para amplificación de gen proteína P de <i>B. subtilis</i> a partir de ADN genómico
BsP5	ATCGGAGTGAGTCATTTGAAGAA	Oligo delantero para amplificación de gen proteína P de <i>B. subtilis</i> a partir de ADN genómico
TmPTEV5	GATCTGGTTCGCGTGGATCCGAAAACCTGTACT TCCAGGGTACCGAGACTTCACCCGC	Oligo delantero para mutagénesis sitio-dirigida de inserto largo que contiene el sitio de corte para proteasa TEV y cambiar sitio de corte de trombina por TEV en el vector pGEX
TmPTEV3	GCGGGTGAAGCTCTCGGTACCCTGGAAGTACAGG TTTTCGGATCCACGCGGAACCAGATC	Oligo reverso para mutagénesis sitio-dirigida de inserto largo que contiene el sitio de corte para proteasa TEV y cambiar sitio de corte de trombina por TEV en el vector pGEX
Bsrnapa3	TTATCCACTTCCAATGACTCTTATATAATGAAGA CTTTCTGAATA	Para clonación independiente de ligación en plásmido pMCSG10
Bsrnapa5	TACTTCCAATCCAATGCTAGTCATTTGAAGAAGC GAAATCGTTTA	Para clonación independiente de ligación en plásmido pMCSG10
Sarnapa3	TTATCCACTTCCAATGACTATTA AAAACTTTGGC AATTTTAAGTAC	Para clonación independiente de ligación en plásmido pMCSG10
Sarnapa5	TACTTCCAATCCAATGCTATGTTATTGAAAAAG CTTACCGAA	Para clonación independiente de ligación en plásmido pMCSG10
TmPHis3	GACATGAAGGTCCCAGTCGTCGTCGTCGTCGCC CTAGGTGGCTCTCG	
TmPHis5	CTGTA CTCCAGGGTCATCATCATCATCATCATG GATCCACCGAGAGC	Inserción de etiqueta de hexa-histidinas en el sitio N-terminal de la proteína P tipo silvestre de <i>T. maritima</i> en el plásmido pGEX
TmPmutHis5	GATCTGGTTCGCGTGGATCCCATCATCATCAT CATCATACCGAGACTTCACCCGCCGC	Oligo delantero para mutagénesis sitio-dirigida para inserto largo inserción de etiqueta de hexa-histidinas a mutantes TmP en pGEX-4T
TmPmutHis	GCGGCGGGTGAAGCTCTCGGTATGATGATGATGA	Oligo reverso para mutagénesis sitio-dirigida para

3	TGATGGGATCCACGCGGAACCAGATC	inserto largo inserción de etiqueta de hexa-histidinas a mutantes TmP en pGEX-4T
SaPsc3	CGTTATCCACTTCCAATGATTTCAAAAACTTTG GCAATTTTAAG	Oligo reverso para mutagénesis sitio-dirigida para inserción de codón de paro en proteína P de <i>S. aureus</i>
SaPsc5	CTTAAAATTGCCAAAGTTTTTTGAAATCATTGGA AGTGGATAACG	Oligo delantero para mutagénesis sitio-dirigida para inserción de codón de paro en proteína P de <i>S. aureus</i>
BsPsc3	GTTATCCACTTCCAATGCTTATATAATGAAGAC	Oligo reverso para mutagénesis sitio-dirigida para inserción de codón de paro en proteína P de <i>B. subtilis</i>
BsPsc5	GTCTTCATTATATAAGCATTGGAAGTGGATAAC	Oligo delantero para mutagénesis sitio-dirigida para inserción de codón de paro en proteína P de <i>B. subtilis</i>

Apéndice B. Métodos suplementarios

Preparación de células competentes. Inicialmente, se inocularon 3 mL de medio LB sin antibiótico con la cepa *E. coli* BL21(DE3)pLysS y se incubaron durante 12 h toda la noche a 37°C en agitación a 250 rpm. Posteriormente, se inocularon 50 mL de medio LB sin antibiótico con 500 µL de cultivo de toda la noche y se creció hasta obtener una O.D.₆₀₀ = 0.4. Se leyó la O.D. en el espectrofotómetro portátil, primero medio LB sin inocular como blanco y después el cultivo a medir. Después, 20 mL del cultivo se transfirieron a tubos Falcon de cultivo por cada cepa y se incubaron en hielo durante 45 min al término de los cuales se procedió a centrifugar a 2.5 rpm durante 3 min a 4°C y se eliminó el sobrenadante. Se resuspendieron las pastillas celulares mediante agitación a 4°C en refrigerador y en agitación orbital en 10 mL de CaCl₂ 50 mM y se incubaron en hielo durante 1 h. Después se centrifugaron a 3 rpm durante 5 min y se removió el sobrenadante. Posteriormente, se resuspendieron en 2 mL de CaCl₂ 50 mM suplementado con 15% de glicerol y se permitió la resuspensión durante 1 h, 30 min. Después, se incubaron en hielo durante 1 h y alícuotas de 150 µL fueron añadidas a tubos para PCR e inmediatamente se congelaron en hielo seco-etanol y se almacenaron a -70°C en REVCO. Típicamente, se usan de 100-150 µL de cepa competente para el proceso de transformación genética por choque térmico. Para hacer gliceroles para almacenamiento de estas cepas, después de crecer a una O.D. de 0.6 en tubos eppendorf de 1.5 mL, se añadieron 700 µL de cepa bacteriana y 300 µL de glicerol estéril, y se congelaron a -72°C. Otra de las metodologías fue mediante una solución de 500 µL de células bacterianas y 500 µL de glicerol. Mezclar bien agitando por inversión y se almacenó a -72°C hasta su uso.

Transformación genética por choque térmico. 2-10 µL de ADN plasmídico se mezclaron con 100 µL de células competentes y se incubaron 30 min en hielo. La transformación por choque térmico en thermoblock se llevó a 42°C durante 2 min, e inmediatamente se incubó durante 5 min en hielo y se plaquearon 100 µL de cultivo en caja Petri. Una colonia de este cultivo fue usada para cultivos de toda la noche en líquido durante 16 h a 37°C. Posteriormente, las células competentes bacterianas fueron derretidas lentamente en hielo. Después, se mezclaron de 2-10 µL de reacción de ligación en 50-100 µL de amortiguador KCM (100 mM KCl, 30 mM CaCl₂, 50 mM MgCl₂, se prepararon a partir de stocks a 1 M, se esterilizó mediante filtración con poro de membrana de 0.22 µM, se alícuotó y se almacenó a -20°C hasta su uso). Gentilmente, se añadieron de 50-100 µL (el mismo volumen usado de KCM anteriormente) de células competentes y se incubaron en hielo durante 30 min. Finalmente, se llevó a cabo un choque térmico de las células a 42°C durante exactamente 2 min e inmediatamente se regresaron a incubar en hielo por 2 min. Después, se añadieron 400 µL de medio SOC a las células y se recuperaron durante 45 min a 37°C en agitación orbital de 250 rpm. Después de este tiempo, 100 µL de células fueron plaqueadas en cajas Petri con 25 mL de medio LB adicionado con un antibiótico adecuado según el fondo genético de la cepa y el plásmido integrado a la célula.

Cálculo de la eficiencia de transformación genética. Para llevar a cabo el cálculo de la eficiencia de transformación genética, se utilizaron las ecuaciones descritas a continuación:

$$TE = (\text{No. Colonias observadas/ng ADN plaqueados}) (1000 \text{ ng}/\mu\text{g}) = \text{CFU}/\mu\text{g ADN} \quad (1)$$

$$\text{ng ADN} = (\text{Conc ADN} \times \mu\text{L usados de ADN} / \text{Vol total}) (\mu\text{L plaqueados}) \quad (2)$$

$$\text{ng} = (100 \text{ ng} \times 1 \mu\text{L} / 25 \mu\text{L}) / 1000 \mu\text{L} = 0.0025 \text{ ng}$$

CFU= 85

TE= $85/0.0025 \mu\text{g} = 34,000 \text{ CFU}/\mu\text{g} = 3.4 \times 10^4$ transformantes por 1 μg de ADN plasmídico usado

Donde:

TE: eficiencia de transformación

CFU: unidades formadoras de colonias

Ecuación (1) eficiencia de transformación

Ecuación (2) nanogramos de ADN plaqueados

Extracción de ADN plasmídico. Usando el kit miniprep QIAGEN, se centrifugó el cultivo de bacterias de toda la noche a 10,000 rpm durante 10 min. Después se resuspendió la pastilla celular de bacterias en 250 μL de amortiguador P1 y se transfirieron a un tubo eppendorf de 1.5 mL. Se agregaron 350 μL de amortiguador N3 y se mezcló inmediatamente mediante inversión. Después, se centrifugó la mezcla a 13,000 rpm durante 10 min y se transfirió el sobrenadante a la columna para inmediatamente centrifugar a 13,000 rpm durante 45 s, y se desechó el sobrenadante. Posteriormente se lavó la columna con 500 μL de amortiguador PB centrifugando durante 45 s a 13,000 rpm. Después se lavó la columna con 750 μL de amortiguador PE centrifugando a 13,000 rpm durante 45 s y se desechó sobrenadante y se centrifugó durante 1 min a 13,000 rpm para remover residuos. Se agregaron 40 μL de amortiguador EB a la membrana y se dejó reposar 5 min, se centrifugó a velocidad máxima en un tubo eppendorf nuevo durante 1 min. Se añadieron 50 μL -100 μL de agua estéril desionizada a la muestra y se incubó durante 30 min a temperatura ambiente para que se hidrate el ADN. Finalmente, se midió en el NanoDrop y se almacenó a -20°C hasta su uso.

Clonación independiente de ligación. Se llevó a cabo una inserción de genes de la proteína P de *S. aureus* y de *B. subtilis* mediante una técnica que no requiere de enzimas de restricción para su incorporación al vector conocida como Clonación Independiente de Ligación (Ligation Independent Cloning, LIC), para lo cual se diseñaron oligonucleótidos para LIC. Los oligonucleótidos para LIC diseñados para el gen de la proteína P de *S. aureus* fueron los siguientes:

Sarnapa5 TACTTCCAATCCAATGCTATGTTATTGGAAAAAGCTTACCGAA

Sarnapa3 TTATCCACTTCCAATG ACTATTA AAAACTTTGGCAATTTTAAGTAC

Los oligonucleótidos para LIC diseñados para el gen de la proteína P de *B. subtilis* fueron los siguientes:

Bsrnapa5 TACTTCCAATCCAATGCTAGTCATTTGAAGAAGCGAAATCGTTTA

Bsrnapa3 TTATCCACTTCCAATGACTCTTATATAATGAAGACTTTCTGAATA

PCR y purificación del producto de PCR. El proceso para reacción en cadena de la polimerasa (PCR), se llevó a cabo en tubos para PCR en donde se llevó a cabo la mezcla de reacción como sigue:

- 10X amortiguador de reacción 5 μL
- Oligonucleótidos 1 μM

- Templado de ADN 1 ng
- ADN polimerasa Taq 1 μL (1.25 U)
- MgSO_4 50 mM 1.5 μL

Las cantidades y concentraciones de esta mezcla de reacción fueron calculadas para un volumen total de mezcla de reacción de 48.5 μL . Se llevó a cabo el proceso de PCR en un equipo termociclador (Applied Biosystems) a temperaturas apropiadas para los blancos específicos a amplificar. Las reacciones de amplificación se llevaron a cabo con una desnaturalización inicial de 94°C durante 3 min, y la amplificación fue de 35 ciclos, donde cada ciclo fue de 30 s a 94°C, 45 s a 55°C, y 1 min a 68°C, con una extensión final de 10 min a 68°C.

Para purificar el producto de PCR, se aplicó el producto de PCR a columnas para micro-centrífuga Qiagen, utilizando el kit para purificación de productos de PCR QIAquick. Se siguió el protocolo determinado por el fabricante y se determinó la concentración final del producto con ayuda del espectrofotómetro ND-1000 (NanoDrop Technologies).

Digestión con SspI. Con el fin de linearizar el vector pMCSG10, se llevó a cabo un tratamiento con la enzima de restricción SspI. Para ello, se preparó la mezcla de reacción correspondiente para cada inserto con amortiguador para enzima de restricción SspI 10X (Neb), 500 ng del vector pMCSG10, y enzima SspI. Posteriormente, esta mezcla de reacción se incubó durante 1.5-2 h a 37°C.

Tratamiento del producto de PCR con ADN polimerasa T4. Para generar los términos cohesivos 5' y 3', se llevó a cabo un tratamiento con la enzima ADN polimerasa T4, por lo que se preparó una mezcla de reacción como sigue:

- Producto de PCR 500 ng
- ADN T4 polimerasa 10X amortiguador de reacción 4 μL
- dCTP (100 mM) 1 μL
- DTT (100 mM) 2 μL
- Agua estéril
- ADN T4 DNA polimerasa 1 μL (se añade al final)

Estas cantidades y concentraciones se calcularon para un volumen total de reacción de 40 μL . Una vez preparada la mezcla de reacción, ésta se incubó durante 30 min a temperatura ambiente y después la actividad de la enzima ADN polimerasa T4 fue inactivada calentando a 75°C durante 20 min.

Tratamiento del vector para LIC con T4 DNA polimerasa. Para inducir la formación de los términos complementarios donde se apareará la secuencia del inserto, se preparó la siguiente mezcla de reacción como se describe:

- Vector DNA pMCSG10 200 ng
- ADN T4 polimerasa 10X amortiguador de reacción 4 μL
- dGTP (100 mM) 1 μL
- DTT (100 mM) 2 μL
- Agua estéril

- ADN polimerasa T4 1 μ L

Estas cantidades y concentraciones se calcularon para un volumen total de reacción de 40 μ L. Una vez preparada la mezcla de reacción, ésta se incubó durante 30 min a temperatura ambiente y después la actividad de la enzima ADN polimerasa T4 fue inactivada calentando a 75°C durante 20 min.

Ensamblaje mediante clonación independiente de ligación. Con el fin de integrar el inserto al vector, se llevó a cabo la mezcla de vector de ADN pMCSG10 así como del producto de PCR de los genes como sigue:

- Vector ADN pMCSG10 3 μ L (~45 ng)
- Producto de PCR *B. subtilis* o *S. aureus* 3 μ L

Estas dos alícuotas se colocaron juntas en una pequeña gota en el fondo de un tubo para PCR con un volumen no menor a 8 μ L y posteriormente se incubó durante 30 min sobre hielo.

Con este protocolo de clonación independiente de clonación, se generaron los plásmidos pMCSG10-SaP y pMCSG10-BsP. Finalmente, con estos vectores se transformaron células *E. coli* XL-GOLD mediante choque térmico para que las enzimas de reparación de ADN de *E. coli* reparen o ligen los *nicks* del vector recién formado. Una vez crecidas, se extrajo el vector y se usó tanto para experimentos de mutagénesis sitio-dirigida o bien, se usaron para transformar células *E. coli* (DH3)pLysS para su subsecuente sobreexpresión.

Mutagénesis sitio-dirigida. El protocolo canónico de QuickChange – Site Directed Mutagenesis (Stratagene) fue modificado para adaptar insertos largos más allá de la inserción de un solo triplete de bases. Para lograr esto, se diseñaron dos oligonucleótidos mutagénicos. El procedimiento tiene dos pasos. En el primero de ellos, dos reacciones de extensión se llevan a cabo en dos tubos por separado; uno que contiene el oligonucleótido *forward* y en el otro el que contiene el oligonucleótido *reverse*. Posteriormente, las dos reacciones son mezcladas y se lleva a cabo el protocolo estándar de PCR. Para el éxito de la inserción de la mutación deseada, ciertas reglas especiales deben seguirse al pie de la letra. Las características de estos oligonucleótidos son:

- La temperatura de alineamiento debe de ser de $T_m = 61^\circ\text{C}$.
- Los oligonucleótidos deben terminar en G o C.
- La longitud de los oligonucleótidos debe ser de entre 25 y 45 pares de bases.
- El contenido de GC debe ser del 30% como mínimo.
- La mutación deseada debe localizarse exactamente a la mitad del oligonucleótido mutagénico diseñado.
- Ambos oligonucleótidos deben de tener la misma temperatura de alineamiento, así como la misma longitud de pares de bases.

1 μ L de la enzima ADN polimerasa Vent fue usada a una concentración de 1 U/ μ L y el ciclo de polimerización en el termociclador fue programado como sigue: la reacción de extensión fue iniciada precalentando a 95°C durante 30 s, posteriormente de 5-19 ciclos fueron llevados a cabo de 95°C

durante 30s, 55°C durante 1 min y 68°C durante 8 min (aunque como regla, se añade 1 min por cada kilobase que mida el plásmido usado). Después de la finalización de la reacción de extensión, 25 µL de cada una de las 2 reacciones de oligo en cada tubo más 1 µL de ADN polimerasa Vent a 1 U/µL se llevó a cabo una extensión final durante 10 min a 72°C y posteriormente se mantuvo indefinidamente a 4°C.

Digestión con enzima de restricción DpnI. Al finalizar los ciclos de amplificación, con el fin de eliminar el plásmido parental el cual se encuentra metilado, se llevó a cabo una destrucción de este plásmido mediada por la enzima de restricción DpnI usando 1 µL a 10 U/µL, mezclando varias veces con la muestra y se centrifugó durante 1 min a 13,000 rpm a T ambiente y después de incubó 1 h, 30 min a 37°C. Como control negativo no se agregó enzima DpnI en una cepa control transformada para comparar.

Transformación genética. Inmediatamente después, se transformaron células *E. coli* XL-GOLD con 25 µL del producto de la digestión con DpnI. Para ello, las células se incubaron durante 30 min en hielo, se indujo un choque térmico durante 4 s a 42°C para inmediatamente incubar en hielo durante 2 min, se agregaron 400 µL de medio SOC, se permitió recuperar a las células durante 1 h a 37°C y se plaquearon de 300-300 µL en caja Petri con medio LB y se incubó a 37°C durante 12-16 h. Una vez crecidas las colonias, se extrajo su ADN plasmídico para secuenciar y verificar la presencia de la mutación y posteriormente se transformó a células *E. coli* BL21(DE3)pLysS para su inducción y subsecuente purificación.

Inserción de una secuencia codificante para una etiqueta de hexa-histidinas mediante mutagénesis sitio-dirigida. Se diseñaron oligonucleótidos mutagénicos para llevar a cabo una mutagénesis sitio dirigida de insertos largos, técnica mediante la cual insertar una etiqueta de seis histidinas en el extremo N-terminal de la secuencia de la proteína P de *T. maritima* y poder conducir experimentos en el sistema Octet RED96 (forteBio) mediante interferometría de biocapas. Los oligonucleótidos fueron diseñados para aparearse al inicio de la secuencia de la proteína P en la región N-terminal, específicamente entre el sitio de corte de la proteasa TEV y la proteína P (TEV-HisTag-proteína P) codificados en el vector de expresión pGEX-4t1 con sitio de corte para trombina y TEV. Se llevó a cabo la mezcla de reacción de los oligonucleótidos mutagénicos usando como templado al vector de expresión pGEX. Se llevaron a cabo 19 ciclos del programa de amplificación:

95°C/1 min | 95°C/30 s – 55°C/1 min – 68°C/ 6 min | 72°C/10 min – 4°C/∞

Oligonucleótido mutagénico delantero:

5'-CTGTACTTCCAGGGTCATCATCATCATCATCATGGATCCACCGAGAGC-3'

T_m= 80.8°C

Oligonucleótido mutagénico reverso:

5'-GACATGAAGGTCCCAGTCGTCGTCGTCGTCGTCCTAGGTGGCTCTCG-3'

T_m= 80.8°C

Inserción de codón de paro en la proteína P de *S. aureus* y *B. subtilis*. Se diseñaron oligonucleótidos mutagénicos para llevar a cabo mutagénesis sitio-dirigida para insertar un codón de paro de la transcripción en la secuencia de los genes que codifican para la subunidad proteica de la RNasa P de *S.*

aureus y de *B. subtilis*, los cuales se encuentran en el vector de expresión pMCSG10-SaP y pMCSG10-BsP. Este procedimiento se llevó a cabo para evitar la transcripción de la secuencia de la etiqueta de histidinas y poder purificar estas proteínas a través de un protocolo de purificación mediante intercambio catiónico. Los oligonucleótidos diseñados para introducir el codón de paro en la proteína P de *S. aureus* fueron los siguientes:

Oligonucleótido delantero:

5'-CTTAAAATTGCCAAAGTTTTTTGAAATCATTGGAAGTGGATAACG-3'

T_m= 68°C

Oligonucleótido reverso:

5'-CGTTATCCACTTCCAATGATTTCAAAAAAGTTTGGCAATTTTAAG-3'

T_m=68°C

Inserción de sitio de corte específico para proteasa TEV mediante mutagénesis sitio-dirigida. Se diseñaron oligonucleótidos mutagénicos para insertar un sitio de reconocimiento para la proteasa TEV en vez de usar el sitio de reconocimiento para trombina. Se precalentó el termociclador a 95°C durante 1 min y se llevaron a cabo 13 ciclos el programa de amplificación:

95°C/1 min | 95°C/30 s – 55°C/1 min – 72°C/ 8 min | 72°C/10 min – 4°C/∞

Oligonucleótido delantero:

5'-GATCTGGTTCCGCGTGGATCCGAAAACCTGTACTTCCAGGGTACCGAGAGCTTCACCCGC-3'

T_m= 82.5°C

Oligonucleótido reverso:

5'-

GCGGGTGAAGCTCTCGGTACCCTGGAAGTACAGGTTTTTCGGATCCACGCGGAACCAGATC-3'

T_m= 82.5°C

Electroforesis en gel de agarosa. Se preparó el gel al 1% (P/V) con 1g de agarosa en polvo grado biología molecular en 100 mL de amortiguador para electroforesis TAE (Tris-acetato, EDTA). Se cargaron 10 µL de muestra en cada pozo adicionadas con amortiguador de carga, se utilizó el marcador molecular y se corrió el gel de 80-150 V de corriente durante aproximadamente 1 h- 1.5 h. El gel se tiñó con EtBr y se visualizó bajo luz UV.

Maxiprep para aislamiento, purificación de plásmido y recuperación del gen M1 ARN. Después de haber llevado a cabo la correspondiente transformación genética por choque térmico, se procedió a la recuperación de plásmido, centrifugando el cultivo a 5000 rpm durante 20 min a 4°C y después se desechó sobrenadante cuidadosamente, se disolvió por vortex y agitación la pastilla celular en 20 mL de solución 1 helada (glucosa 9 mg/mL, Tris 25 mM pH 8.0, EDTA pH 8.0 10 mM). Se mezclaron ambas pastillas celulares en un tubo. Se añadió en un tubo 20 mL, se disolvió y se transfirieron esos 20 mL al otro tubo. Solo 20 mL en total. Posteriormente, se añadieron 100 mg de lisozima en 40 mL de solución 1 y se mezcló gentilmente durante 30 min a T ambiente. Se añadieron 80 mL/L de solución 2 (NaOH 200 mM, SDS 2%) fresca, se mezcló por inversión 6 veces durante 20 s y se incubó en hielo durante 5 min. Se añadieron 40 mL/L de solución 3 (KOAc 3M, CH₃COOH 11.75%) y se mezcló gentilmente por

inversión, girando y agitando al mismo tiempo de manera horizontal toroidalmente 20 veces y se incubó en hielo durante 20 min. Posteriormente, se centrifugó durante 40 min a 8000 rpm a 4°C y se transfirió el sobrenadante a un tubo de centrifuga nuevo sin mezclar con la pastilla y se añadieron 0.6 volúmenes de isopropanol. Se mezcló una vez por inversión y se incubó en hielo durante 10 min. Al término de este tiempo, se centrifugó durante 20 min a 8000 rpm a 4°C, se desechó el sobrenadante y se permitió secar la pastilla. Se disolvió la pastilla con 10 mL de amortiguador frío TE y se resuspendió, transfiriendo este volumen a un tubo de propileno de 50 mL. Luego, se añadieron 10 mL de solución fenol:cloroformo 1:1 y se agitó vigorosamente para después centrifugar durante 15 min a 5000 rpm 4°C y se recuperó la fase superior en un tubo nuevo. Posteriormente, se dividió el volumen final en dos tubos de 50 mL y se añadió 1 mL por tubo de NaOAc 3 M pH7 y 2.5 volúmenes de EtOH absoluto y se incubaron los tubos a -20°C durante 40 min. Después se centrifugó a 5000 rpm durante 25 min a 4°C, se desechó el sobrenadante, se removió todo el líquido restante y se secaron las pastillas. Por último, se disolvió cada pastilla de muestra en 2.64 mL de amortiguador TE (Tris-HCl 10 mM pH 8.0, EDTA 1 mM pH 8.0) frío y se mezclaron ambos volúmenes en uno.

Gradiente de Cloruro de Cesio. Para llevar a cabo el gradiente de cloruro de cesio, se pesaron 4.37 g CsCl por tubo y se agregó ADN disuelto hasta que pesara 3.696 g y se disolvió a temperatura ambiente. Después se añadió amortiguador TE hasta un peso de 3.696 g y se mezcló vigorosamente con vortex hasta asegurar que toda la sal estuviera perfectamente disuelta. Se añadieron 424 µL de bromuro de etidio a 10 mg/mL y se mezcló inmediatamente. Posteriormente, se transfirió la solución a dos tubos de plástico QuickSeal Polyallomer (13.5 mL, 16 x 76 mm, Beckman-Coulter). Se sellaron con las tapas de presión y se sometieron a ultracentrifugación usando el rotor NVT-90K en una ultracentrífuga tipo 90 Ti, de ángulo fijo (titanio, 8 x 13.5 mL, 90000 rpm, Beckman-Coulter) y se selló cada cápsula aplicando 100 pulgadas/libra de torque con ayuda de un dinamómetro. Los parámetros de ultracentrifugación fueron 90000 rpm a 22°C durante 19-24 h.

Extracción de ADN de tubos. Para llevar a cabo la extracción de tubos de la ultracentrífuga, se retiró el rotor de la ultracentrífuga con extremo cuidado y se abrió cada cápsula. Los tubos se mantuvieron en una caja de Petri con pinzas y se localizaron las bandas rojas en los tubos. Éstas se pincharon y fueron aspiradas con ayuda de una jeringa y se recuperaron en tubos eppendorf de 2.0 mL.

Remoción de EtBr. Para llevar a cabo la remoción del bromuro de etidio de las muestras, se añadió 1 volumen de agua saturada con n-butanol y se mezcló vigorosamente. Después la muestra se centrifugó a 5000 rpm por 5 min a temperatura ambiente y se desechó la fase rojiza al contenedor de bromuro de etidio y se repitieron los pasos 1 y 2 hasta que no quedó color rojizo observable.

Lavado. Para llevar a cabo el lavado de la muestra, se transfirió la solución de ADN a un tubo falcon nuevo de 15 mL y se diluyó con el amortiguador frío TE para triplicar el volumen original (1 a 3 mL). Posteriormente, se añadieron 3 volúmenes de EtOH absoluto frío y se mezcló gentilmente por inversión. Se incubó por 40 min a -20°C (3 mL para 12 mL). El EtOH debe estar almacenado a -20°C. Se centrifugó la muestra durante 20 min a 5000 rpm y se desechó el sobrenadante. Se centrifugó una vez más y se desechó todo el líquido remanente. Después se resuspendió la pastilla en 300 µL de amortiguador TE frío y se dividió el volumen en dos tubos eppendorf de 1.5 mL. Se añadió 30 µL de NaOAc 3 M pH 7 a cada uno. Se mezcló gentilmente y se añadieron 550 µL de EtOH frío. Se mezcló por inversión y se mantuvo a -20°C durante 10 min. Se centrifugó durante 10 min a 14000 rpm y se

desechó el sobrenadante con pipeta (vortexear hasta que el pellet se desprenda). Posteriormente, se añadió 1 mL de EtOH 80%, se resuspendió la pastilla de la muestra y se centrifugó durante 5 min a 14000 rpm. Se desechó el sobrenadante y se repitió el paso 8 una vez más, removiendo todo líquido remanente. Se secó la pastilla de la muestra y se resuspendió en 100 μ L de amortiguador frío TE. Finalmente, se cuantificó la muestra en un Nanodrop y posteriormente se continuó con el proceso de transcripción *in vitro*.

Transcripción *in vitro* del gen M1 ARN de *Thermotoga maritima*.

Digestión de ADN plasmídico. Se utilizó el ADN plasmídico pUC19 que porta el gen M1 ARN que codifica para la subunidad de ARN de *T. maritima* y porta sitios de restricción para la enzima de restricción FokI. Para llevar a cabo la digestión de 1-1.2 mg de ADN plasmídico, se necesita un volumen final de reacción de 3.5 mL. Se añadieron los reactivos siguientes en el orden especificado:

- a. H₂O DEPC a 3.5 mL
- b. NEBuffer4 10X 350 μ L
- c. ADN 10 μ g/ μ L 135 μ L (o volumen necesario para 1.2 mg)
- d. FokI 4U/ μ L 300 μ L (o volumen suficiente para 1,200 U (1U/ μ g ADN))

Estos reactivos se mezclaron y se incubaron a 37°C durante 1 h. Se añadieron 60 μ L de EDTA 0.5 M, 600 μ L de NaOAc 3 M y se mezcló. Después se añadieron 3.5 mL de fenol:cloroformo y se agitó vigorosamente. Se centrifugó durante 10 min a 5000 rpm y se recuperó la fase acuosa. Se repitió 3 veces más el paso de la extracción fenol:cloroformo. Luego, se añadieron 3 volúmenes de EtOH y se incubó durante 1 h a -20°C; al término de este tiempo la muestra se centrifugó durante 30 min a 5000 rpm y se desechó el sobrenadante. Se añadieron 3 mL EtOH 80% y se vortexeó para resuspender la pastilla de muestra. Por último, se centrifugó durante 5 min y se desechó el sobrenadante. Se disolvió la pastilla de muestra con cuidado en 150 μ L de agua DEPC. Este es el ADN que se utilizó para llevar a cabo la transcripción *in vitro* del gen M1 ARN.

*Transcripción *in vitro*.* Para llevar a cabo la transcripción *in vitro*, se selló el tubo con una película plástica y se preparó el amortiguador de transcripción T7 5X (H₂O DEPC 10 mL, amortiguador de transcripción 3 X 3.33 mL, NTP 5mM, 1 mL, RNase inhibitor 40,000 U/mL, 50 μ L, ADN pUC19 1.5 mg 100 μ L, PPase 1 U/mL 60 μ L y enzima RNAPol T7 25 μ g/ μ L 100 μ L). Primeramente, se selló el tubo con película plástica y se incubó durante 6 h a 37°C. Posteriormente, se añadieron 3.5 volúmenes de EtOH absoluto, se mezcló gentilmente por inversión y se incubó toda la noche a -70°C para precipitar el ADN. Al término de este periodo, la muestra se descongeló y se centrifugó durante 40 min a 5000 rpm y 4°C. Después se removió y se desechó el sobrenadante. Se disolvió la pastilla blanca en 2-2.5 mL con amortiguador de carga para gel.

Purificación y recuperación del ARN P.

Purificación de ARN en gel de poliacrilamida. Para llevar a cabo la purificación de ARN se procedió a ensamblar la cámara grande de electroforesis (Kit ajustable de secuenciación de ácidos nucleicos, 33 cm ancho, modelo SG-400-33, CBS Scientific Co., 3-4 mm de espesor, 30X30 cm) y se pegaron los dos

cristales con cinta adhesiva. Se aseguró que el fondo y los límites laterales estuvieran sellados y sin fugas, vertiendo un poco de agua en medio. Se vertió gel de acrilamida al 6% y se colocó en plástico teflón en la parte superior para hacer un solo pozo grande. Después de polimerizar se llenaron las cámaras con amortiguador de corrida TBE 1X. Previamente al inicio de la corrida, se procedió a limpiar el pozo con jeringa y se precalentó el gel con 400 V (15 W) durante 10 min. Se cargó el gel con la muestra preparada durante la última sección (Transcripción *in vitro*) y las condiciones de corrida fueron 350-400 V (15 W) durante 1 - 1.5 h. Al término de la corrida, se desensambló la cámara electroforética y se transfirió el gel cuidadosamente sobre una película plástica transparente; después se visualizó la banda principal clara bajo luz UV sobre el *fosfoimager* y se señaló el área a recuperar sobre la película plástica.

Recuperación de ARN. Para llevar a cabo la recuperación del ARN, se cortó la banda en condiciones libres de RNasas y se transfirió a un tubo falcon de 50 mL. Inmediatamente, el gel se maceró y se mezcló con 20 mL de amortiguador de elución para ARN y se incubó en agitador orbital a 4°C toda la noche. Posteriormente, se filtró la masa con filtro Millipore Steriflip Filter Unit de 50 mL y un dispositivo de vacío. Se dividió el volumen de elución en dos tubos de 50 mL y se añadieron 3.5 volúmenes de EtOH absoluto frío a cada uno. Se mezcló gentilmente por inversión y se mantuvo a -70°C durante 1 h. Al término de este tiempo, se descongelaron los tubos en hielo y fueron centrifugados durante 30 min a 5000 rpm y 4°C, después de lo cual se removió y se desechó el sobrenadante. Luego se añadieron 20 mL de EtOH 80% y se resuspendió la pastilla. Se centrifugó durante 20 min a 5000 rpm y 4°C; se repitió este proceso una vez más. Se retiró la pastilla de muestra y se disolvió en 100-150 µL de agua DEPC. Después se transfirió la solución a tubos de 1.5 mL y se congelaron durante 10 min a -70°C. Se descongeló y se centrifugó 10 min a 12000 rpm y 4°C, para después transferir la solución a un tubo de 200 µL. Finalmente, se cuantificó en Nanodrop y se calcularon la concentración de ARN, registrando además el radio 260/280.

Visualización de ARN en gel de poliacrilamida. Con el fin de visualizar la muestra de ARN, se preparó un PAGE desnaturante al 6, 8 ó 12% (aproximadamente 40 mL) y se ensambló la cámara de electroforesis mediana Sturdier (15x15 cm, Vertical Slab gel unit, The Sturdier, Model SE400, Hoefer Scientific Instruments). Se vertió el gel y, después de polimerizar, se llenaron las cámaras con amortiguador de corrida TBE 1X. Se limpiaron los pocillos con jeringa y precalentar el gel a 400 V (15 W) durante 10 min. Se cargó el gel y se corrió a 350-400 V (15 W) por 1.5 h – 2 h. Se desensambló la cámara electroforética y transfirió el gel cuidadosamente a una película plástica y se visualizaron las bandas en gel bajo luz UV sobre el *fosfoimager*.

Sobreexpresión y purificación de la ARN polimerasa T7. La purificación de la ARN polimerasa T7 fue necesaria con el fin de llevar a cabo subsecuentes protocolos para obtener la subunidad de ARN P mediante transcripción *in vitro* y se logró con dos pasos de purificación: primeramente, se trató la muestra con una cromatografía por afinidad en columna de Ni-NTA y posteriormente se sometió a otra cromatografía por afinidad en columna de Heparina. La sobreexpresión de la ARN polimerasa T7 fue inducida con 1 mM IPTG y resistencia a ampicilina. La pastilla celular se colectó y se lisaron mediante sonicación en el amortiguador de lisis (50 mM Tris pH 8.0, 100 mM NaCl, 5% glicerol y 1 mM imidazol). Se añadió lisozima a una concentración final de 100 µg/mL, detergente Brij-58 al 10% y se mezcló gentilmente, se añadió PMSF 1 mM, pepstatina A y leupeptina a una concentración final de 1

$\mu\text{g/mL}$. Se mezcló de nuevo gentilmente durante 30 min a 4°C . Posteriormente, las células se lisaron mediante sonicación 5 s ON, 5 s OFF hasta que la suspensión celular se observó clara. Se centrifugó a 13,000 rpm durante 30 min y el sobrenadante se cargó a una columna HisTrap FF (GE Healthcare) equilibrando previamente la columna en el amortiguador de lisis. Después se lavó la columna con 5 volúmenes de columna con amortiguador de lisis y después se lavó la columna con 5 volúmenes de amortiguador de lavado (amortiguador de lisis + 10 mM imidazol) y por último se eluyó la muestra con 5 volúmenes de columna en amortiguador de elución (amortiguador de lisis + 100 mM imidazol).

Posteriormente, la muestra se preparó para el siguiente paso de purificación por columna de Heparina, llevando a cabo un cambio de amortiguador para remover el imidazol de la muestra y tener a la proteína totalmente en amortiguador de unión para la columna de heparina usando tubos de centrifuga Amicon con el apropiado límite de porosidad ($M_{\text{RNAPol T7}} = 99 \text{ kDa}$). La operación de la columna HiTrap Heparin HP (GE Healthcare), con dimensiones de columna de 5 x 1 mL, se llevó a cabo a una velocidad de flujo de 2 mL/min, equilibrando con el amortiguador de unión (fosfato de sodio 10 mM, pH 7.2, filtrado por membrana de $0.45 \mu\text{M}$ antes de uso) con 10 volúmenes de columna (10 mL). Se inyectó la muestra y después de llegar de nuevo a una señal basal de mAU, se llevó a cabo un lavado con 10 volúmenes de columna, posterior al cual la proteína se eluyó con 5 volúmenes de columna empleando varios amortiguadores de elución (10 mM fosfato de sodio pH 7.2), adicionados con diferentes concentraciones de NaCl (250 mM, 500 mM, 1 M y 2 M de NaCl), para determinar la concentración de NaCl necesaria para eluir la polimerasa, aplicando una estrategia de elución de gradiente escalonado de 250-2 M NaCl. La fracción eluida se concentró en tubos para centrifuga Amicon, se midió concentración en espectrofotómetro ND-1000 (Nanodrop Technologies) y después se cambió por el amortiguador de almacenamiento (50 mM fosfato de sodio, pH 7.2, 50% glicerol, 1 mM EDTA y 1 mM DTT). Finalmente, una muestra de la proteína concentrada pura se visualizó y analizó mediante un gel SDS-PAGE al 12%.

Sobreexpresión y purificación de la proteasa TEV con etiqueta de histidinas. Se transformó la cepa *E. coli* BL21(DE3)pLysS con resistencia a cloranfenicol con el vector de expresión pET23 que porta el gen codificante para la proteasa TEV y el gen de resistencia a ampicilina. Se crecieron dos preinóculos con 3 ml de LB más ampicilina y cloranfenicol de BL21(DE3)pLysS con el vector de TEV toda la noche a 37°C y 250 rpm. Se crecieron 4 L de medio LB adicionados con ampicilina y cloranfenicol, inoculando 500 mL de LB en matraces de 2 L más 0.5 mL del precultivo. Las células se incubaron a 37°C y 250 rpm hasta una $\text{O.D.}_{600} = 0.6$ y se indujo la sobreexpresión con IPTG a 0.1 g/L, momento en el cual se disminuyó la temperatura a 15°C y se incubó el cultivo durante 16 h a 250 rpm. Posteriormente se cosecharon las bacterias centrifugando a $6,000 \times g$ durante 15 min a 4°C , se desechó el sobrenadante y se congelaron las pastillas celulares a -20°C . Una vez congeladas, las pastillas celulares fueron resuspendidas celulares en 50 ml de amortiguador A (NaH_2PO_4 50 mM, NaCl 300 mM, imidazol 50 mM, pH 8.0) y se lisaron mediante sonicación con 8 repeticiones de 1 min cada una con incubaciones de 1 min en hielo por repetición. Se centrifugó el producto del proceso de sonicación a 12,000 rpm durante 30 min a 4°C . Previamente al inicio de la purificación, se cargaron las bombas A y B del equipo de FPLC con amortiguador A y B (NaH_2PO_4 50 mM, NaCl 300 mM, imidazol 500 mM, pH 8.0), respetivamente. Se equilibró la columna de afinidad por Níquel con amortiguador A con 5 volúmenes de columna a un flujo de 5 mL/min y una presión máxima de 0.5 MPa (volumen de columna=5 mL). La muestra estuvo incubada en hielo y se inyectó a la columna a través de bomba A. Después se lavó la

columna con 5 volúmenes de columna con amortiguador A. La elución de TEV se llevó a cabo con un gradiente de imidazol a 500 mM con amortiguador B. Se llevó a cabo un gradiente lineal con amortiguador B hasta 100% durante 25 min. La proteína eluyó a 30% de amortiguador B y se colectaron las fracciones con proteasa. Posteriormente, se llevó a cabo el intercambio del amortiguador de las fracciones utilizando una columna de exclusión de tamaño molecular (columna HiPrep Desalting 26/10, Sephadex G-25, GE Healthcare) con un volumen de columna de 30 mL, utilizando el amortiguador de desalado (Tris 50 mM, EDTA 0.5 mM, DTT 1 mM, pH 8.0). La presión máxima fue de 0.35 MPa y se cargaron 10-15 mL de muestra. El flujo de operación fue de 7 mL/min. Se colectaron las fracciones con proteasa TEV y se concentró mediante centrifugación con un filtro, con límite de porosidad de 10 kDa, de millipore. Finalmente, se midió la concentración de proteasa TEV mediante Nanodrop y la muestra se almacenó a -70°C en amortiguador de desalado adicionado con 50% de glicerol.

Clonación, sobreexpresión y purificación de la proteína P de *T. maritima* tipo silvestre mediante intercambio catiónico. El plásmido empleado para la sobreexpresión de la proteína P fue el vector de expresión pGEX-4Ta (Amersham) inducible con isopropil β -D tiogalactósido (IPTG); contiene el promotor de la ARN polimerasa T7 y el gen codificante para la proteína de fusión GST. El plásmido derivado, el cual se obtuvo de Krivenko y col. (114), porta el inserto del gen *rnpA* codificante para la subunidad proteica de la RNasa P bacteriana de *T. maritima*. Células de *E. coli* BL21(DE3) pLysS fueron transformadas con ADN plasmídico pGEX-4Ta (114) y crecidas a 37°C en medio LB adicionado con ampicilina a 100 μ g/ μ L y cloranfenicol a 34 μ g/ μ L. La expresión de la proteína fue inducida añadiendo 0.25 mM de IPTG en medio LB, cuando las células alcanzaron una O.D.₆₀₀ de 0.4, se cosecharon después de 6 h por centrifugación (30 min, a 6000 rpm, 4°C) y la pastilla celular fue almacenada a -20°C. Las células descongeladas provenientes de 3 L de cultivo, fueron resuspendidas en 40 mL de amortiguador de lisis (100 mM Tris-HCl, 4 mM EDTA, 10% glicerol a pH 7.5) y sonicadas con un equipo Branson Sonifier 250 (30 s encendido, 40 s apagado y control de salida al 70%), con el objeto de remover la proteína de fusión GST se añadieron 600 unidades NIH de trombina bovina al extracto y se dejó reposar la solución por 16 h a temperatura ambiente y los restos celulares fueron removidos por centrifugación (30 min, 13000 rpm, 4°C). El sobrenadante se recuperó y se trató con 5 M de urea; la solución resultante se sometió a una cromatografía de intercambio catiónico con una columna de matriz de sulfopropil Sefarosa, (SP SepharoseTM, GE Healthcare) de 13 cm de alto por 2 cm de diámetro interno. Se inyectaron 40 mL de muestra y el flujo de trabajo fue de 7 mL/min con amortiguador de unión 50 mM de Tris-HCl, pH 7.5, 0.2 mM EDTA y 5 M Urea; la proteína fue eluida con un gradiente lineal de 0.1 a 3 M de NaCl con 50 mM de Tris-HCl, pH 7.5, 0.2 mM EDTA y 5 M Urea. El gradiente se llevó a cabo en un volumen de 150 mL y la proteína P eluyó entre 1.3 y 1.8 M de NaCl. En seguida se llevó a cabo un cambio de amortiguador mediante una cromatografía de exclusión de tamaño molecular (columna HiPrep Desalting 26/10, Sephadex G-25, GE Healthcare). Las dimensiones de columna fueron de 10 cm de alto por 2.6 cm de diámetro interno, con un volumen máximo de inyección de 15 mL y un volumen de columna de 53 mL. Se inyectaron 10 mL dos veces y el flujo de operación fue de 10 mL/min usando como amortiguador: 10 mM Tris-HCl, pH 7.5, 0.2 mM EDTA y 1 M NaCl. Las fracciones obtenidas del paso de purificación anterior se juntaron y se hizo una dilución a una concentración final de 0.2 M NaCl. Esta dilución se destinó a una cromatografía de intercambio catiónico en una columna de matriz de metilsulfonato, de 10 cm de alto por 1.5 cm de diámetro interno empacada con resina SourceTM 15S (GE Healthcare). Se inyectaron 20 mL de muestra y el flujo de trabajo fue de 7 mL/min con el amortiguador de unión (50 mM de Tris-HCl, pH 7.5 y 0.2

mM EDTA), mientras que la muestra fue eluida en un gradiente lineal de 0.1 a 3 M NaCl con amortiguador de elución (50 mM de Tris-HCl, pH 7.5 y 0.2 mM EDTA). La longitud del gradiente se llevó a cabo en un volumen de 150 mL y la proteína P eluyó entre 1.0 y 1.2 M de NaCl. El tren de purificación fue llevado a cabo en un equipo ÄKTAprime plus (GE Healthcare). La proteína se concentró a 13.23 mg/mL usando un filtro de centrifugación Amicon Ultra 10K (Millipore), cuantificada por un espectrofotómetro NanoDrop 1000 UV-Vis (Thermo Scientific) y almacenada a 4°C hasta su uso.

Sobreexpresión y purificación de la proteína P de *T. maritima* con etiqueta de hexa-histidinas. El gen codificante para la proteína P de *T. maritima* (síntesis del gen llevada a cabo por la compañía DNA 2.0 ATUM), fue clonado en el vector de expresión pD454-TmP-GST como una proteína de fusión a Glutación S-transferasa (GST), con un sitio de corte para proteasa TEV y una etiqueta de seis histidinas en el extremo N-terminal. El vector de expresión fue transformado en células *E. coli* BL21(DE3)pLysS. Las células fueron usadas para inocular 3 L de medio LB e incubado con una agitación orbital a 250 rpm a 37°C hasta que alcanzó una O.D.₆₀₀=0.6, momento en el cual la expresión de proteína P fue inducida con 1 mM de IPTG durante 6 h. La pastilla celular fue colectada y sometida a lisis celular por sonicación en 40 mL de amortiguador de lisis (20 mM fosfato de sodio monobásico, pH 8.0, 500 mM NaCl, 20 mM imidazol, 10% glicerol y 1 mM TCEP). La proteasa TEV (700 µg) fue añadida directamente al lisado celular e incubado durante 16 h a 34°C. El lisado celular fue centrifugado a 13,000 g y urea sólida fue añadida al sobrenadante para alcanzar una concentración final de 5 M. Posteriormente, la muestra fue purificada mediante cromatografía de afinidad en una columna de Ni-NTA (HisTrap FF crude, GE Healthcare), equilibrada con amortiguador de unión (20 mM fosfato de sodio monobásico, 500 mM NaCl, 5 M urea, 30 mM imidazol, pH 8.0). Posteriormente, un lavado de 30 min con un flujo de 5 mL/min fue llevado a cabo después del cual la proteína P fue eluida con 20 mM fosfato de sodio monobásico, 500 mM NaCl, 5 M urea, 2 M imidazol, pH 8.0. La elución fue cargada a una columna de exclusión de tamaño molecular (HiPrep 26/10, GE Healthcare) y equilibrada con amortiguador de desalado (20 mM fosfato de sodio monobásico, pH 8.0 y 500 mM NaCl). La proteína eluida fue concentrada a 5 mg/mL y almacenada a 19°C hasta su uso.

Sobreexpresión y purificación de la proteína P con la proteína de fusión GST. Con el fin de llevar a cabo experimentos de cristalogénesis para inducir otro grupo espacial diferente se sobreexpresó la proteína P junto con la proteína de fusión GST propia del vector de expresión pD454-TmP-GST, por lo que se evitó el paso de digestión con la proteasa TEV para obtener intacta toda la fusión. Se llevó a cabo un protocolo de purificación que involucró primeramente inyectar 40 mL del lisado celular que se llevó a cabo en amortiguador de unión (PBS 1X, pH 7.3 y 1 mM DTT) a una columna GSTrap (GSTrap FF crude, GE Healthcare) de 5 mL de volumen, con presión máxima de 0.3 MPa aplicando un flujo de 5 mL/min, la cual se equilibró con 4 volúmenes de columna con amortiguador de unión previamente a la inyección de la muestra. Después de un lavado de 10 volúmenes de columna, se procedió a eluir la proteína de fusión aplicando un golpe del 100% de amortiguador B (50 mM Tris-HCl, 10 mM glutatión reducido, pH 8.0 y 1 mM DTT), usando aproximadamente 5 volúmenes de columna. Posteriormente, esta muestra eluida se inyectó en una columna para intercambio catiónico SP Sefarosa. La fusión se eluyó aplicando un gradiente lineal de 0 – 100% de amortiguador de elución (50 mM Tris-HCl pH 7.5, 0.2 mM EDTA y 3 M NaCl) durante 20 min con un flujo de 6 mL/min. Las fracciones se colectaron 10 µL de cada una fueron cargadas en un gel SDS-PAGE 15% para visualizar las muestras.

Sobreexpresión y purificación por intercambio catiónico de la proteína P de *B. subtilis*. Células *E. coli* BL21(DE3)pLysS fueron transformadas con el vector de expresión pWT1 que porta el gen codificante para la proteína P de *B. subtilis*. Con estas transformantes se prepararon 50 mL de un preinóculo que fue usado en la inoculación de 3 L de medio LB. Después, se indujo la expresión de la proteína P de *B. subtilis* cuando la O.D.₆₀₀=0.7-0.8 usando 0.5 mM de IPTG, a 37°C, 250 rpm durante 4-5 h. Posteriormente, la pastilla celular se sometió a lisis celular mediante sonicación en amortiguador de unión (50 mM NaOAc, 5 mM EDTA, 100 mM NaCl, 5 M urea, pH 6.5) adicionando con ¼ pastilla *complete protease inhibitors* (ROCHE), después de lo cual la muestra se centrifugó durante 20 min a 13,000 rpm. El sobrenadante de este lisado celular fue inyectado a una columna CM Sefarosa (carboximetil celulosa) para intercambio catiónico débil (ICD) aplicando un flujo de 6 mL/min y teniendo como límite una presión máxima de 0.5 MPa. Posteriormente, la proteína se eluyó aplicando un gradiente lineal de 0-1 M NaCl con amortiguador de elución (50 mM NaOAc, 5 mM EDTA, 5 M urea, 1 M NaCl, pH 6.5). Después se procedió a remover la urea de la muestra usando una columna de exclusión de tamaño molecular (HiPrep Desalting 26/10, GE), en amortiguador de desalado (50 mM NaOAc, 5 mM EDTA, 100 mM NaCl, pH 6.5) con un flujo de 10 mL/min. Finalmente, esta muestra sin urea se sometió a un intercambio catiónico re-naturalizante inyectando en una columna CM Sefarosa equilibrada con amortiguador de unión (50 mM NaOAc, 5 mM EDTA, 100 mM NaCl, pH 6.5) y se eluyó la proteína en amortiguador de elución (50 mM NaOAc, 5 mM EDTA, 100 mM NaCl, 1 M NaCl, pH 6.5) con un gradiente lineal de 0-1 M NaCl (0-100% amortiguador de elución) durante 25 min y un flujo de 6 mL/min, con límite de presión máxima de 0.5 MPa. La muestra eluida de esta manera, se concentró con ayuda de tubos Amicon con un límite de porosidad de 10 kDa en amortiguador 10 mM Tris-HCl, pH 8.0 y fue visualizada con gel SDS-PAGE al 15%. La proteína P de *B. subtilis* purificada mediante este protocolo se almacenó a -20°C con 50% de glicerol.

Sobreexpresión y purificación de la proteína P de *S. aureus*. Células *E. coli* BL21(DE3)pLysS fueron transformadas con el vector de expresión pMCSG10 el cual porta el gen *rnpA* de *S. aureus* con inserción de codón de paro. Estas células se usaron para preparar un precultivo de 50 mL el cual fue usado para la inoculación de 3 L de medio LB con ampicilina (100 mg/mL), y antibióticos adecuados, que se incubó durante 1 h 40 min, momento en el cual se indujo la sobreexpresión con 0.5 mM de IPTG durante 6 h a 37°C y agitación orbital de 250 rpm. Posteriormente, la pastilla celular fue colectada por centrifugación y se llevó a cabo la lisis celular mediante sonicación en hielo usando también ¼ pastilla *complete protease inhibitor* (ROCHE). La proteína de fusión GST se eliminó incubando con 700 µg de proteasa TEV durante 16 h a 34°C. Al término de este tiempo, el lisado celular se centrifugó durante 20 min a 13,000 rpm. El sobrenadante de la lisis celular se inyectó a una columna SP Sefarosa para intercambio catiónico fuerte que fue previamente equilibrada con amortiguador de unión (20 mM MES, pH 6.0). La muestra se eluyó con un gradiente lineal de 0-1 M MgCl₂ (0-100% de amortiguador de elución: 20 mM MES, pH 6.0 y 1 M MgCl₂) durante 20 min a un flujo de 6 mL/min. Posteriormente, se llevó a cabo el desalado de la muestra con una columna de exclusión de tamaño molecular (HiPrep Desalting 26/10, GE) en amortiguador de desalado (20 mM MES, pH 6.0) a un flujo de 9 mL/min. Esta muestra eluida se sometió a un intercambio catiónico fuerte usando una columna SP Sefarosa en amortiguador de unión (20 mM MES, pH 6.0) aplicando un gradiente lineal de 0 - 1 M MgCl₂ (0-100% de amortiguador de elución: 20 mM MES, pH 6.0) durante 25 min a un flujo de 6 mL/min. Después, la muestra se sometió a una filtración en gel con ayuda de una columna Superdex 75 (10 – 100 kDa de intervalo de resolución, un límite de presión máxima de columna de 1.8 MPa y flujo máximo de 1.5 mL/min) usando el

amortiguador 20 mM MES, 500 mM NaCl, pH 6.0 a un flujo de 0.8 mL/min y se equilibró durante 30 min (20 mL). Se colectaron fracciones de 1 mL. Finalmente, la muestra eluida se concentró en tubos de centrifuga Amicon con límite de porosidad de 10 kDa en amortiguador 500 mM NaCl pH 6.0, y se visualizó en gel SDS-PAGE al 15%. La proteína P de *S. aureus* purificada mediante este protocolo se almacenó a -20°C con 50% de glicerol.

Sobreexpresión y purificación de la proteína P de *B. subtilis* y *S. aureus* con etiqueta de histidinas.

Células *E. coli* (DH3)pLysS transformadas mediante choque térmico por separado con vectores de expresión pD861-GST con resistencia a kanamicina y que portan el gen de la proteína P de *S. aureus* y de *B. subtilis*, fueron utilizadas para inocular 3 L de medio LB con antibióticos adecuados, los cuales fueron incubados a 37°C y 250 rpm hasta alcanzar una O.D.₆₀₀=0.8. A partir de este momento, se indujo su expresión con 0.02% de ramnosa durante 6-7 h a 37°C en régimen de agitación orbital, después de lo cual la pastilla celular fue colectada mediante centrifugación. Las células se lisaron mediante sonicación en amortiguador de lisis (20 mM fosfato de sodio monobásico, 1 mM TCEP, 20 mM imidazol, pH 8.0.), con tiempos de sonicación de 20 s ON y 20 s OFF, durante 10 min, incubando en hielo y controlando la temperatura. Posteriormente, el lisado sin centrifugar, se dejó incubando durante 16 h con 500 µL de proteasa TEV a 0.9 mg/mL a 34°C sin agitación. Al término de este periodo, se centrifugó el lisado celular y el sobrenadante fue tratado con urea sólida para lograr una concentración final de 5 M urea. Una vez que la urea se diluyó completamente, 40 mL de lisado fue sometido a una cromatografía de afinidad con una columna HisTrap FF 5 x 1 mL (GE Healthcare), en amortiguador de unión (20 mM fosfato sodio monobásico pH 8.0, 5 M urea, 30 mM imidazol y 500 mM NaCl), a un flujo de 5 mL/min. Una vez que la muestra fluyó completamente por la columna, se llevó a cabo un lavado de 30 min después de lo cual se eluyó la muestra con amortiguador de elución (20 mM fosfato sodio monobásico pH 8.0, 5 M urea, 30 mM imidazol, 500 mM NaCl y 2 M imidazol). Posteriormente, se procedió a remover la urea con ayuda de una columna de exclusión de tamaño molecular HiPrep Desalting (GE Healthcare de 30 mL) en amortiguador de desalado (20 mM fosfato sodio monobásico pH 8.0) a un flujo de 6 mL/min. Las muestras se concentraron en un tubo para centrifuga Amicon con un límite de porosidad de 10 kDa donde las proteínas se cambiaron a un amortiguador de almacenamiento PBS 1X. Finalmente, las muestras concentradas se analizaron mediante gel SDS-PAGE al 18% y se almacenaron a 19°C.

Purificación de la proteína P mediante co-expresión con ARN P como chaperona.

3 L de medio de cultivo LB adicionado con antibióticos adecuados, fue inoculado con el precultivo de células *E. coli* BL21(DE3)pLysS transformadas con un vector de expresión que porta el gen codificante de la proteína P con etiqueta de histidinas y resistencia a kanamicina, así como con otro vector que porta el gen de la subunidad catalítica ARN P con resistencia a ampicilina. Para ello, se agregó la mitad de cada concentración standard para antibiótico. Estos cultivos fueron incubados a 37°C y 250 rpm de agitación orbital hasta alcanzar una O.D.₆₀₀=0.8 y se indujo primeramente la expresión de la subunidad ARN P con 1 mM de IPTG durante 1 h y, posteriormente, se indujo la expresión de la proteína P con ramnosa al 0.02% durante 6-7 h. Al término de este tiempo, se colectó la pastilla celular por centrifugación y se procedió a lisar las células mediante sonicación con 40-50 mL de amortiguador de lisis (20 mM Fosfato de sodio monobásico, 500 mM NaCl, 20 mM imidazol, 10% glicerol, 1 mM TCEP, pH 8.0). Después se agregaron 500-700 µL de proteasa TEV a 1 mg/mL y se mezcló bien el lisado, dejando incubar durante 16 h a 34°C. Finalizado este tiempo, se centrifugó el lisado y al sobrenadante se le añadió urea sólida

para alcanzar una concentración final de 5 M. Una vez que se diluyó la urea, la muestra se sometió a cromatografía por afinidad, para lo cual se cargó mediante bomba a una columna Ni-NTA de 5 mL que fue previamente equilibrada con amortiguador de unión (20 mM fosfato de sodio monobásico, 500 mM NaCl, 5 M urea, 30 mM imidazol, pH 8.0.). Una vez que la muestra pasó totalmente por la columna, se llevó a cabo un lavado de 30 min y se eluyó con 100% de amortiguador de elución (20 mM fosfato de sodio monobásico, 500 mM NaCl, 5 M urea, 2 M imidazol, pH 8.0). Después se eliminó la urea y se replegó la proteína en una columna de exclusión de tamaño molecular equilibrando previamente con amortiguador de desalado (20 mM fosfato de sodio monobásico, 500 mM NaCl, pH 8.0). Por último, la proteína P se concentró y se almacenó a 19°C.

Electroforesis en gel de poliacrilamida dodecil sulfato de sodio (SDS-PAGE). Al final del tren de purificación, la pureza de la proteína fue monitoreada mediante separación electroforética en geles desnaturalizantes de poliacrilamida (SDS-PAGE) al 18%. Se llevó a cabo en un sistema de electroforesis Mini-PROTEAN 3 Multi-Casting Chamber de Bio-Rad. El voltaje fue constante de 200 V y el tiempo de corrida fue de 1 h, 45 min. El amortiguador de corrida fue Tris-HCl 25 mM, Glicina 200 mM y SDS 0.1 %. El volumen de amortiguador de muestra fue de 2 μ L, El volumen que se cargó de muestra fue de 8 μ L, cargando 3, 6 y 12 μ g de proteína y se cargó un volumen de 10 μ L de marcador de peso molecular SeeBlue™ Pre-stained Standard (Life Technologies). El gel se reveló tiñendo con azul de coomassie en inmersión calentando en horno de microondas en 5 ciclos de periodos de 50 s y destiñendo en agua destilada toda la noche. Las condiciones de la corrida de electroforesis fueron 200 V durante 45 min.

Búsqueda de condiciones de cristalización para nuevo grupo espacial. Con el fin de inducir otros grupos espaciales y otro empaquetamiento diferente en el cual compuestos optimizados pudieran difundir a través de los espacios intersticiales del cristal, se llevó a cabo una búsqueda de condiciones e cristalización diferentes a la formulación estándar desarrollada para la proteína P de *T. maritima* en este trabajo. Para lograr esto, se llevó a cabo una búsqueda en colecciones de reactivos de cristalización de Hampton Research como: Silver bullets, Salt Rx 1, Salt Rx 2, Index Screen, Crystal screen 1, Crystal screen 2, PEG-pH screen, PEG-Ion screen, así como el Hampton additives. Asimismo, se llevaron a cabo experimentos de cristalización utilizando proteínas de fusión con la proteína P como la GST y la etiqueta de hexa-histidinas y también se empleó *macroseeding* con cristales de buena calidad previamente obtenidos. Esta búsqueda se llevó a cabo en placas de cristalización Swissci™ de 96 pocillos con un volumen de licor madre de 40 μ L y 3 gotas de 2 μ L de volumen total, donde 1 μ L correspondió a la proteína P a una concentración de 3 mg/mL y el restante a la condición probada, misma que se encontraba en el reservorio de 40 μ L. Las placas se incubaron para su monitoreo en tiempo indefinido o hasta observar cristales a 19°C.

Apéndice C. Resultados suplementarios

-- Principio del ensayo de medición de actividad en tiempo real por fluorescencia

Desarrollo de una nueva técnica para la medición de la actividad de la RNasa P bacteriana mediante espectroscopía de fluorescencia

Esta sección trata sobre el desarrollo de un novedoso ensayo en tiempo real mediante espectroscopía de fluorescencia que fue ideado y desarrollado por el M. en C. Carlos Díaz Tufinio y el Dr. Alfredo Torres Larios. El desarrollo de esta técnica forma parte de este mismo proyecto en colaboración y se describe a continuación dada su importancia en la interpretación de los resultados obtenidos en el presente trabajo.

Sustrato mínimo artificial fluorescente minihélice

Anteriormente, Branvall y col. (84), desarrollaron un pequeño sustrato artificial de ARN de cadena sencilla llamado pMini3pb, el cual fue identificado como el sustrato de menor tamaño que puede ser reconocido y procesado por la RNasa P. A este sustrato (llamado en este trabajo minihélice o Mh), se le aplicaron modificaciones químicas para añadir un fluoróforo y un apagador en sus extremos 5' y 3', respectivamente (Figura S1), para llevar a cabo ensayos de actividad por fluorescencia.

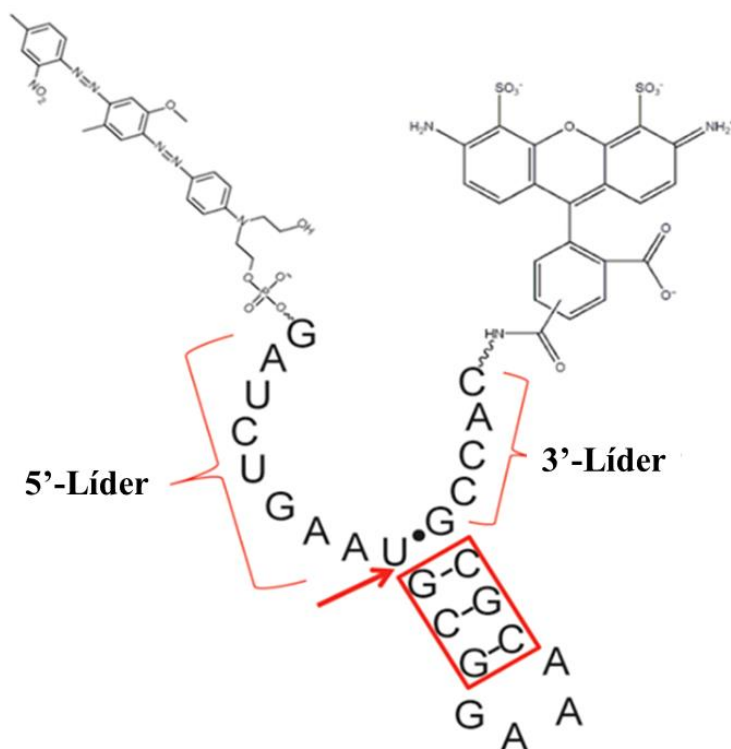


Figura S1. Estructura secundaria del sustrato artificial minihélice fluorescente. En corchetes se indican las ubicaciones del extremo 5' líder y del extremo 3' líder. Al término del extremo 5' líder se localiza el apagador Black Hole Quencher (BHQ), mientras que en el extremo 3' líder se localiza el fluoróforo Alexa Fluor 488. La distancia entre ambos extremos 5' y 3' es de 42 Å, distancia que se encuentra dentro de los límites para el efecto FRET de 100 Å. La flecha roja indica el sitio de corte canónico por la RNasa P. Señalado en el recuadro rojo, se localiza el tallo de tres pares de bases para el reconocimiento por la RNasa P. El tamaño del fluoróforo y el apagador no están a escala con respecto a las bases para mejor claridad.

La adición del fluoróforo Alexa Fluor 488 y el apagador BHQ no son perjudiciales para la actividad de la RNasa P

Se optimizó el ensayo de actividad por fluorescencia en gel en presencia y ausencia de fluoróforo en el sustrato Mh para corroborar que el fluoróforo no inhibiera la reacción (Figura S2). Con ambos sustratos, a medida que aumentó el tiempo de incubación de la reacción, el porcentaje de sustrato Mh disminuyó a medida que se incrementaba la cantidad de producto (Figura S2).

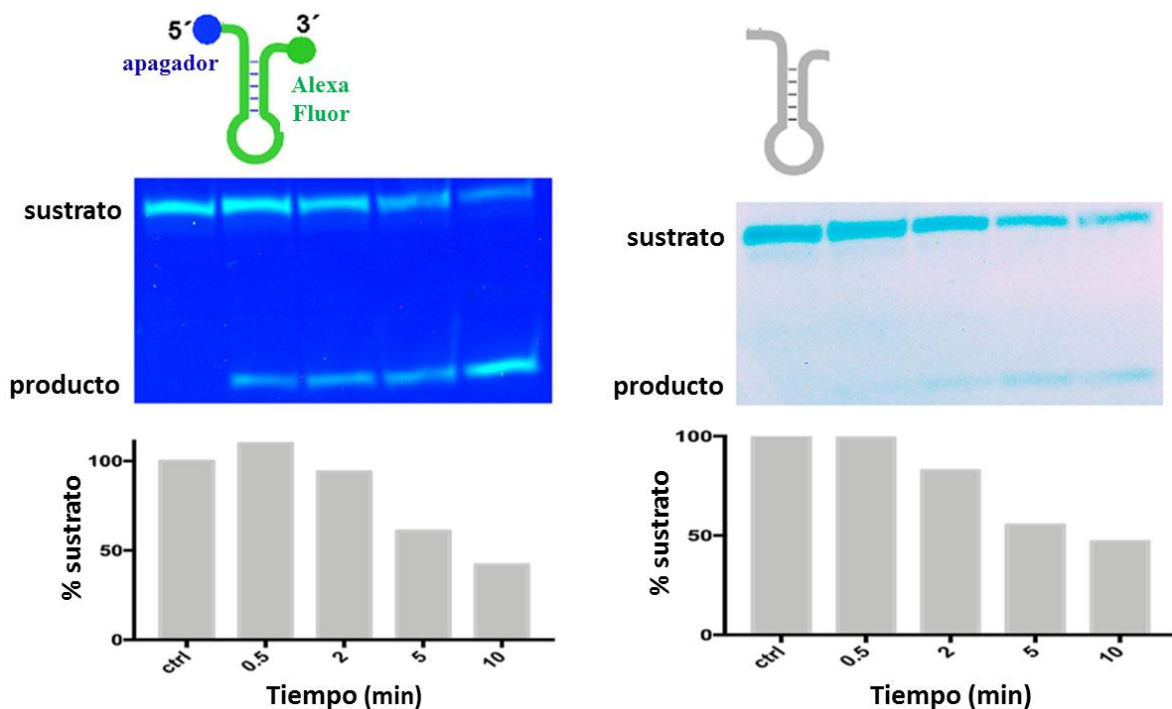


Figura S2. Ensayo de actividad de RNasa P en gel con sustrato Mh. Panel izquierdo: ensayo de actividad en gel usando el sustrato artificial Mh fluorescente. En el panel de en medio se muestra la fluorescencia del sustrato Mh escindido y el producto 5'-líder en gel urea-PAGE al 20% visualizado bajo luz UV. **Panel derecho:** ensayo de actividad en gel correspondiente al sustrato Mh no fluorescente. En el panel de en medio se muestra la fluorescencia del sustrato Mh escindido y el producto 5'-líder en gel PAGE al 20% con tinción de EtBr bajo UV. En la parte inferior se muestran las gráficas de densitometría de las bandas del gel del panel del medio en % de sustrato procesado vs tiempo. En la parte superior se muestra el esquema del sustrato Mh fluorescente a la izquierda y el sustrato Mh no fluorescente a la derecha. En la parte inferior se muestran las gráficas de barra que representan el análisis por densitometría de las bandas del gel del panel del medio como % de sustrato procesado vs tiempo.

Asimismo, ya que era necesario corroborar la detección de nuestros hits con la reacción en presencia del sustrato canónico de pre-tRNA, se empleó otro sustrato artificial similar a un pre-tRNA llamado pre-tRNA-mitades, que consta de dos oligonucleótidos, uno con un fluoróforo (5') y otro con un apagador (3') (Figura S3).

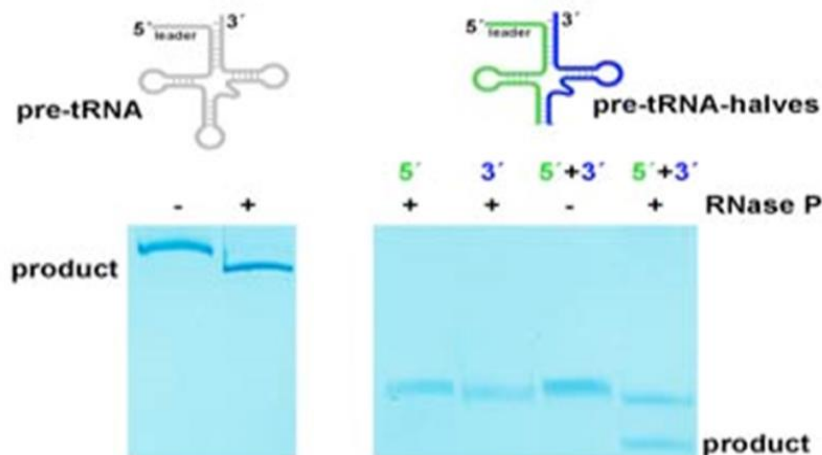


Figura S3. Ensayo de actividad de RNasa P en gel con sustrato pre-tRNA canónico y sustrato artificial pre-tRNA-mitades. **Izquierda.** En la parte superior se muestra el esquema del pre-tRNA canónico y en la parte inferior se muestra un gel urea-PAGE al 19% donde se observa actividad de RNasa P (signo – y +, ensayo sin y con RNasa P, respectivamente). **Derecha.** En la parte superior se muestra el esquema del sustrato artificial pre-tRNA-mitades que no cuenta con la secuencia del asa del anticodón y en la parte inferior se muestra un gel urea-PAGE al 19% donde se observa actividad de RNasa P cuando se combinan las dos mitades 5' y 3' para formar el sustrato artificial pre-tRNA-mitades en presencia de RNasa P (signo – y +, ensayo sin y con RNasa P, respectivamente).

Con el fin de determinar si este sustrato artificial con fluoróforo y apagador en sus extremos se unía a la RNasa P, se determinaron los parámetros cinéticos de unión del sustrato artificial y natural en las mismas condiciones por interferometría de biocapas (ver establecimiento del método más abajo). Como resultado se obtuvo que la afinidad calculada para ambas interacciones está en buena concordancia, ya que el valor de K_D de la interacción con su sustrato canónico pre-tRNA por la RNasa P fue de 1.8 nM, mientras que la K_D calculada para la interacción del sustrato artificial fue de 3.5 nM (Figura S4).

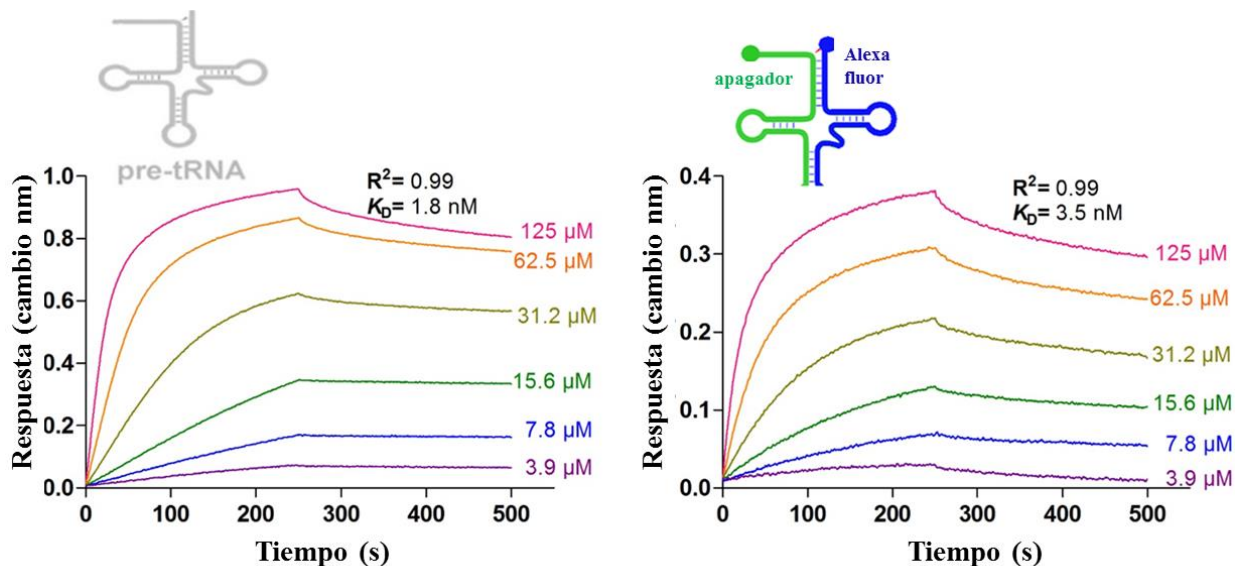


Figura S4. El pre-tRNA canónico y el pre-tRNA-mitades se unen a la RNasa P con afinidad similar. **Izquierda.** Titulación del pre-tRNA canónico contra la RNasa P con una $K_D = 1.8 \text{ nM}$. **Derecha.** Titulación del sustrato artificial pre-tRNA-mitades contra la RNasa P con una $K_D = 3.5 \text{ nM}$.

Empleando el espectrofotómetro lector de placas Synergy Mx Multimode Detector (BioTek) y el sustrato Mh, los parámetros cinéticos de la reacción pudieron ser determinados con buenos valores de reproducibilidad y confianza (Figura S5).

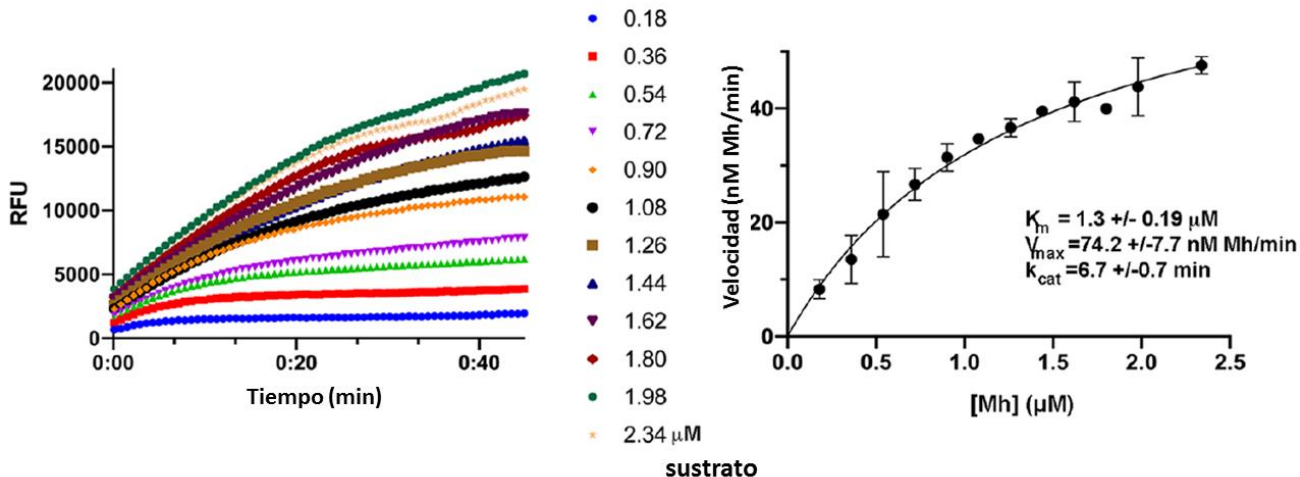


Figura S5. Caracterización de parámetros cinéticos de la RNasa P a partir del ensayo por espectroscopía de fluorescencia. Izquierda. Curvas de ensayos cinéticos de señal de fluorescencia vs tiempo, donde se monitoreó 11 nM de holoenzima RNasa P constante contra 0.18 μM – 3.24 μM de concentración de sustrato Mh fluorescente. **Derecha.** Modelo de Michaelis-Menten de la reacción de la RNasa P obtenido a partir de las velocidades iniciales de las gráficas del panel izquierdo contra concentraciones crecientes de sustrato Mh. En el lado extremo derecho se muestran los parámetros cinéticos de reacción. El monitoreo se llevó a cabo en un instrumento Synergy MX Multimode Detector (BioTek).

-- Principio y desarrollo del ensayo de medición de interacciones en tiempo real por interferometría de biocapas

Validación del uso de biosensores Ni-NTA para experimentos específicos con la proteína P mediante BLI

Se utilizó a la proteína P con etiqueta de histidinas explotando la tecnología de interferometría de biocapas ya que este tipo de construcción es muy práctica de emplear. Sin embargo, el uso de biosensores Ni-NTA no está recomendado para aplicaciones de detección de unión de molécula pequeña, por lo que era necesario comprobar el principio de prueba y validar esta metodología.

Como veremos más adelante, los parámetros cinéticos de unión determinados para los compuestos optimizados C-7 y C-10 pudieron ser confirmados mediante MST, así como también la afinidad del 5'-líder por la proteína P.

Se utilizó un instrumento para interferometría de biocapas (BLI) Octet RED96 de la compañía forteBio. La configuración del instrumento se compone por una estación para el brazo robótico y detector, una bandeja para biosensores y una bandeja para microplacas (Figura S6).

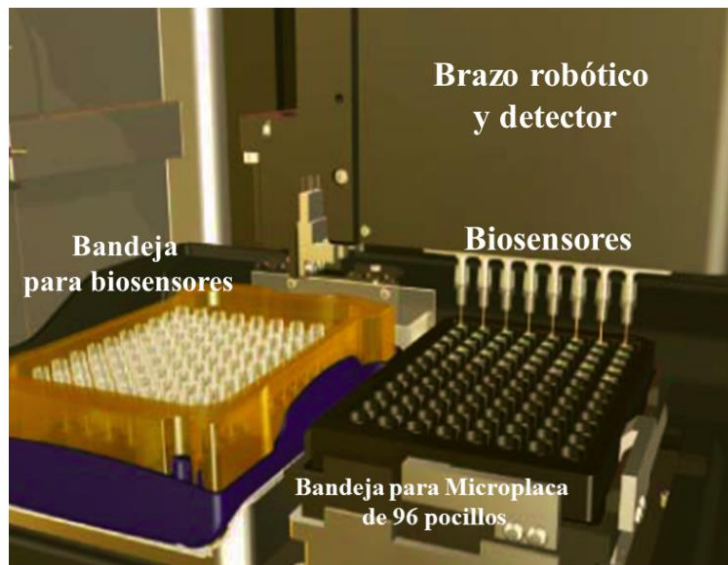


Figura S6. Configuración de la instrumentación del sistema OCTET RED96. La configuración de la instrumentación para interferometría de biocapas se compone por una bandeja para 96 biosensores (lado izquierdo), de una bandeja para microplaca de 96 pocillos (lado derecho) con capacidad para 200 μ L que además tiene la capacidad de generar agitación orbital de hasta 2000 rpm y de una placa térmica que calienta de 25°C-40°C las muestras. Un brazo robótico (parte superior derecha), el cual alberga el detector óptico y la lámpara, se encarga de transportar los biosensores desde la bandeja de biosensores hasta la bandeja para microplacas de 96 pocillos, donde hace un recorrido de una columna de pocillos a la siguiente de acuerdo a los requerimientos experimentales.

El experimento típico en este tipo de instrumento se puede dividir en cinco pasos muy generales, aunque en la práctica la técnica es muy flexible a necesidad del investigador. Primeramente, es necesario inmovilizar el blanco a un biosensor adecuado a la química de éste, posteriormente se lleva a cabo un lavado y una línea base antes de comenzar la asociación del ligando para finalmente llevar a cabo el paso de la disociación del mismo. La región de análisis a partir del cual se calculan los parámetros cinéticos, es la zona experimental que involucra solamente asociación y disociación (Figura S7).

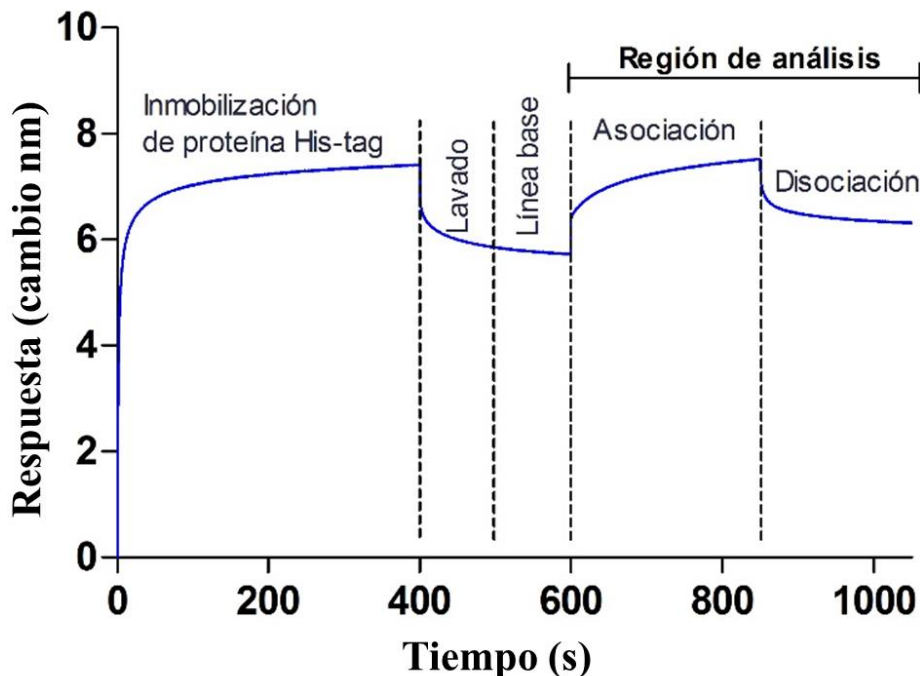


Figura S7. Experimento típico mediante interferometría de biocapas con la proteína P. Inicialmente, una proteína de interés con etiqueta de histidinas se inmoviliza en biosensores de Ni-NTA. Posteriormente, se hace un lavado para eliminar el exceso de proteínas que no lograron ser inmovilizadas al biosensor usando un amortiguador cuya composición debe ser constante a lo largo de todos los pocillos del ensayo. Inmediatamente después, se genera una línea base antes de seguir con la asociación del ligando y, posteriormente, su disociación en un pocillo con amortiguador. La región de análisis se usa para obtener los parámetros cinéticos de unión, manteniendo constante la concentración de proteína inmovilizada y variando la concentración de ligando.

Inmovilización de la proteína P con etiqueta de histidinas a biosensores Ni-NTA

Se llevaron a cabo cuatro repeticiones de este experimento para asegurar la reproducibilidad de inmovilización con pasos que involucraron: línea base-asociación-disociación-reasociación con TmP solamente y un lavado final con amortiguador de imidazol 250 mM para comprobar la disociación completa de la proteína P con etiqueta de histidinas de los biosensores Ni-NTA (Figura S8).

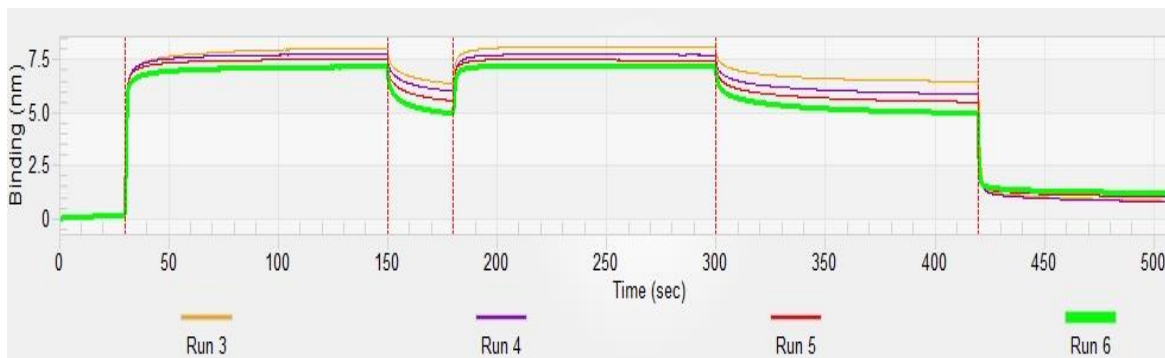


Figura S8. Comprobación de inmovilización de la proteína P con etiqueta de histidinas a biosensores Ni-NTA. Se llevaron a cabo cuatro repeticiones denotadas como Run 3 – Run 6. La concentración de la proteína P con etiqueta de histidinas fue constante a 0.05 mg/mL (3.7 μ M) en amortiguador CHAPS 0.1%, 170 mM NaCl adicionado con 10% DMSO. Se utilizó un biosensor para cada corrida.

Validación de biosensores Ni-NTA mediante BLI de la interacción de proteína P y RNasa P con sus sustratos canónicos

Experimentos que midieron los parámetros cinéticos de la unión del 5'-líder de ADN y de ARN, revelaron que el líder de ADN se une con mucha menor afinidad que el líder de ARN a la proteína P (Figura S9).

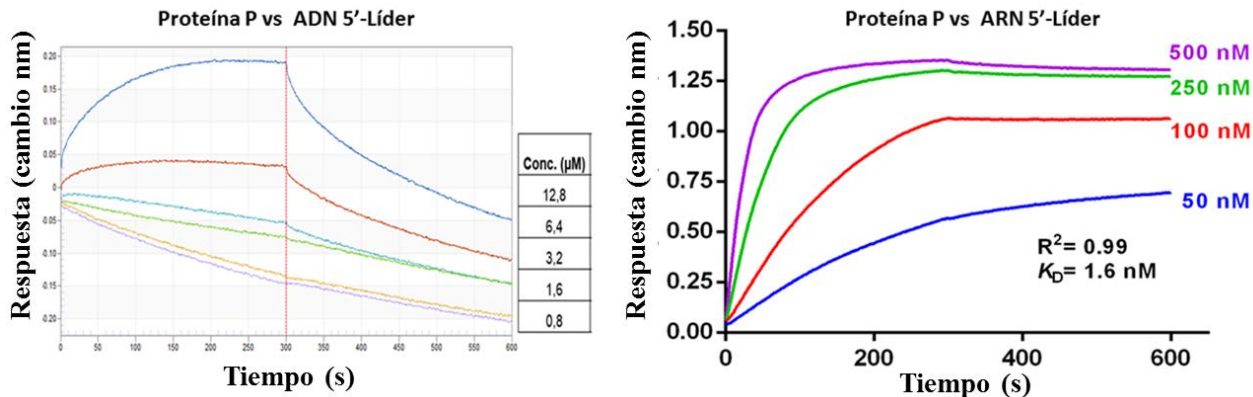


Figura S9. La afinidad del 5'-Líder de ADN por la proteína P es muy baja en comparación con la del 5'-líder de ARN. Izquierda. Curva de titulación del ADN 5'-líder en un rango de concentraciones de 0.8 μM – 12.8 μM . No pudo ser calculada la K_D para esta interacción. **Derecha.** Sensorgramas de las curvas de titulación del ARN 5'-Líder en un rango de concentraciones de 50 nM – 500 nM. Una K_D de 1.6 nM fue determinada para esta interacción. En ambos experimentos, el ensayo fue llevado a cabo en amortiguador PBS 1X, pH 7.4 adicionado con 5% DMSO.

Los parámetros cinéticos de unión obtenidos experimentalmente en nuestro laboratorio, estuvieron en buena concordancia con valores obtenidos en la literatura en cuanto al valor de K_D de la interacción RNasa P-proteína P (4), RNasa P-pre-tRNA (5) y a nuestro conocimiento, es la primera vez que se mide directamente la interacción entre una secuencia 5'-líder consenso con la proteína P en experimentos dosis-respuesta por BLI (Figura S10) y confirmado por MST.

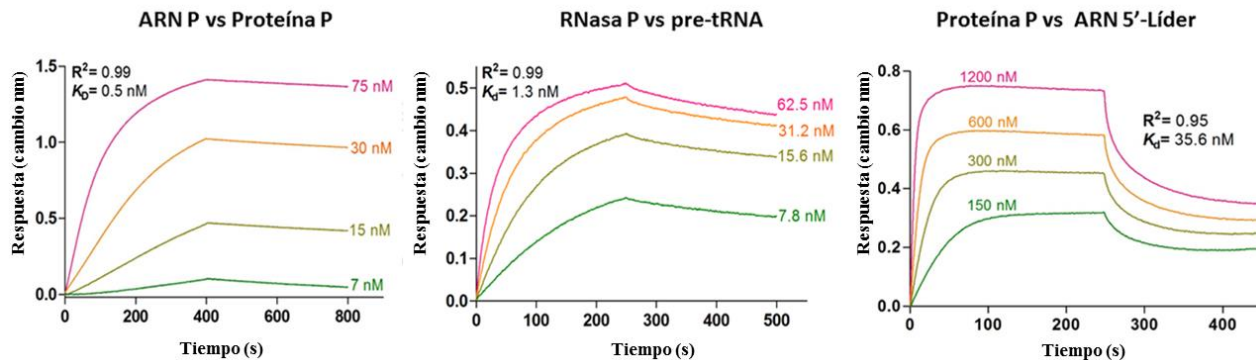


Figura S10. Validación de biosensores Ni-NTA para la RNasa P mediante interferometría de biocapas. Izquierda. Sensorgramas de las curvas de titulación de la proteína P contra la subunidad ARN P en un intervalo de concentraciones de 7 nM – 75 nM. Una K_D de 0.5 nM fue determinada para esta interacción. **Panel central:** sensorgramas de las curvas de titulación del pre-tRNA contra la RNasa P en un rango de concentraciones de 7.8 nM – 62.5 nM. **Derecha.** Estos tres experimentos se llevaron a cabo en amortiguador HAAM (HEPES 20 mM, acetato de amonio 100 mM, $MgCl_2$ 10 mM, pH 7.4).

Validación del uso de biosensores Ni-NTA para detección de molécula pequeña mediante BLI

De acuerdo a los fundamentos teóricos de la técnica de interferometría de biocapas, el límite inferior de detección es de 150 Da. El fragmento detectado por cristalografía de rayos X, el 2-MBX (Figura 7), tiene un peso molecular de 151.19 Da que se encuentra muy cerca de dicho límite. El fragmento pudo ser detectado y se observó una señal dosis-respuesta (Figura 10), lo que es una prueba de principio de la medición de interacciones de molécula pequeña con la proteína P.

Controles de los experimentos de termoforesis en microescala (MST)

El control positivo consistió en medir la afinidad del extremo 5' líder de ARN del pre-tRNA con la proteína P, la cual previamente obtuvimos mediante BLI con una $K_D = 1.6$ nM. El resultado de este experimento fue el cálculo de una K_D en buen acuerdo con nuestros datos previos (Figura S11).

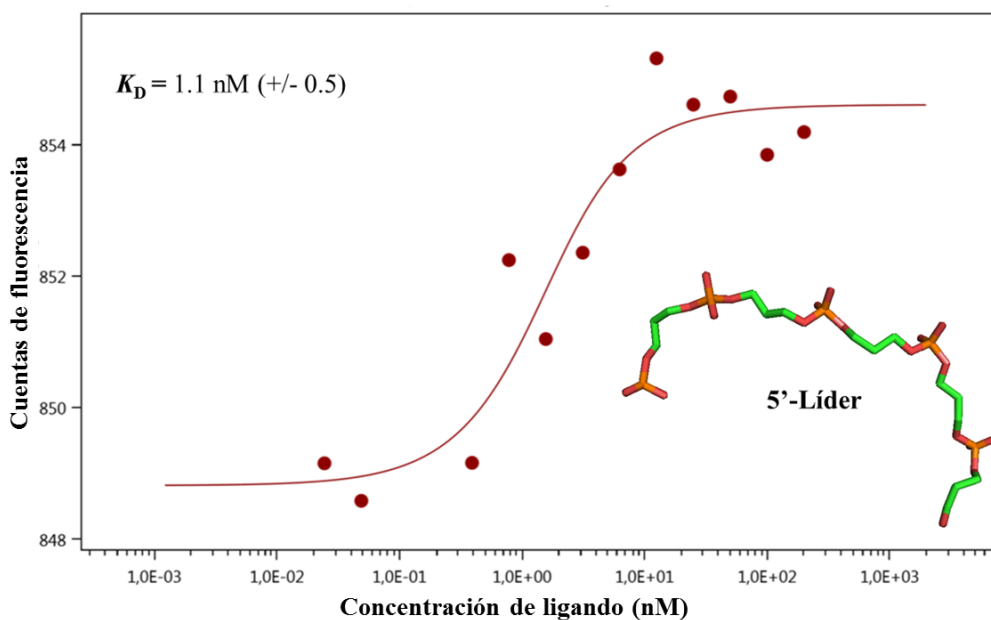


Figura S11. El extremo 5' líder interacciona con la proteína P mediante MST con una afinidad similar a la obtenida por BLI. Para llevar a cabo experimentos con la proteína P, la etiqueta fluorescente NT-647 fue covalentemente unida a la proteína P (vía acoplamiento de grupos NHS). En el experimento por MST, el experimento se mantuvo constante la proteína P etiquetada con NT-647, mientras que la concentración del extremo 5'-líder de ARN no etiquetado con sonda fluorescente fue variado de 0.02 nM – 500 nM. El ensayo se llevó a cabo en amortiguador de fosfatos, pH 7.4, adicionado con 500 mM NaCl y 1 mg/mL BSA. Después de una corta incubación las muestras fueron cargadas dentro de capilares MST NT.115 *premium glass* y la medición por MST fue llevada a cabo usando usando el equipo Monolith NT.115^{Pico}. Las concentraciones en el eje X están graficadas en escala logarítmica en nM. Una K_D de 1.1 nM +/- 0.5 nM fue determinada para esta interacción.

El control negativo se realizó con uno de los compuestos (C-4) de nuestra micro-biblioteca de 11 compuestos derivados del 2-MBX para el cual no fue detectada ninguna interacción con la proteína P mediante BLI (Figura S12).

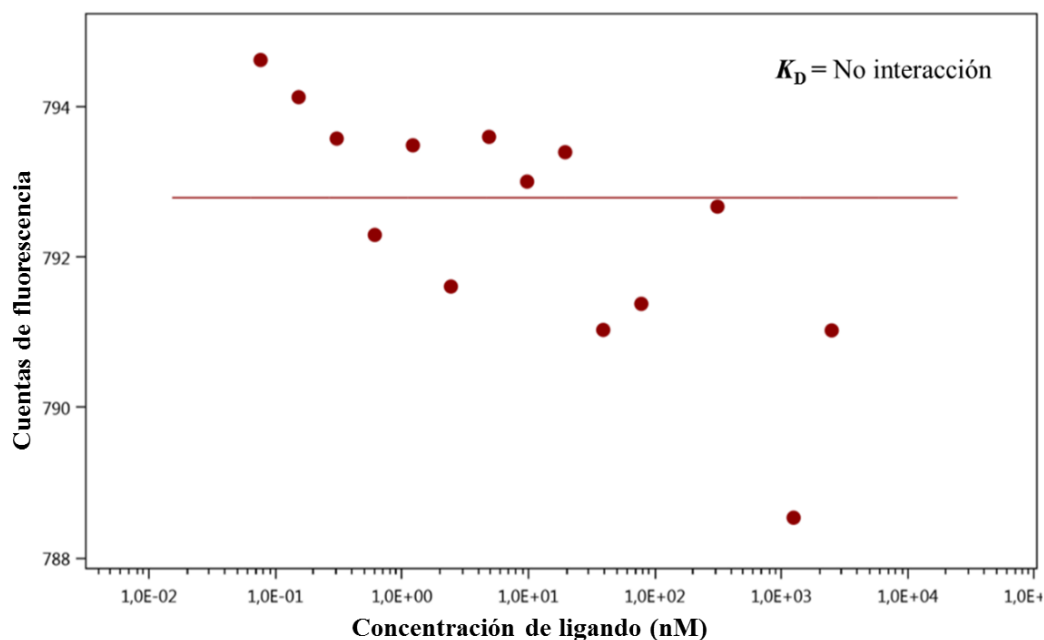


Figura S12. El compuesto control negativo C-4 no interacciona con la proteína P a través de experimentos por termoforesis a microescala. Para llevar a cabo experimentos con la proteína P, la etiqueta fluorescente NT-647 fue covalentemente unida a la proteína P (vía acoplamiento de grupos NHS). En el experimento por MST, el experimento se mantuvo constante la proteína P etiquetada con NT-647, mientras que la concentración del Compuesto 4 no etiquetado con sonda fluorescente fue titulado en 14 puntos de 0.1 nM – 2.5 μ M. El ensayo se llevó a cabo en amortiguador de fosfatos, pH 7.4, adicionado con 500 mM NaCl y 1 mg/mL BSA. Después de una corta incubación las muestras fueron cargadas dentro de capilares MST NT.115 *premium glass* y la medición por MST fue llevada a cabo usando usando el equipo Monolith NT.115. Las concentraciones en el eje X están graficadas en escala logarítmica en nM. Ninguna interacción fue determinada para este experimento.

Cristalización de la proteína P

El protocolo de purificación de la proteína P sin etiqueta de hexa-histidinas resultó en la obtención de cristales con un hábito cristalino regular y reproducible (Figura S13, izquierda) que difractaron a una resolución de 1.5 Å (Figura S13, centro) en una fuente de luz en sincrotrón. La resolución de la estructura mediante reemplazo molecular, resultó en un modelo de la proteína P de *T. maritima* de buena calidad como se puede observar en su mapa de omisión (Figura S13, panel derecho).

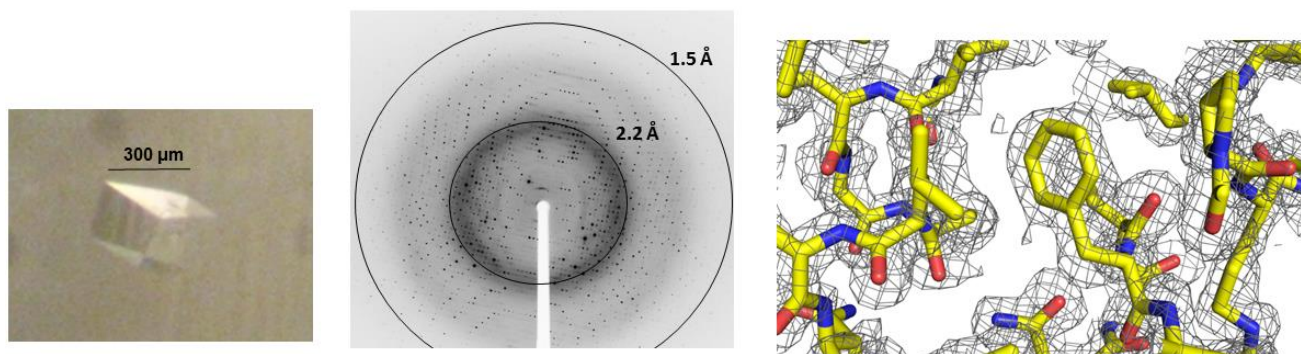


Figura S13. Cristal obtenido a partir de la proteína P silvestre de *T. maritima*. **Izquierda.** Los cristales se obtuvieron en la condición con polietilenglicol (PEG) 1000 al 12%, acetato de sodio trihidratado (NaOAc), pH 4.8 a 100 mM y sulfato de potasio (K_2SO_4) a 200 mM. Las dimensiones aproximadas de los cristales fueron 300 μm x 200 μm x 90 μm . **Centro.** Patrón de difracción típico a 1.5 Å. **Derecha.** Mapa de omisión (2mFo-Fc) de recocido simulado (gris) sobre el modelo (colores) de la estructura cristalográfica de la proteína P.

Apéndice D. Recomendaciones para experimentos futuros

Las siguientes son recomendaciones para experimentos y estrategias que posiblemente se puedan considerar como vertientes del proyecto y que no necesariamente tienen relación directa con los objetivos principales de esta tesis. Asimismo, estas recomendaciones se encuentran fuera de la mira o del contexto de las perspectivas de este estudio. Sin embargo, estas recomendaciones se engloban perfectamente dentro de los esfuerzos que se deben tomar con el objetivo de desarrollar nuevas clases de antibióticos que tengan como blanco a la RNasa P bacteriana por lo que es importante explorar estas propuestas experimentales.

1. Desarrollar una biblioteca de compuestos enfocada en quimiotipos que se hayan observado como ligandos redundantes de la proteína P con base en análogos de los “hits” detectados en este estudio. Dicha colección de compuestos podría estar conformada por análogos que tengan como núcleo estructural constante a las antraquinonas o hidroxiantraquinonas, que son la familia de compuestos a partir de los cuales se deriva la purpurina.
2. Determinar las estructuras cristalográficas de alta resolución de la proteína P en complejo con la hemateína y la juglona. Solamente cuando se cuenta con una estructura de un complejo proteína-ligando es que realmente pueden comenzar los esfuerzos para progresar el “hit” a un compuesto líder. La hemateína y juglona son compuestos orgánicos de bajo peso molecular que se unieron a la proteína P en el intervalo de decenas de micromolar durante experimentos de unión dosis-respuesta por BLI. De obtenerse la estructura de estos complejos, ambos “hits” pueden ser optimizados mediante diseño basado en estructura para generar compuestos líder, por lo que es importante considerar estos dos ligandos como valiosos puntos de partida para diseñar inhibidores de la RNasa P bacteriana más potentes y selectivos.
3. Determinar la estructura del complejo de la proteína P con iriginol hexaacetato (Ir6Ac) e irigenol para luego optimizar estos ligandos a través del diseño guiado por estructura y de estudios de relaciones estructura-cinética y de estructura-actividad. Al igual que la hemateína y la juglona, el irigenol y su derivado el Ir6Ac, son dos “hits” que pueden ser prometedores en la búsqueda de nuevos inhibidores. Además, como ya se discutió en el trabajo, el iriginol hexaacetato es un inhibidor de la RNasa P bacteriana de *B. subtilis* que ya se encuentra validado por otro grupo de investigación (6), y el cual nosotros determinamos que se une a la proteína P de manera dosis-respuesta en el intervalo de decenas de nanomolar ($K_D= 80$ nM y 159 nM en condiciones diferentes) mediante experimentos por BLI. Con base en estas observaciones, sería interesante estudiar esta interacción para poder generar compuestos líder con mejor biodisponibilidad y características tipo fármaco.

4. Realizar una búsqueda de nuevos inhibidores en un formato de alta eficiencia mediante un ensayo basado en células que expresen la RNasa P bacteriana y PRORP en condiciones diferenciales, usando un sistema de condiciones permisivas (arabinosa) y no permisivas (glucosa).
5. Determinar el efecto de la purpurina en la inhibición de la RNasa P de bacterias patógenas modelo (p. ej.: *K. pneumoniae*, *S. aureus*, *E. coli* y *M. tuberculosis* MRSA). Este análisis podría derivar en conocimiento importante acerca de la actividad de la purpurina o de sus derivados en el espectro bacteriano.
6. Determinar la estructura cristalográfica de alta resolución de la proteína P en complejo con el extremo 5'-líder. Esto es importante ya que de tener conocimiento sobre los contactos específicos de las bases de ARN con residuos que forman parte del surco de unión del líder en la proteína P, es factible el diseño de compuestos análogos y lograr un impedimento óptimo del reconocimiento de la proteína P por el 5'-líder.
7. Optimizar los compuestos C-7 y C-10 hacia compuestos tipo fármaco, mediante la determinación de la estructura cristalográfica de complejos con la proteína P con el objetivo de iniciar el proceso del diseño basado en estructura y de tal manera obtener compuestos líder con parámetros mejorados de biodisponibilidad y selectividad.
8. Ya que recientemente se determinaron las asignaciones de resonancia para NMR de la proteína P de *T. maritima* por un grupo de investigación independiente (7), es factible y deseable determinar la estructura tridimensional por NMR de la proteína P con los compuestos C-7 y C-10 como una metodología alternativa para lograr vencer el obstáculo técnico de obtener cristales de estos complejos, los cuales han sido elusivos hasta la fecha.

Apéndice E. Glosario de términos

La siguiente es una lista no exhaustiva de los términos especializados (8) más relevantes de este trabajo.

Ångström

Unidad métrica de distancia igual a 10^{-10} m o 0.1 nm. La unidad lleva este nombre en honor al físico sueco del siglo 19 Anders Jonas Ångström. Esta unidad no es parte del Sistema Internacional de Unidades, sin embargo, se usa extensamente en cristalografía de rayos X debido a su practicidad en la medición de distancias entre átomos y moléculas.

Afinamiento

Es el proceso iterativo de mejorar el acuerdo entre el modelo molecular y los datos cristalográficos. Un elemento importante en afinamiento es un ajuste computacional masivo de mínimos cuadrados de 1) las posiciones atómicas en el modelo, 2) ocupaciones y, 3) factores de temperatura con el fin de mejorar su acuerdo con: a) las intensidades de las reflexiones y b) el criterio de parámetros estructurales como ángulos y longitudes de enlaces.

Apilamiento

En química supramolecular, el apilamiento se refiere a un tipo de interacción intermolecular que se basa en el ordenamiento en pila, frecuentemente de moléculas aromáticas, que es adoptado debido a interacciones interatómicas.

Co-cristalización

En descubrimiento de fármacos por cristalografía de rayos X, se refiere a la cristalización de un complejo de una biomacromolécula y una molécula orgánica de bajo peso molecular a partir de una solución incubada de ambos componentes.

Compleción de los datos

Es el número de reflexiones cristalográficas medidas en un conjunto de datos, expresado en porcentaje de las reflexiones realmente colectadas, comparado con los datos únicos para una dada simetría cristalina y a una resolución específica.

Compuesto líder

En descubrimiento de fármacos, un compuesto líder es un compuesto químico que posee actividad biológica o farmacológica que es factible de ser útil terapéuticamente, pero que tiene una estructura sub-óptima que requiere modificación para acoplarse de mejor manera con el blanco. Su estructura química sirve como un punto de partida para modificaciones químicas con el fin de mejorar potencia, selectividad o parámetros farmacocinéticos y así convertirse en un compuesto tipo fármaco.

Compuesto tipo fármaco

Concepto en diseño de fármacos para referirse a un compuesto químico que tiene buenos parámetros de biodisponibilidad como: solubilidad en agua y grasa, alta potencia, alta selectividad por el blanco terapéutico, permeabilidad de barrera sangre-cerebro, baja o nula toxicidad, disponibilidad oral, alta eficiencia de ligando y un peso molecular entre 200 y 600 Da.

Conector o linker

Una vez que se han identificado dos fragmentos que se han unido a dos cavidades distintas del blanco terapéutico, éstos pueden ser unidos mediante una molécula conectora para obtener un compuesto líder con mayor potencia, eficiencia, selectividad y afinidad. Estos conectores pueden ser derivados de grupos funcionales alquilo, conectores

funcionalizados o inspirados en ligandos conocidos. Un ejemplo de un fármaco que fue desarrollado usando esta estrategia es venetoclax, un fármaco oral en investigación para el tratamiento de la leucemia linfática crónica y la leucemia mieloide aguda.

Diseño de fármacos basado en fragmentos

Metodología en la cual se usan compuestos orgánicos de bajo peso molecular (fragmentos) para interrogar a un blanco terapéutico. Se explora mayor espacio químico, se pueden obtener ligandos eficientes y selectivos que se adaptan o acoplan de manera óptima a una cavidad de interés del blanco. Una vez detectado un “hit” por diversas técnicas, éste se modifica químicamente para mejorar sus características de biodisponibilidad.

Drogabilidad

Anglicismo de la palabra “druggability”. Este término se usa para referirse a la posibilidad de un blanco terapéutico de ser modulado por una molécula orgánica de bajo peso molecular y es crucial para determinar si un proyecto de descubrimiento de fármacos puede avanzar de “hit” a compuesto líder.

Eficiencia de ligando

Es una medida de la energía de unión por cada átomo de un ligando a su blanco y es usada en programas de investigación en descubrimiento de fármacos para ayudar a la selección de compuestos líder con combinaciones óptimas de propiedades fisicoquímicas y farmacológicas. Matemáticamente, la eficiencia de ligando puede ser definida como el radio de la energía libre de Gibbs (o valor de afinidad o actividad) dividido entre el número de átomos pesados (diferentes al hidrógeno) del compuesto.

Ensayabilidad

Anglicismo de la palabra “assayability”. Con el fin de apoyar un proyecto de búsqueda de inhibidores, es necesario contar con ensayos bioquímicos y/o celulares para medir la unión de compuestos al blanco y la función del blanco. A este conjunto de parámetros, nos referimos bajo el término “ensayabilidad” del blanco terapéutico. La probabilidad de establecer un ensayo significativo depende de la clase e información del blanco.

Espacio químico

Se estima que el número teórico de compuestos químicos que es posible sintetizar es de 10^{60} y que para sintetizar tan solo 10 mg de cada compuesto posible necesitaríamos más materia que la que tiene el universo observable. El diseño de fármacos basado en fragmentos es una estrategia para explorar este espacio químico de manera más eficiente.

Factor B

Es una medida de cuanto un átomo oscila o vibra alrededor de una posición especificada en el modelo. Se espera que los átomos en los términos de cadenas laterales de aminoácidos exhiban una mayor libertad de movimiento que aquellos átomos que forman la cadena principal de la proteína. La difracción es afectada por esta variación en la posición atómica, por lo que es realista asignar un factor de temperatura a cada átomo e incluir el factor entre los parámetros a optimizar durante el afinamiento de mínimos cuadrados.

Factor residual R_{sym} o R_{merge}

Es una medida del ajuste entre mediciones múltiples e independientes de la misma reflexión relacionadas por simetría en un set de datos cristalográficos. Estas diferentes mediciones se encuentran en diferentes imágenes de datos. Las reflexiones relacionadas por simetría deberían tener intensidades idénticas. Si no lo son, esto sugiere algún tipo de error en la medición.

Fragmento

Dentro del área del diseño de fármacos basado en fragmentos, el término fragmento se refiere a un compuesto químico orgánico, estructuralmente simple, con un peso molecular de entre 150-300 Da, a partir del cual se diseña un compuesto tipo fármaco.

Grupo espacial

Designación de la simetría de la celda unitaria de un cristal. La simetría de la celda unitaria guía al cristalógrafo en el desarrollo de una estrategia de colección de datos que medirá todas las reflexiones únicas con la redundancia deseada.

Hit

Compuesto orgánico de bajo peso molecular que presenta una actividad deseada durante una búsqueda de ligandos y cuya actividad es confirmada al re-analizarlo en experimentos de dosis-respuesta.

I/σ

Es una medida de la señal sobre ruido del pixelaje de la intensidad de las reflexiones en el patrón de difracción.

Mapa de densidad electrónica

El mapa es una imagen tridimensional de las nubes de electrones presente en la molécula. En un proceso llamado interpretación del mapa, el cristalógrafo construye un modelo para encajar esta imagen.

Mean (I) half-set correlation $CC_{1/2}$

Coefficiente de correlación entre la intensidad estimada a partir de la mitad del conjunto de datos. Es el indicador primario para la selección del límite de alta resolución para el procesamiento de datos. Se relaciona con la señal sobre ruido efectiva de los datos.

Mosaicidad

Es la medida en ángulos del grado del desorden de largo alcance la celda unitaria dentro de un cristal. Menor mosaicidad es indicativo de cristales más ordenados y por lo tanto mejor difracción.

Ocupación

Uno de los muchos parámetros incluidos en el afinamiento. La ocupación de un átomo es una medida de la fracción de moléculas en el cristal en el cual un átomo ocupa realmente la posición especificada en el modelo. Si todas las moléculas en el cristal son precisamente idénticas, entonces las ocupaciones para todos los átomos son de 1.00. La ocupación se incluye en los parámetros de afinamiento ya que ocasionalmente dos o más conformaciones distintas se observan en una cadena lateral de superficie. El modelo puede afinarse mejor si a los átomos en esta región se les asignan ocupaciones iguales a la fracción de las cadenas laterales en cada conformación. Por ejemplo, si las dos conformaciones ocurren con igual frecuencia, los átomos involucrados reciben ocupaciones de 0.5 en cada de sus dos posiciones posibles.

Redundancia

La medición de regiones traslapadas del patrón de difracción tridimensional, el cual se hace para asegurar cobertura completa, casi siempre resulta en múltiples mediciones de la misma reflexión. La redundancia es calculada como el número de reflexiones medidas entre el número de reflexiones únicas. El cálculo da el número promedio de mediciones independientes de cada reflexión en un conjunto de datos cristalográfico.

Relación R_{work}/R_{free}

El factor R es la medida de la calidad del modelo atómico obtenido de los datos cristalográficos. Cuando se resuelve una estructura de una proteína, primeramente, se construye un modelo atómico y entonces se calcula un patrón de

difracción simulado basado en ese modelo. El valor R mide que tan bien el patrón de difracción simulado se ajusta al patrón de difracción observado experimentalmente. Un conjunto de átomos totalmente aleatorios resultará en un valor R_{work} de 0.63, mientras que un ajuste perfecto tendría un valor de 0. Típicamente, los valores son de 0.20. Un ajuste puede no ser perfecto por muchas razones. Una de las más grandes razones es que los cristales de proteína y ácidos nucleicos contienen canales de agua. El agua no tiene una estructura definida y no está incluida en el modelo atómico. Otras razones incluyen el desorden y vibración que no se toma en cuenta en el modelo. Existe un potencial problema con el uso de valores R_{work} para evaluar la calidad de una estructura. El proceso de afinamiento casi siempre es usado para mejorar el modelo atómico de una estructura dada para ajustarlo mejor con los datos experimentales y mejorar el valor R_{work} . Desafortunadamente, esto introduce un sesgo en el proceso, ya que el modelo atómico es usado junto con el patrón de difracción para calcular la densidad electrónica. El uso del valor R_{free} es una manera menos sesgada de hacer esto. Antes de que el afinamiento comience, alrededor de 10% de las observaciones experimentales son removidas del conjunto de datos y, después, el afinamiento se lleva a cabo con el restante 90%. El valor R_{free} es calculado al determinar que tan bien el modelo predice el 10% que no fue usado en el afinamiento. Para un modelo ideal que no está sobre interpretando los datos, el R_{free} será similar al valor R_{work} (9).

Remojo

Proveniente de la palabra anglosajona “soaking”, se refiere a la metodología para incorporar átomos pesados, ligandos u otros compañeros pequeños de unión hacia un cristal ya crecido mediante la adición de esa sustancia al licor madre. La concentración de ligando que se debe añadir depende de la afinidad de unión al blanco y debe ser varias veces más grande que la K_D .

Resolución

Es el límite en el ángulo de difracción hasta el cual los rayos X difractados por un cristal pueden ser detectados. Está directamente relacionado con la definición óptica en la cual la distancia mínima que dos objetos pueden estar separados y aún poder verse como dos objetos separados. La resolución normalmente se describe en Ångströms (Å). Entre más alta la resolución, menor espacio interplanar y más detallado el mapa de densidad electrónica. Así, “un modelo a 2 Å de resolución”, significa que el modelo toma en cuenta difracción de conjuntos de planos paralelos de átomos equivalentes que tienen un espacio entre ellos de 2 Å en la celda unitaria. Los planos de átomos que se encuentran más cercanamente dan origen a reflexiones más lejanas del centro del patrón de difracción. Las reflexiones más lejanas que la resolución establecida no pueden ser obtenidas, o bien, la señal es muy débil para ser confiables. Aunque un mapa de densidad electrónica a 2 Å, visto como una superficie de contorno vacía, puede no permitir discernir átomos adyacentes a una distancia de 2 Å o menos, las restricciones estructurales en el modelo incrementan de manera importante la precisión de las posiciones de átomos.

Bibliografía de la sección de apéndices

1. Gößringer M, Kretschmer-Kazemi Far R, Hartmann RK (2006) Analysis of RNase P protein (rnpA) expression in *Bacillus subtilis* utilizing strains with suppressible rnpA expression. *J Bacteriol.* 188(19):6816-6823.
2. Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol.* 5,189(1):113-30.
3. Wegscheid B, Hartmann RK (2006) The precursor tRNA 3'-CCA interaction with *Escherichia coli* RNase P RNA is essential for catalysis by RNase P *in vivo*. *RNA.* 12:2135-2148.
4. Talbot SJ, Altman S (1994) Gel retardation analysis of the interaction between C5 protein and M1 RNA in the formation of the ribonuclease P holoenzyme from *Escherichia coli*. *Biochem.* 33:1399-1405.
5. Buck AH, Dalby AB, Poole AW, Kazantsev AV, Pace NR (2005) Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J.* 24:3360-3368.
6. Liu X, Chen Y, Fierke CA (2014) A real-time fluorescence polarization activity assay to screen for inhibitors of bacterial ribonuclease P. *Nucl Ac Res.* 42(20):e159.
7. Zeng D, Brown BP, Voehler MW, Cai S, Reiter NJ (2018) NMR resonance assignments of RNase P protein from *Thermotoga maritima*. *Biomol NMR Assign.* 12:183.
8. Rupp, Bernhard (2009) *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology.* 1ST Edition. Garland Science.
9. Kleywegt GJ, Jones TA (1997) Model building and refinement practice. *Methods in Enzymology* 277, 208-230.

13. ANEXOS

Anexo A. Artículo Nucleic Acids Research 2019 (primer autor).

A screening platform to monitor RNA processing and protein-RNA interactions in ribonuclease P uncovers a small molecule inhibitor

Ezequiel-Alejandro Madrigal-Carrillo^{1,†}, Carlos-Alejandro Díaz-Tufinio^{1,2,†},
Hugo-Aníbal Santamaría-Suárez¹, Marcelino Arciniega¹ and Alfredo Torres-Larios^{1,*}

¹Department of Biochemistry and Structural Biology, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Mexico City, Mexico and ²Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Mexico City, Mexico

Received October 26, 2018; Revised April 08, 2019; Editorial Decision April 09, 2019; Accepted April 10, 2019

ABSTRACT

Ribonucleoprotein (RNP) complexes and RNA-processing enzymes are attractive targets for antibiotic development owing to their central roles in microbial physiology. For many of these complexes, comprehensive strategies to identify inhibitors are either lacking or suffer from substantial technical limitations. Here, we describe an activity-binding-structure platform for bacterial ribonuclease P (RNase P), an essential RNP ribozyme involved in 5' tRNA processing. A novel, real-time fluorescence-based assay was used to monitor RNase P activity and rapidly identify inhibitors using a mini-helix and a pre-tRNA-like bipartite substrate. Using the mini-helix substrate, we screened a library comprising 2560 compounds. Initial hits were then validated using pre-tRNA and the pre-tRNA-like substrate, which ultimately verified four compounds as inhibitors. Biolayer interferometry-based binding assays and molecular dynamics simulations were then used to characterize the interactions between each validated inhibitor and the P protein, P RNA and pre-tRNA. X-ray crystallographic studies subsequently elucidated the structure of the P protein bound to the most promising hit, purpurin, and revealed how this inhibitor adversely affects tRNA 5' leader binding. This integrated platform affords improved structure-function studies of RNA processing enzymes and facilitates the discovery of novel regulators or inhibitors.

INTRODUCTION

Regulatory RNAs, ribozymes, and RNA-protein complexes are appealing antibiotic targets due to their essential functions in microbial metabolism (1–3). This clinical importance is exemplified by the ribosome, which is currently the target of roughly 50% of known antibiotics (4). The focus of this work is Ribonuclease P (RNase P), the only ribozyme other than the ribosome that is present in all three domains of life (see (5–8) for some reviews). This essential ribonucleoprotein complex remains to be exploited as a target for much-needed novel antibacterial agents (9).

The composition of RNase P varies across the three domains of life (10) and therefore may afford high selectivity in drug targeting (11,12). While in archaea and eukaryotes, RNase P is comprised of one RNA subunit and four to ten proteins (13), in bacteria, this complex is formed by an RNA subunit (P RNA, 350–400 nucleotides, 110–125 kDa) and a single protein (P protein, ~110 amino acids, 13 kDa) (14). In all species, the P RNA serves as the primary biocatalyst (15) for the cleavage of the 5'-leader sequence of pre-tRNAs during tRNA maturation (16). The P protein, on the other hand, binds the distal 5'-leader region of the pre-tRNA substrate, enhances the affinity of metal ions, and assists in product release (17–20). RNase P is dependent on divalent metal ions (Mg²⁺ is needed for proper folding and activity (21–23)) and *in vitro*, high concentrations of Mg²⁺ are sufficient to allow P RNA catalysis, even in the absence of the protein subunit (15,24,25).

There are several known inhibitors of RNase P (reviewed in (12,26,27)), with iriginol hexaacetate and methylene blue being the most recently discovered (26,28). However, to date, neither natural nor highly specific RNase P inhibitors have been identified. The dynamic and flexible nature of the RNase P complex may have hindered the evolution of specific binders (29) and moreover, the protein subunit has some features of intrinsically disordered proteins, which

*To whom correspondence should be addressed. Tel: +52 55 5622 5640; Fax: +52 55 5622 5630; Email: torres@ifc.unam.mx

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

could hamper the stable binding of ligands (30). There also may be an inherent limitation for small molecules to compete with the protein for binding to the P RNA, depending on the relative affinity of the P RNA:P protein pair.

Monitoring RNase P activity has traditionally relied on using radiolabeled RNA substrates (31). In this method, radiolabeled RNA is mixed with RNase P. The reaction is then monitored by denaturing polyacrylamide gel electrophoresis (Urea-PAGE) followed by auto-radiography (31). Very recently, alternatives to measure the kinetics of RNase P on pre-tRNA substrates have emerged. A recent paper developed a fluorescence polarization-based activity assay to monitor RNase P activity on a full-length substrate (26).

To expand the repertoire of inhibitors that may serve as clinical leads as antibiotics, a robust platform for assessing RNase P activity in high throughput is essential. Current activity protocols are either inhibited by 1% (v/v) DMSO or not amenable to high-throughput screening. Additionally, validation methods to verify efficacious RNase P inhibition are lacking. Recognizing these shortcomings, we aimed to create improved methods for discovering small molecule inhibitors of RNase P. To accomplish this goal, we focused on the following strategy: (i) developing a novel method to measure real-time RNase P activity in solution; (ii) obtaining real-time measurements of the binding of the RNase P holoenzyme, P RNA and P protein to their native substrate and inhibitors and (iii) obtaining structural information, *in silico* and through X-ray crystallography, on the binding of potential inhibitors to the P protein.

In this work, we present an RNase P activity assay that exploits a previously reported minimal model substrate (pMini3bpUG, herein referred to as Minihelix or Mh) (32,33). This substrate utilizes a FRET mechanism in which the RNase P substrate couples both a 3' fluorophore and a 5' non-fluorescent quencher. Cleavage and release of the quencher molecule by RNase P enables the detection of enzymatic activity by measuring fluorescence emission over time, which is amenable for monitoring steady-state kinetics and for high-throughput screening assays.

We then implemented this method to assess a compound library of 2560 small molecules and found four compounds that inhibit RNase P activity. These inhibitors were effective in the presence of both a canonical pre-tRNA substrate and a novel pre-tRNA-like substrate (herein referred to as 'bipartite pre-tRNA') that is composed of two RNA oligonucleotides and monitors the reaction in an analogous way to the Mh substrate. To avoid the sensitivity of RNase P processing to organic solvents using the bipartite pre-tRNA substrate, we dissolved the hits in PEG 200 rather than DMSO. This procedure allowed us to validate the inhibitory properties of these molecules under varied conditions.

Positive hits were then verified and characterized using biolayer interferometry (34), which allowed us to perform the following tasks: (i) define the affinity parameters of the RNase P holoenzyme, P RNA and P protein to the RNA substrates, (ii) discriminate between the interactions of a given compound with the holoenzyme, P RNA, P protein or the substrate and (iii) determine if a given compound hinders the binding of the holoenzyme to pre-tRNA or the P protein to the 5'-leader.

Once validated, we performed docking and molecular dynamics simulations with each hit and identified putative binding sites for two inhibitors on the P protein. Furthermore, purpurin, a competitive inhibitor which behaved as the most consistent hit across our assays, was shown to bind the P protein by X-ray crystallography, with its binding site corresponding to part of the 5'-leader binding site.

MATERIALS AND METHODS

Selection of bacterial RNase P

We chose to work with RNase P from the thermophilic bacterium *Thermotoga maritima* for several reasons. First, it represents the ancestral and most common 'type A' ribozyme, the same that is present in *Escherichia coli* (35). Second, the crystal structure of the RNase P holoenzyme from *T. maritima* in complex with tRNA (19) as well as the apo-structure of the RNA subunit (36) are known, thus facilitating structure-function studies. Third, the folding protocol of the RNase P holoenzyme from *T. maritima* (described below) leads to crystals that present a functional and homogeneous molecular complex (19). Fourth, although the RNase P from *T. maritima* has optimal activity at 50°C (37), the holoenzyme is active in a temperature range of 20–55°C.

RNase P production

The preparation of P RNA and pre-tRNA from *T. maritima* was carried out by *in vitro* transcription and purified by denaturing 6% urea-PAGE with standard protocols as described previously (19,38). The P protein was purified according to (39). The A260/A280 ratio was between 0.6 and 0.7 for each preparation of purified protein.

RNase P reconstitution

The reconstitution of the RNase P holoenzyme was performed as described in (40,19). This protocol produces a homogeneous and functional ribozyme that can be crystallized by mixing P RNA and P protein in a 1:1.1 molar ratio in 1× TH buffer (33 mM Tris-HCl, 66 mM HEPES, pH 7.4) and 400 mM ammonium acetate (AmOAc). The mixture was heated to 95°C for 2 min in a water bath, then chilled on ice for 2 min, after which MgCl₂ was added to a final concentration of 100 mM. Although RNase P from *T. maritima* is catalytically active at 10 mM MgCl₂ (19), we found that for our minihelix substrate, 100 mM MgCl₂ demonstrated more robust activity. This higher concentration was used in all experiments described herein. The holoenzyme was further incubated at 50°C for 10 min and 37°C for 40 min. The final concentration of holoenzyme was 333 nM (35.6 μg/ml of P RNA and 4.7 μg/ml of P protein).

Substrate minihelix (Mh)

The minihelix (Mh) substrate (Figure 1A) was adopted from a previously characterized RNase P substrate, pMini3bpUG (32,33) and coupled to either Black Hole Quencher 1 (BHQ1) or Black Hole Quencher 2 (BHQ2) at the 5'-end and Alexa Fluor 488 at the 3'-end. This probe was

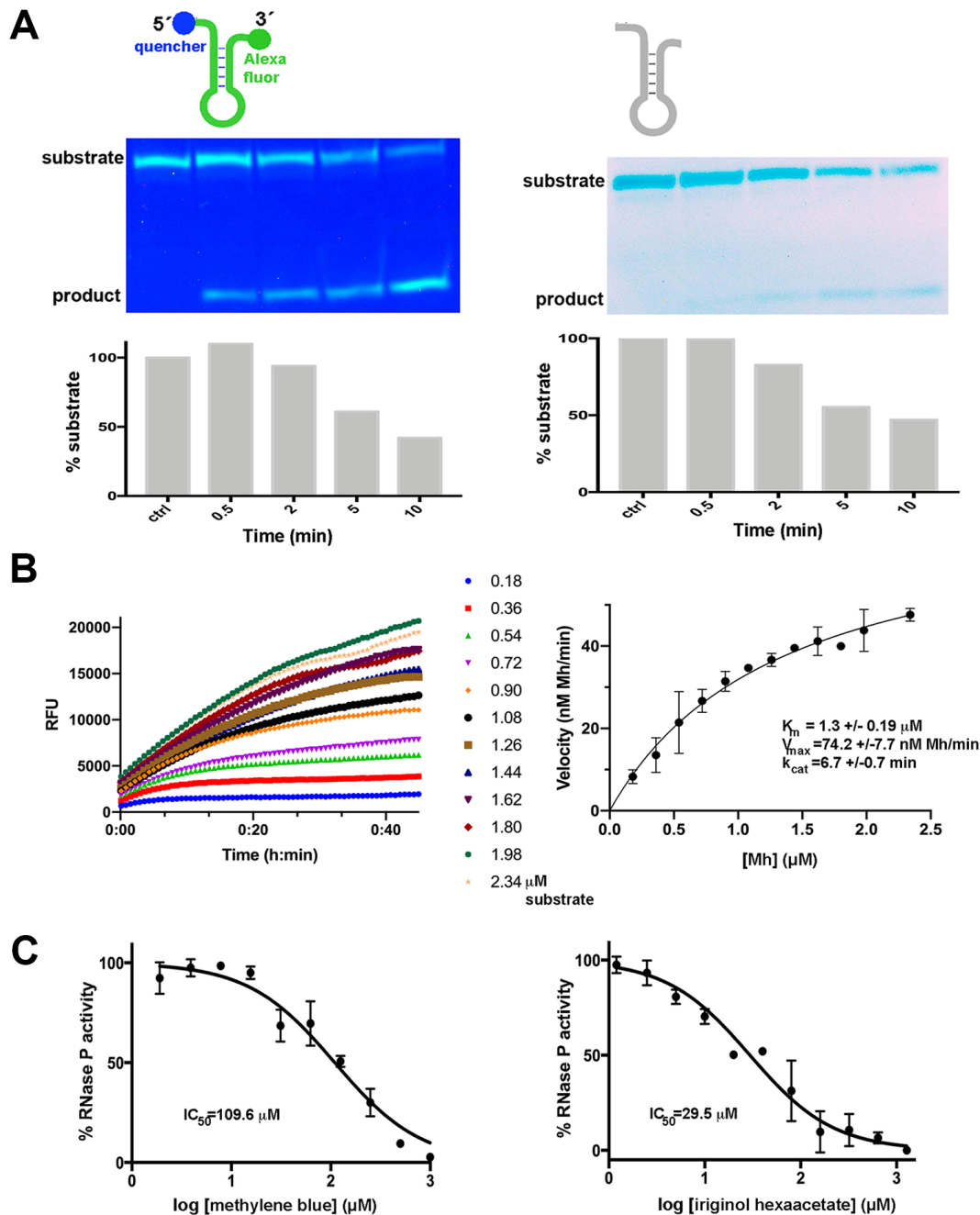


Figure 1. Validation of the fluorescent minihelix (Mh) substrate. (A) Cleavage by bacterial RNase P. Top. Schematic representation of the 24-nucleotide fluorescent (left) or non-fluorescent (right) minihelix (Mh) substrate with the predicted secondary structure according to Brannvall et al. (32). Middle left. RNase P holoenzyme from *Thermotoga maritima* cleaves the fluorescent Mh substrate, as visualized by 20% urea-PAGE under UV light. The fluorescence detected from both the substrate (upper band) and the product (lower band) is due to the unfolding of the Mh probe under the denaturing gel conditions. The time-dependent disappearance at 55°C of the 24 nt Mh fluorescent substrate (5 ng), with a simultaneous accumulation of the 15 nt product, confirms the catalytic cleavage of RNase P on the proposed minihelix. Middle right. FRET-pair addition does not affect the cleavage of RNase P on the fluorescent probe as visualized by UV after PAGE and ethidium bromide staining. Bottom. Similar time courses of substrate conversion are observed for the fluorescent (left) and non-fluorescent (right) Mh substrates. (B) Kinetic assays of RNase P with the fluorescent substrate show a typical Michaelis-Menten enzymatic behavior. Left. Representative plots of kinetic assays are shown. RNase P holoenzyme was incubated at 20°C at 11 nM, with increasing Mh concentrations as indicated. Right. Michaelis-Menten model fitting ($R^2 = 0.9368$). Initial velocities from three independent trials were plotted against substrate concentration; error bars indicate the standard error for each plot. (C) The reported RNase P inhibitors methylene blue and iriginol hexaacetate inhibit bacterial RNase P from *Thermotoga maritima*. Left. The inhibition of RNase P by methylene blue. Dose-response curve ($R^2 = 0.9543$) calculated from the normalized slope values from the linear portions of the kinetic assays in triplicate shown in Supplementary Figure S5. The reported K_i value using a conventional pre-tRNA substrate and the RNA subunit of different bacterial RNase P's is between 14 and 28 μM (28). Right. Inhibition of RNase P by iriginol hexaacetate. Dose-response curve ($R^2 = 0.9556$) calculated from the normalized slope values from assays in triplicate. The reported IC_{50} value using a fluorescent pre-tRNA substrate is 0.8 μM for *Bacillus subtilis* RNase P (26).

purchased either from Sigma-Aldrich or Integrated DNA Technologies (IDT). A requested scale of 0.05 mmol yielded between 6.3 and 20 nmol (57.5 to 181 μg , enough for 1300–4000 reactions depending on the experimental substrate concentration). Before use, the minihelix was diluted from a 100 μM stock (dissolved in water) to 3.6 μM in activity buffer H (33 mM Tris-HCl, 66 mM HEPES, pH 7.4, 400 mM AmOAc and 100 mM MgCl_2).

Substrate bipartite pre-tRNA

The bipartite pre-tRNA substrate derived from pre-tRNA^{Phe} from *T. maritima* (Supplementary Figure S13) was designed for the ease of chemical synthesis and conceived from the fact that the anticodon loop is not required for RNase P recognition (Reiter *et al.*). As for the Mh substrate, fluorescent bipartite pre-tRNA was coupled to Black Hole Quencher 1 (BHQ1) at the 5'-end of the 5' oligo, and to Alexa Fluor 488 at the 3'-end of the 3' oligo (see Supplementary Table S1). The probes were purchased from Integrated DNA Technologies (IDT). A requested scale of 250 nmol yielded 27.4 nmol (0.36 mg) for the 5' oligo and 3.8 nmol (50 μg) for the 3' fluorescent oligo. Before use, the probes were diluted from a 100 μM stock (dissolved in water) and assembled by mixing and heating to 3.6 μM in activity buffer L (Tris-HCl pH 7.4 33 mM, HEPES 66 mM, 100 mM AmOAc and 10 mM MgCl_2).

Fluorescence-based activity assay

Reactions were performed in a total volume of 20 μl . Pipetting was done by hand using multichannel pipettes. First, an 18 μl reaction mixture containing 16.5 μl of activity buffer (H for Mh or L supplemented with 10% PEG 200 for bipartite pre-tRNA) and 270 nM (Mh) or 135 nM (bipartite pre-tRNA) (1.5 or 0.7 μl of the 3.6 μM stock) of the fluorescent substrate probe was pre-incubated on ice for at least 5 min in a black polystyrene 384-well microtiter microplate with a flat transparent bottom (Corning ref 3540). The holoenzyme was then added at a final concentration of 33 nM (2 μl of a 333 nM stock) in 1 \times TH buffer and mixed. Kinetic activity was monitored in a Synergy MX Multimode Detector (BioTek) at 37°C with rapid agitation. At least 20 readings were taken per well every 15 s. The probe was excited at 485 nm, and its emission was monitored at 535 nm with an integration time of 1 s. From the plots of relative fluorescent units (RFU) versus time, linear regression was used to determine the slopes in the linear portion of the curves, using at least 12 data points to obtain initial velocities. Non-linear fitting to obtain the Michaelis–Menten parameters was performed using GraphPad Prism[®] 6.01 software.

Gel-based activity assay

For gel-based activity assays, reactions were prepared as described in the previous section (fluorescence activity assay), incubated for 1–15 min at 37–50°C, then loaded onto a denaturing 7 M urea–20% acrylamide gel in the case of the Mh substrate or 19% acrylamide (40% acrylamide/bisacrylamide 38.5:1.5) for the bipartite pre-tRNA substrate. The activity conditions for the gel reported

in Figure 3A, to monitor the cleavage of the holoenzyme and P RNA, were under 100 mM AmOAc and 10 mM MgCl_2 . The gels (13 \times 16 \times 0.15 cm) were electrophoresed at 20 V/cm for 2–3 h. UV shadowing or ethidium bromide staining was used to visualize RNA cleavage. Detection of minihelix fluorescence was achieved without staining using a UV lamp with excitation at \sim 254 nm.

Library screening

Screening was performed with the Spectrum Collection (MicroSource Discovery Systems, Inc), which comprises 2560 compounds, each dissolved in DMSO at 10 mM. For our initial screening, compounds from this collection were tested in cocktails; eight compounds were tested per well in 384-well format. The final concentration of each compound in the cocktail was 1.25 mM. For the activity assay, reactions were set up in a total volume of 20 μl (13.5 μl of activity buffer, 1.5 μl of fluorescent Mh at 3.6 μM , 3 μl of DMSO or cocktail). This set-up resulted in a final concentration of 187.5 μM for each compound. The control wells included for each experiment series were: (i) reaction with twice (3 μl) the fluorescent Mh amount to avoid a signal saturation for the rest of the wells, (ii) two reactions with 10% DMSO (no cocktail), (iii) Mh substrate only, with no ribozyme, (iv) reaction with 100 mM CaCl_2 , which it is used as a negative control as it competes with MgCl_2 (EDTA was not used because metal ions are essential to maintaining the structure of P RNA (41)). The reaction was initiated by adding 2 μl of holoenzyme at 333 nM and monitored as described above. Each round of experiments contained twenty different cocktails corresponding to one row of a 384-well plate (Supplementary Figures S6 and S24). Cocktails that showed inhibition (those with a slope value lower than any of the positive controls in the same plate) were verified (Supplementary Figure S7) and deconvoluted (Supplementary Figure S8) to determine the individual fragment or fragments responsible for the inhibition. At this stage, each compound from the inhibitory cocktails was tested at a final concentration of 1 mM (2 μl of the 10 mM stock). We performed a gel-based assay for the nine selected compounds to qualitatively verify the results obtained by the fluorescence assay (Supplementary Figure S9). After deconvolution, each of the final four potential positive compounds was purchased individually from a different provider (Sigma-Aldrich) and tested inhibition using the *T. maritima* Mh, pre-tRNA and bipartite pre-tRNA substrates (Figure 2 and Supplementary Figures S10–S12 and S14).

Expression and purification of His-tagged P protein for binding assays

The gene encoding the P protein from *T. maritima* (gene synthesis performed by DNA 2.0 ATUM) was cloned into the expression vector TmP-pD454-GST as a Glutathione S-transferase (GST) fusion protein, with a TEV protease cleavage site and a His₆-tag at the N-terminus. This vector was transformed into *E. coli* BL21(DE3)pLysS cells. Cells were then used to inoculate 3 l of Luria-Bertani media and incubated with orbital shaking at 250 rpm and 37°C until

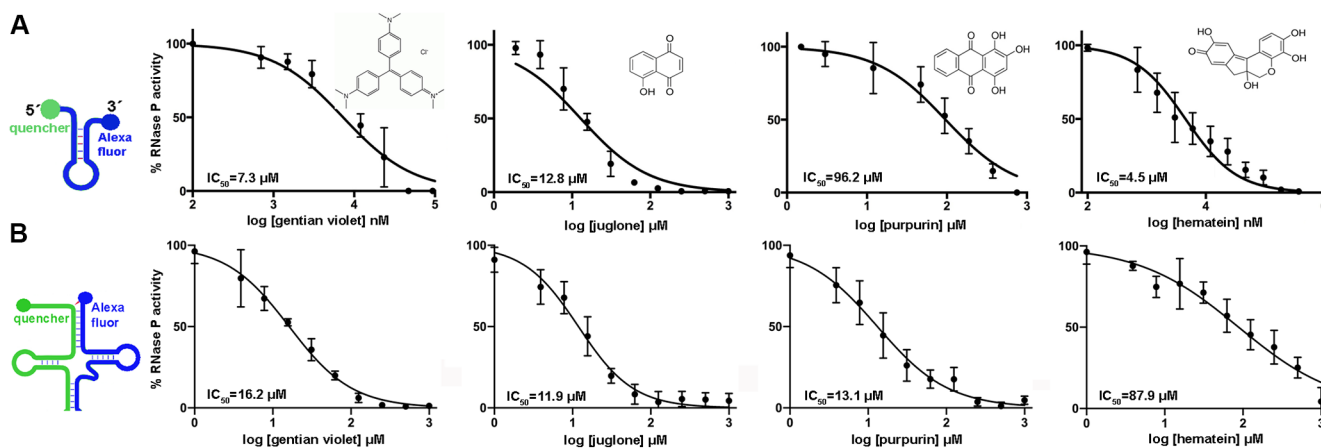


Figure 2. Four validated hits inhibit RNase P holoenzyme from *Thermotoga maritima* in the micromolar range using the substrates Mh and bipartite pre-tRNA. After the screening step (Supplementary Figures S6–S9), four compounds inhibit RNase P using substrates Mh (A) and bipartite pre-tRNA (B). See Supplementary Figure S13 for a description and validation of the bipartite pre-tRNA substrate. Experiments were performed using compounds from Sigma Aldrich in standard conditions and as described in Supplementary Figures S10 and S14. The four hits are gentian violet (tris(4-(dimethylamino)phenyl)methylum chloride), juglone (5-hydroxy-1, 4-naphthoquinone), purpurin (1,2,4-trihydroxyanthraquinone) and hematein (3,4,6a,10-tetrahydroxy-6a,7-dihydroindeno[2,1-*c*]chromen-9(6H)-one). A brief description of some properties found of these compounds is given in the Supplementary Information.

the OD₆₀₀ reached 0.6. Protein expression was then induced with IPTG at 1 mM for 6 h. The cell pellet was collected and sonicated in 40 ml of lysis buffer (20 mM sodium phosphate monobasic, pH 8.0, 500 mM NaCl, 20 mM imidazole, 10% glycerol, 1 mM TCEP). TEV protease (700 μg) was then added directly to the lysate and incubated for 16 h at 34°C. The lysate was centrifuged at 13 000g and solid urea was added to the supernatant to reach a final concentration of 5 M. The sample was then purified by Ni-NTA affinity chromatography column (HisTrap FF crude, GE Healthcare) equilibrated with binding buffer (20 mM sodium phosphate monobasic, 500 mM NaCl, 5 M urea, 30 mM imidazole, pH 8.0). A 30 min wash at 5 ml/min was performed after which the protein was eluted in 20 mM sodium phosphate monobasic, 500 mM NaCl, 5 M urea, 2 M imidazole, pH 8.0. The elution was then loaded onto a gel filtration column (HiPrep 26/10 desalting, GE Healthcare) equilibrated with desalting buffer (20 mM sodium phosphate monobasic, pH 8.0, 500 mM NaCl). The eluted protein was concentrated to 5 mg/ml and stored at 19°C until needed.

Biolayer interferometry assays (P RNA and holoenzyme)

Real-time binding assays of the holoenzyme, P RNA, P protein, the substrate minihelix and the 5'-leader to individual compounds were performed by biolayer interferometry using an Octet RED96 system (ForteBio). For the holoenzyme or P RNA, a 5' biotinylated RNA oligonucleotide with the sequence 5'biotin-CAUUCAGACGAU was immobilized on Super Streptavidin (SSA) biosensors. Then, a modified P RNA or holoenzyme was added. The P RNA modification consisted of a sequence with part of the P1 stem sequence complementary to the biotinylated oligonucleotide (AUCGUCUGGA) (Supplementary Figure S23A). P1 is far from the active site or other regions involved in specific interactions (Supplementary Figure S23B) (19). The activity of this construct was indistinguishable from the wild-type holoenzyme. Alternatively,

the holoenzyme was immobilized through the His-tagged P protein from *T. maritima* at 0.05 mg/ml (3.4 μM) on Ni-NTA biosensors. Control biosensors were used to subtract non-specific binding of compounds to the biosensor surface. All assay steps (immobilization, wash, baseline, association, and dissociation) were performed in 20 mM HEPES pH 7.4, 100 mM ammonium acetate, 10 mM MgCl₂, 2.5% DMSO. The assays were performed on black bottom 96-well microplates (Greiner Bio-One 655209) in a total volume of 200 μl, at 30°C with orbital shaking at 1000 rpm. Experiments were controlled with the software Data Acquisition 8.2 (ForteBio, Inc.). Kinetic binding parameters were calculated using Data Analysis 8.2 (ForteBio, Inc.). After subtraction of reference biosensors, the binding curves were aligned to the X- and Y-axis and the association-dissociation inter-step curve in order to get a common baseline for the association and dissociation phases.

Biolayer interferometry assays (P protein)

Immobilization of the His-tagged P protein from *T. maritima* at 0.05 mg/ml (3.4 μM) was performed on Ni-NTA biosensors. For the initial binding tests to the P protein, a compound concentration of 1 mM was used. Periods for each step of the assay were as follows: immobilization (400 s), wash (100 s), baseline establishment (100s), association (250 s), dissociation (250 s). For each titration curve, six points were used corresponding to the 125, 62.5, 31.25, 15.6 and 7.8 μM compound concentrations. Control biosensors were used to subtract non-specific binding of compounds to the biosensor surface. All assay steps (immobilization, wash, baseline, association, and dissociation) were performed in 20 mM HEPES pH 7.4, 100 mM ammonium acetate, 10 mM MgCl₂, 5% DMSO. The assays were performed on black bottom 96-well microplates (Greiner Bio-One 655209) in a total volume of 200 μl, at 30°C with orbital shaking at 1000 rpm. Experiments were controlled with the software Data Acquisition 8.2 (Forte-

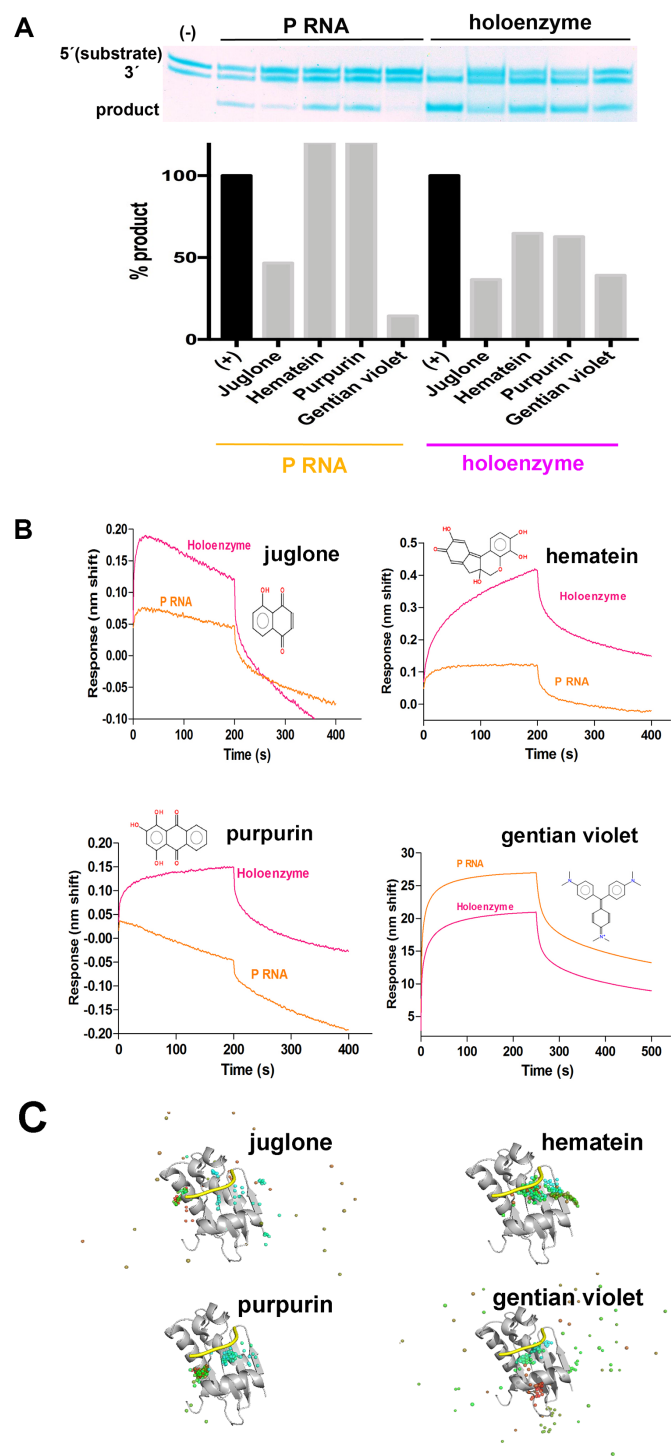


Figure 3. The inhibitors purpurin and hematein bind to the holoenzyme. (A) Top. 19% urea-PAGE demonstrates bipartite pre-tRNA substrate is processed with either P RNA or the RNase P holoenzyme from *T. maritima*. Bottom. Comparison of RNase P activity using P RNA (left) or the holoenzyme (right), indicates that purpurin and hematein do not inhibit the P RNA reaction, as derived by densitometry analysis of the lower band (product). (B) The relative affinities of the inhibitors for the P RNA or the holoenzyme differ. Binding sensorgrams obtained by biolayer interferometry (BLI). P RNA or RNase P holoenzyme was immobilized on super streptavidin (SSA) biosensors through a 5'-biotinylated RNA oligonucleotide complementary to a region on the P1 stem of P RNA (see Materials and Methods). Remarkably, purpurin does not bind to P RNA (or-

Bio, Inc.). Kinetic binding parameters were calculated using Data Analysis 8.2 (ForteBio, Inc.). After subtraction of reference biosensors, the binding curves were aligned to the X- and Y-axis and the association-dissociation inter-step curve in order to get a common baseline for the association and dissociation phases.

Biolayer interferometry assays (minihelix substrate)

3'-Biotinylated minihelix was purchased from Integrated DNA Technologies (IDT) (24-mer RNA sequence, 5'-GAUCUGAAUGCGGAAACGCGCCAC-3'-Biotin). The biotinylated minihelix was immobilized on Super Streptavidin (SSA) biosensors at a concentration of 10 μ M. An initial screening of compounds at 1 mM yielded a binding signal only for two compounds: gentian violet and methylene blue. For the titration experiments, 10 μ M of 3'-biotinylated minihelix was immobilized for 300 s on Super Streptavidin (SSA) biosensors, followed by a wash step of 100 s and association and dissociation steps of 200 s in the presence of 6 different concentrations of compound ranging from 500 to 15.6 μ M. SSA biosensors with no minihelix were used to measure non-specific binding of compounds. All other parameters (buffer, instrument control, analysis) are as described above.

Binding of the P protein to the 5' RNA leader and binding hindrance

A 5' RNA leader sequence with 10-bases 5'-GGAAAAAGAU-3' was purchased from IDT and used at a concentration of 5 μ M. This sequence is the same used to obtain the crystal structure of the RNase P holoenzyme (19). The His-tagged P protein was immobilized on Ni-NTA biosensors at 0.05 mg/ml (3.4 μ M). The 5' RNA leader does not bind to the Ni-NTA biosensor surface. The steps performed on the assay and their periods were: P protein immobilization (400 s), wash (100 s), baseline (100ms) (1000 rpm), compound association (800 s), compound dissociation (200 s) (500 rpm), baseline (100 s), 5' leader association (200 s), 5' leader dissociation (200 s) (1000 rpm). The assay was performed at 30°C. The compounds were used at 100 μ M throughout. All assay steps were performed in 20 mM HEPES pH 7.4, 100 mM ammonium acetate, 10 mM MgCl₂, 5% DMSO.

← In contrast, gentian violet binds not only to P RNA but also to the Mh substrate (Supplementary Figure S16). (C) Molecular Dynamics (MD) simulations of RNase P protein in complex with the validated inhibitors. In each panel, semitransparent spheres represent the geometric center, over a 200 ns trajectory, of the corresponding tested compound. The 5'-leader (yellow) is shown for reference purposes only; it was not included in the simulation. Sphere color coding indicates the starting and final positions in blue and red, respectively. The starting complex conformation for gentian violet, juglone and hematein correspond to the best pose suggested by docking simulations. Purpurin's starting conformation corresponds to that observed in the crystal structure in complex with the P protein (PDB code: 6MAX. See Figure 5). Hematein and purpurin remain bound to the P protein throughout the dynamics protocol.

Molecular docking

In silico molecular docking of gentian violet, juglone, purpurin and hematein were performed on the RNase P protein from *Thermotoga maritima* (PDB code 1NZ0). The analysis was conducted using AutoDockVina 1.2 (42). Protein and small molecule compounds' structures were prepared by employing scripts from AutoDockTools (43). Two search boxes, with sizes 30 Å × 30 Å × 30 Å, were explored, aiming to cover the tRNA 5'-leader binding region (PDB code 3Q1R). These two boxes were centered at ($x = 51.45, y = -23.71, z = -3.4$) and ($x = 49.01, y = -18.63, z = -8.04$). When indicated, side chain flexibility was considered for F17, K51, K53 and K90. Search parameters were set to their default values. The best-suggested pose from the two searches for each ligand was selected as a starting position to perform Molecular Dynamics simulations.

Molecular dynamics simulation

Molecular dynamics simulations were performed on the best binding modes (according to the docking score), for gentian violet, juglone and hematein. Purpurin's starting conformation corresponded to that observed in the crystal structure (PDB code: 6MAX). All systems were simulated with GROMACS 5.0.3 (44) using the AMBER99SB-ILDN all-atom force field (45) in an explicit water solvent scheme with periodic boundary conditions. The topology and force field parameters files for the gentian violet and purpurin compounds were generated with ACYPE (46). The protein protonation was performed with Molprobit (47). The water molecule model used was based on the transferable intermolecular potential with three sites (48). The solvent also included sodium chloride (NaCl) at 0.15 mM plus a small surplus of ions for electric charge neutralization. The long-range electrostatic interactions were computed using fast Particle-Mesh Ewald (PME) (49), using a grid spacing of 1.2 Å. The van der Waals and short-range electrostatic interactions were computed using a cutoff of 1.1 Å for both parameters. The neighbor list was updated every 20 steps using the Verlet cut-off scheme. All bonds were constrained using LINCS (50), enabling a simulation time step of 0.002 ps. The following startup protocol was applied to all modeled systems: Firstly, two following minimization processes with (i) 1000 steps under the steepest descendant algorithm and (ii) 500 steps using the conjugate gradient algorithm. Subsequently, the following two simulations were carried out to equilibrate the system at 300 K: (i) 500 ps in the canonical ensemble at 300 K using a V-rescale thermostat ($\tau = 0.1$ ps) (51) and (ii) 500 ps in the isobaric-isothermal ensemble using Berendsen (52) pressure coupling at 1 bar ($\tau = 2.0$ ps). During these processes, a restraint constant of 1000 kJ mol⁻¹ nm⁻² was applied to all heavy protein atoms and all main chain heavy atoms for the temperature and pressure equilibrations, respectively. An additional 500 ps were simulated within the isobaric-isothermal ensemble without restraints. Finally, 100 ns of molecular dynamics simulation was performed in the isobaric-isothermal ensemble using the Parrinello-Rahman barostat with no restraints (53). For all systems, the protocol was repeated three times starting from the minimized structures with a new assignment of velocities.

Protein crystallization, fragment screening, and structure determination

P protein from *T. maritima* in 50 mM Tris-HCl pH 7.5, 0.2 mM EDTA was crystallized at 3 mg/mL in 12% PEG-1000, 100 mM sodium acetate, pH 4.8, and 200 mM potassium sulfate by vapor diffusion in a sitting drop configuration in 96-well crystallization plates (Swissci™), where 1 µl of P protein plus 1 µl of mother liquor were mixed thoroughly. Crystals appeared in 3 days at 19°C with approximate dimensions of 300 × 200 × 90 µm. Due to the frequent appearance of salt crystals of similar morphology, the presence of protein crystals was checked under a CrystaLight™ 100 UV Source (Molecular Dimensions). P protein crystals were severely damaged at concentrations of DMSO as low as 5%. Thus, purpurin was solubilized in 50% PEG-400, which proved useful in dissolving this compound entirely and maintaining the integrity of P protein crystals. From this stock, a cryoprotectant solution was prepared at a final concentration of 20 mM purpurin, which additionally contained 35% PEG-1000, 100 mM sodium acetate, pH 5.2, 20 mM potassium sulfate and 1 mM DTT. Crystals were soaked for 3 months using the solution with purpurin. Crystals were flash-frozen and stored in liquid nitrogen until needed. Diffraction data were collected at 100 K at a wavelength of 0.9785 Å at the Life Sciences Collaborative Access Team (LS-CAT) 21-ID-F and 21-ID-G beamlines at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL, USA). Data were indexed and processed with XDS (54) and reduced with Aimless (55).

For fragment screening, the reference structure was solved by molecular replacement using Phaser (56), with a search model based on the structure of the P protein from *Thermotoga maritima* (PDB entry: 1NZ0). There is one monomer in the asymmetric unit, with a crystal solvent content of 40%. The presence of purpurin was initially confirmed by difference Fourier maps calculated using the structure of the wild-type enzyme.

The ligand was manually fit, and the restraints for refinement were generated using the eLBOW program of the PHENIX suite (57) using the SMILES string C1 = CC = C2C(= C1)C(= O)C3 = C(C2 = O)C(= C(C = C3O)O)O. Refinement was alternated with manual building/refinement in COOT (58), PHENIX and the PDB_REDO server (59).

The B-factor of the ligand is 42.5 versus 16.0, 25.4 and 21.1 Å² for the surrounding residues Val33 Gln28 and Arg89, respectively and the density of the ligand is evident in the 2mF_o - DF_c map at a contour level of 1.5σ. σA-weighted, 2F_o - F_c simulated annealing omit maps (Supplementary Figure S19) were used to validate the presence of the ligand further.

The model presents no Ramachandran outliers. Data collection and refinement statistics are summarized in Supplementary Table S2. Figures were prepared with Pymol (The PyMOL Molecular Graphics System, Version 1.7.4, Schrödinger, LLC). The atomic coordinates and structure factors of the RNase P protein from *T. maritima* in the complex with purpurin have been deposited in the Protein Data Bank (PDB) with the accession code 6MAX.

RESULTS

A new fluorescent probe as an RNase P substrate

In this work, we designed a synthetic oligonucleotide substrate probe with a fluorophore and a non-fluorescent quencher (NFQ) at the two termini (Figure 1A, Supplementary Table S1) based on a FRET detection mechanism. If the substrate is cleaved, physical separation of the quencher from the fluorophore allows light emission that can be detected with a fluorometer (60–62). A short RNase P substrate candidate that fulfills the assay design requirements was the minihelix (Mh) pMini3pUG (Figure 1A), whose cleavage kinetics were previously characterized (32,33) and confirmed in the present study. Substrate functionality was first tested in a polyacrylamide gel-based assay (Figure 1A) to ascertain uniform cleavage of the substrate by the RNase P holoenzyme and ensure that the appended fluorophore and quencher did not hinder catalysis. After testing the RNase P activity on the modified FRET substrate and confirming the specific cleavage (Figure 1A), we also tested the minihelix construct without the fluorescent label. The same RNase P processing pattern was observed in both cases (Figure 1A) thus supporting that the fluorophore and the quencher do not affect the ability of this substrate to be enzymatically processed by RNase P. Finally, the titration of the Mh fluorescent substrate using unlabeled pre-tRNA or Mh (Supplementary Figure S1) in solution activity assays (described below) led to a decrease in the relative fluorescence in time-course measurements.

The Mh substrate can be used in a novel, real-time in-solution RNase P fluorescence-based assay.

The in-solution activity assays were performed at 37°C in a total volume of 20 μ l on a fluorescence microplate reader in 384-well plates (Figures 1B, C, 2 and Supplementary Figures S1–S10). Supplementary Figure S2A demonstrates the stability over time and the FRET quenching efficiency of the substrate (fluorescent substrate or holoenzyme with CaCl₂ added as a negative control). Using the Mh substrate, we did not observe RNase P activity (Supplementary Figure S2A) at 10 mM MgCl₂. The kinetic analysis further demonstrated that this is indeed a viable and useful assay for monitoring RNase P activity (Figure 1B, left and Supplementary Figure S2B). A Michaelis–Menten fit was performed, and the kinetic parameters were calculated (Figure 1B, right). The obtained K_M and k_{cat} values are 1.3 μ M and 6.7 min⁻¹, respectively. The only previously reported values for this substrate are $K_D = 2.8 \mu$ M and $k_{obs} = 0.86 \text{ min}^{-1}$ using the RNase P from *E. coli* (33).

The novel fluorescence assay using Mh is reproducible, tolerant of DMSO, sensitive to known RNase P inhibitors and amenable for high-throughput screening

Before performing this assay in a large-scale setup, we first assessed its reproducibility in 10–12 reactions usually performed in a single run (Supplementary Figure S3). To evaluate the robustness of the assay, we used a Z' -factor scoring method (63) for both the end-point of the assay (Supplementary Figures S3A, C) and the initial velocities of the re-

action (Supplementary Figure S3B, D) (Z' -factor = 0.73). The linear responses in all cases were 3–5 min long (Figure 1B and Supplementary Figures S3–S10). In subsequent experiments, we considered the kinetics information given by the linear parameter to calculate the IC₅₀ values for the inhibitors described in Figures 3 and 4. Most notably, we found that DMSO does not decrease the activity of RNase P on the cleavage of the Mh substrate up to a concentration of 30% (v/v) (Supplementary Figure S4). For the assays involving inhibitors (Figure 2A and Supplementary Figures S5–S10), we used a concentration of 10% DMSO.

To benchmark our assay against known inhibitors, we tested the activity of two newly-described RNase P inhibitors, methylene blue (33) (Figure 1C, left and Supplementary Figures S5A, B) and iriginol hexaacetate (26) (Figure 1C, right and Supplementary Figures S5C, D). Notably, we detected inhibition for both compounds, and IC₅₀ values are in reasonable agreement with the published data (109 μ M versus 62 μ M (reported) for methylene blue and 29.5 μ M versus 0.8 μ M (reported) for iriginol hexaacetate). As expected, we did not detect inhibition by neomycin B or kanamycin B, as it is known that high Mg²⁺ concentrations nullify their effects on RNase P (64). These results corroborate the validity of the assay as we were able to detect inhibition in the same range as previously reported inhibitors and exclude non-specific interactors. Our chosen RNase P system and the developed assay (regarding substrate, real-time monitoring, volume, buffer composition, [E]/[S] ratio and DMSO tolerance) are a reliable platform to search for new inhibitors in a high-throughput manner.

Library screening and hit validation identifies four potential inhibitors of RNase P

We performed a screen of the Spectrum Collection library (Discovery Systems, Inc.) comprising 2560 compounds (Supplementary Figure S6). A schematic view of the overall workflow and results is presented in Supplementary Figure S22. In our setup, the screening strategy involved initial testing of cocktails with eight compounds per assay/well. To select a positive hit, we considered the inhibition of the initial velocity of the reaction based on its Z -factor robustness (Supplementary Figures S3B and D). Several cocktails exhibited initial velocity changes that were significantly decreased (Supplementary Figure S6). In general, we further analyzed the compounds that demonstrated an initial velocity slope lower than the positive control or a negative slope (Supplementary Figure S6. Red bar corresponds to the positive control). We verified the selected cocktails by a second round of selection (Supplementary Figure S7) and finally deconvoluted to identify active compounds (Supplementary Figure S8). The potential hits were further verified at this stage by the correspondence between the fluorescence-based assay and the canonical gel-based inhibition test using lower concentrations of compounds (Supplementary Figure S9).

Using this strategy, after the initial screening (Supplementary Figure S6), 37 cocktails were selected and verified again for inhibition (Supplementary Figure S7). Of these, ten cocktails with evident inhibition were retained and then deconvoluted to deduce the 10 active compounds in each

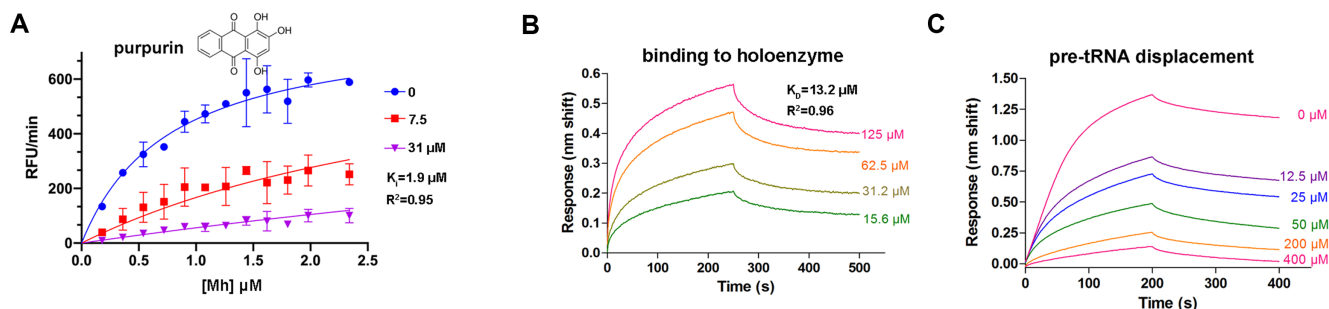


Figure 4. Characterization of the inhibition of *T. maritima* RNase P by purpurin. (A) Best global fit for inhibition of RNase P in the presence of varying concentrations of fluorescent Mh substrate. A competitive inhibition model best fit the data ($R^2 = 0.9549$) with $K_i = 1.9 \pm 0.5 \mu\text{M}$. Symbols represent the mean \pm standard error as determined from three independent experiments at each concentration. See also Supplementary Figure S18. (B) The affinity of purpurin for the RNase P holoenzyme comparable to its K_i . Binding sensorgrams obtained by biolayer interferometry (BLI). RNase P holoenzyme at 400 nM was immobilized on Ni-NTA. (C) Purpurin displaces bound pre-tRNA in a dose-response behavior. Binding sensorgrams obtained by biolayer interferometry (BLI). RNase P holoenzyme was immobilized on Ni-NTA biosensors, incubated with purpurin at varying concentrations (0–400 μM), and then the association and dissociation pre-tRNA was monitored. Of the four inhibitors found in this study, only purpurin exhibited a dose-response behavior in this experiment. See also Supplementary Figure S17.

(Supplementary Figure S8). One compound, ethidium bromide, was discarded due to its known non-specific effects on nucleic acids. We thus proceeded with 9 compounds for confirmation of inhibition at a lower concentration (Supplementary Figure S9).

A gel-based inhibition assay was performed in parallel for the nine selected compounds to qualitatively verify the results obtained by the fluorescence-based assay of Supplementary Figure S9A for the nine selected compounds (Supplementary Figure S9B) (250 and 500 μM at this stage versus 1 mM for the deconvolution stage shown in Supplementary Figure S8). In this assay, the compound iodoquinol resulted in a false positive (the inhibition does not correlate in both assays, Supplementary Figures S9A and B). Pararosaniline and emodin did not inhibit RNase P under the conditions tested in these assays (Supplementary Figure S9), nifursol is a weak inhibitor (Supplementary Figure S9), and alexidine hydrochloride seems to precipitate the holoenzyme (Supplementary Figure S9B and Supplementary Figure S8C). These results excluded five more compounds.

After this more restrictive analysis, four remaining compounds were pursued further: gentian violet (Tris(4-(dimethylamino)phenyl)methylmethyl chloride), juglone (5-hydroxy-1, 4-naphthoquinone), purpurin (1,2,4-trihydroxyanthraquinone) and hematein (3,4,6a,10-tetrahydroxy-6a,7-dihydroindeno[2,1-c]chromen-9(6H)-one). We purchased these reagents from a different provider (Sigma Aldrich) and used these compounds to generate dose-response curves (Figure 2A and Supplementary Figure S10). The range of IC_{50} values was between 4.5 and 96.2 μM (Figure 2A).

Pre-tRNA and a pre-tRNA-like substrate allow validation of the four identified hits using gel and fluorescence-based assays

First, we qualitatively validated the results obtained using the fluorescent Mh substrate by gel-based assays using a canonical pre-tRNA substrate (Supplementary Figures S11 and S12). Next, to mimic a substrate similar to the natu-

ral pre-tRNA, we designed a new probe consisting of two individual oligonucleotides (39 and 38 nucleotides for the 5' and 3' components, respectively) (Figure 2B, right, Supplementary Figure S13 and Supplementary Table S1). This probe, named bipartite pre-tRNA, differs from a canonical pre-tRNA in the sense that it lacks the anticodon loop and it is composed of two oligonucleotides (Supplementary Figure S13A). Binding experiments of pre-tRNA and bipartite pre-tRNA substrates to the holoenzyme using biolayer interferometry show that their affinity is comparable (K_D values of 1.8 and 3.5 nM, Supplementary Figure S13B). For the binding experiments, we tested the agreement of the affinity values obtained by BLI with previous K_D determinations of the RNase P holoenzyme:pre-tRNA and the P RNA:P protein pairs (65,40). These K_D values are in excellent agreement with the published data (Supplementary Figure S15A, B). Finally, titration of fluorescent bipartite pre-tRNA, with a similar design to fluorescent pre-tRNA, in time-course fluorescence emission experiments using unlabeled substrates shows highly similar profiles (Supplementary Figure S13C). Thus, we used fluorescent bipartite pre-tRNA as an analog that is highly similar to pre-tRNA, and we obtained dose-response curves using this substrate in the presence of the four hits obtained using fluorescent Mh (Figure 2B and Supplementary Figure S14). As the sensitivity to DMSO for this assay using the bipartite pre-tRNA substrate was very high (no more than 2% DMSO was tolerated, similar to (26,17), we used 10% PEG 200 instead.

Overall, the fluorescent screening strategy—using both a model substrate, tolerant to high DMSO concentrations and a pre-tRNA-like substrate, useful to demonstrate the inhibition properties under low monovalent and Mg^{2+} conditions—demonstrates its utility to find new small molecule inhibitors directed against RNase P.

Purpurin and hematein bind to the RNase P holoenzyme

Gel-based assays performed using the bipartite pre-tRNA substrate (Figure 3A) under a single buffer condition where both P RNA and the holoenzyme display activity demon-

strated that the hits purpurin and hematein inhibit the reaction catalyzed by the RNase P holoenzyme but not the P RNA alone (Figure 3A). The gel-assay using the bipartite pre-tRNA substrate allowed us to distinguish this behavior clearly. In agreement with this result, binding assays with immobilized P RNA or holoenzyme using a 5' biotinylated RNA oligonucleotide (Figure 3B) show that purpurin does not bind to P RNA and hematein binds only slightly to P RNA (Figure 3B). In contrast with the behavior of these two hits, gentian violet inhibits the reaction *via* P RNA (Figure 3A) and also binds more strongly to it (Figure 3B). Moreover, unlike the other three hits, gentian violet binds to the biotinylated minihelix substrate in the micromolar range (Supplementary Figure S16). The hit juglone displayed a less clear behavior in these assays (Figure 3A, B).

Given that the main difference between P RNA and the holoenzyme is the presence of the P protein, we performed a preliminary interrogation of potential inhibitor interactions between the P protein and each of the four hits using docking and Molecular Dynamics (MD) (Figure 3C and Supplementary Figure S18). Small molecule docking studies of gentian violet, juglone and hematein were performed on the model obtained from crystallographic data of the RNase P protein from *T. maritima* (PDB codes 1NZ0, 3Q1R). Next, Molecular Dynamics (MD) simulations were performed on the four compounds with the best-scoring docking poses or the purpurin position as observed in complex with the P protein by X-ray crystallography (Figure 5C, PDB code: 6MAX). While the dynamics preserve the overall protein structure, the binding modes of the tested compounds were affected (Figure 3C and Supplementary Figure S20). While hematein and purpurin rearranged themselves into a new binding mode around the initial docking or crystallographic pose around the P protein (Figure 3C and Supplementary Figure S20), juglone and gentian violet did not adopt a stable arrangement (Figure 3C and Supplementary Figure S20).

Overall, the activity, binding, docking and MD experiments point to gentian violet as a non-specific inhibitor that binds to RNA, juglone as an inhibitor with an ambiguous mechanism and hematein and purpurin as inhibitors of *T. maritima* RNase P holoenzyme which may specifically target the P protein.

Purpurin inhibits RNase P holoenzyme through a competitive mechanism and shows a dose-response relationship on its binding to the holoenzyme and pre-tRNA displacement

We tested whether any of the four validated hits were capable of displacing the canonical substrate pre-tRNA from the holoenzyme in a binding assay. Purpurin displays a dose-response relationship and remarkably, none of the other three hits demonstrated this behavior (Figure 4C, Supplementary Figure S17). Together, all of the data collected for purpurin are in agreement with the activity and binding criteria to validate an RNase P inhibitor (Table 1). Further characterization shows that the K_i value of purpurin (Figure 4A, Supplementary Figure S18), obtained from an adjustment to a competitive inhibition model, is in general agreement with the K_D value obtained for RNase P holoenzyme-purpurin (Figure 4B, Table 1).

Purpurin has a binding site on the P protein in a region important for substrate binding

To further explore the possible binding of purpurin to the P protein, we tested the binding interference of an RNA oligonucleotide with the sequence of a 5' leader in the presence of this compound (Figure 5A). Remarkably, purpurin interferes with the binding of the leader to the P protein (Figure 5A). The three other hits were also tested in this manner (Supplementary Figure S19). Juglone and gentian violet do not impede the leader binding (Supplementary Figure S19). We discounted hematein because, although it displayed leader-binding hindrance (Supplementary Figure S19), unlike purpurin, it lacked a dose-response relationship on pre-tRNA displacement in the presence of the holoenzyme (see Figure 4C for purpurin). We next determined that purpurin binds to the P protein with micromolar affinity in a dose-response relationship (Figure 5B). Finally, the crystal structure of *T. maritima* P protein in complex with purpurin (Figure 5C, D Supplementary Figure S21 and Table S2) shows a binding region of the compound located in a hydrophobic patch formed by residues Val 33, Leu 35 and Ile 87. Additional hydrogen-bonding interactions are observed between purpurin and Gln 28 and Arg 89. Arg 89 is a conserved protein residue, essential for efficient *T. maritima* RNase P activity, that potentially contacts the pre-tRNA leader binding region (Figure 5D) (38). Taken together, the binding and X-ray crystallography results indicate the existence of a region in the P protein that is sensitive to the binding of inhibitory compounds of *T. maritima* RNase P and that this region is also crucial for substrate binding.

DISCUSSION

In this work, we developed and implemented novel methods to discover and validate inhibitory compounds against bacterial RNase P and gain insight into the structure and function of this ribozyme. Although FRET assays have been used for high-throughput screening of RNA-based targets (66,61,62), this is the first description of this type of method tailored for RNase P. The most noteworthy differences between the proposed method compared to the recent advances described by (26) are the DMSO tolerance (provided by the use of the Mh substrate), the salt tolerance (which reduces non-specific binding) and the development of a novel bipartite pre-tRNA substrate (which allows hit validation). In particular, the DMSO tolerance of the new assay allows for additional chemical space to be probed beyond what was previously explored. It should be noted however that, given the assay design, inhibitors that alter the fidelity of processing might appear as false negatives in the fluorescence-screening step. It should also be considered that there might be a quenching effect from cleaved 5' leader that could lead to an underestimation of the catalytic activity. Despite these potential issues, the presented method offers a powerful way to discover new inhibitors under a variety of conditions and substrates.

Additionally, this work presents the first binding screening method amenable for high-throughput for RNase P. The method may be useful for enhanced characterization of known inhibitors (26) or to demonstrate specific binding to the holoenzyme, P RNA or P protein (67,68). Notably,

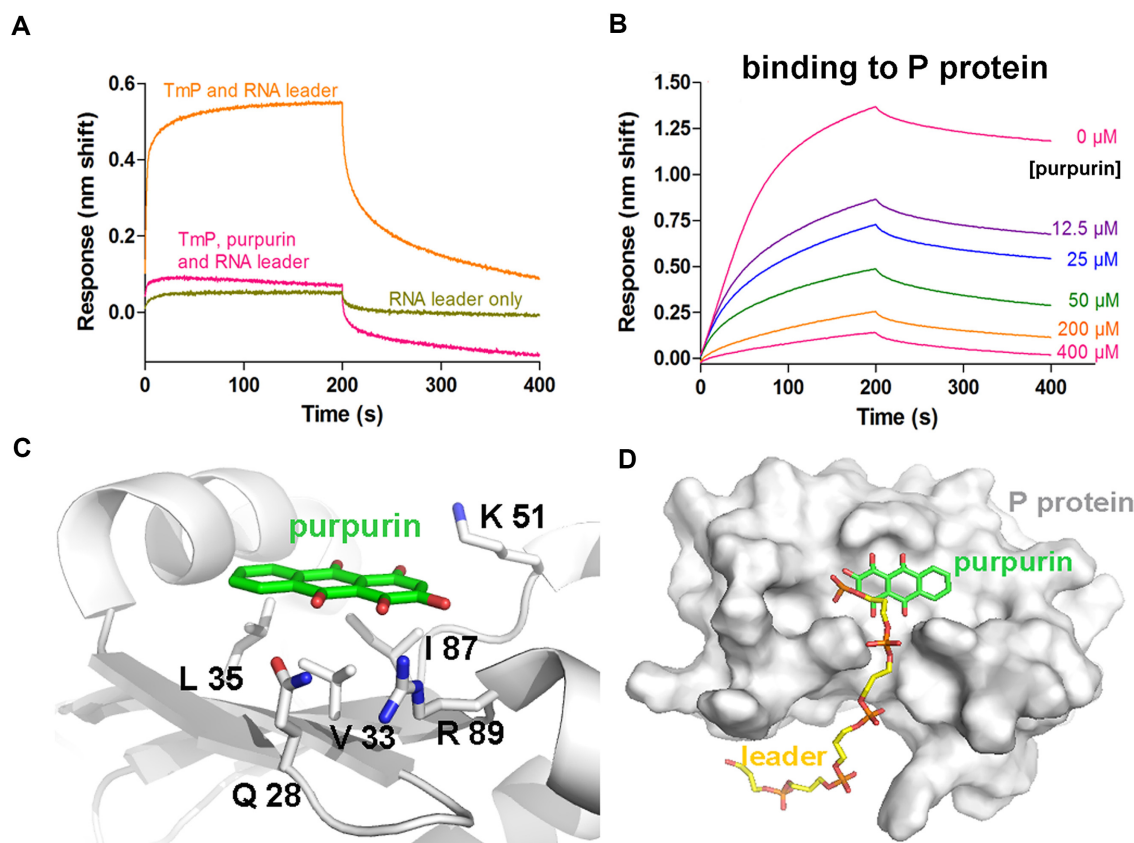


Figure 5. Purpurin binds the P protein. (A) Purpurin hinders the binding of the P protein to the 5' RNA leader. Association and dissociation of the 5' RNA leader to the P protein (TmP) is measured over time. P protein was immobilized on Ni-NTA biosensors while purpurin was assayed at 100 μM and the 5' RNA leader at 5 μM . A 10-nt leader was used for these experiments. See Supplementary Figure S10 for the experimental design and testing of the other three inhibitors. (B) Purpurin binds to the P protein from *T. maritima* in a dose-response manner ($K_D = 6.8 \mu\text{M}$). Binding sensorgrams obtained by BLI. P protein was immobilized on Ni-NTA biosensors after which its association and dissociation from indicated compounds were monitored. Sensorgrams were fit to a 1:1 Langmuir binding model. (C) Location of purpurin (green) on the *T. maritima* P protein as determined by X-ray crystallography (PDB code: 6MAX; see Supplementary Figure S19 and Supplementary Table S2). (D) Location of the 5' RNA leader (yellow) on the P protein in the context of the RNase P holoenzyme in complex with tRNA (19) (PDB code: 3Q1R). In both cases, residue Arg 89 (labeled in panel C) forms interactions with the different ligands.

Table 1. Summary of inhibition and affinity parameters for the four identified hits

Compound name	Mh IC_{50} (μM)	Bipartite pre-tRNA IC_{50} (μM)	Targets RNA	Dose-response displacement of pre-tRNA	K_i and K_d (holoenzyme) (μM)
Gentian violet	7.3 ± 1.6	16.2 ± 2.1	++ (substrate)	No	ND
Juglone	12.8 ± 1.9	11.9 ± 1.7	+	No	ND
Purpurin	96.2 ± 21.7	13.1 ± 2.3	-	Yes	1.9 ± 0.4 , 13.2 ± 2.6
Hematein	4.5 ± 0.9	87.9 ± 19.7	-	No	ND

this method provided insights into the inhibitory mode of action via monitoring the pre-tRNA or the leader binding hindrance to the holoenzyme or the P protein. Overall, the results obtained point to the P protein as a promising target for the yet undiscovered inhibitors.

We also used molecular docking, molecular dynamics, and X-ray crystallography to elucidate the possible site of action for the validated hits. P protein is an ideal target for these techniques. Due to its small size (117 aa, 14 kDa), efficient MD protocols are executed successfully. The protein crystallizes in 3 days, the crystals diffract to high resolution ($\sim 1.5 \text{ \AA}$), a structure of the holoenzyme is known, and the critical binding sites have been identified (19). Most im-

portantly, the crystallization precipitant (PEG 1000) is fully compatible with the solvent used to solubilize fragments (PEG 400 50%). However, it has been challenging to obtain complexes of the P protein with different ligands, probably due to steric hindrance in the crystal packing. Nonetheless, we succeeded in obtaining the crystal structure of the P protein in complex with purpurin, which binds in a region important for the 5'-leader interaction with the protein (38). This region includes R89, a highly conserved residue that makes putative contacts with the phosphate backbone of N-5 of the pre-tRNA leader. The mutation of this residue has substantial effects on RNase P activity (38). These re-

sults point to the P protein as an attractive component for targeted inhibition.

Table 1 shows a summary of the four validated hits and some parameters analyzed in this work. Most importantly, purpurin is the only compound that shows similar IC_{50} values using the fluorescent substrates Mh and bipartite pre-tRNA and presented a dose-response relationship on the pre-tRNA displacement assay. Moreover, the obtained K_D value to the holoenzyme and the K_i are in reasonable agreement. Holistically, this work demonstrates the necessity to consider several corroborating approaches before committing to an initial small-molecule inhibitor hit. While purpurin binds somewhat weakly, its relatively low molecular weight (256 Da), may allow its further optimization.

Three of the found hits—juglone, hematein, and purpurin—have characteristics of known non-specific inhibitors in drug screens, specifically catechol and quinone groups that are known to undergo addition reactions, complex with divalent metal ions, or act as redox cyclers (69–71). On the other hand, several lines of evidence point to purpurin as a potential inhibitor amenable to further optimization. This compound maintains inhibition even at remarkably high concentrations of $MgCl_2$ (100 mM) and ammonium acetate (400 mM), shows a dose-response relationship in binding the holoenzyme, binding P protein and displacing pre-tRNA, and it presents a specific binding site on the P protein, as revealed by the co-crystal structure.

Taken together, this work offers the following technological advances that may be more widely applied to the study of RNA-protein complexes: (i) establishment of bilayer interferometry as a real-time technique for measuring ligand interactions with all the components of RNase P (P RNA, P protein and substrate); (ii) demonstrating the utility of a novel, pre-tRNA-like substrate, which may be useful for the study of other tRNA-processing enzymes and (iii) highlighting the use of low molecular PEGs as substitutes for DMSO in the study of small ligands in complex with macromolecules. This work offers a multi-technique platform for inhibitor discovery and indicates, for the first time, that the P protein is a promising target for future antibiotics.

DATA AVAILABILITY

Atomic coordinates and structure factors for the reported crystal structure have been deposited in the Protein Data Bank under accession number 6MAX.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Jonathan Ipsaro and Alfonso Mondragón for critical reading of the manuscript. Elizabeth Nallely Cabrera-González and María Teresa Lara Ortiz are thanked for their technical support. Adela Rodríguez-Romero is acknowledged for her support in the mass determination of purpurin. Gabriel del Río (G.D.R.) is thanked for supporting us in the use of the Synergy

MX (BioTek) equipment. Chi-Hao Luan from the High-Throughput Analysis Laboratory, Department of Molecular Biosciences, Northwestern University is greatly acknowledged for his support and guidance. E.A.M.C. is a recipient of a grant from CONACyT to perform graduate studies. Part of this work was performed to fulfill the requirements of a doctoral dissertation (E.A.M.C.) in the Posgrado en Ciencias Bioquímicas, Universidad Nacional Autónoma de México.

Authors contributions: C.A.D.T., E.A.M.C., M.A.C. and A.T.L. conceived and designed the experiments; C.A.D.T., E.A.M.C., H.A.S.S., M.A.C. and A.T.L. performed the experiments; C.A.D.T., E.A.M.C., M.A.C. and A.T.L. analyzed the data; A.T.L. contributed reagents/materials/analysis tools; C.A.D.T., E.A.M.C., M.A.C. and A.T.L. wrote the paper.

FUNDING

Fundación Marcos Moshinsky, México, Consejo Nacional de Ciencia y Tecnología (CONACyT) [CB-82308, CB-252316 to G.D.R. and CB-253323, PDCPN2014-47543 to A.T.L.]; Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, Universidad Nacional Autónoma de México (PAPIIT-DGAPA) [IN202416 and IA202917, to A.T.L. and M.A.C.]; HP Cluster Platform 3000SL ‘Miztli’ [LANCAD-UNAM-DGTIC-320]; Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE [DE-AC02-06CH11357]; use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor [085P1000817]. Funding for open access charge: Instituto de Fisiología Celular, Universidad Nacional Autónoma de México.

Conflict of interest statement. None declared.

REFERENCES


- Deigan, K.E. and Ferre-D’Amare, A.R. (2011) Riboswitches: discovery of drugs that target bacterial gene-regulatory RNAs. *Acc. Chem. Res.*, **44**, 1329–1338.
- Dersch, P., Khan, M.A., Muhlen, S. and Gorke, B. (2017) Roles of regulatory RNAs for antibiotic resistance in bacteria and their potential value as novel drug targets. *Front. Microbiol.*, **8**, 803.
- Matzov, D., Bashan, A. and Yonath, A. (2017) A bright future for antibiotics? *Annu. Rev. Biochem.*, **86**, 567–583.
- McCoy, L.S., Xie, Y. and Tor, Y. (2011) Antibiotics that target protein synthesis. *Wiley Interdiscip. Rev. RNA*, **2**, 209–232.
- Kazantsev, A.V. and Pace, N.R. (2006) Bacterial RNase P: a new view of an ancient enzyme. *Nat. Rev. Microbiol.*, **4**, 729–740.
- McClain, W.H., Lai, L.B. and Gopalan, V. (2010) Trials, travails and triumphs: an account of RNA catalysis in RNase P. *J. Mol. Biol.*, **397**, 627–646.
- Mondragon, A. (2013) Structural studies of RNase P. *Annu. Rev. Biophys.*, **42**, 537–557.
- Klemm, B.P., Wu, N., Chen, Y., Liu, X., Kaitany, K.J., Howard, M.J. and Fierke, C.A. (2016) The diversity of ribonuclease P: protein and RNA catalysts with analogous biological functions. *Biomolecules*, **6**, E27.
- Brown, E.D. and Wright, G.D. (2016) Antibacterial drug discovery in the resistance era. *Nature*, **529**, 336–343.
- Hernandez-Cid, A., Aguirre-Sampieri, S., Diaz-Vilchis, A. and Torres-Larios, A. (2012) Ribonucleases P/MRP and the expanding ribonucleoprotein world. *IUBMB Life*, **64**, 521–528.

11. Willkomm,D.K., Gruegelsiepe,H., Goudinakis,O., Kretschmer-Kazemi Far,R., Bald,R., Erdmann,V.A. and Hartmann,R.K. (2003) Evaluation of bacterial RNase P RNA as a drug target. *Chembiochem*, **4**, 1041–1048.
12. Drainas,D. (2016) Antibiotics and RNase P. *Antibiotics (Basel)*, **5**, 15.
13. Esakova,O. and Krasilnikov,A.S. (2010) Of proteins and RNA: the RNase P/MRP family. *RNA*, **16**, 1725–1747.
14. Kole,R. and Altman,S. (1981) Properties of purified ribonuclease P from *Escherichia coli*. *Biochemistry*, **20**, 1902–1906.
15. Guerrier-Takada,C., Gardiner,K., Marsh,T., Pace,N. and Altman,S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849–857.
16. Altman,S. (2011) Ribonuclease P. *Philos. Trans. R Soc. Lond. B Biol. Sci.*, **366**, 2936–2941.
17. Kurz,J.C., Niranjanakumari,S. and Fierke,C.A. (1998) Protein component of *Bacillus subtilis* RNase P specifically enhances the affinity for precursor-tRNA^{Asp}. *Biochemistry*, **37**, 2393–2400.
18. Kurz,J.C. and Fierke,C.A. (2002) The affinity of magnesium binding sites in the *Bacillus subtilis* RNase P x pre-tRNA complex is enhanced by the protein subunit. *Biochemistry*, **41**, 9545–9558.
19. Reiter,N.J., Osterman,A., Torres-Larios,A., Swinger,K.K., Pan,T. and Mondragon,A. (2010) Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*, **468**, 784–789.
20. Niland,C.N., Anderson,D.R., Jankowsky,E. and Harris,M.E. (2017) The contribution of the C5 protein subunit of *Escherichia coli* ribonuclease P to specificity for precursor tRNA is modulated by proximal 5' leader sequences. *RNA*, **23**, 1502–1511.
21. Smith,D. and Pace,N.R. (1993) Multiple magnesium ions in the ribonuclease P reaction mechanism. *Biochemistry*, **32**, 5273–5281.
22. Christian,E.L., Smith,K.M., Perera,N. and Harris,M.E. (2006) The P4 metal binding site in RNase P RNA affects active site metal affinity through substrate positioning. *RNA*, **12**, 1463–1467.
23. Hsieh,J., Koutmou,K.S., Rueda,D., Koutmos,M., Walter,N.G. and Fierke,C.A. (2010) A divalent cation stabilizes the active conformation of the *B. subtilis* RNase P x pre-tRNA complex: a role for an inner-sphere metal ion in RNase P. *J. Mol. Biol.*, **400**, 38–51.
24. Pannucci,J.A., Haas,E.S., Hall,T.A., Harris,J.K. and Brown,J.W. (1999) RNase P RNAs from some Archaea are catalytically active. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 7803–7808.
25. Kikovska,E., Svard,S.G. and Kirsebom,L.A. (2007) Eukaryotic RNase P RNA mediates cleavage in the absence of protein. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 2062–2067.
26. Liu,X., Chen,Y. and Fierke,C.A. (2014) A real-time fluorescence polarization activity assay to screen for inhibitors of bacterial ribonuclease P. *Nucleic Acids Res.*, **42**, e159.
27. Willkomm,D.K., Pfeffer,P., Reuter,K., Klebe,G. and Hartmann,R.K. (2010) In: Liu,F. and Altman,S. (eds). *Ribonuclease P*. Springer, Protein Reviews, pp. 235–256.
28. Wu,S., Mao,G. and Kirsebom,L.A. (2016) Inhibition of bacterial RNase P RNA by phenothiazine Derivatives. *Biomolecules*, **6**, E38
29. Kirsebom,L.A. (2007) RNase P RNA mediated cleavage: substrate recognition and catalysis. *Biochimie*, **89**, 1183–1194.
30. Henkels,C.H., Kurz,J.C., Fierke,C.A. and Oas,T.G. (2001) Linked folding and anion binding of the *Bacillus subtilis* ribonuclease P protein. *Biochemistry*, **40**, 2777–2789.
31. Gossringer,M., Helmecke,D. and Hartmann,R.K. (2012) Characterization of RNase P RNA activity. *Methods Mol. Biol.*, **848**, 61–72.
32. Brannvall,M., Kikovska,E., Wu,S. and Kirsebom,L.A. (2007) Evidence for induced fit in bacterial RNase P RNA-mediated cleavage. *J. Mol. Biol.*, **372**, 1149–1164.
33. Wu,S., Chen,Y., Mao,G., Trobro,S., Kwiatkowski,M. and Kirsebom,L.A. (2014) Transition-state stabilization in *Escherichia coli* ribonuclease P RNA-mediated cleavage of model substrates. *Nucleic Acids Res.*, **42**, 631–642.
34. Abdiche,Y., Malashock,D., Pinkerton,A. and Pons,J. (2008) Determining kinetics and affinities of protein interactions using a parallel real-time label-free biosensor, the Octet. *Anal. Biochem.*, **377**, 209–217.
35. Masquida,B. and Westhof,E. (2011) RNase P: at last, the key finds its lock. *RNA*, **17**, 1615–1618.
36. Torres-Larios,A., Swinger,K.K., Krasilnikov,A.S., Pan,T. and Mondragon,A. (2005) Crystal structure of the RNA component of bacterial ribonuclease P. *Nature*, **437**, 584–587.
37. Paul,R., Lazarev,D. and Altman,S. (2001) Characterization of RNase P from *Thermotoga maritima*. *Nucleic Acids Res.*, **29**, 880–885.
38. Reiter,N.J., Osterman,A.K. and Mondragon,A. (2012) The bacterial ribonuclease P holoenzyme requires specific, conserved residues for efficient catalysis and substrate positioning. *Nucleic Acids Res.*, **40**, 10384–10393.
39. Krivenko,A.A., Kazantsev,A.V., Adamidi,C., Harrington,D.J. and Pace,N.R. (2002) Expression, purification, crystallization and preliminary diffraction analysis of RNase P protein from *Thermotoga maritima*. *Acta Crystallogr. D Biol. Crystallogr.*, **58**, 1234–1236.
40. Buck,A.H., Dalby,A.B., Poole,A.W., Kazantsev,A.V. and Pace,N.R. (2005) Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J.*, **24**, 3360–3368.
41. Chen,Y., Liu,X., Wu,N. and Fierke,C.A. (2017) Fluorescence-based real-time activity assays to identify RNase P inhibitors. *Methods Mol. Biol.*, **1520**, 201–225.
42. Trott,O. and Olson,A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.*, **31**, 455–461.
43. Morris,G.M., Huey,R., Lindstrom,W., Sanner,M.F., Belew,R.K., Goodsell,D.S. and Olson,A.J. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.*, **30**, 2785–2791.
44. Pronk,S., Pall,S., Schulz,R., Larsson,P., Bjelkmar,P., Apostolov,R., Shirts,M.R., Smith,J.C., Kasson,P.M., van der Spoel,D. et al. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, **29**, 845–854.
45. Lindorff-Larsen,K., Piana,S., Palmo,K., Maragakis,P., Klepeis,J.L., Dror,R.O. and Shaw,D.E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, **78**, 1950–1958.
46. Sousa da Silva,A.W. and Vranken,W.F. (2012) ACPYPE - AnteChamber PYthon Parser interface. *BMC Res. Notes*, **5**, 367.
47. Chen,V.B., Arendall,W.B. 3rd, Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
48. Jorgensen,W., Chandrasekhar,J. and Madura,J. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
49. Darden,T., York,D. and Pedersen,L. (1993) Particle mesh Ewald - an N,LOG(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089.
50. Hess,B., Bekker,H., Berendsen,H.J.C. and Fraaije,J. (1997) LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, **18**, 1463–1472.
51. Bussi,G., Donadio,D. and Parrinello,M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.*, **126**, 14101.
52. Berendsen,H.J.C., Postma,J.P.M., Vangunsteren,W.F., Dinola,A. and Haak,J.R. (1984) Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684–3690.
53. Parrinello,M. and Rahman,A. (1981) Polymorphic transitions in single-crystals—a new molecular dynamics method. *J. Appl. Phys.*, **52**, 7182–7190.
54. Kabsch,W. (2010) Xds. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
55. Evans,P.R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.*, **67**, 282–292.
56. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
57. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
58. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.

59. Joosten,R.P., Long,F., Murshudov,G.N. and Perrakis,A. (2014) The PDB_REDO server for macromolecular structure model optimization. *IUCrJ*, **1**, 213–220.
60. Thakur,C.S., Xu,Z., Wang,Z., Novince,Z. and Silverman,R.H. (2005) A convenient and sensitive fluorescence resonance energy transfer assay for RNase L and 2',5' oligoadenylates. *Methods Mol. Med.*, **116**, 103–113.
61. Wiseman,R.L., Zhang,Y., Lee,K.P., Harding,H.P., Haynes,C.M., Price,J., Sicheri,F. and Ron,D. (2010) Flavonol activation defines an unanticipated ligand-binding site in the kinase-RNase domain of IRE1. *Mol. Cell*, **38**, 291–304.
62. Di Grandi,M., Olson,M., Prashad,A.S., Beberitz,G., Luckay,A., Mullen,S., Hu,Y., Krishnamurthy,G., Pitts,K. and O'Connell,J. (2010) Small molecule inhibitors of HIV RT Ribonuclease H. *Bioorg Med. Chem. Lett.*, **20**, 398–402.
63. Zhang,J.H., Chung,T.D. and Oldenburg,K.R. (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.*, **4**, 67–73.
64. Mikkelsen,N.E., Brannvall,M., Virtanen,A. and Kirsebom,L.A. (1999) Inhibition of RNase P RNA cleavage by aminoglycosides. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 6155–6160.
65. Talbot,S.J. and Altman,S. (1994) Gel retardation analysis of the interaction between C5 protein and M1 RNA in the formation of the ribonuclease P holoenzyme from *Escherichia coli*. *Biochemistry*, **33**, 1399–1405.
66. Klostermeier,D. and Millar,D.P. (2001) RNA conformation and folding studied with fluorescence resonance energy transfer. *Methods*, **23**, 240–254.
67. Olson,P.D., Kuechenmeister,L.J., Anderson,K.L., Daily,S., Beenken,K.E., Roux,C.M., Reniere,M.L., Lewis,T.L., Weiss,W.J., Pulse,M *et al.* (2011) Small molecule inhibitors of *Staphylococcus aureus* RnpA alter cellular mRNA turnover, exhibit antimicrobial activity, and attenuate pathogenesis. *PLoS Pathog.*, **7**, e1001287.
68. Lounsbury,N., Eidem,T., Colquhoun,J., Mateo,G., Abou-Gharbia,M., Dunman,P.M. and Wayne,E.C. (2018) Novel inhibitors of *Staphylococcus aureus* RnpA that synergize with mupirocin. *Bioorg. Med. Chem. Lett.*, **28**, 1127–1131.
69. Baell,J. and Walters,M.A. (2014) Chemistry: Chemical con artists foil drug discovery. *Nature*, **513**, 481–483.
70. Dahlin,J.L., Nissink,J.W., Strasser,J.M., Francis,S., Higgins,L., Zhou,H., Zhang,Z. and Walters,M.A. (2015) PAINS in the assay: chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *J. Med. Chem.*, **58**, 2091–2113.
71. Dahlin,J.L. and Walters,M.A. (2016) How to triage PAINS-full research. *Assay Drug Dev. Technol.*, **14**, 168–174.

Anexo B. Artículo Proteins 2018 (coautor).

Mimicking a p53-MDM2 interaction based on a stable immunoglobulin-like domain scaffold

Pedro Jimenez-Sandoval¹ | Ezequiel A. Madrigal-Carrillo² |
 Hugo A. Santamaría-Suárez² | Daniel Maturana³ | Itzel Rentería-González¹ |
 Claudia G. Benitez-Cardoza⁴ | Alfredo Torres-Larios² | Luis G. Brieba¹ 

¹Laboratorio Nacional de Genómica para la Biodiversidad, Centro de Investigación y de Estudios Avanzados del IPN, Apartado Postal 629, Irapuato, Guanajuato, CP 36821, México

²Departamento de Bioquímica y Biología Estructural, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, Apartado Postal 70-243, Mexico City 04510, México

³NanoTemper Technologies GmbH, Floessergasse 4, Munich 81369, Germany

⁴Laboratorio de Investigación Bioquímica, Programa Institucional en Biomedicina Molecular ENMyH-Instituto Politécnico Nacional, Guillermo Massieu Helguera No. 239, La Escalera Ticoman, D.F, Mexico City 07320, Mexico

Correspondence

Luis G. Brieba, Km. 9.6 Libramiento Norte Carretera. Irapuato-León, 36821 Irapuato Guanajuato, México.
 Email: luis.brieba@cinvestav.mx

Funding information

Consejo Nacional de Ciencia y Tecnología, Fronteras de la ciencias #11

Abstract

Antibodies recognize protein targets with great affinity and specificity. However, posttranslational modifications and the presence of intrinsic disulfide-bonds pose difficulties for their industrial use. The immunoglobulin fold is one of the most ubiquitous folds in nature and it is found in many proteins besides antibodies. An example of a protein family with an immunoglobulin-like fold is the Cysteine Protease Inhibitors (ICP) family I42 of the MEROPs database for protease and protease inhibitors. Members of this protein family are thermostable and do not present internal disulfide bonds. Crystal structures of several ICPs indicate that they resemble the Ig-like domain of the human T cell co-receptor CD8 α . As ICPs present 2 flexible recognition loops that vary accordingly to their targeted protease, we hypothesize that members of this protein family would be ideal to design peptide aptamers that mimic protein-protein interactions. Herein, we use an ICP variant from *Entamoeba histolytica* (EhICP1) to mimic the interaction between p53 and MDM2. We found that a 13 amino-acid peptide derived from p53 can be introduced in 2 variable loops (DE, FG) but not the third (BC). Chimeric EhICP1-p53 form a stable complex with MDM2 at a micromolar range. Crystal structure of the EhICP1-p53(FG)-loop variant in complex with MDM2 reveals a swapping subdomain between 2 chimeric molecules, however, the p53 peptide interacts with MDM2 as in previous crystal structures. The structural details of the EhICP1-p53(FG) interaction with MDM2 resemble the interaction between an antibody and MDM2.

KEYWORDS

crystal structure, inhibitor cysteine proteases, p53-MDM2 interaction, peptide aptamer, peptide mimetics

1 | INTRODUCTION

Protein-protein interactions (PPIs) are pivotal for most cellular processes and the ability to block or promote PPIs is inherent to medical therapies and biotechnological applications. One approach of promoting PPIs is by mimicking or “copying” the recognition patterns or binding elements in a molecular system. Although proteins are marginally stable, they can tolerate multiple amino acid substitutions and maintain their native structure.¹ In a seminal work Colas and coworkers designed new loops in *Escherichia coli*'s thioredoxin to “mimic the recognition function of the complementarity-determining regions of immunoglobulins” and dubbed these engineered proteins as

peptide aptamers (PAs). PAs are a class of recognition molecules based on a protein scaffold that constrains and displays a peptide sequence, which in turns binds to the target element.² The inserted sequence is constrained at its amino and carboxy termini, protecting the aptamer from proteolysis and promoting an increase in binding affinity by reducing conformational freedom. PAs resemble antibodies, in which a peptide sequence is constrained within a rigid protein scaffold. Antibodies are naturally engineered proteins, historically used as high-specificity reagents for molecular recognition applications. However, antibodies are thermo sensitive large multi-chain molecules of >100 kDa that require disulfide cross-links for proper folding.³

Although antibodies cannot be easily purified from bacteria,⁴ expression and purification of recombinant antibodies in bacteria has been achieved by the use of *in vitro* selection methods and manipulation of bacterial strains.^{5,6} During the last years, through protein engineering techniques, several proteins have been used as scaffolds to mimic PPIs. These proteins include the fibronectin type III domain, the A domain (affibody), the PDZ and SH3 domains, the ankyrin repeat domain (darpins), thioredoxin, green fluorescent protein, and T-cell receptors, among others.^{2,7-12} Numerous studies use diverse proteins as PAs to introduce peptide sequences typically from 5 to 20 amino acids.^{13,14} The desired characteristics of a PA scaffold are: tolerant to amino acid substitutions, thermostability, abundant expression in bacteria, and lack of toxicity.

The Ig-fold is one of the most common structural motifs in nature.^{15,16} This fold consists of 2 β sheets formed by 7 or 8 antiparallel β strands.¹⁷ Although the large number of proteins used as PA scaffolds,^{2,7-12} just few examples of proteins with an Ig-fold as PA exist. Domains with an Ig-like fold are found in many proteins that are not antibodies. For instance, the outer membrane protein of the genus *Leptospira* (LigB),¹⁷ type III domain of human fibronectin (FN3 domain),¹⁸ and inhibitors of cysteine proteases (ICPs) from the MEROPS family I42¹⁹⁻²³ among many others. The most remarkable example of a PA with an Ig-fold is the FN3 domain. This domain has been employed to develop PAs that bind to a variety of targets like integrins, lysozyme, Src SH3, and the ARVCF peptide.^{9,24-28}

Crystal structures of diverse ICPs indicate that members of this protein family contain structural features advantageous for an PA. These proteins are thermostable, fold into an Ig-like domain and contain 3 loops exposed to the solvent that are determinant in protease recognition.^{21,22,29,30} Chagasin is the archetypical member of the ICP family and this protein has been crystalized alone or in complex with several cysteine proteases.^{23-25,34} The genome of the protozoan parasite *Entamoeba histolytica* encodes for 2 ICPs (EhICP1 and EhICP2) and the crystal structure of 1 of them has been solved.^{16,35} ICPs also bear resemblance to the structure of the Ig-like domain of the human T cell co-receptor CD8 α that can recognize a broad range of peptides.^{22,31,32} Based on its structural properties we postulate that ICPs would be a proper scaffold to be used as a PA.^{19,21}

2 | MATERIALS AND METHODS

2.1 | Homology model, rational design, and *in silico* folding prediction

The final homology model for EhICP1 was refined from 25 intermediate models under the default parameters of the MOE homology modeling protocol using the CHARMM27 force field for energy minimization. To mimic the interface between p53 and MDM2, we selected a synthetic peptide that resembles a part of the transactivation domain (α -helical) of p53 which is target of MDM2, and is able to form a MDM2-complex with very high affinity.³³ This sequence (N-TSFAEYWNLLSP) was independently grafted in each EhICP1 loop to generate 3 possible binding chimeras against MDM2. A simple

structural analysis was carried out to select the residues that could be mutated without altering the folding. Briefly, It was taken into account the hydrophobic core (shared among many others Ig-like domains³⁴) and the proximity of the surrounding residues that could be forming hydrogen bonds, salt bridges and other interactions in a radius of 6 Å, where van der Waals forces become relatively weak. Predicted structures of the EhICP1 variants (BC, DE and FG) were generated based of the standard I-TASSER pipeline.

2.2 | Cloning, expression, and purification

DNA fragments containing the open reading frames of p53 binding domain of MDM2 (residues 1-150), EhICP1-p53(BC), EhICP1-p53(DE), and EhICP1-p53(FG) were synthesized as gBlocks (IDT, Coralville, Iowa). gBlocks were cloned into NdeI and BamHI restriction sites of a modified pET19b vector (Courtesy of Professor Tom Ellenberger, Washington University School of Medicine). Positive clones were confirmed by Sanger DNA sequencing. Positive pET19b-MDM2, pET19b-EhICP1-p53(BC), DE, and FG plasmids were transformed in *E. coli* BL21 (DE3) separately. For each construction, 1 L of bacterial culture was grown at 37°C in LB medium containing 100 μ g/mL of ampicillin. The proteins were induced at an OD₆₀₀ = 0.7 by adding IPTG to a final concentration of 0.5 mM, and the bacterial culture was grown overnight at 16°C. Proteins were purified by immobilized-metal affinity chromatography (IMAC) using pre-packed Ni-Sepharose columns. Cell pellet was resuspended in 30 mL of 25 mM Tris pH 8.0 and 150 mM NaCl (lysis buffer). The resuspended cell culture was sonicated and centrifuged at 13 000 rpm for 45 min at 4°C. The clarified lysate was loaded into equilibrated Ni-Sepharose columns, after extensive washing with lysis buffer (containing 50 mM imidazole) proteins were eluted with elution buffer (25 mM Tris pH 8.0, 150 mM NaCl, and 500 mM imidazole). The N-terminal histidine tag was removed with PreScission Protease (GE Healthcare Life Sciences). EhICP1 mutants and MDM2 were further purified by gel filtration chromatography on a Superdex 75 10/300 GL column (GE Healthcare Life Sciences) in 25 mM Tris pH 8.0 and 150 mM NaCl.

2.3 | Determination and isolation of the protein complex

Purified EhICP1 mutants were mixed (individually) with a molar excess of purified MDM2 and incubated for 1 h at 4°C with gentle shaking. Protein complexes were determined by gel filtration on a Superdex 75 10/300 GL column (GE Healthcare Life Sciences, Tris buffer pH 8.0, 150 mM NaCl), fractions containing the protein complex between variants of EhICP1 and MDM2 were collected and analyzed by SDS-PAGE. Molecular mass of the complex was calculated based on the peaks and a molecular mass standard.

2.4 | Complex crystallization and data collection

Crystal screening was performed using the hanging-drop vapor diffusion technique by mixing equal volumes of pure protein complex (EhICP1-p53(FG)-MDM2) at 5.0 mg/mL with precipitants from a sparse

matrix screen. Crystals suitable for X-ray diffraction analysis appeared under 2 crystallization conditions. Condition 1 consisted of 0.1 M NaCl, 0.1 M Hepes pH 7.5, 1.6 M $(\text{NH}_4)_2\text{SO}_4$ whereas condition 2 consisted of 1.5 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M Tris pH 8.5, 12% v/v glycerol. Crystals were soaked in mother liquor supplemented with 20% w/v of glycerol prior to flash cooling in liquid nitrogen. A single crystal from both conditions were used to collect diffraction data at 100 K using a MarMosaic 225 CCD detector at APS beam line LS-CAT 21-ID-F with 0.5° rotation per image.

2.5 | Determination of dissociation constants

The determination of the dissociation constant was carried using microscale thermophoresis (MST) and bio-layer interferometry. For MST analysis, MDM2, EhICP1, and mutants were purified by IMAC, digested with PPS followed by gel filtration. MST was carried out at room temperature (Monolith NT.115 Nano Temper) with 1.5×10^{-7} M of EhICP1 proteins as targets and with 2.67×10^{-9} M to 8.75×10^{-5} M of labeled MDM2 (Monolith Protein Labeling Kit RED-NHS (Amine Reactive), MO-L001) as a ligand in PBS buffer pH 7.4.

For bio-layer interferometry, binding experiments were made on an Octet Red96 System (ForteBio) at 30°C in PBS Buffer pH 7.4, Triton X-100 0.05% (v/v) and 5 mM TCEP. First, MDM2 was biotinylated using biotinamido hexanoic acid N-hydroxysuccinimide ester (Sigma, B2643) according with the manufacturer's protocol and dialyzed to remove biotin excess. The biotinylated protein MDM2 was immobilized on Streptavidin Biosensors. Samples were agitated at 1000 rpm. SA biosensors were saturated using 4.7×10^{-7} M of biotinylated MDM2, for 10 min. EhICP1 variant was prepared at 4.16, 8.31, 12.5, 16.6, 20.8, and 24.9×10^{-6} M in triplicate. Association was monitored for 500 s and dissociation was followed for 400 s into buffer alone. The light shift distance or spectral shift ($\Delta\lambda$), which is directly correlated with the change in the thickness (nm) of the biological layer on the SA biosensor tip, is plotted against time (seconds). The association and dissociation responses were baseline corrected. The individual signal responses at each concentration as well as the maximum responses (R_{max}) were calculated as an average of 3 independent measurements. Interferometry data were globally fitted to a 2:1 heterogeneous ligand model (2:1 HL fit) (Octet Software Version 8.2, ForteBio).

2.6 | Determination of thermal stability

Thermal unfolding of EhICP1 mutants was measured in 20 mM Tris pH.80, by monitoring changes in ellipticity at 218 and 232 nm. Protein samples were subject to a gradual increase in temperature (heating rate of 1.0 K min^{-1}) while the protein concentration was constant. Changes in the CD signal were analyzed after normalization of the transition curves to the fraction of unfolded molecules.

Thermal stability assay was carried out by differential scanning fluorimetry, on a Prometheus NT.48 (NanoTemper Technologies). Every sample was measured by triplicate and subjected to a temperature gradient, $20\text{--}95^\circ\text{C}$ with $1.5^\circ\text{C min}^{-1}$. Data analysis was performed using NT Melting Control software. The melting temperature (T_m) was

determined by fitting experimental data using a polynomial function, in which the maximum slope is indicated by the peak of its first derivative.

2.7 | Structure determination, refinement, and analysis

Diffraction intensities were integrated and scaled with tetragonal P4_32_12 lattice symmetry using MOSFLM and SCALA.³⁵ A molecular replacement solution was found with Phaser-MR³⁶ using as a search templates for molecular replacement an alanine-truncated model of a single monomer from the crystal structure of the EhICP2 (PDB ID: 3M86²³) and a model of the human Mdm2-p53 binding domain from the structure PDB ID: 3EQS,³³ using multiple ensembles search as strategy. Data were analyzed with phenix.xtriage to detect outliers, twinning and other conditions.³⁷ Crystallographic refinement was performed with phenix.refine.³⁸ A total of 5.14% of the reflections were used as the test sets for the complex. Simulated annealing, individual atomic coordinate and individual atom isotropic displacement parameter refinement strategies were performed with non-crystallographic symmetry (NCS) restraints. The model was adjusted manually to improve the fit to likelihood weighted electron density maps using Coot.³⁹ Solvent molecules were added where supported by both chemistry and geometry, and difference electron density was $>3\sigma$. The quality and stereochemistry of the model were evaluated using Coot validation tools, Molprobit and the wwPDB Validation Server.⁴⁰

3 | RESULTS

3.1 | ICPs present 2 variable and 1 conserved solvent exposed loops

A structural comparison between chagasin and EhICP2 highlights that ICPs adopt an Ig-fold with 2 β -sheets and 3 solvent exposed loops dubbed BC, DE and FG. In ICPs, loop BC contains the conserved NPTTG motif and folds into a 3_{10} helix, whereas loops DE and FG are flexible and vary in amino acid sequence (Figure 1A^{21-23,29}). Loops DE and FG are solvent exposed and their structures vary among ICPs, suggesting that they would be able to interchange with other amino acids sequences to mimic PPIs.

3.2 | A cysteine-less mutant of EhICP1 is more stable than EhICP2 and chagasin

The amino acid sequences of EhICP2 and chagasin do not contain a single cysteine, whereas the amino acid sequence for EhICP1 contains 3 cysteines (Figure 1B). EhICP1 purification in the absence of reducing agents yield high molecular weight aggregates that may reflect intermolecular disulfide-bond formation (data not shown). Because of this observation we decided to work with a triple cysteine to serine mutant that eliminates all EhICP1's cysteines (C39S, C58S, and C60S). Herein-after we refer to this triple mutant as EhICP1. Peptide substitutions in PAs decrease their thermal stability,⁴¹ thus a thermostable scaffold will increase the possibilities to have a stably folded chimera. To test the feasibility of their use as PAs, we measured the thermal denaturation

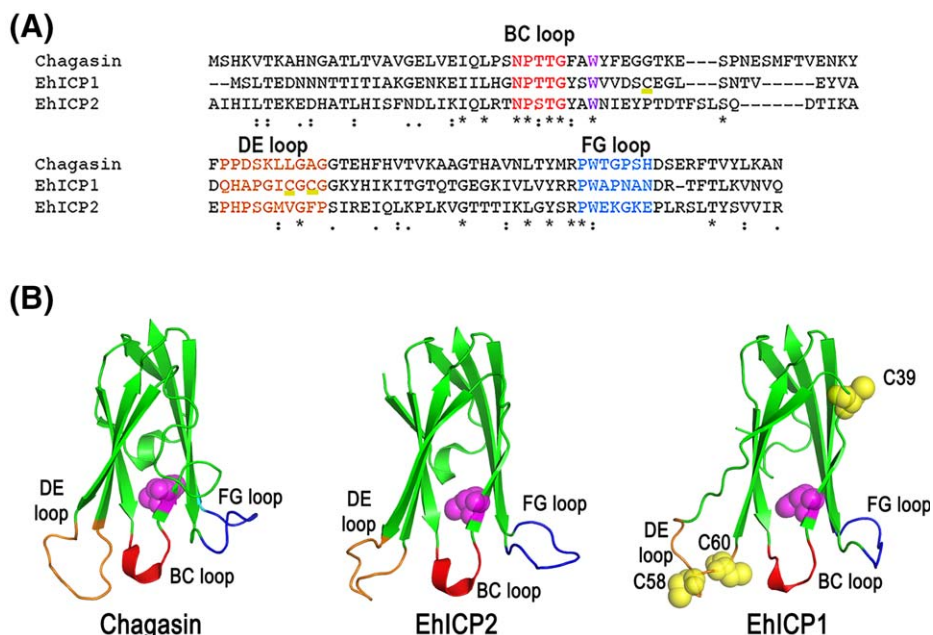


FIGURE 1 Structural features of ICPs as PA. A, Amino acid sequence alignment of EhICP1, EhICP2, and chagasin. The alignment highlights the amino acid conservation and the BC loop and the divergence at DE and FG loops. The conserved Trp residue at the hydrophobic core is colored in magenta and the 3 cysteines present in EhICP1 are highlighted in yellow. B, Structural comparison between the crystal structures of chagasin and EhICP2 and a homology model of EhICP1. The BC, DE, and FE loops of are colored in red, orange, and blue, respectively. The regions of major structural difference between both proteins are at the FG and DE loops. In contrast to the presence of disulfide bonds that stabilize Igs, ICPs harbor a conserved Trp residue (magenta) at their hydrophobic core. The cysteine residues present in EhICP1 are colored in yellow in a sphere representation [Color figure can be viewed at wileyonlinelibrary.com]

of 3 of ICPs: chagasin, EhICP1 and EhICP2. We measured thermal denaturation using circular dichroism (CD) at 218 and 232 nm in order to determine the contribution of the peptidic bond and aromatic amino acids to the overall structure (Table 1). The thermal denaturation profile of ICPs, shown that their melting temperatures (T_m) rank between 69.9 and 72.9°C when measured at 218 nm and rank between 81.74 and 83.32 nm when measured at 232 nm (Table 1). Although the melting temperatures of the 3 inhibitors are similar, we choose as a scaffold the triple EhICP1 mutant because it presents a slightly higher T_m and because this protein is reversible upon chemical and thermal denaturation (data not shown). The T_m of ICPs is similar to the T_m of type III domain of human fibronectin (Fn3) (84°C) that also folds as Ig-domain and has been successfully employed as an AP.^{28,42}

TABLE 1 Thermal stability of ICPs measured at 218 and 232 nm

Protein	Wavelength (nm)	T_m ap (°C)
EhICP2	218	71.2 ± 0.18
EhICP1-cys	218	72.9 ± 0.21
Chagasin	218	69.9 ± 0.16
EhICP2	232	81.74 ± 0.17
EhICP1-cys	232	83.34 ± 0.12
Chagasin	232	82.34 ± 0.29

*Melting curves were obtained by monitoring the circular dichroism ellipticity at 218 and 232 nm. T_m represents the change in the midpoint of the thermal denaturation curve.

3.3 | EhICP1 presents exposed loops likely to be substituted

As a model system, we decided to study the recognition interface between the tumor suppressor p53 and its principal antagonist, the human E3 ubiquitin-protein ligase MDM2. The latter because of the wealth of structural information and the relevance of this interaction.^{33,43} To mimic the interface between p53 and MDM2, we selected a modified peptide that corresponds to the α -helical transactivation domain of p53.³³ The N-TSFAEYWNLLSP peptide binds to MDM2 at a nanomolar range. This binding affinity is approximately 100 times more tighter than the affinity present for the canonical p53 peptide sequence (N-ETFSDLWKLLPE).³³ We aimed to construct 3 PAs containing a modified peptide mimicking the interacting α -helix of p53 (N-TSFAEYWNLLSP) by independently replacing each of the 3 solvent accessible loops of EhICP1. Thus, we sought to construct 3 EhICP1-p53 chimeras: EhICP1-p53(BC), EhICP1-p53(DE), and EhICP1-p53(FG).

To select for positions to insert the p53 peptide in EhICP1, we generated a homology model of this inhibitor. With basis on this structural model, we selected the boundaries for amino acid replacement (Figure 2). From this structural model, we selected 1 position for each of 3 accessible loops, in which the p53 peptide could be introduced (Figure 2). The resulting insertion places were the residues with the highest B-factors in the structure of the analogous EhICP2, as high B-factors often correlate which regions of high flexibility.²³ The region selected for amino acid insertion at the DE and FG loops correlate with the regions responsible for the hinge motion in those loops observed

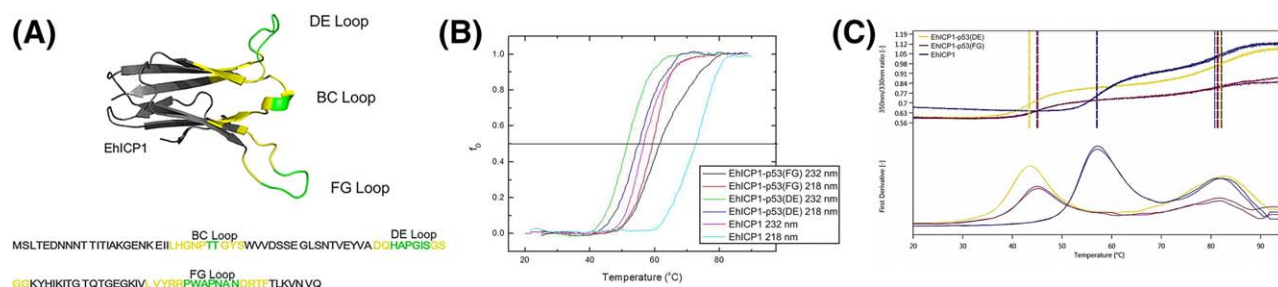


FIGURE 2 Structural analysis of the EhICP1 as a putative PA. A, Structural model of EhICP1 showing the suggested insertion sites for the p53 peptide in the BC, DE, and FG loops. Residues near to the conserved hydrophobic core are colored in yellow, whereas residues that do not stabilize interactions and are likely to be mutated are shown in green. Residues colored in green for each loop (BC (29–30), DE (53–59), and FD (85–91)) were replaced with the p53 (N-TSFAEYWNLSP) peptide. B, Normalized fraction of unfolding for EhICP1, EhICP1-p53(DE), and EhICP1-p53(FG) upon temperature increase. CD measurements were taken at 218 and 232 nm. C, Thermal unfolding measured by Differential Scanning Fluorimetry of EhICP1 and EhICP1-p53 chimeras upon temperature increase [Color figure can be viewed at wileyonlinelibrary.com]

between the 2 crystal structures of EhICP2.²³ To mimic the p53/MDM2 interactions, a p53 peptide of 13 amino acids was introduced instead of the corresponding sequence of EhICP1 colored in green (Figure 2A). We were curious to model the chimeric proteins using the I-TASSER server.⁴⁴ The chimeric proteins in which the p53 peptide was inserted in the BC, DE, and FG-loops presented C-scores (confidence score) near 0.5 suggesting the probability that the chimeras would fold with an Ig-like domain (Supporting Information Figure S1 and Table 2).

3.4 | EhICP1 tolerates substitutions in loops DE and FG but not in BC

Recombinant proteins carrying the p53 peptide at the DE and FG-loops were soluble in bacteria, while expression of the chimera in the BC-loop resulted in the formation of inclusion bodies (data not shown). Thus, the introduction of the p53 peptide in the BC-loop of EhICP1 creates a distortion in the folding of the enzyme that promotes protein aggregation. Thermal unfolding of the EhICP1-p53(DE) and EhICP1-p53(FG) chimeras were analyzed by changes in tryptophan fluorescence by circular dichroism (CD) and Differential Scanning Fluorimetry (DSF). In the CD experiment, the thermal stability of EhICP1-p53(DE) and EhICP1-p53(FG) chimeras was measured at 218 and 232 nm. The melting temperatures at 218 nm were of 49 and 59°C for EhICP1-p53 (DE) and EhICP1-p53(FG), respectively, whereas the melting temperatures at 232 nm were 51 and 60°C (Figure 2B). Thus, the introduction

of a p53 derived peptide in place of the canonical sequence at the DE and FG loops of EhICP1 decrease the T_m by approximately 12°C in the case of EhICP1-p53(FG) and 14°C for EhICP1-p53(DE). The thermal stability of the chimeras, compared to EhICP1, measured by CD present similar changes as assessed by DSF. Thermal unfolding by DSF results in the presence of 2 peaks that are associated with 2 T_m 's. The first transition peak by DSF may correspond to the Trp residue in the p53 peptide, whereas the second transition may correspond to the contribution of the hydrophobic core Trp that indicates total protein unfolding. In the thermal unfolding of the 3 proteins, EhICP1 present a T_{m1} of 56.50°C and a T_{m2} of 83°C. The p53 chimera at the DE-loops presented T_{m1} of 43°C and a T_{m2} of 82.7°C, whereas the FG chimera presented a T_{m1} of 45°C and a T_{m2} of 82.4°C (Figure 2C and Table 2). Thus, the data obtained by DSF indicates that introduction of a p53 peptide in the DE and FG loops alters protein stability by 12 and 11°C, respectively. An analysis of the Gibbs free energy of unfolding for EhICP1-p53(DE) and EhICP1-p53(FG) using the data obtained by DSF⁴⁵ indicates that the introduction of the p53 peptide reduces the stability by 23.85 and 19.62 kJ mol⁻¹, respectively (Table 2).

3.5 | DE and FG-loop constructs interact with MDM2

As the modified α -helical transactivation domain of p53 binds to MDM2 with an affinity constant in the nM range, we were curious to test if EhICP1 chimeras carrying the p53 modified peptide would be able to establish an interaction with MDM2. Each EhICP1-p53 chimera

TABLE 2 Correlation between the confidence of the structural model with the measured melting points and changes of Gibbs free-energy for the EhICP1 construct and the chimeric constructs

Protein	EhICP1-CXS	EhICP1-p53(BC)	EhICP1-p53(DE)	EhICP1-p53(FG)
C-score	0.54	0.40	0.65	0.67
TM-score	0.79 ± 0.09	0.77 ± 0.10	0.80 ± 0.09	0.80 ± 0.09
T_m , °C	56.50 ± 0.00	No data	43.65 ± 0.07	45.10 ± 0.00
ΔG_{N-U} , kJ mol ⁻¹	42.47 ± 0.65	No data	18.62 ± 0.34	22.85 ± 0.33

ΔG_{N-U} is the Gibbs free energy of unfolding. Reported values represent the average of 3 replicates, and the uncertainty is reported as 1 standard deviation. Calculated based on the methodology described by Wright et al.⁴⁵

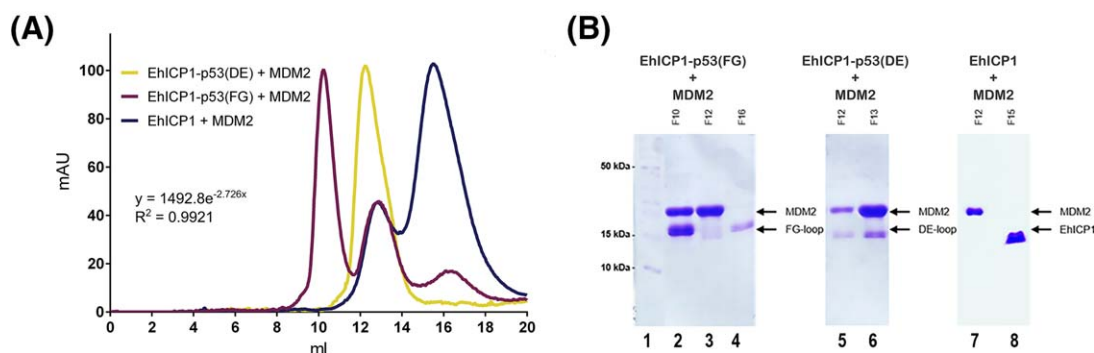


FIGURE 3 EhICP1-p53 interaction with MDM2 assessed by gel filtration. A, Protein complex containing EhICP1 and chimeras of EhICP1-p53 were incubated with MDM2 and loaded into a S75 gel filtration column. The blue line corresponds to the control, EhICP1 incubated with MDM2. The 2 discrete protein peaks that migrate as proteins of 15 and 36 kDa correspond to EhICP1 and dimeric MDM2, respectively. The chromatogram of EhICP1-p53(FG) incubated with MDM2 sample (magenta line) present a peak of high molecular weight (~ 67 kDa), whereas EhICP1-p53(DE) incubated with MDM2 (yellow line) shows a peak of approximately 36 kDa. B, SDS-PAGE of representative fractions eluted in gel filtration. Lane 1 indicates the molecular weight standard. Lanes 7 and 8 correspond to the 2 peaks obtaining by incubating EhICP1 and MDM2. Lane 2 correspond to the 67 kDa peak of EhICP1p53(FG) incubated with MDM2, lane 3 corresponds to MDM2 and lane 4 corresponds to EhICP1-p53(FG). Lanes 5 and 6 corresponds to the fraction incubated with EhICP1-p53(DE) and MDM2. The eluted fraction is labeled in top of the SDS-PAGE [Color figure can be viewed at wileyonlinelibrary.com]

was mixed with recombinant MDM2 and incubated. By passing the incubated proteins throughout a gel filtration column we observed that EhICP1 is not able to interact with MDM2 with a sufficient strength, as we observed 2 protein peaks in the elution profile. One peak corresponds to a protein with a molecular mass of EhICP1 (15 kDa) and the other peak corresponds to a protein of the molecular mass of a dimer of MDM2 (35 kDa) (Figure 3A (blue trace) and Figure 3B, lanes 7 and 8). On the other hand, the incubation of EhICP1-p53(DE) with MDM2 results in the appearance of a protein peak of approximately 36 kDa. This protein peak is consistent with the formation of a 1:1 complex between MDM2 and EhICP1-p53(DE) (Figure 3B, lanes 5 and 6).

Incubation of MDM2 with EhICP1-p53(FG) resulted in the appearance of a protein peak that elutes at a volume that corresponds to approximately 67 kDa. This may indicate the presence of heterodimers in a 2:2 ratio (Figure 3, lanes 2 to 4).

3.6 | The designed FG-loop shows a near K_D to that of p53/MDM2

We investigated if EhICP1 constructs could mimic the N-terminal transactivation domain of p53 by quantifying their interaction with MDM2. As a first approach, we measured the dissociation constant

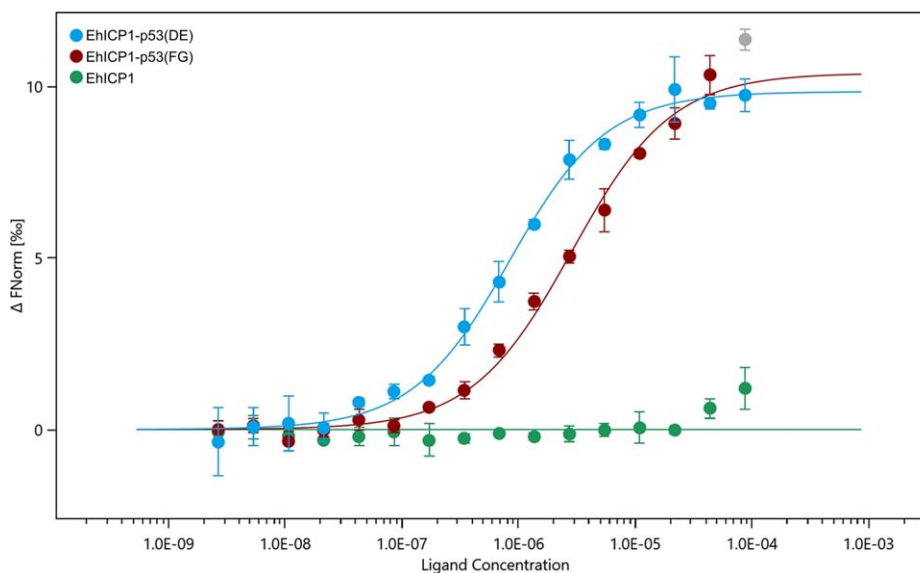


FIGURE 4 Binding affinities between EhICP1-p53 chimeras and MDM2. Binding isotherms determined by MST showing in green the titration curve of the EhICP1 and EhICP1-p53 chimeras toward MDM2. As expected there is no dose-response effect when EhICP1 and MDM2 were incubated (green lane). The binding isotherms of EhICP1-p53(DE) and EhICP1-p53(FG) chimeras toward MDM2 are colored in blue and brown, respectively. The data show that the EhICP1-p53(FG) chimera presents a slightly higher binding affinity toward MDM2 than the EhICP1-p53(DE) chimera [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Dissociation constants of EhICP1-p53 chimeras and MDM2 measured by biolayer interferometry

	$K_D(M)$	K_{Derror}	$k_{on}(1/Ms)$	$k_{onError}$	$k_{off}(1/s)$	$k_{offError}$	R^2
FG 2:1 HL fit	2.57×10^{-6}	5.89×10^{-8}	1.83×10^2	2.07	4.7×10^{-4}	9.37×10^{-6}	0.9956
DE 2:1 HL fit	1.44×10^{-6}	2.37×10^{-8}	1.12×10^3	1.63	1.62×10^{-3}	1.28×10^{-5}	0.9961

Association (k_{on}) and dissociation (k_{off}) rate constants were measured using an Octet instrument, and the apparent binding constants (K_D) were calculated as k_{off}/k_{on} .

between EhICP1 p53 chimeras and MDM2 by MicroScale Thermophoresis (MST). As expected no interaction between EhICP1 and MDM2 was detected (Figure 4, green trace). On the other hand, EhICP1-p53(DE) and EhICP1-p53(FG) interact with MDM2 in a concentration dependent manner (Figure 4, blue and red trace, respectively).

The K_D between MDM2 and EhICP1-p53(DE) was 0.76×10^{-6} M, while the K_D with the EhICP1-p53(FG) variant was 2.79×10^{-6} M, indicating that placing the p53 peptide in the DE instead of the FG loop results in a binding affinity that is 3.7-fold higher. In order to corroborate the affinity constants of EhICP1-p53(FG) and EhICP1-p53(DE) against MDM2 we re-measured them by bio-layer interferometry

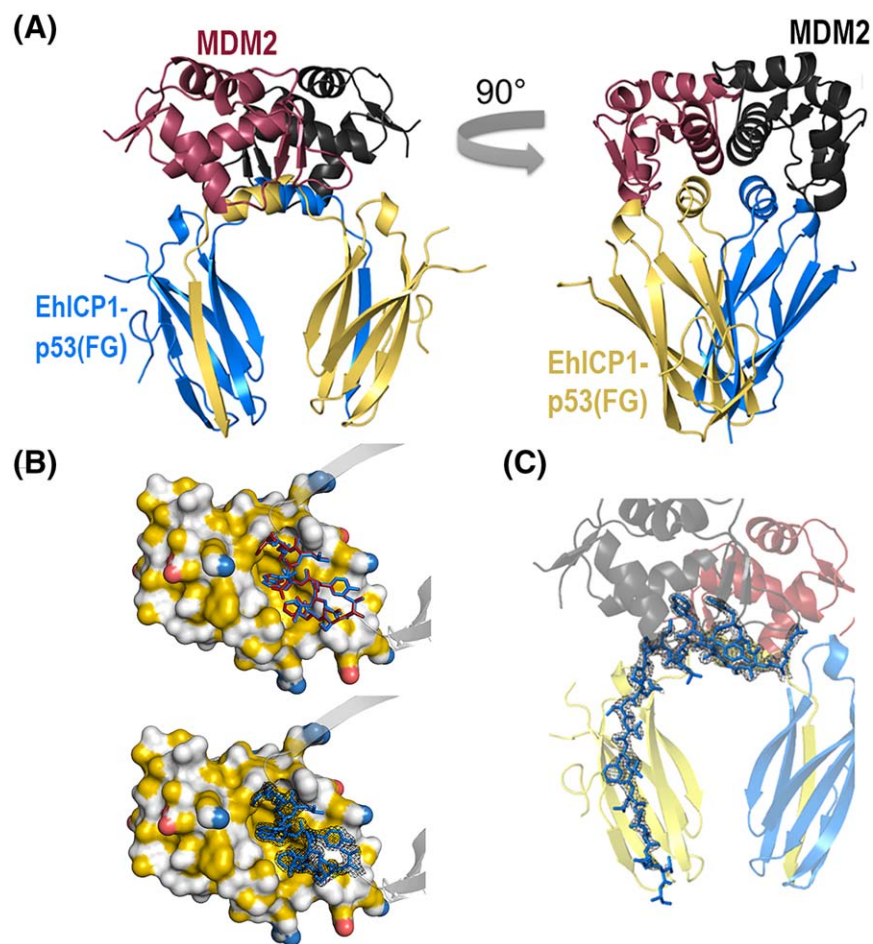


FIGURE 5 Crystal structure of the complex EhICP1-p53(FG) in complex with MDM2. A, Ribbon diagram showing the interaction between the p53 peptide in the FG-loop of EhICP1 and MDM2 and its 90° rotation view. The heterotetramer is stabilized by a swapping domain (by the entire swap of a β sheet of each monomer into the other) formed of 2 FG-loop molecules (blue and yellow). Each EhICP1-p53(FG) protein interacts with 1 molecule of MDM2 (black and magenta). B, The introduced peptide in the FG-loop construct adopts an α -helix conformation that its able to interact with MDM2 hydrophobic cavity (MDM2 residues are colored in yellow (hydrophobic), magenta (negative charged), and cyan (positive charged) according to the Pymol YRB script.⁵² The upper panel show a structural alignment of the EhICP1-p53(FG)-MDM2 complex superimposed with p53 peptide-MDM2 complex (PDB ID:3EQS). EhICP1-p53(FG) peptide is represented as blue sticks, and p53 peptide is colored in red. The lower panel shows an omit map of the p53 sequence of the EhICP1-p53(FG) protein into the cavity of MDM2. C, Omit map of the EhICP1-p53(FG) interacting α -helix and the β sheet swapped into the other EhICP1-p53(FG) molecule

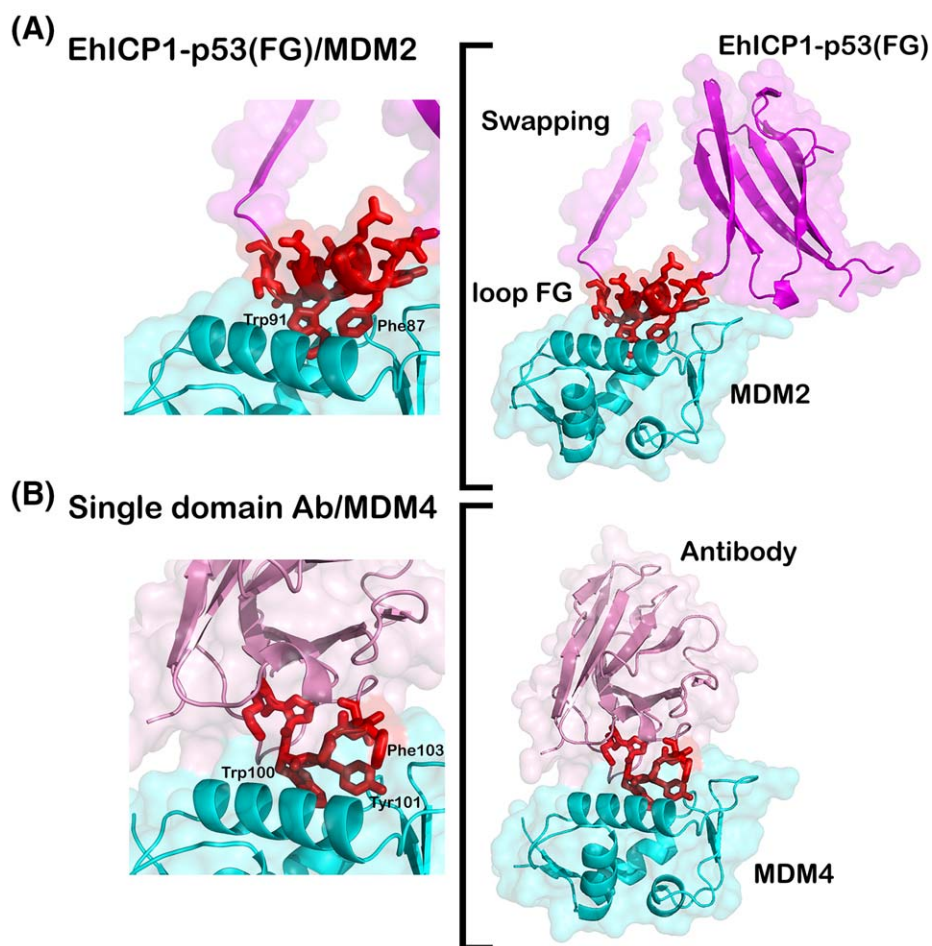


FIGURE 6 Structural comparison between EhICP1(FG) p53 in complex with MDM2 to a single-domain antibody against MDM4. A, Crystal structure of EhICP1-p53(DE) in complex with MDM2. Residues Phe87 and Trp 87 form hydrophobic interactions with MDM2. B, Crystal structure of MDM4 in complex with a single domain antibody. The recognition loop of the antibody is a ball-stick representation. Residues Trp100, Tyr101 and Phe103 of the antibody establish stacking interactions with MDM4. In both structures the main interactions are mediated by a Trp [Color figure can be viewed at wileyonlinelibrary.com]

(Table 3). The K_D between MDM2 and EhICP1-p53(DE) was 1.44×10^{-6} M and for EhICP1-p53(FG) was 2.57×10^{-6} . These data indicate that the p53 peptide substituting the DE loop of EhICP1 only presents a 2-fold increase in binding affinity with respect with the substitution at the FG loop. Although there is a discrepancy in the binding affinities between methods, both MST and bio-layer interferometry yield affinity constants in the low micromolar range.

3.7 | Crystal structure of the EhICP1-p53(FG)/MDM2 complex reveals a swapping subdomain arrangement, while the EhICP1-p53(DE)/MDM2 complex failed to crystallize

We attempted to crystallize EhICP1-p53(DE) and EhICP1-p53(FG) in complex with MDM2. However, we only obtained crystals for the EhICP1-p53(FG)-MDM2 complex. The crystal structure of the complex was determined by molecular replacement using the crystal structures of EhICP2 and MDM2 as search models. Molecular replacement revealed the presence of 4 molecules in the asymmetric unit, the

solution revealed clear electron density for 2 MDM2 molecules and 2 EhICP1-p53(FG) molecules in the space group $P4_32_12$. (Figure 5A and Supporting Information Table S1). Residues 1–24 and 111–150 of MDM2 were not modeled as no electron density was presented for those residues, whereas electron density for all EhICP1-p53(FG) chimera residues was observed. The crystal structure of the complex revealed that the α -helical transactivation domain of p53 acquires a proper secondary structure in the EhICP1-(FG) scaffold that fits into MDM2 (Figure 5B, lower panel). Interestingly, the introduced p53 derived peptide (upper panel) interacts with MDM2 exactly as observed in the crystal structure of p53 in complex with MDM2.³³ The crystal structure also reveals a swapping subdomain between both molecules. Domain swapping occurs by exchanging β -strand 7 of adjacent subunits to interact with the β -strands of the neighboring subunit (Figure 5C). Within the heterotetrameric structure are 2 interacting sites: 1 for each pair of FG-loop/MDM2 molecules. The structural details of the interaction are shown in Figure 6.

In order to determine if the structure of EhICP1-p53(FG) with MDM2 mimics an antibody-mediated interaction, we compared the

crystal structure with the structure of a single-domain antibody against the MDM2 ortholog MDM4 (also dubbed MDMX). This structural comparison highlights the structural convergence between both complexes.⁴⁶

In the single domain antibody, 2 conserved aromatic residues Trp100 and Tyr101 interact in the same crevice in which p53 interacts with MDM2. EhICP-p53(FG) exerts all its interactions with MDM2 by the introduced p53 amino acid sequence, whereas the single domain antibody also exerts additional interactions, like the hydrogen bond between residue Glu110 and Try99 of MDM4 or Phe 111 that sits into a hydrophobic cleft of MDM4 (Figure 6, bottom).

4 | DISCUSSION

Protein scaffolds with an Ig-like fold represent a valuable platform for the design of proteins with desirable specificities. For instance, nature has chosen this domain to create antibodies that can accept an overwhelming number of variants in this scaffold. Antibodies been used for protein engineering, but there are remaining problems due, in part, to their disulfide cross-linking propensity. Those limitations have been partially solved by the use of specific bacterial strands and selection methods.^{5,6} Here we present a strategy for overcoming disulfide cross-linking propensity of the Ig-like fold: the use of the EhICP1 protein, which can tolerate large mutations in 2 of its loops (DE and FG loop) and remain properly folded.

MDM2 has been targeted by aptamer proteins using thioredoxin and the unstructured protein 4EBP1 as scaffolds.^{41,47} Thioredoxin and 4EBP1 scaffolds containing a p53 sequence N-MPRFMDYWEGLN that was optimized by phage display and binds with 2 orders of magnitude more strongly than the canonical p53 sequence.⁴⁸ The optimized p53 peptide in a thioredoxin scaffold has a binding constant that is 2-fold the binding affinity for the peptide alone.⁴¹ Thus, the PA scaffold increases the affinity of the interaction with MDM2. The increased affinity of the PA scaffold may be related to a lower conformational entropy of the p53 peptide constrained into the PA. As mentioned before, the p53 optimized peptide binds to MDM2 with a 3.36×10^{-9} M affinity constant, that is approximately 100 times more tighter than the binding of the canonical p53 peptide (N-ETFSDLWKLPE).³³ The binding of the optimized p53 peptide grafted onto the EhICP1 scaffold decreases the binding affinity to micromolar range. The high binding of the p53 peptide is attributed to the hydrogen bond network that expands its binding surface on MDM2.³³ The formation of the swapping subdomain in EhICP1 may explain the decrease in the binding affinities, as unfavorable interactions between EhICP1-p53(FG) and MDM2 are not observed. Spatial limitations are likely responsible for the swapping subdomain present in the crystal structure of the EhICP-p53(FG)-MDM2 complex (Figure 5).

In order to increase the binding affinity of EhICP1-p53(FG) to MDM2 further studies and constructs are needed to promote the correct position of the p53 peptide replacing the FG loop. This strategy may involve random selection as in other PA.⁴⁹

Even though the K_D obtained was considerably higher than the K_D of the optimized p53 peptide selected for the design,³³ it has a similar value to the interaction between wild-type p53 and MDM2, indicating that our construct is able to mimic the p53-MDM2 interaction with the same strength and order of magnitude.⁴³

ICPs from *Plasmodium* form large molecular weight aggregates of >100 kDa. The latter suggests that ICPs have an intrinsic disposition for swapping subdomain arrangements that led to the formation of fibrils.⁵⁰ As the formation of swapping subdomain structures is dependent of the incubation period in which proteins can test partially folded states, it is possible that other inhibitors like chagasin or EhICP2 would assemble as large molecular weight aggregates originated by swapping subdomains.⁵¹ In summary, we report that an ICP from *E. histolytica* is an effective scaffold for aptamer construction against MDM2. The successful use of this protein as a PA owes to its structural resemblance with the Ig-like domain of the human T cell co-receptor CD8 α

ACKNOWLEDGMENTS

This work was supported by grant "Fronteras de la Ciencia-CONACYT" # 11 to L.G.B. We thank Corina-Diaz Quezada for technical support. P.J.S. thanks CONACYT for graduate studies fellowships. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817).

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.

ORCID

Luis G. Briebe  <http://orcid.org/0000-0001-7199-5779>

REFERENCES

- [1] Dill KA. Dominant forces in protein folding. *Biochemistry*. 1990;29(31):7133–7155.
- [2] Colas P, Cohen B, Jessen T, Grishina I, McCoy J, Brent R. Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature*. 1996;380(6574):548–550
- [3] Honegger A. Engineering antibodies for stability and efficient folding. *Handb Exp Pharmacol*. 2008;(181):47–68.
- [4] Skerra A. Bacterial expression of immunoglobulin fragments. *Curr Opin Immunol*. 1993;5(2):256–262.
- [5] Mazor Y, Van Blarcom T, Mabry R, Iverson BL, Georgiou G. Isolation of engineered, full-length antibodies from libraries expressed in *Escherichia coli*. *Nat Biotechnol*. 2007;25(5):563–565.
- [6] Makino T, Skretas G, Kang TH, Georgiou G. Comprehensive engineering of *Escherichia coli* for enhanced expression of IgG antibodies. *Metab Eng*. 2011;13(2):241–251.

- [7] Chlewicki LK, Holler PD, Monti BC, Clutter MR, Kranz DM. High-affinity, peptide-specific T cell receptors can be generated by mutations in CDR1, CDR2 or CDR3. *J Mol Biol.* 2005;346(1):223–239.
- [8] Hiipakka M, Poikonen K, Saksela K. SH3 domains with high affinity and engineered ligand specificities targeted to HIV-1 Nef. *J Mol Biol.* 1999;293(5):1097–1106.
- [9] Koide A, Bailey CW, Huang X, Koide S. The fibronectin type III domain as a scaffold for novel binding proteins. *J Mol Biol.* 1998; 284(4):1141–1151.
- [10] Reina J, Lacroix E, Hobson SD, et al. Computer-aided design of a PDZ domain to recognize new target sequences. *Nat Struct Mol Biol.* 2002;9(8):621–627.
- [11] Binz HK, Amstutz P, Kohl A, et al. High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotechnol.* 2004; 22(5):575–582.
- [12] Abedi MR, Caponigro G, Kamb A. Green fluorescent protein as a scaffold for intracellular presentation of peptides. *Nucleic Acids Res.* 1998;26(2):623–630.
- [13] Reverdatto S, Burz DS, Shekhtman A. Peptide aptamers: development and applications. *Curr Top Med Chem.* 2015;15(12):1082–1101.
- [14] Skerra A. Alternative non-antibody scaffolds for molecular recognition. *Curr Opin Biotechnol.* 2007;18(4):295–304.
- [15] Bork P, Holm L, Sander C. The immunoglobulin fold: structural classification, sequence patterns and common core. *J Mol Biol.* 1994; 242(4):309–320.
- [16] Bork P, Doolittle RF. Proposed acquisition of an animal protein domain by bacteria. *Proc Natl Acad Sci U S A.* 1992;89(19):8990–8994.
- [17] Mei S, Zhang J, Zhang X, Tu X. Solution structure of a bacterial immunoglobulin-like domain of the outer membrane protein (LigB) from *Leptospira*. *Proteins.* 2015;83(1):195–200.
- [18] Batori V, Koide A, Koide S. Exploring the potential of the monobody scaffold: effects of loop elongation on the stability of a fibronectin type III domain. *Protein Eng.* 2002;15(12):1015–1020.
- [19] Jimenez-Sandoval P, Lopez-Castillo LM, Trasvina-Arenas CH, Briebe LG. Cysteine proteases inhibitors with immunoglobulin-like fold in protozoan parasites and their role in pathogenesis. *Curr Protein Pept Sci.* 2017;18(10):1035–1042.
- [20] Rawlings ND, Barrett AJ, Finn R. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 2016;44(D1):D343–D350.
- [21] Costa TF, Lima AP. Natural cysteine protease inhibitors in protozoa: fifteen years of the chagasin family. *Biochimie.* 2016;122:197–207.
- [22] Wang SX, Pandey KC, Scharfstein J, et al. The structure of chagasin in complex with a cysteine protease clarifies the binding mode and evolution of an inhibitor family. *Structure.* 2007;15(5):535–543.
- [23] Casados-Vazquez LE, Lara-Gonzalez S, Briebe LG. Crystal structure of the cysteine protease inhibitor 2 from *Entamoeba histolytica*: functional convergence of a common protein fold. *Gene.* 2011;471 (1–2):45–52.
- [24] Gilbreth RN, Esaki K, Koide A, Sidhu SS, Koide S. A dominant conformational role for amino acid diversity in minimalist protein-protein interfaces. *J Mol Biol.* 2008;381(2):407–418.
- [25] Lipovsek D, Lippow SM, Hackel BJ, et al. Evolution of an interloop disulfide bond in high-affinity antibody mimics based on fibronectin type III domain and selected by yeast surface display: molecular convergence with single-domain camelid and shark antibodies. *J Mol Biol.* 2007;368(4):1024–1041.
- [26] Karatan E, Merguerian M, Han Z, Scholle MD, Koide S, Kay BK. Molecular recognition properties of FN3 monobodies that bind the Src SH3 domain. *Chem Biol.* 2004;11(6):835–844.
- [27] Huang J, Koide A, Makabe K, Koide S. Design of protein function leaps by directed domain interface evolution. *Proc Natl Acad Sci U S A.* 2008;105(18):6578–6583.
- [28] Hackel BJ, Kapila A, Witttrup KD. Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling. *J Mol Biol.* 2008;381(5):1238–1252.
- [29] Figueiredo da Silva AA, de Carvalho Vieira L, Krieger MA, Goldenberg S, Zanchin NIT, Guimarães BG. Crystal structure of chagasin, the endogenous cysteine-protease inhibitor from *Trypanosoma cruzi*. *J Struct Biol.* 2007;157(2):416–423.
- [30] Redzynia I, Ljunggren A, Bujacz A, Abrahamson M, Jaskolski M, Bujacz G. Crystal structure of the parasite inhibitor chagasin in complex with papain allows identification of structural requirements for broad reactivity and specificity determinants for target proteases. *FEBS J.* 2009;276(3):793–806.
- [31] Leahy DJ, Axel R, Hendrickson WA. Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 Å resolution. *Cell.* 1992;68(6):1145–1162.
- [32] Garcia KC, Degano M, Pease LR, et al. Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science.* 1998;279(5354):1166–1172.
- [33] Pazgier M, Liu M, Zou G, et al. Structural basis for high-affinity peptide inhibition of p53 interactions with MDM2 and MDMX. *Proc Natl Acad Sci U S A.* 2009;106(12):4665–4670.
- [34] Fowler SB, Clarke J. Mapping the folding pathway of an immunoglobulin domain: structural detail from phi value analysis and movement of the transition state. *Structure.* 2001;9(5):355–366.
- [35] Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallograph D.* 2011;67(4):271–281.
- [36] McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007;40(4):658–674.
- [37] Adams PD, Afonine PV, Bunkóczi G, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallograph D.* 2010;66(2):213–221.
- [38] Afonine PV, Grosse-Kunstleve RW, Echols N, et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallograph D.* 2012;68(4):352–367.
- [39] Vagin AA, Steiner RA, Lebedev AA, et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallograph D.* 2004;60(12):2184–2195.
- [40] Chen VB, Arendall WB, Headd JJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallograph D.* 2010;66(1):12–21.
- [41] Brown CJ, Dastidar SG, See HY, et al. Rational design and biophysical characterization of thioredoxin-based aptamers: insights into peptide grafting. *J Mol Biol.* 2010;395(4):871–883.
- [42] Parker MH, Chen Y, Danehy F, et al. Antibody mimics based on human fibronectin type three domain engineered for thermostability and high-affinity binding to vascular endothelial growth factor receptor two. *Protein Eng Des Sel.* 2005;18(9):435–444.
- [43] Yu GW, Rudiger S, Veprintsev D, Freund S, Fernandez-Fernandez MR, Fersht AR. The central region of HDM2 provides a second binding site for p53. *Proc Natl Acad Sci U S A.* 2006;103(5):1227–1232.
- [44] Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods.* 2015; 12(1):7–8.

- [45] Wright TA, Stewart JM, Page RC, Konkolewicz D. Extraction of thermodynamic parameters of protein unfolding using parallelized differential scanning fluorimetry. *J Phys Chem Lett.* 2017;8(3):553–558.
- [46] Yu GW, Vaysburd M, Allen MD, Settanni G, Fersht AR. Structure of human MDM4 N-terminal domain bound to a single-domain antibody. *J Mol Biol.* 2009;385(5):1578–1589.
- [47] See HY, Lane DP. A novel unstructured scaffold based on 4EBP1 enables the functional display of a wide range of bioactive peptides. *J Mol Biol.* 2010;404(5):819–831.
- [48] Bottger V, Bottger A, Howard SF, et al. Identification of novel mdm2 binding peptides by phage display. *Oncogene.* 1996;13(10):2141–2147.
- [49] Wright CF, Christodoulou J, Dobson CM, Clarke J. The importance of loop length in the folding of an immunoglobulin domain. *Protein Eng Des Sel.* 2004;17(5):443–453.
- [50] Sundararaj S, Saxena AK, Sharma R, et al. Cross-talk between malarial cysteine proteases and falstatin: the BC loop as a hot-spot target. *PLoS One.* 2014;9(4):e93008
- [51] Blancas-Mejia LM, Ramirez-Alvarado M. Systemic amyloidoses. *Annu Rev Biochem.* 2013;82:745–774.
- [52] Hagemans D, van Belzen IA, Luengo TM, Rüdiger SG. A script to highlight hydrophobicity and charge on protein surfaces. *Front Mol Biosci.* 2015;2:56.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Jimenez-Sandoval P, Madrigal-Carrillo EA, Santamaría-Suárez HA, et al. Mimicking a p53-MDM2 interaction based on a stable immunoglobulin-like domain scaffold. *Proteins.* 2018;86:802–812. <https://doi.org/10.1002/prot.25519>