



VNIVERSITAT
E VALÈNCIA



INSTITUTO DE
BIOMEDICINA DE
VALENCIA CSIC

Doctorado en Biomedicina y Biotecnología
Departamento de Bioquímica y Biología Molecular

**Estudios estructurales y funcionales de
dUTPasas, una amplia familia de enzimas
metabólicas con funciones reguladoras.**

Carla Sanz Frasquet

Tesis Doctoral - Febrero 2023

Memoria presentada por Carla Sanz Frasquet para optar al grado
de Doctora por la Universidad de Valencia

Dirigida por el Dr. Alberto Marina Moreno

Unidad de Cristalografía de Macromoléculas

Departamento de Genómica y Proteómica

Instituto de Biomedicina de Valencia

Consejo Superior de Investigaciones Científicas

D **Alberto Marina Moreno**, Profesor de Investigación de OPI del Dpto. de Genómica y Proteómica del Instituto de Biomedicina de Valencia (CSIC)

CERTIFICA/N:

Que la presente memoria, titulada “**Estudios estructurales y funcionales de dUTPasas, una amplia familia de enzimas metabólicas con funciones reguladoras**”, corresponde al trabajo realizado bajo su dirección por D. **Carla Sanz Frasquet**, para su presentación como Tesis Doctoral en el Programa de Doctorado en Biomedicina y Biotecnología de la Universitat de València.

Y para que conste firma/n el presente certificado en Valencia, 21 de Febrero de 2023,

FDO. Alberto marina Moreno

AGRADECIMIENTOS

A Vicent, con quien entrelacé mi mano a los 15 años y ahí sigue,
Gracias por todo, pero sobre todo por ser tú, y por elegirme a mí. Yo te elegiría una y mil veces.

A mis padres, quienes me enseñaron que las cosas se consiguen con esfuerzo y sacrificio. Muchas gracias por confiar en mí y transmitirme apoyo moral en todo este tiempo. Gracias por la ayuda en forma de *tuppers* y viajes en coche hasta y desde la estación de tren, casi igual de cruciales que los ánimos.

A mis suegros, quienes son casi como mis padres. Muchas gracias por todo el apoyo moral, la confianza en mí, y las charlas multitema. Siempre es un placer hablar con vosotros. Gracias también por todos los *tuppers*.

A mis hermanas, la biológica y la política. Muchas gracias por los cotilleos y los cafés matutinos. A estos últimos, los echo mucho de menos.

A mis amigas Encarna y Sara, muchas gracias por el apoyo moral, los cafés y las videollamadas. Habéis hecho más llevadero este trance.

A la gente del IBV, gracias por toda la ayuda y los buenos momentos. Siento no poder nombraros a cada uno de vosotros, pero sabéis que os llevo en el corazón y os agradezco lo bien que me he sentido formando parte de la pequeña familia que es el IBV. Gracias especiales a Mónica, con quien he compartido muchos cruasanes reparadores de preocupaciones.

A la gente del laboratorio,

Dr. Rafa y Dr. Christian, mis predecesores en las Duts,
muchas gracias por acogerme tan bien cuando llegué al lab, ha sido un verdadero placer trabajar con vosotros. Filosofías de vida distintas, pero geniales trabajando juntas, y de las que he intentado quedarme con lo mejor. No obstante, tengo que hacer una referencia especial a Rafa, mi tutor de TFM y quién sentó las bases de cómo trabajar en el lab, y de quién aprendí gran parte de las cosas que sé. Hiciste fácil trabajar, aprender y reír al mismo tiempo. Muchas gracias por todo, nunca olvidaré lo crucial que resultaste ser en mi inicio científico, ni tampoco los sistemas de 3 ecuaciones para preparar las placas de cristalización.

Dr. Paqui,

Consejera, madre y gran científica. Muchas gracias por todo, por ser la mami del grupo, y por todos los buenos momentos de almuerzo. Te doy las gracias por los mejores consejos ante cualquier cuestión, ya fuera científica o personal.

Dr. Sara,

Gracias por las mejores tardes de trabajo. No solo por la “hora-loca” que nos ha definido, sino por todo lo aprendido, los consejos que me has dado y las risas que hemos compartido.

Cuasi Dr. Anmol,

Creo que fui la primera persona española en saludarte con dos besos cuando llegaste aquí desde la India...ninguno de los dos olvidará ese momento. Muchas gracias por los buenos momentos, por los consejos y las tostadas compartidas. Jamás olvidaré tus conocimientos informáticos que me han salvado en varias ocasiones, pero tampoco los cafés después de comer.

Javi y Elena,

Las últimas incorporaciones al laboratorio, muchas gracias por seguir propulsando el buen rollo en el laboratorio. Es un placer trabajar con vosotros. Echaré de menos las charlas en los almuerzos.

Alonso,

ya te acercas al final de la Tesis. Mucho ánimo para acabar tu etapa predoctoral, y mucha suerte para seguir con la posdoctoral, aunque realmente sé que no la necesitarás.

Y por último, al Dr. Alberto Marina, y no por ser el menos importante sino el más crucial, como en los artículos científicos. Gracias por haberme dado la oportunidad de realizar la Tesis en tu laboratorio, y por guiarme en el proceso. No podría haber “caído” en mejor sitio, con un gran científico al que admirar.

LISTA DE ABREVIATURAS

ADN: ácido desoxirribonucleico

A. E.: Actividad enzimática

APS: amonio persulfato

ARN: ácido ribonucleico

ATP: adenosina trifosfato

dATP: desoxiadenosina trifosfato

BLI: interferometría de biocapa (*BioLayer Interferometry*)

DBD: dominio de unión al ADN (*DNA Binding Domain*)

dITP: desoxiinosina trifostato

dGTP: desoxiguanosina trifosfato

dn/dc: incremento del índice de refracción

DrMazG: MazG de *Deinococcus radiodurans*

DTT: ditioteitrol

dTTP: desoxitimidina trifosfato

dUDP: desoxiuridina difosfato

dUMP: desoxiuridina monofosfato

Dut: dUTPasa

dXTP: desoxixantosina trifosfato

EDTA: ácido etilendiaminetetraacético (*Ethylenediaminetetraacetic acid*)

EMSA: ensayos electroforéticos de retraso en gel (*Electrophoretic mobility shift assay*)

EnMazG: MazG de fago de *Enterococcus faecalis*.

eV: electronvoltios

FW: cebador *forward*, que se une al ADN en dirección 5'→3'

FPLC: cromatografía líquida rápida de proteínas (*Fast protein liquid chromatography*)

hDut: dUTPasa humana

IPTG: Isopropil- β - D - 1 - tiogalactopiranosido

ITC: calorimetría de titración isotérmica (*Isothermal Titration Calorimetry*)

KeV: Kiloelectronvoltio

LB: medio de cultivo Luria-Bertani

IvDut: dUTPasa de *Litopenaeus vannamei*

MCS: *Multiple Cloning Site*

mDut: dUTPasa de *Mycobacterium tuberculosis*

Native-PAGE: electroforesis en gel de poliacrilamida en condiciones nativas

NDP quinasa: nucleósido-difosfato quinasa

NEB: New England Biolabs

NF- κ B: factor nuclear potenciador de las cadenas ligeras kappa de las células B activadas

NTPs: oxinucleótidos trifosfato

dNTPs: desoxinucleótidos trifosfato

ORF: marco de lectura abierta (*Opening Reading Frame*)

pb: pares de bases

PBS: solución salina tamponada con fosfatos

PCR: reacción en cadena de la polimerasa (*Polymerase Chain Reaction*)

PEG: polietilenglicol

PICl: islas cromosómicas inducidas por fago

PPAR α : receptor α activado por el proliferador del peroxisoma

PPi: pirofosfato

ϕ 11Dut: dUTPasa trimérica del fago ϕ 11 de *Staphylococcus aureus*

ϕ 80 α Dut: dUTPasa trimérica del fago ϕ 80 α de *S. aureus*

ϕ O11Dut: dUTPasa trimérica del fago ϕ O11 de *S. aureus*

PMSF: inhibidor de proteasas de serina (fluoruro de fenilmetilsulfonilo, *phenylmethylsulfonyl fluoride*)

Res.: resolución del juego de datos cristalográficos

RMSD: raíz cuadrada de la media de los cuadrados de las desviaciones para un conjunto de n puntos v y w (de sus siglas en inglés Root Mean Square Deviation)

$$rmsd(v, w) = \sqrt{\frac{1}{n} \sum_{i=1}^n ||v_i - w_i||^2}$$

RNR: ribonucleótido reductasa

RV: cebador *reverse*, que se une al ADN en dirección 3'→5'

RXR: receptor X retinoide

SeMazG: MazG del fago IPLA7 de *Staphylococcus epidermidis*

Stl^{FL}: proteína represora Stl completa

Stl^{N-ter}: fragmento N-terminal de la proteína represora Stl (residuos 1 a 156)

TEMED: tetrametiletilendiamina

U. A.: contenido de la unidad asimétrica en el juego de datos cristalográficos

UDG: uracilo-ADN-glicosilasa

UTP: uridina trifosfato

VIH: virus de la inmunodeficiencia humana

MPV: virus del mono Mason-Pfizer

WT: no mutante (wild-type)

Introducción.....	1
1. Estabilidad genómica y dUTPasas.....	3
2. Clasificación de las Duts.....	5
2.1. Duts triméricas.....	5
2.1.1. Centro activo.....	8
2.1.2. Mecanismo de hidrólisis.....	11
2.1.3. El motivo V.....	13
2. 2. Duts monoméricas.....	14
2. 3. Duts diméricas.....	17
2.3.1. Centro activo.....	21
2.3.2. Mecanismo de hidrólisis.....	24
3. Funciones <i>moonlighting</i> : ¿Qué son?	26
4. Islas de patogenicidad de <i>Staphylococcus aureus</i> (SaPIs).....	28
5. Inducción de las SaPIs por Duts.....	33
5.1. El StI de SaPIbov1 interacciona con Duts.....	34
5.2. La interacción de StI de SaPIbov1 con Duts diméricas y triméricas como fracaso del intento de escape del fago de la inducción de la SaPI.....	38
6. Proteínas MazG.....	45
6. 1. ¿Qué son las proteínas MazG?.....	45
6.1.1. Organización de dominios MazG.....	45
6. 2. Estructura tridimensional de las MazG.....	46
6.2.1. Estructura tridimensional de MazG de fago de <i>Enterococcus faecalis</i>	48
6.2.2. Organización del centro activo de MazG del fago ϕ FL2A de <i>Enterococcus faecalis</i> (EnMazG).....	50
6.2.3. El nucleótido dUTP puede adoptar dos conformaciones en el centro activo de EnMazG y SeMazG.....	54

6. 3. Las proteínas MazG degradan más de un nucleótido, y utilizan cationes distintos al Mg ²⁺	56
6. 4. Las proteínas MazG del tipo I están relacionadas con las Duts diméricas.....	57
Objetivos.....	61
Material y métodos.....	65
1. Técnicas en biología molecular.....	67
1. 1. Construcción de vectores.....	67
1.1.1. Sistema de expresión génica.....	67
1.1.2. Sistema de clonación	68
1. 2. Transformación de células competentes y selección positiva de clones.....	75
1. 3. Comprobación de células positivas.....	76
1. 4. Preparación de agarosa.....	77
1. 5. Sistemas de expresión bacteriana.....	77
1. 6. Mutagénesis dirigida.....	78
2. Expresión y purificación de proteínas.....	81
2. 1. Pruebas de expresión proteica.....	81
2. 2. Prueba de solubilidad proteica.....	82
2. 3. Purificación a gran escala.....	87
2.3.1. Expresión proteica y lisis bacteriana.....	67
2.3.2. Cromatografía de afinidad.....	89
2.3.3. Electroforesis en gel de poliacrilamida en condiciones desnaturizantes (SDS-PAGE).....	89
2.3.4. Digestión de la etiqueta unida a la proteína.....	91
2.3.5. Cromatografía de exclusión molecular.....	93
2.3.6. Diálisis.....	94
2.3.7. Concentración centrífuga de proteínas.....	94
2.3.8. Cuantificación proteica por el método de Bradford.....	95

2. 4. Ensayo de termoestabilidad (termofluor).....	95
2. 5. Western-Blot.....	96
2. 6. MALS y DLS.....	97
2. 7. Ensayos de interacción de proteínas.....	98
2.7.1. Geles de acrilamida en condiciones nativas (Native-PAGE).....	98
2.7.2. Interferometría de biocapa.....	99
2. 8. Ensayos de interacción proteína-ADN.....	103
2. 9. Medida de actividad enzimática por reactivo verde malaquita.....	103
3. Cristalogénesis, congelación y difracción de cristales de proteína.....	105
3. 1. Puesta de placas de <i>screening</i>	105
3. 2. Congelación de cristales.....	107
3. 3. Difracción de rayos X de cristales de proteína.....	108
4. Resolución de la estructura tridimensional.....	109
4. 1. Indexado y refinado de la celda.....	109
4. 2. Integración.....	110
4. 3. Traslación, escalado y mezcla de datos.....	110
4. 4. Cálculo de la fase.....	111
4. 5. Trazado y refinado de la estructura.....	113
4. 6. Cálculo de picos de dispersión anómala.....	114
Resultados.....	115
Capítulo I.....	117
1. 1. El fragmento N-terminal de StI es capaz de unirse a las Duts humana y de <i>Mycobacterium tuberculosis</i>	119
1. 2. Determinación de las estructuras tridimensionales de StI ^{N-ter} unido a mDut y a hDut.....	124
1.2.1. Estructura tridimensional del complejo mDut-StI ^{N-ter}	125
1.2.2. Estructura tridimensional del complejo hDut-StI ^{N-ter}	129
1. 3. La diana de StI es el trímero de Dut.....	132

1. 4. Stl mimetiza al nucleótido para interactuar con las Duts triméricas..	136
1. 5. El motivo V de las Duts triméricas es un elemento activo en la unión con Stl.....	141
1. 6. El represor Stl tiene capacidad universal para unirse a las Duts.....	143
Capítulo II.....	149
2. 1. Análisis molecular de las diferencias cinéticas entre SeMazG Y EnMazG.....	151
2.1.1. La participación del tercer monómero en el centro activo es necesario para una actividad catalítica óptima.....	152
2.1.2. La tetramerización es necesaria para la actividad catalítica.....	154
2. 2. El uso de otros cationes divalentes en EnMazG y SeMazG.....	161
2.2.1. El magnesio no puede ser sustituido por otro catión divalente para que EnMazG y SeMazG puedan hidrolizar dUTP.....	161
2.2.2. La actividad enzimática de EnMazG y SeMazG es inhibida por cationes divalentes.....	163
2.2.3. Modulación de la actividad de EnMazG y SeMazG por cationes divalentes.....	166
2.2.4. El análisis de las estructuras tridimensionales de EnMazG en presencia de cationes divalentes permite proponer un mecanismo de modulación por el catión.....	173
2.2.4.1. El uso de la difracción anómala para completar la información obtenida con la difracción nativa.....	175
2.2.4.2. El análisis de las estructuras tridimensionales de EnMazG en presencia del catión Mn^{2+} permite proponer un mecanismo de modulación...	176
2.2.4.3. El análisis de las estructuras tridimensionales de EnMazG resueltas a partir de la co-cristalización con los cationes Ca^{2+} , Zn^{2+} y Ni^{2+} permiten proponer una hipótesis sobre una modulación diferencial del mecanismo de inhibición según el catión implicado.....	191
Capítulo III.....	201
3. 1. Duts de <i>Mycobacterium</i> tienen actividad moonlighting.....	203
3. 2. Una ADN-helicasa dependiente de ATP (Rv3202c).....	205
3. 3. Una proteína de membrana que incluye un β -propeller (Rv3524).....	208

3. 4. Un regulador transcripcional de la familia PucR (Rv1429).....	212
3. 5. Una proteína de función desconocida (Rv3354).....	215
Discusión.....	219
1. El represor StI de SaPI _{bov1} como un inhibidor de Duts universal.....	221
1. 1. El mimetismo como garantía de no escape.....	221
1. 2. StI interacciona con las tres familias de Duts.....	222
1. 3. El motivo V de las Duts triméricas juega un papel activo en la interacción con el represor StI de SaPI _{bov1}	226
1. 4. ¿Es StI un inhibidor universal absoluto de Duts?.....	229
2. El análisis del mecanismo molecular de MazG, con similitudes al de Duts dimericas, podría revelar variabilidad en el mecanismo catalítico.....	233
2. 1. El catión que ocupa la posición 3 podría ser determinante en la catálisis del sustrato.....	243
3. La modulación de EnMazG y SeMazG por cationes divalentes diferentes al Mg ²⁺ podría señalar al fago en la bacteria huésped.....	244
4. Toda investigación es útil, aunque no se consigan los resultados previstos.....	246
Conclusiones.....	247
Anexos.....	251
Anexo I. Tabla detallada de los intentos de clonación de la ADN helicasa dependiente de ATP de <i>M. tuberculosis</i> realizados en esta Tesis.....	253
Anexo II. Tampones generales utilizados en esta Tesis.....	254
Anexo III. Tablas recopilatorias de los valores RMSD (Å) tras la superposición de los monómeros de Dut y StI ^{N-ter} de la estructura tridimensional del complejo mDut-StI ^{N-ter}	255
Anexo IV. Tablas recopilatorias de los valores RMSD (Å) tras la superposición de los monómeros de Dut y StI ^{N-ter} de la estructura tridimensional del complejo hDut-StI ^{N-ter}	256
Anexo V. Tabla recopilatoria de los valores RMSD (Å) de superposición de los monómeros de StI ^{N-ter} de las estructuras en complejo con hDut y mDut, y con la estructura de referencia ($\phi 11$ Dut-StI ^{N-ter}).....	257

Anexo VI. Interacciones intermoleculares en el complejo BovI-StI ^{N-ter} – mDut.....	259
Anexo VII. Interacciones intermoleculares en el complejo BovI-StI ^{N-ter} – hDut.....	260
Anexo VIII. Residuos de reconocimiento e interacción de StI ^{N-ter} en los complejos con ϕ 11Dut, mDut, hDut y EBVDut.....	261
Anexo XIX. Participación en proyectos relacionados con el COVID-19.....	262
Bibliografía.....	263

INTRODUCCIÓN

Figura I. 1. Reacción enzimática catalizada por dUTPasa. El sustrato dUTP es hidrolizado a dUMP y PPi.

Figura I. 2. Vías de producción y consumo del dUTP en la célula.

Figura I. 3. Estructuras representativas de Duts de las tres familias.

Figura I. 4. Alineamiento de secuencias de Duts triméricas.

Figura I. 5. Localización estructural de las estructuras secundarias y motivos conservados en Duts triméricas.

Figura I. 6. Estructura tridimensional de Dut de *E. coli*.

Figura I. 7. Participación de las tres subunidades del trímero de Dut en cada uno de los centros activos.

Figura I. 8. Especificidad por uracilo en Dut de *M. tuberculosis* y participación de los motivos conservados en el centro activo.

Figura I. 9. Vista detallada del centro activo de Dut de humano.

Figura I. 10. Esquema de la reacción de hidrólisis catalizada por Duts triméricas.

Figura I. 11. Alineamiento de secuencia de Duts triméricas y la Dut monomérica representativa del virus de Epstein-Barr.

Figura I. 12. Estructura tridimensional de la Dut del virus de Epstein-Barr.

Figura I. 13. Diagrama representativo de la organización tridimensional de las subunidades de Duts triméricas y de los dominios de Duts monoméricas.

Figura I. 14. Alineamiento de secuencias de Duts diméricas.

Figura I. 15. Monómero de Dut de *Trypanosoma cruzi*.

Figura I. 16. Dímero de Dut de *T. cruzi*.

Figura I. 17. Estructuras tridimensionales de *T. cruzi* en ausencia y presencia de nucleótido.

Figura I. 18. Localización del centro activo en Duts diméricas (*T. cruzi*).

Figura I. 19. Vista detalla del centro activo de Dut de *T. cruzi*.

Figura I. 20. Mecanismo de hidrólisis del dUTP en Dut dimérica.

Figura I. 21. Dos ejemplos de estrategias de regulación mediada por Duts.

Figura I. 22. Ciclo ERE e infección de las SaPIs.

Figura I. 23. Organización génica de PICIs encontradas en diferentes especies.

Figura I. 24. Represión y desrepresión de SaPI.

Figura I. 25. El fago 80 α es explotado por distintas SaPIs para iniciar su ciclo ERE.

Figura I. 26. Alineamiento de Duts triméricas para la localización del motivo VI específico de Duts de fago.

Figura I. 27. Modelo de interacción represor StI – Dut trimérica de fago de *S. aureus*.

Figura I. 28. Estructura secundaria y terciaria del represor StI dividida en los dos fragmentos utilizados para la resolución estructural.

Figura I. 29. Estructura tridimensional del complejo de Dut trimérica del fago ϕ 11 de *S. aureus* con StI^{N-ter}.

Figura I. 30. Vista detallada del centro activo del complejo ϕ 11Dut-StI^{N-ter}.

Figura I. 31. StI mimetiza las interacciones del sustrato dUTP con ϕ 11Dut.

Figura I. 32. Estructura tridimensional del complejo de StI^{C-ter} con la Dut dimérica del fago ϕ O11 de *S. aureus*.

Figura I. 33. Estrategia de mimetización por parte del StI para unirse a Dut dimérica del fago de *S. aureus* ϕ O11.

Figura I. 34. Estructuras tridimensionales de proteínas MazG de diferentes organismos.

Figura I. 35. Estructuras tridimensionales de la proteína EnMazG.

Figura I. 36. Selectividad y acomodación de la base nitrogenada uracilo y la desoxirribosa en el centro activo de EnMazG.

Figura I. 37. Posicionamiento de los tres fosfatos del nucleótido y los tres cationes Mg²⁺ en el centro activo de EnMazG.

Figura I. 38. Diferencias entre las dos conformaciones adoptadas por el nucleótido en el centro activo de EnMazG y SeMazG.

Figura I. 39. Doble conformación del nucleótido dUpNHpp en el centro activo de MazG de *D. radiodurans* (DR2231).

Figura I. 40. Superposición de las estructuras tridimensionales de EnMazG y Dut dimérica de *C. jejuni*.

Figura I. 41. Alineamiento y superposición de las secuencias aminoacídicas y las estructuras de las MazG de *D. radiodurans* y de fago de *E. faecalis*.

Tabla T. I. 1. Clasificación de las proteínas MazG en función de la organización de sus dominios.

MATERIAL Y MÉTODOS

Figura M.1. Estructura común simplificada de las construcciones génicas contenidas en los vectores de expresión usados en esta Tesis.

Figura M. 2. Diagrama de los 3 sistemas de clonación utilizados en esta Tesis.

Figura M. 3. Diagrama de la técnica interferometría de biocapa.

Figura M. 4. Imagen de difracción.

Tabla M. 1. Tabla detallada de las clonaciones exitosas preparadas en esta Tesis.

Tabla M. 2. Tabla recopilatoria de los cebadores utilizados en la mutagénesis dirigida.

Tabla M. 3. Concentraciones de antibiótico usadas en los cultivos bacterianos.

Tabla M. 4. Tabla recopilatoria de todos los tampones probados sin éxito para solubilizar las proteínas de esta Tesis.

Tabla M. 5. Tabla recopilatoria de las condiciones de expresión y solubilidad seleccionadas para cada proteína utilizada en esta Tesis.

Tabla M. 6. Tabla recopilatoria de los tampones usados en la purificación de las proteínas utilizadas en esta Tesis.

Tabla M.7. Relación de los porcentajes de acrilamida usados en los geles para SDS-PAGE de las proteínas utilizadas en esta Tesis.

Tabla M. 8. Relación de proteasas utilizadas en el corte de la etiqueta de las proteínas utilizadas en esta Tesis.

Tabla M. 9. Tabla recopilatoria de las condiciones utilizadas en las medidas de interferometría de biocapa.

Tabla M. 10. Tabla recopilatoria de las concentraciones de Dut, sustrato, solución tampón y tiempos de toma de muestra de los experimentos de actividad enzimática realizados en esta Tesis.

Tabla M. 11. Tabla recopilatoria de las condiciones de cristalización y crioprotección de los cristales de proteína cuya estructura tridimensional se ha obtenido en este trabajo de Tesis.

RESULTADOS

Figura R.1.1. Geles al 8 % de acrilamida en condiciones nativas para la visualización del complejo de Stl^{N-ter} con mDut, hDut y ϕ 11Dut.

Figura R.1.2. Determinación de las constantes de afinidad en el equilibrio de las interacciones de Stl^{N-ter} con mDut, mDut H145F, hDut y ϕ 11Dut por interferometría de biocapa (BLI).

Figura R. 1. 3. Disrupción del complejo ADN-Stl en presencia de ϕ 11Dut, mDut, o hDut por ensayos EMSA.

Figura R. 1. 4. Representación del complejo Dut de *M. tuberculosis* con el represor Stl^{N-ter}.

Figura R. 1. 5. Superposición de los dos complejos mDut-Stl^{N-ter} presentes en la unidad asimétrica de la celda unidad.

Figura R. 1. 6. Superposición de las estructuras de tres de los monómeros Stl^{N-ter} presentes en el complejo mDut.

Figura R. 1. 7. Representación del complejo Dut humana con el represor Stl^{N-ter}.

Figura R. 1. 8. Las Duts triméricas presentan una conformación competente para la unión de Stl^{N-ter}.

Figura R. 1. 9. La estructura del trímero de mDut en complejo con el Stl es más similar a la forma unida al sustrato que a la conformación *apo*.

Figura R. 1. 10. Superposición de las estructuras de Stl^{N-ter} en su conformación libre y unido a la mDut.

Figura R. 1. 11. Superposición de las estructuras de Stl^{N-ter} unidas a mDut, hDut y ϕ 11Dut.

Figura R. 1. 12. El represor Stl mimetiza las interacciones con el sustrato para interaccionar con las Duts triméricas.

Figura R. 1. 13. El represor Stl mimetiza las interacciones del sustrato dUTP con las Duts triméricas.

Figura R. 1. 14. Motivos catalíticos de las Duts triméricas.

Figura R. 1. 15. El motivo V de hDut está implicado en el reconocimiento y unión de Stl.

Figura R. 1. 16. Comparación estructural de las Duts trimérica y monomérica.

Figura R. 1. 17. Modelo del complejo Dut monomérica - Stl^{N-ter}.

Figura R. 1. 18. StI inhibe la actividad enzimática de $_{EBV}Dut$.

Figura R. 1. 19. Análisis de secuencia de los residuos de los motivos catalíticos de las Duts triméricas y monoméricas implicados en el reconocimiento y unión de StI^{N-ter}.

Figura R. 2. 1. Perturbación de la actividad enzimática de los mutantes puntuales EnMazG R42A y SeMazG R48A.

Figura R. 2. 2. Cambios conformacionales de las hélices $\alpha 2$ y $\alpha 3$ del dímero de EnMazG en el cambio de conformación abierta a cerrada.

Figura R. 2. 3. Localización estructural de los residuos I30 e Y36 en la zona de tetramerización de EnMazG y SeMazG, respectivamente.

Figura R. 2. 4. Superposición estructural de las enzimas EnMazG y SeMazG.

Figura R. 2. 5. Perturbación de la actividad enzimática de los mutantes puntuales EnMazG I30A/Y y SeMazG Y36A/I.

Figura R. 2. 6. Perturbación de la actividad enzimática de EnMazG y SeMazG en presencia de Mg^{2+} u otros cationes divalentes.

Figura R. 2. 7. Perturbación de la actividad enzimática de EnMazG en presencia de Mg^{2+} y cationes divalentes alternativos en ratios molares crecientes.

Figura R. 2. 8. Perturbación de la actividad enzimática de SeMazG en presencia de Mg^{2+} y cationes divalentes alternativos en ratios molares crecientes.

Figura R. 2. 9. Actividad enzimática de EnMazG en presencia de Mg^{2+} en combinación con otros cationes divalentes en ratios molares crecientes.

Figura R. 2. 10. Actividad enzimática de SeMazG en presencia de Mg^{2+} en combinación con otros cationes divalentes en ratios molares crecientes.

Figura R. 2. 11. Inhibición de la actividad catalítica de EnMazG en presencia de distintos cationes divalentes.

Figura R. 2. 12. Inhibición de la actividad catalítica de SeMazG en presencia de distintos cationes divalentes.

Figura R. 2. 13. Doble conformación del fosfato y observada en el nucleótido dUTP en algunos centros activos de la estructura tridimensional de EnMazG en presencia de Mg^{2+} .

Figura R. 2. 14. Fotografías de los cristales de EnMazG en presencia de Mn^{2+} en distintas ratios molares con Mg^{2+} .

Figura R. 2. 15. Espectro de emisión de fluorescencia del Mn al excitar el cristal de proteína con un barrido de energías del haz desde 6,48 a 6,6 KeV en el sincrotrón.

Figura R. 2. 16. Superposición de las cuatro estructuras de EnMazG en presencia de los cationes Mg^{2+} y Mn^{2+} .

Figura R. 2. 17. Conformaciones activa e inactiva del nucleótido dUTP en el centro activo de EnMazG.

Figura R. 2. 18. Superposición de las dos conformaciones del nucleótido en EnMazG.

Figura R. 2. 19. Mecanismo de inhibición de EnMazG por el catión divalente Mn^{2+} .

Figura R. 2. 20. Morfologías adoptadas por los cristales de EnMazG en presencia de Ca^{2+} y Zn^{2+} .

Figura R. 2. 21. Superposición de los dímeros de las estructuras de EnMazG en conformación *apo*, en complejo con el nucleótido no hidrolizable dUpNHpp, y en presencia de los cationes divalentes Ca^{2+} , Ni^{2+} y Zn^{2+} .

Figura R. 2. 22. Coordinación de los iones Mn^{2+} en el centro activo de EnMazG.

Figura R. 3.1. La inserción del género *Mycobacterium* no está presente en otras Duts.

Figura R. 3.2. Localización estructural del *loop* específico del género *Mycobacterium*.

Figura R. 3.3. Distribución de dominios de la proteína Rv3202c según Pfam.

Figura R. 3.4. Modelado *in-silico* de la proteína Rv3524 obtenido en el servidor *Phyre2*.

Figura R. 3. 5. Perfil de termofluor de la proteína Rv3524 (β -*propeller*).

Figura R. 3. 6. Espectro de dicroísmo circular de la proteína Rv3524 (β -*propeller*) y resumen de las estructuras secundarias encontradas.

Figura R. 3.7. Resultado de la electroforesis en gel de acrilamida en condiciones desnaturalizantes para las muestras de la prueba de solubilidad desde 2 M urea de la proteína Rv1429.

Figura R. 3. 8. Perfil de termofluor de las muestras del regulador transcripcional de la familia PucR a distintas concentraciones de urea.

Figura R. 3. 9. Predicción de la estructura secundaria de la proteína Rv3354 por el servidor *Phyre2*.

Figura R. 3. 10. Membrana revelada tras someter las muestras de la prueba de solubilidad de la proteína Rv3353 (N33 a S129) a Western-Blot.

Tabla R. 1. 1. Afinidades de unión entre Stl^{N-ter} y Duts de *M. tuberculosis* (salvaje y mutante (mDut y mDut H145F, respectivamente), humana (hDut) y de fago ϕ 11 de *S. aureus* (ϕ 11Dut).

Tabla R. 1. 2. Colección de datos cristalográficos y estadísticos de refinado de la estructura del complejo mDut-Stl^{N-ter}.

Tabla R. 1. 3. Colección de datos cristalográficos y estadísticos de refinado de la estructura del complejo hDut-Stl^{N-ter}.

Tabla R. 1. 4. Tabla de valores RMSD en Å derivados de la superposición de los monómeros de Stl^{N-ter} unidos a ϕ 11Dut, mDut y hDut con Stl^{N-ter} en conformación *apo*.

Tabla R. 1. 5. Interacciones residuo-residuo en el modelo del complejo SaPIBov1-Stl^{N-ter} - EBVDut.

Tabla R. 2. 1. Parámetros cinéticos de EnMazG y SeMazG.

Tabla R. 2. 2. Pesos moleculares de EnMazG y SeMazG no mutantes y mutantes de arginina calculados por SEC-MALS.

Tabla R. 2. 3. Residuos equivalentes en EnMazG y SeMazG que median la tetramerización.

Tabla R. 2. 4. Pesos moleculares de EnMazG y SeMazG no mutantes y mutantes de la zona de tetramerización calculados por SEC-MALS.

Tabla R. 2. 5. Tabla de colección de datos cristalográficos y estadísticos de refinado.

Tabla R. 2. 6. Tabla de valores de RMSD en Å para la superposición de los tetrámeros de EnMazG en complejo con el dUpNHpp y cationes Mg²⁺, Mn²⁺, y las ratios molares Mn²⁺: Mg²⁺ 1: 2,5, 1:4, y 1:8.

Tabla R. 2. 7. Tablas recopilatorias de la señal anómala de cada catión en las cuatro estructuras de EnMazG en presencia de Mn²⁺ resueltas en esta Tesis.

Tabla R. 2. 8. Resultados agrupados del análisis de señal anómala del catión Mn²⁺ en los centros activos de las cuatro estructuras de EnMazG resueltas en esta Tesis y los centros activos de la estructura de EnMazG en complejo con dUpNHpp y Mg²⁺.

Tabla R. 2. 9. Parámetros cristalográficos de la estructura de EnMazG en presencia de Mg²⁺ y Ca²⁺.

Tabla R. 2. 10. Tabla de valores de RMSD en Å para la superposición del tetrámero de EnMazG en conformación *apo* con el tetrámero de EnMazG en presencia de los cationes divalentes de Ca, Ni y Zn, y con la estructura en complejo con el dUpNHpp y el Mg²⁺.

Tabla R. 2 .11. Geometrías moleculares de los distintos cationes divalentes utilizados en esta Tesis, junto con las distancias de los enlaces establecidos entre éstos y átomos de carbono u oxígeno, y entre moléculas de agua.

Tabla R. 3. 1. Tabla recopilatoria de los tampones probados sin éxito para solubilizar la proteína ADN helicasa dependiente de ATP en las tres aproximaciones.

Tabla R. 3. 2. Tabla recopilatoria de los tampones en la prueba de solubilidad sin éxito para la proteína Rv1429 (regulador transcripcional de la familia PucR).

Tabla R. 3. 3. Tabla recopilatoria de los tampones en la prueba de solubilidad sin éxito para la proteína Rv3354.

DISCUSIÓN

Figura D. 1. El Stl^{N-ter} mimetizaría el nucleótido en el centro activo de la EBVDut.

Figura D.2. Modelados *in-silico* de las Duts de *D. virilis* y *C. elegans*.

Figura D. 3. Alineamientos de las 3 copias de los genes de Dut de *C. elegans* y *D. virilis*.

Figura D. 4. Localización del extremo C-terminal en la estructura de tres complejos de Stl^{N-ter} con tres Duts triméricas.

Figura D.5. Localización del *lb-loop*.

Figura D.6. La Dut de *E. faecalis* ordena el motivo V sobre el centro activo de la misma subunidad, y el *lb-loop* podría suponer un impedimento estérico para la interacción con Stl.

Figura D.7. Superposición de las estructuras tridimensionales de la hDut y la Dut del MPV en complejo con el sustrato no hidrolizable α,β -imido-dUTP.

Figura D. 8. Localización del catión 3 en Duts diméricas.

Figura D. 9. Alineamiento de secuencias de Duts diméricas (de fago de *Staphylococcus aureus* ϕ DI y ϕ O11) y MazG (de fago de *E. faecalis* y *S. epidermidis*, EnMazG y SeMazG).

Figura D.10. Localización del catión 3 en Duts diméricas y en EnMazG.

Figura D. 11. Localización de los cationes en Duts diméricas y diferentes MazG.

Figura D.12. Superposición de los análogos no hidrolizables del sustrato dUTP en la estructura en complejo con DrMazG con la estructura en complejo con EnMazG, y con la estructura en complejo con Dut dimérica del fago ϕ DI de *S. aureus*.

Figura D. 13. MazZ tiene una mayor similitud con EnMazG que con Dut ϕ DI.

Figura D. 14. Superposición del análogo de sustrato no hidrolizable dUpNHp, cationes y residuos coordinantes de éstos últimos de las estructuras de Dut de *C. jejuni* y MazG del cianofago S-2L.

Figura D. 15. La disposición del centro activo en Duts diméricas en ambos sustratos difiere en la participación del tercer catión en el caso de dUDP y en un pequeño desplazamiento de la molécula de agua catalítica (Acat).

RESUMEN

El dUTP es el nucleótido no canónico mayoritario en la célula. Su errónea incorporación en el ADN dispara los sistemas de reparación, lo cuales pueden causar la muerte celular por una actuación ininterrumpida.

Con el objetivo de prevenir esta situación, y conocida la incapacidad de las ADN polimerasas de distinguir entre el dTTP y el dUTP, las células poseen las proteínas dUTPasas (Duts), capaces de degradar este nucleótido no canónico y, en consecuencia, aumentar la ratio dTTP/dUTP en el *pool* de nucleótidos celular.

A parte de la hidrólisis del dUTP, las Duts poseen otras funciones *moonlighting*, entre las cuales destaca la capacidad de Duts de fagos de *Staphylococcus aureus* de interaccionar con el principal represor de sus islas de patogenicidad (SaPIs, de sus siglas en inglés *Staphylococcus aureus* Patogenicity Islands), StI, desreprimiendo la isla y permitiendo su transducción.

En Tesis anteriores desarrolladas en el laboratorio donde se ha realizado ésta, se ha descifrado el mecanismo molecular de estas Duts, tanto diméricas como triméricas, en su interacción con el represor StI de la SaPI_{bov1}, el cual utiliza una estrategia de mimetización del sustrato de la Dut. En esta Tesis se ha confirmado la utilización de un mecanismo de interacción común de StI a Duts triméricas de fago, Duts triméricas de procariontes, y Duts triméricas de eucariotes. Para ello se ha resuelto la estructura tridimensional de StI^{N-ter} en complejo con Dut de *Mycobacterium tuberculosis* y con Dut humana.

Además, se ha confirmado la interacción de StI con la tercera familia de Duts, las monoméricas, de las cuales aún no se había descrito la interacción y con la cual se ratifica el carácter universal del StI como inhibidor de Duts. Se propone, asimismo, un mecanismo de interacción muy similar al utilizado con las Duts de fago triméricas y en el que participarían residuos de los 5 motivos conservados.

Por otro lado, la Dut de *M. tuberculosis* posee un *loop* específico de género de 5 aminoácidos, cuya presencia determina la supervivencia de la bacteria. En esta Tesis se ha intentado identificar proteínas del propio organismo capaces de interaccionar con la Dut y el mecanismo molecular utilizado, con la finalidad de ampliar las funciones *moonlighting* descritas para las Duts y descifrar el posible papel de este *loop* en el proceso.

Finalmente, junto a las Duts monomérica y triméricas, se ha hecho hincapié en una familia de proteínas íntimamente relacionadas, las MazG, propuestas como el ancestro de las Duts diméricas. En esta Tesis, se ha analizado el mecanismo molecular de la inhibición en presencia de diversos cationes divalentes sustituyentes al Mg²⁺, con el objetivo de descifrar si estas proteínas podrían poseer alguna función *moonlighting*.

INTRODUCCIÓN

1. Estabilidad genómica y dUTPasas.

La estabilidad del genoma depende en parte de la fidelidad en la replicación del ADN. Por ello, es de elevada importancia un fino ajuste de los niveles del *pool* de nucleótidos en la célula, que prevendrán una incorporación errónea de nucleótidos no canónicos en el ADN o el ARN. Para ello, las células han desarrollado un vasto surtido de enzimas capaces de degradar y sintetizar desoxinucleótidos.

El metabolismo celular constantemente genera desoxinucleótidos trifosfato no-canónicos como el dUTP, dITP, dXTP, 8-oxo-dGTP o 2-oxo-dATP (Tormo-Más et al. 2013). El dUTP es el desoxinucleótido no canónico mayoritario en la célula, y para mantener un nivel adecuado existen, en todos los reinos de la vida, una variedad de enzimas con actividad dUTPasa que se agrupan en familias estructural y funcionalmente distintas, y que juegan un papel de sanitización celular (Varga et al. 2007).

La concentración de dUTP en la célula se establece a través de un equilibrio entre reacciones enzimáticas que lo producen y otras que lo consumen. El dUTP se puede formar por la acción de una NDP quinasa sobre el dUDP; por la acción de la ribonucleótido reductasa (RNR) utilizando el UTP como sustrato; y por desaminación espontánea o catalizada del dCTP. Asimismo, el dUTP es consumido en la reacción catalizada por la dUTPasa (Dut) que lo hidroliza a dUMP y PPi, (Figura I.1) proporcionando dUMP como precursor para la síntesis de los nucleótidos de timina (Figura I. 2).

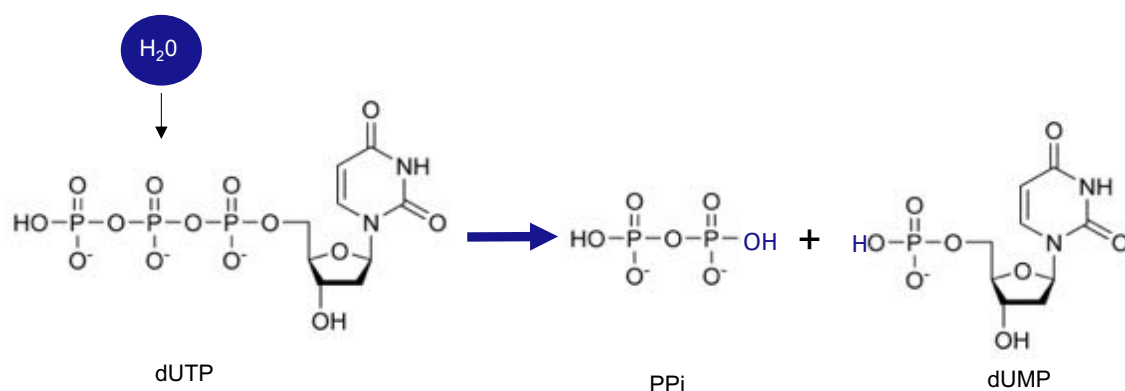


Figura I.1. Reacción enzimática catalizada por dUTPasa (Dut). El sustrato dUTP es hidrolizado a dUMP y PPi.

INTRODUCCIÓN

En el proceso de replicación del ADN, el dUTP compete con el dTTP, debido a la naturaleza inespecífica de las ADN polimerasas implicadas, incapaces de diferenciar entre dUTP y dTTP (Nyman 2001). De ahí el papel crucial jugado por las Duts para mantener una ratio dTTP : dUTP elevada que evita o minimiza la incorporación errónea del uracilo en el ADN (Mol et al. 1996; Kerepesi et al. 2016; Nyman 2001). No obstante, junto a la prevención por parte de las Duts de la incorporación de uracilo al ADN, también existe un mecanismo alternativo de sanitización a través de la enzima uracilo-ADN-glicosilasa (UDG), que escinde el uracilo ya incorporado en el ADN e inicia el sistema de reparación por escisión de base. Sin embargo, la incorporación excesiva de uracilo provoca una sobrecarga de los sistemas de reparación del ADN, causando roturas en las cadenas del ácido desoxirribonucleico y, en los casos más graves, ocasiona la muerte celular.

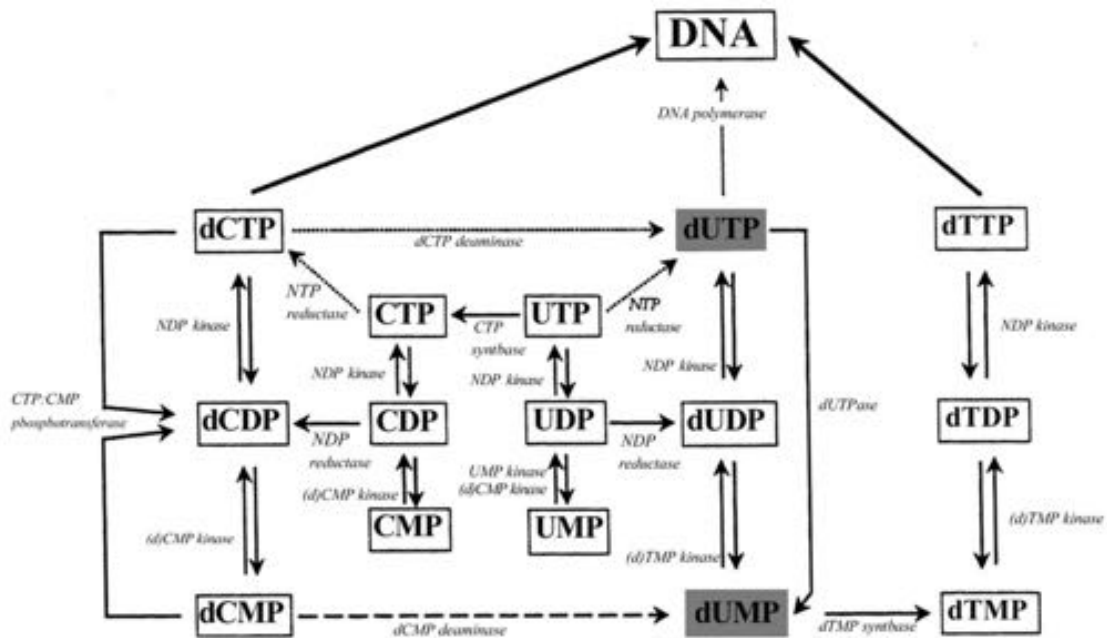


Figura I.2. Vías de producción y consumo del dUTP en la célula. Se identifican las enzimas individuales junto a las correspondientes flechas. Las flechas punteadas indican actividades enzimáticas que solo se han encontrado en *E. coli* u otras enterobacterias, mientras que las flechas en guiones refieren a actividades encontradas en la mayoría de organismos pero no en *E. coli*. Figura tomada de Nyman 2001.

2. Clasificación de las Duts.

Las Duts son consideradas enzimas ubicuas, pero no todas ellas son iguales. Bajo este nombre se agrupan enzimas (EC 3.6.1.23) que presentando la misma actividad enzimática poseen características funcionales y estructurales diferentes. Este hecho ha llevado a clasificar las Duts en tres familias en función de su estado de oligomerización: monoméricas, diméricas y triméricas. A continuación ahondaremos en cada una de las familias de Duts, pero a modo introductorio, en la Figura I.3. se muestra una Dut representativa de cada familia. Resulta evidente la mayor similitud estructural entre las Duts monoméricas y triméricas respecto a las diméricas.

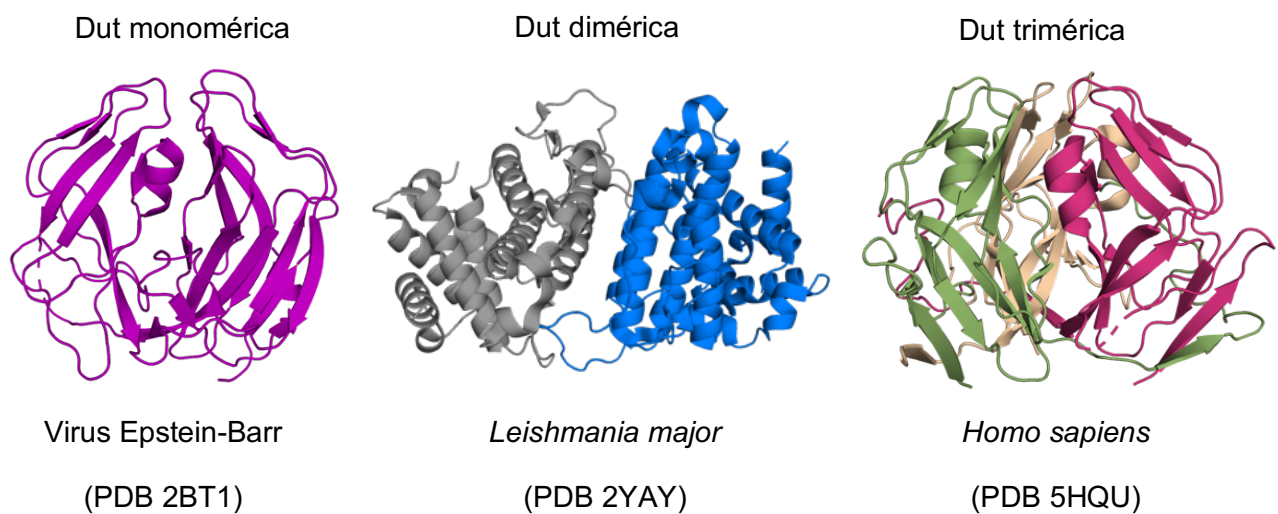


Figura I.3. Estructuras representativas de Duts de las tres familias. En cada estructura de Dut se indica la familia perteneciente (arriba) y el organismo origen junto al código PDB (abajo). Cada monómero está distintamente coloreado.

2.1. Duts triméricas.

Las Duts triméricas son las más extendidas en la naturaleza, y se encuentran en animales, plantas, hongos, bacterias y virus (tanto de ADN como de ARN) (Nyman 2001), por ello las describimos en primer lugar. La secuencia aminoacídica de estas Duts presenta 5 motivos distribuidos a lo largo de la estructura, lo cuales están ampliamente conservados entre las Duts de esta familia y constituyen su firma identificativa (motivos I a V) (Figura I. 4 e I. 5). Además de los cinco motivos consenso, algunas Duts triméricas presentan un motivo extra, de longitud y secuencia muy variable, denominado motivo VI.

INTRODUCCIÓN

Existe una amplia caracterización estructural (más de 120 estructuras en el PDB) de esta familia de Duts, con múltiples representantes de los diferentes reinos. Esto ha mostrado que todas ellas presentan un plegamiento básico conservado, con una estructura cuaternaria caracterizada por formar homotrimeros donde cada subunidad presenta un plegamiento fundamentalmente en hebras β . Específicamente, cada monómero está constituido por un conjunto de nueve hebras β ($\beta 1$ - $\beta 9$) que se agrupan en 2 conjuntos de 4 hojas, además de una única hélice α entre las hojas $\beta 5$ y $\beta 6$, y un extremo C-terminal muy flexible sin estructura secundaria (Figura I. 5).

Los monómeros se agrupan formando el trímero alrededor de un canal central (Figura I. 6) cuya naturaleza es distinta entre Duts y puede incluir la participación de moléculas de agua e/o iones, jugando un papel relevante en la actividad y estabilidad del enzima. Por ejemplo, la naturaleza hidrofóbica del canal central de la Dut de *E. coli*, frente a la polaridad del canal de las Duts humana o de *Drosophila*, se ha relacionado con su mayor estabilidad al calor o agentes químicos (Vértessy and Tóth 2009).

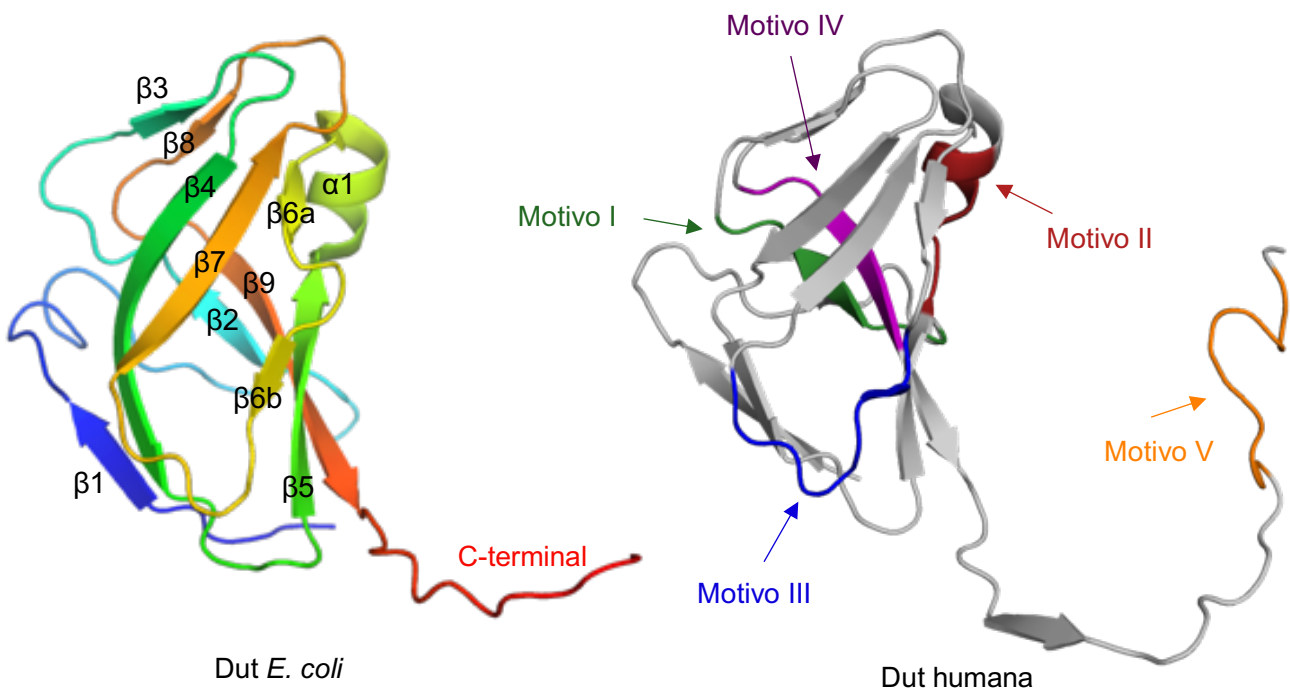


Figura I. 5. Localización estructural de las estructuras secundarias y motivos conservados en Duts triméricas. (Izquierda) Estructura del monómero de Dut de *E. coli* donde se identifican las estructuras secundarias. (Derecha) Estructura del monómero de Dut de humano donde se identifican los motivos conservados en distintos colores. Códigos PDB: 1EUW *E. coli*, 1Q5U humano.

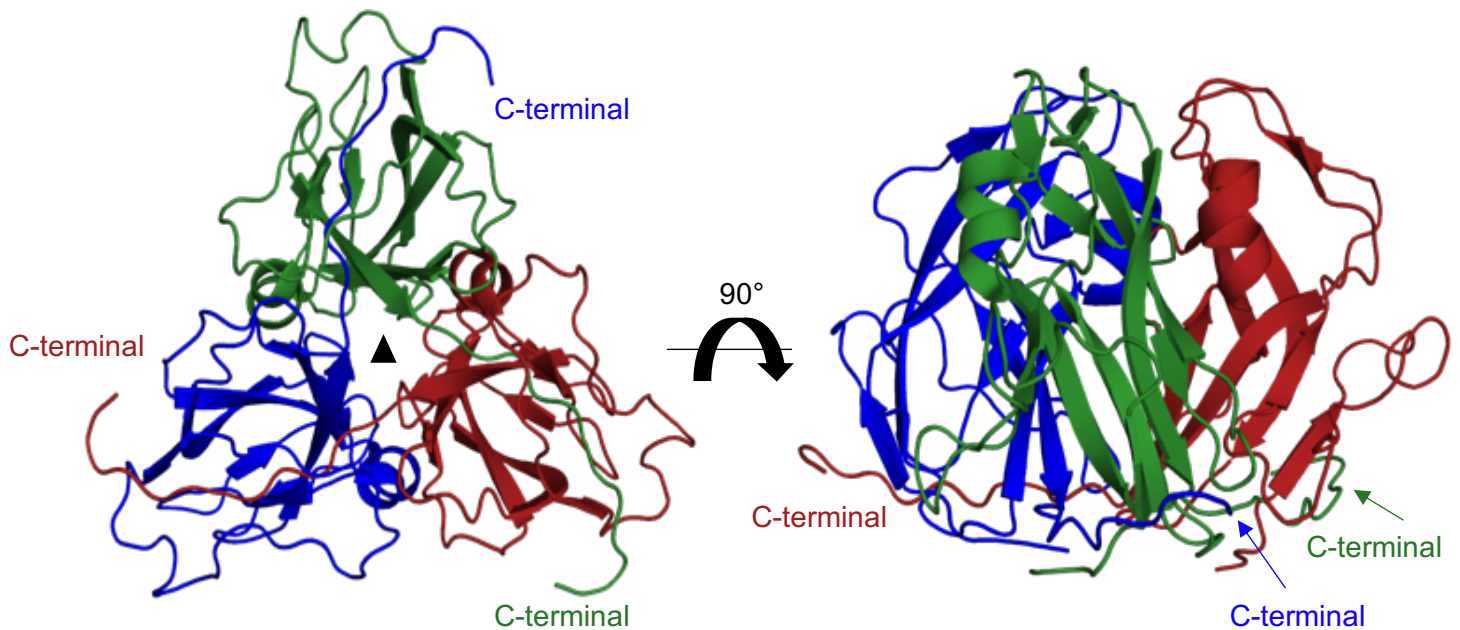


Figura I.6. Estructura tridimensional de Dut de *E. coli*. Se muestran dos vistas ortogonales. Se identifican los segmentos C-terminales de Dut, para la visualización del intercambio de este fragmento entre subunidades, así como el canal central del trímero con un triángulo negro (izquierda). Código PDB 1EUW.

2.1.1. Centro activo.

Estructuras de Duts en complejo con análogos del sustrato o producto muestran que el trímero de Dut presenta 3 centros activos independientes, cada uno de ellos localizado en la interfase de dos monómeros. En cada centro activo contribuyen los tres monómeros. Uno de los monómeros contribuye con los motivos I, II y IV; un segundo monómero aporta el motivo III; y un tercer monómero aporta el motivo V (Figura I. 7).

Los motivos I, II y IV participan en la unión del trifosfato del sustrato (dUTP) y el ión metálico. Específicamente, los motivos II y IV mantienen interacciones con los fosfatos α y β , mediante un residuo de Ser y Arg del motivo II respectivamente; y una Gln conservada del motivo IV que también interacciona con un oxígeno del fosfato α , y coordina la molécula de agua que llevará a cabo el ataque nucleofílico (agua catalítica). Por su parte, dos residuos de Asp del motivo I coordinan indirectamente al catión requerido como cofactor a través de la interacción con dos de las 3 moléculas de agua que lo coordinan. Además, el

INTRODUCCIÓN

catión es coordinado directamente por un oxígeno de cada uno de los tres fosfatos.

El motivo III participa en la unión del nucleósido y aporta un residuo de Asp clave en el inicio de la catálisis. Este motivo determina la especificidad por el uracilo y la consecuente discriminación de otras bases nitrogenadas mediante dos mecanismos. Por una parte, mediante exclusión estérica que impide la entrada de purinas y ribosa. Los residuos del motivo III tapizan el fondo del centro activo donde se acomoda el anillo del nucleótido y el azúcar, e impiden por restricción del espacio que un anillo de purina pueda acomodarse. Además, el residuo de Tyr conservada se dispone frente al azúcar del nucleótido apilando sobre él el anillo de su cadena lateral, descartando a la ribosa por la presencia del oxígeno extra (Figura I. 8). Por otra parte, una red de enlaces de hidrógeno específicos establecidos por los residuos de este mismo motivo con el uracilo lo discrimina frente a otras pirimidinas.

El motivo V, situado en el extremo C-terminal de la Dut, se sitúa sobre el centro activo a modo de tapa mediante su estructura flexible en forma de *P-loop*, creando un recinto cerrado para la posterior hidrólisis del sustrato. Específicamente, un residuo aromático (habitualmente una fenilalanina o histidina) posiciona su anillo sobre el anillo de uracilo (Figura I.7 e I. 8) (Tormo-Más et al. 2013; Mol et al. 1996; Tóth et al. 2007; Vértessy and Tóth 2009; Chan et al. 2004).

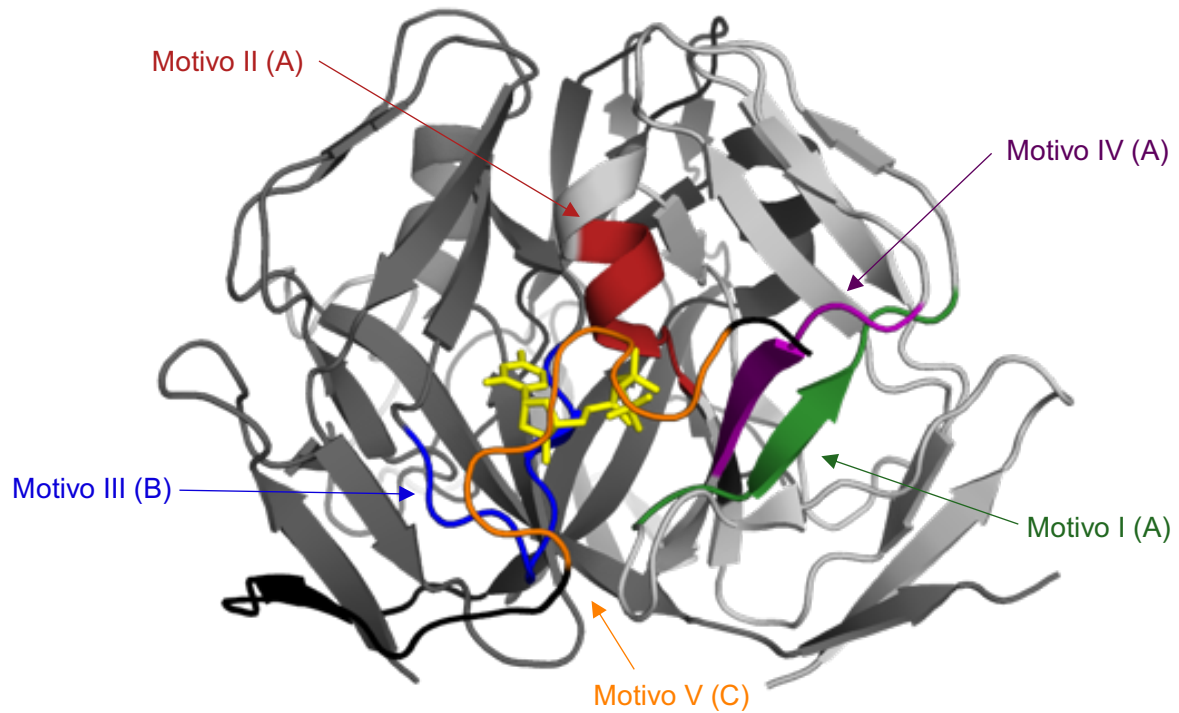


Figura I. 7. Participación de las tres subunidades del trímero de Dut en cada uno de los centros activos. Se localizan sobre la estructura trimérica de Dut de humano, en uno de los tres centros activos, los motivos conservados en distintos colores. Entre paréntesis, se identifica a qué monómero corresponden. Los motivos I, II y IV son proporcionados por el monómero A, el motivo III es proporcionado el monómero colindante (B), y el motivo V es proporcionado por el monómero restante (C). Código PDB: 2HQU.

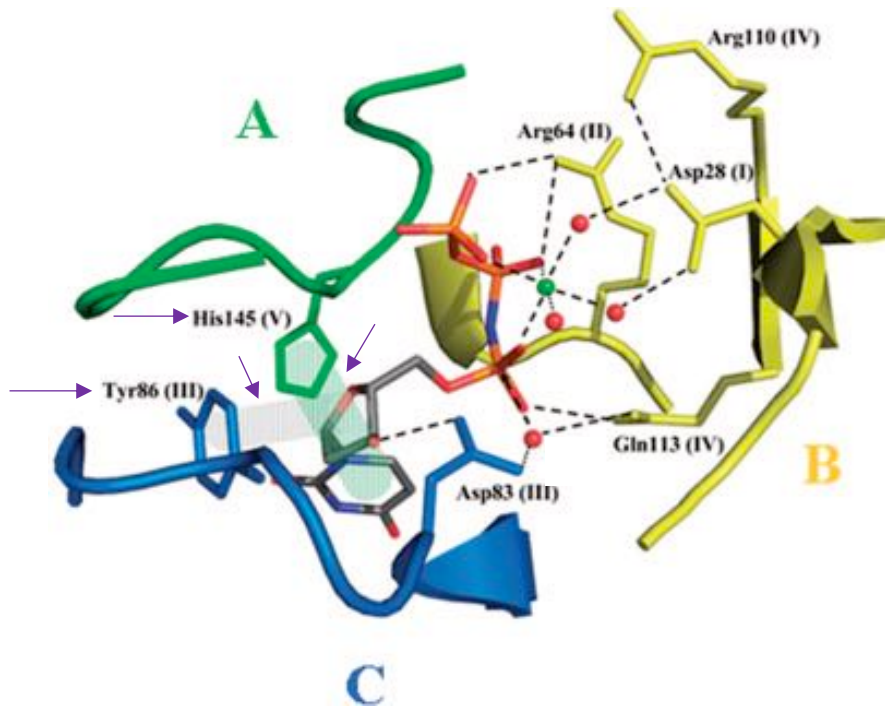


Figura I. 8. Especificidad por uracilo en Dut de *M. tuberculosis* y participación de los motivos conservados en el centro activo. En flechas moradas se identifican dos residuos, Y86 del motivo III y H145 del motivo V, claves en la discriminación de otros nucleótidos a favor del uracilo. La Y86 establece interacciones de apilamiento con la desoxirribosa y la H145 las establece con el uracilo, que se representan con rectángulos de color gris y verde respectivamente, y se señalan también con flechas moradas. En cada centro activo participan residuos de los 5 motivos conservados procedentes de las distintas subunidades: motivos I, II y IV del monómero B, motivo III del monómero C, y motivo V del monómero A. Se especifican los motivos a los que corresponden los residuos entre paréntesis, y se representan en varillas coloreadas según la subunidad a la que pertenecen (A en verde, B en amarillo, C en azul). Se representa en varillas el nucleótido coloreado por átomos (C en gris, O en rojo, N en azul y P en naranja), en esferas rojas las moléculas de agua, el ión Mg^{2+} en una esfera verde, y las interacciones entre los elementos en líneas discontinuas. Código PDB 2PY4. Figura adaptada de *Vértessy and Tóth 2009*.

2.1.2. Mecanismo de hidrólisis.

La arquitectura del sitio activo está altamente interconectada, con todos los motivos conservados proporcionados por las distintas subunidades del trímero. Una vez ha sido seleccionado el dUTP y es posicionado correctamente por las interacciones selectivas, un residuo de aspártico conservado del motivo III establece interacciones con el grupo 3' OH de la desoxirribosa y a la vez coordina una molécula de agua que actuará durante la catálisis (agua catalítica). Cuando todos los participantes del centro catalítico ocupan su posición, el extremo C-terminal que incluye al motivo V cierra el centro activo, y con ello, acerca los

INTRODUCCIÓN

átomos de los reactivos para la catálisis, ya que el ión Mg^{2+} está unido a la cadena de fosfato pero débilmente.

La catálisis es iniciada por el ataque nucleofílico de la molécula de agua catalítica, que se posiciona en línea con el fosfato α y es activada por el residuo de ácido aspártico del motivo III (Figura I. 9). Tras el ataque nucleofílico sobre el fosfato α , la molécula de agua une el ión hidroxilo al fosfato α , rompiendo en dirección $\alpha \rightarrow \beta$ el enlace fosfoanhídrido que une ambos fosfatos, y formándose como productos dUMP y PPI. Tras la rotura del enlace, se produce la salida de los productos, la cual se ha propuesto como un proceso secuencial. Primeramente se liberaría el producto $PPI-Mg^{2+}$, que implicaría una pérdida de contactos entre ambos fosfatos y el motivo V, llevando al desorden de éste y en consecuencia a la pérdida de la interacción del residuo clave fenilalanina (o histidina) con el anillo de uracilo, liberándose entonces el producto dUMP. La liberación de los productos dejaría el centro activo libre y disponible para empezar un nuevo ciclo de hidrólisis (Figura I. 10) (Nyman 2001; Barabás et al. 2004; Varga et al. 2007; Tóth et al. 2007).

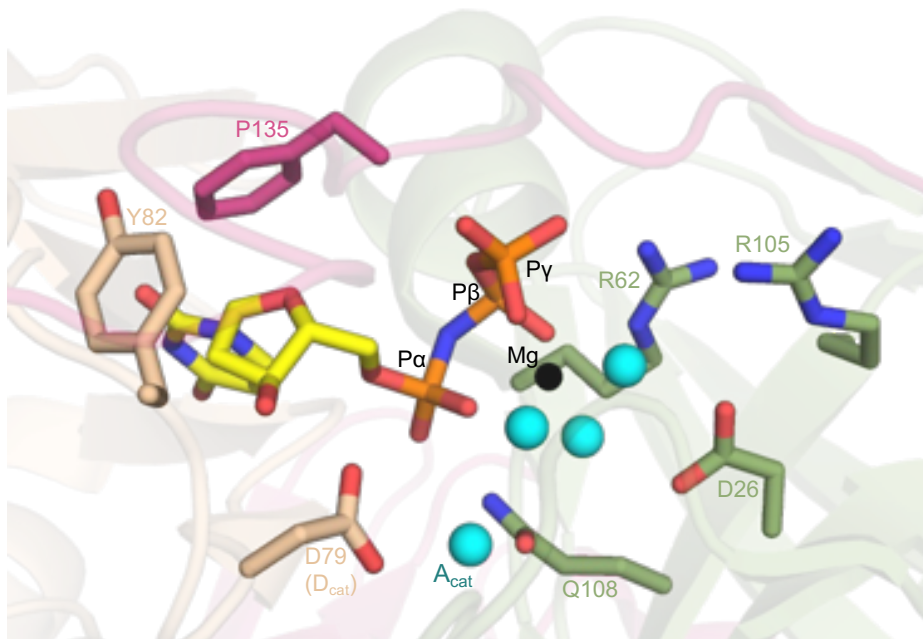


Figura I. 9. Vista detallada del centro activo de Dut de humano. La fracción del nucleótido correspondiente con el uracilo y la desoxirribosa es posicionada mediante interacciones con dos residuos aromáticos (Y82 y P135), mientras que los tres fosfatos coordinan el ión Mg^{2+} (esfera negra) que actúa como cofactor, el cual también es coordinado por tres moléculas de agua (esferas cian), coordinadas a su vez por residuos de arginina y aspartato (R62, D26). Para la catálisis, el agua catalítica (A_{cat}) se posiciona en línea con el fosfato α ($P\alpha$) gracias a la coordinación con el aspártico catalítico (D_{cat}) y un residuo de glutamina (Q108), e inicia el ataque nucleofílico. Tanto el nucleótido como las cadenas laterales se representan en varillas, con los carbonos en color verde, magenta o trigo según corresponden a una u otra subunidad del trímero de Dut, o en amarillo en caso del sustrato. Los átomos de O aparecen en rojo y los de N en azul. Código PDB: 2HQU.

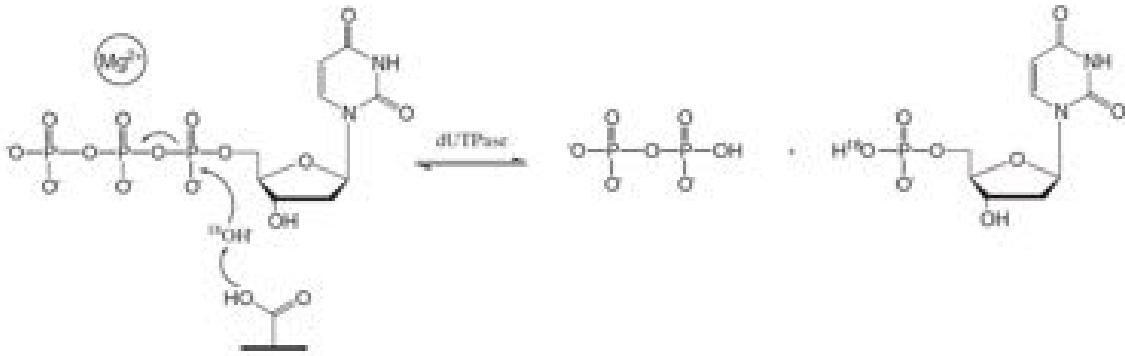


Figura I. 10. Esquema de la reacción de hidrólisis catalizada por Duts triméricas. El aspártico catalítico polariza a la molécula de agua, que inicia el ataque nucleofílico sobre el fosfato α y en dirección $\alpha \rightarrow \beta$. El dUMP y el PPI son generados como productos de la hidrólisis. Figura tomada de *Hemsworth, González-Pacanowska, and Wilson 2013*.

2.1.3. El motivo V.

El motivo V en el extremo C-terminal es el único fragmento de las Duts triméricas que sufre cambios conformacionales de importancia durante el ciclo enzimático (Dubrovay et al. 2004; Vértessy and Tóth 2009). Este fragmento C-terminal es un polipéptido rico en glicinas, y esto es una característica relativamente común en proteínas que unen nucleótidos, como las quinasas, aunque su secuencia consenso, Arg-Gly-X-Gly-Gly-Phe-Gly, es propia de esta familia de enzimas. Se le denomina habitualmente *P-loop* por ser un motivo homólogo en secuencia al motivo denominado con ese nombre que está presente en proteínas ATPasas y GTPasas para la interacción con fosfatos (conocidas como *P-loop* NTPasas) (G.S. Prasad 2001). Tanto en Duts como en las quinasas este *loop* C-terminal cubre el nucleótido unido casi completamente, pero mientras en las Duts un residuo de fenilalanina conservado (mayormente, aunque se han descrito con histidina o tirosina) se apila sobre el uracilo, en las quinasas hay una valina conservada que se apila sobre el anillo hidrofóbico de adenina del ATP. La presencia de múltiples residuos de glicina responde a diferentes características de este *P-loop* como aportar flexibilidad y eliminar problemas estéricos con los fosfatos en su reconocimiento y estabilización en el centro activo mediante su cadena principal (Mol et al. 1996).

En cuanto a la disposición relativa del motivo V, en Duts triméricas éste se posiciona sobre el centro activo de la subunidad del monómero más distal a modo de tapa, haciendo la contribución del tercer monómero al centro activo. No obstante, ha sido descrito que en la Dut de betaretrovirus el centro activo solo requiere la participación de dos subunidades del trímero. Esto es debido a que el motivo V se ordena sobre el centro activo de la misma subunidad que aporta el motivo III y no en el de la subunidad adyacente (Németh-Pongrácz et al. 2007). Esta reordenación es debida al acortamiento de residuos entre los motivos IV y

V, el cual induce un cambio en la direccionalidad del extremo C-terminal volviendo sobre la misma subunidad en lugar de proyectarse hacia las adyacentes. Este acortamiento no afectaría a la parte final del fragmento donde se encuentra el motivo V, manteniendo idéntica la disposición e interacciones en el centro activo (Németh-Pongrácz et al. 2007).

Otras proteínas que unen nucleótidos y homólogas a las Duts, como las superfamilias dCTP desaminasas (DCD) y la dCTP desaminasa/Dut (DCD-DUT), no poseen motivo V como las Duts y su constante de especificidad (k_{cat}/K_M) es 3 órdenes de magnitud menor a la calculada en Duts. En el caso de las enzimas bifuncionales dUDP/dUTPasa tampoco poseen *P-loop* ni otro elemento estructural de la coordinación del fosfato γ (Pécsi et al. 2011).

2.2. Duts monoméricas.

Las Duts monoméricas constituyen la familia de Duts minoritaria, y solo están presentes en herpesvirus, como el virus Epstein-Barr. Como su nombre indica, las Duts monoméricas no forman oligómeros. Estas Duts también poseen 5 motivos conservados en su secuencia, idénticos a los motivos de las Duts triméricas pero dispuestos en la cadena polipeptídica en distinto orden (Figura I.11). Aunque entre ambas familias de Duts la homología de secuencia es baja, entre el 14 y el 21 %, la estructura tridimensional de Dut monomérica ha mostrado que esta familia de enzimas posee 3 dominios y dos de ellos se pliegan de forma similar a Duts triméricas (Figura I.12). Los dominios I y II tienen el mismo plegamiento que dos monómeros de una Dut trimérica mientras que el dominio III es bajo en estructura secundaria y tendría una función estructural para mantener una disposición adecuada para la catálisis de los otros dos dominios. (Figura I.13) (Tarbouriech et al. 2005).

La homología estructural y conservación de secuencia es reflejo de un centro activo que es extremadamente similar al de las Duts triméricas (Tarbouriech et al. 2005). En el caso de Duts monoméricas, el centro activo no se dispone en la interfaz entre dos subunidades, si no entre dos dominios, I y II, y en su formación participan del mismo modo los motivos I a IV, y el motivo V resulta también esencial en la catálisis al posicionarse sobre el centro activo con su residuo de fenilalanina interaccionado con el anillo del nucleótido. La molécula de agua catalítica inicia el ataque nucleofílico para la hidrólisis del dUTP (Tarbouriech et al. 2005).



Figura I. 11. Alineamiento de secuencias de Duts triméricas y la DUT monomérica representativa del virus de Epstein-Barr. Se encuadran los 5 motivos conservados, tanto en Duts triméricas como monoméricas. Motivo I en verde, motivo II en azul, motivo III en rojo, motivo IV en amarillo y motivo V en negro. Se resaltan en color gris claro, gris oscuro o amarillo los residuos conservados en un 40, 60 o 80 %, respectivamente. Los acrónimos corresponden con *Homo sapiens* (Hsap), *E. coli* (Ecol), *M. tuberculosis* (Mtub), virus de la anemia infecciosa equina (EIAV), dominio 1 de la DUT del virus de Epstein-Barr (N), y dominios II y III de la misma (C). Figura adaptada de *Tarbouriech et al. 2005*.

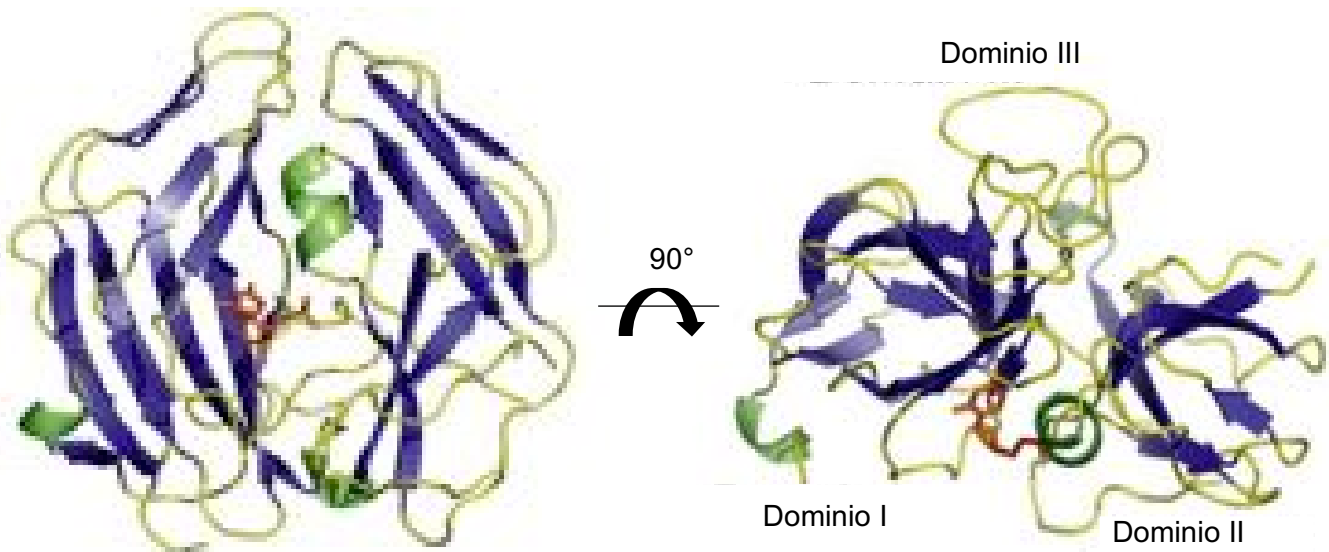


Figura I. 12. Estructura tridimensional de la DUT del virus de Epstein-Barr. Se muestran dos vistas ortogonales y se localizan los tres dominios (*derecha*). Los dominios I y II tienen el mismo plegamiento que un monómero de DUT trimérica, mientras que el dominio III solamente posee un poco de estructura secundaria. Las hojas β , hélices α y los *loops* están coloreados en azul, verde y amarillo respectivamente. El sustrato se representa en varillas y en color naranja. Figura adaptada de *Tarbouriech et al. 2005*.

INTRODUCCIÓN

La homología estructural y la conservación de secuencia en los 5 motivos que constituyen el único centro activo han llevado a proponer que el gen de la Dut monomérica proviene de duplicación o triplicación y fusión del gen de Dut trimérica. En consecuencia, los otros cinco motivos que son redundantes se localizan en la superficie expuesta de la enzima, y debido a que no están sometidos a selección por no participar en la formación de centros activos, muestran diferencias considerables en cuanto a sustituciones y longitud de secuencia (Tarbouriech et al. 2005). El centro activo formado por los motivos I a IV ajustan al tamaño del dUTP y discriminan contra otras moléculas similares como el UTP o el dCTP (Freeman et al. 2009).

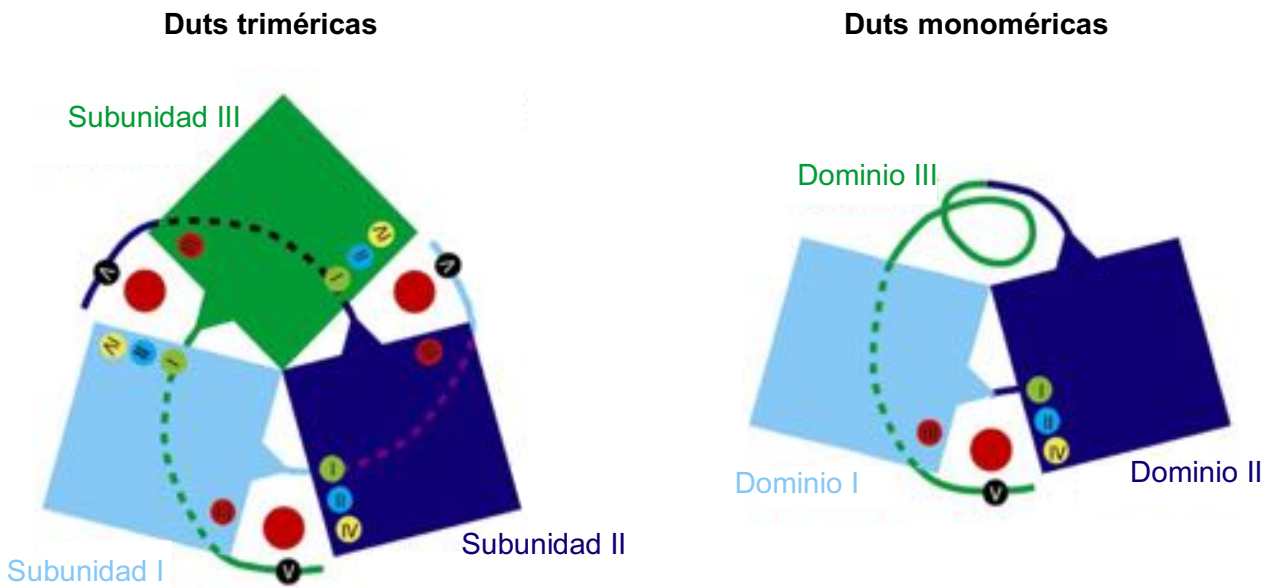


Figura I.13. Diagrama representativo de la organización tridimensional de las subunidades de Duts triméricas (izquierda) y de los dominios de Duts monoméricas (derecha). Se localizan los motivos conservados I - V representados en esferas de distintos colores alrededor del centro activo (representado con una esfera roja). Figura adaptada de Tarbouriech et al. 2005.

2.3. Duts diméricas.

Las Duts diméricas representan un grupo reducido presente en protozoos como *Leishmania* y *Trypanosoma*, algunas especies bacterianas como *Campylobacter jejuni*, o en bacteriófagos como el T4, y fagos de *Staphylococcus aureus* como ϕ DI y ϕ O11 (Nyman 2001; Donderis et al. 2017; Rafael Ciges-Tomas et al. 2019).

La caracterización estructural de representantes de esta familia muestra que su plegamiento está totalmente formado por hélices α , siendo por ello completamente diferentes a las Duts triméricas y monoméricas formadas mayoritariamente por hojas β . El centro activo de las Duts diméricas también está formado por cinco motivos característicos de esta familia, aunque no guardan relación con los motivos de las enzimas triméricas (Figura I. 14) (Tormo-Más et al. 2013).

Cada monómero de Dut dimérica está formada por 10 a 12 hélices α (10 en el caso de la Dut de *Leishmania major*, 11 en *C. jejuni* y 12 en *Trypanosoma cruzi*). Todas ellas se organizan en dos dominios, uno rígido y otro móvil, según si se mantiene invariable tras la unión del sustrato en el centro activo o lleva a cabo cambios estructurales sustanciales para acomodar los nucleótidos (pudiendo llegar a ser hasta de 20 Å respecto de la disposición *apo*), respectivamente (Figura I. 15). Ambos dominios se disponen en 60 ° uno respecto del otro.

A parte de esta estructura común entre Duts diméricas en las que se distribuyen los 5 motivos conservados, existen algunas Duts de esta familia que al igual que las Duts triméricas poseen además una parte central variable entre 50 y 60 residuos denominada motivo VI, flanqueada por los motivos IV y V, y que se pliega igualmente en hélices α (Donderis et al. 2017).

Cada monómero oligomeriza formando homodímeros a través del dominio rígido de ambas subunidades, organizadas alrededor de un eje de simetría 2, y dejando los dominios móviles hacia el exterior. La zona de dimerización implica principalmente a las interacciones entre los dos dominios rígidos, y a la inserción del *loop* de una subunidad en la hendidura de la otra subunidad (Figura I. 16) (Harkiolaki et al. 2004).

INTRODUCCIÓN

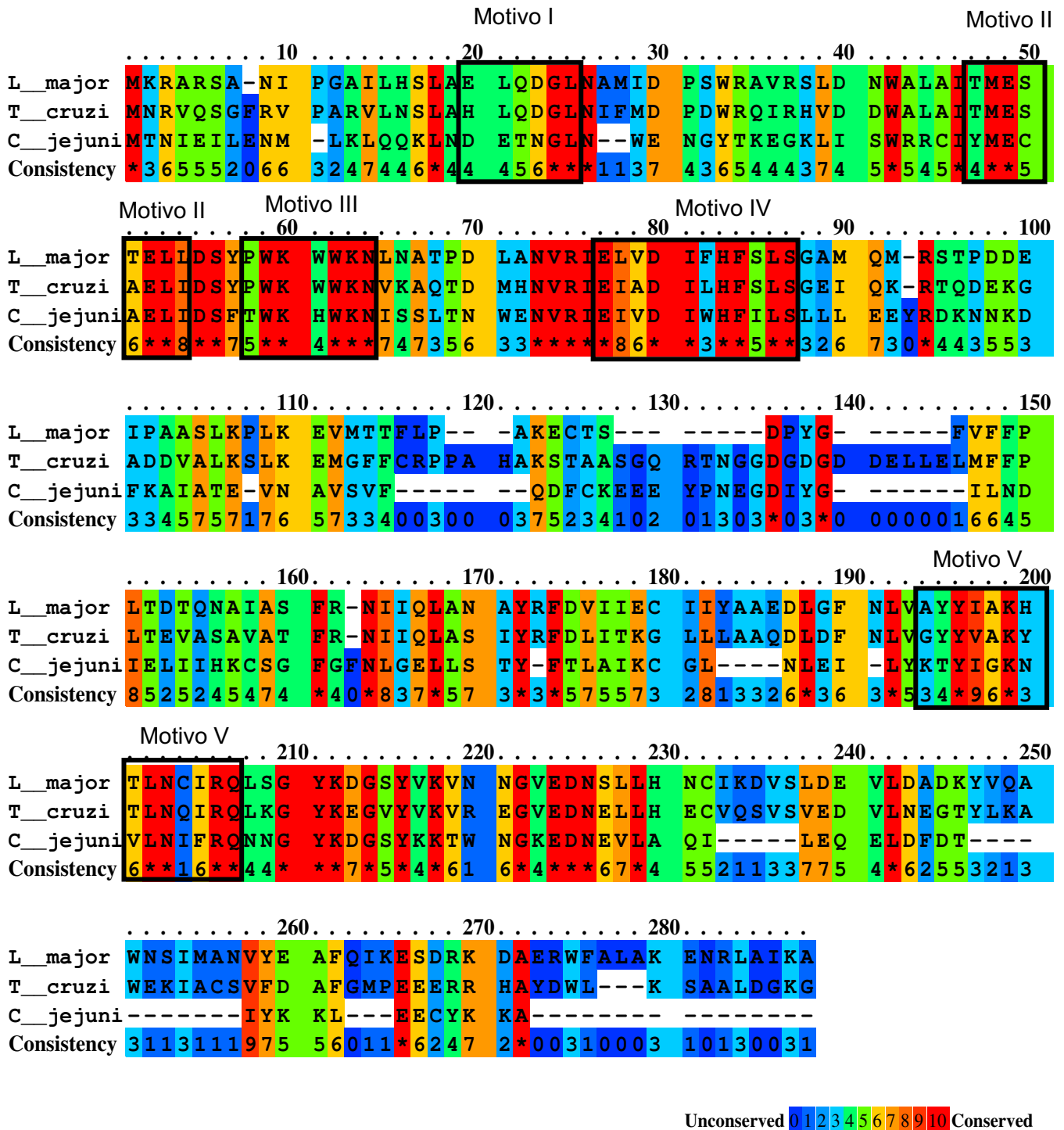


Figura I. 14. Alineamiento de secuencias de Duts diméricas. Se incluyen las secuencias aminoácidas de Duts de *L. major*, *T. cruzi* y *C. jejuni*. Se encuadran y numeran los motivos conservados I – V. La conservación de residuos se indica con una gradación de colores desde rojo a azul de mayor a menor grado de conservación. Alineamiento realizado en el servidor PRALINE.

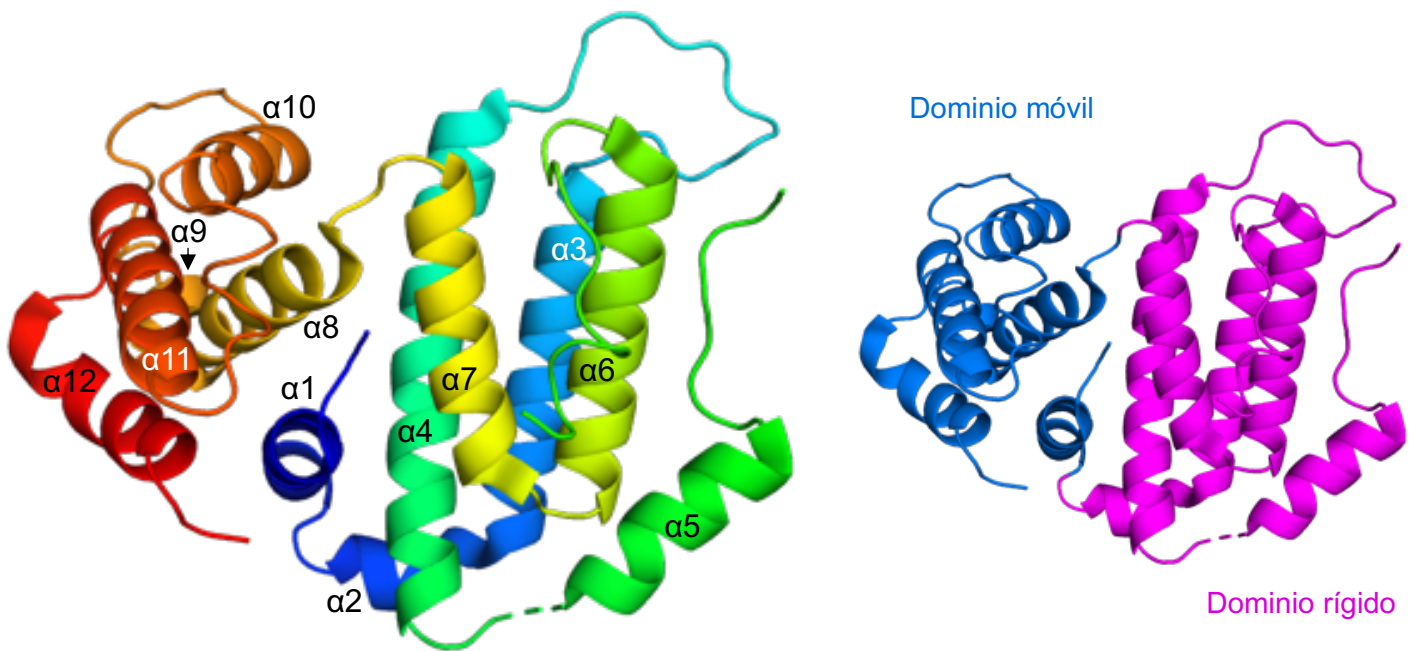


Figura I. 15. Monómero de Dut de *T. cruzi*. Formada por 12 hélices α (izquierda), que se agrupan en un dominio rígido (derecha, magenta) y un dominio móvil (derecha, azul). PDB 1OGL.

Estas Duts poseen dos centros activos, independientes y sin efectos alostéricos por la entrada del sustrato. Aunque las dos subunidades participan en la formación de ambos centros activos, cuando uno de ellos es ocupado el dominio móvil desplazado es el de la subunidad que aporta la mayoría de los residuos al centro activo, mientras que el dominio móvil de la otra subunidad permanece impassible (Figura I. 17).

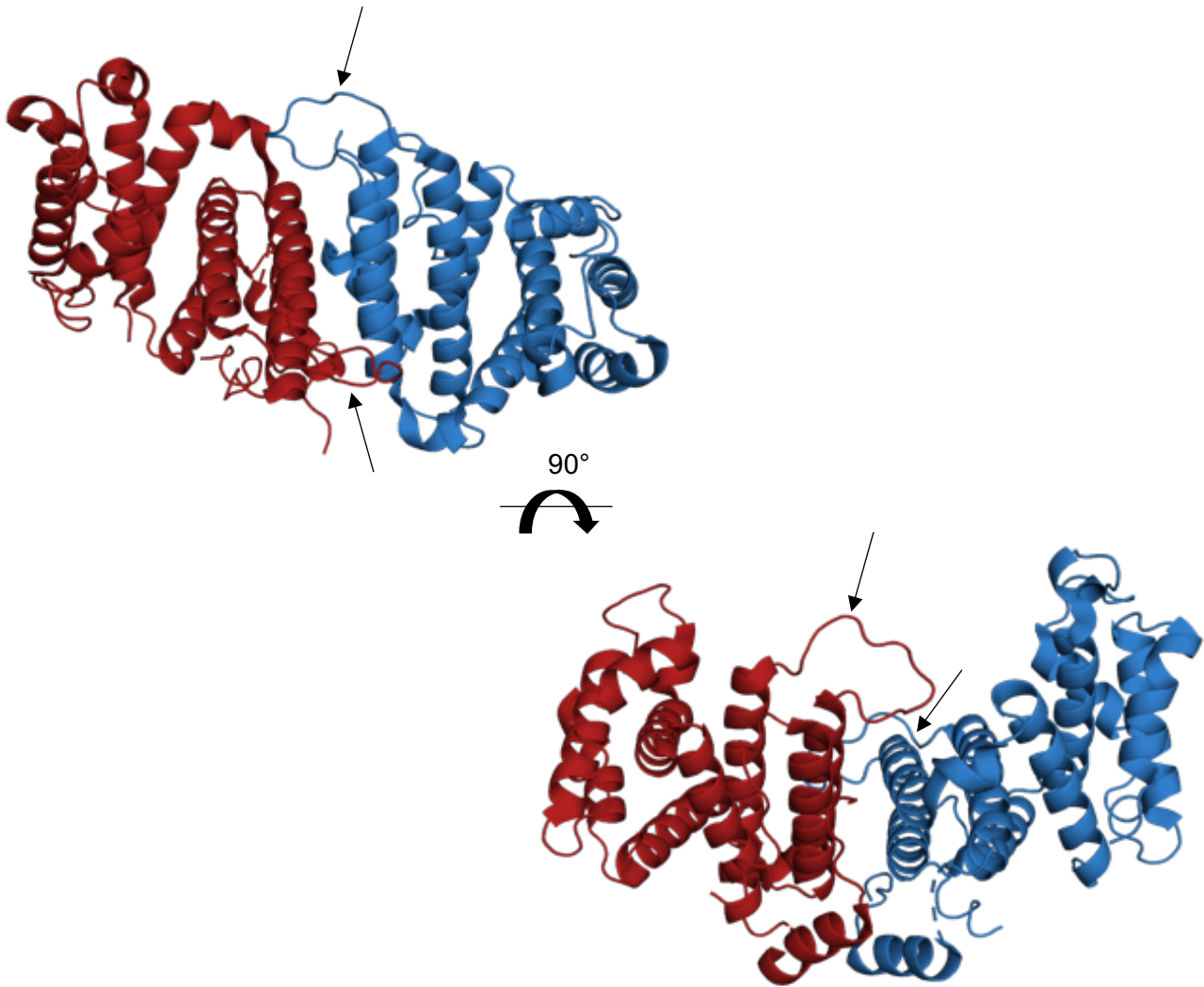


Figura I. 16. Dímero de Dut de *T. cruzi*. Se muestran dos vistas ortogonales de la zona de dimerización. El *loop* de cada subunidad se inserta en la hendidura de la otra subunidad. Los *loops* se identifican con flechas negras, y cada subunidad en un color. Código PDB 1OGL.

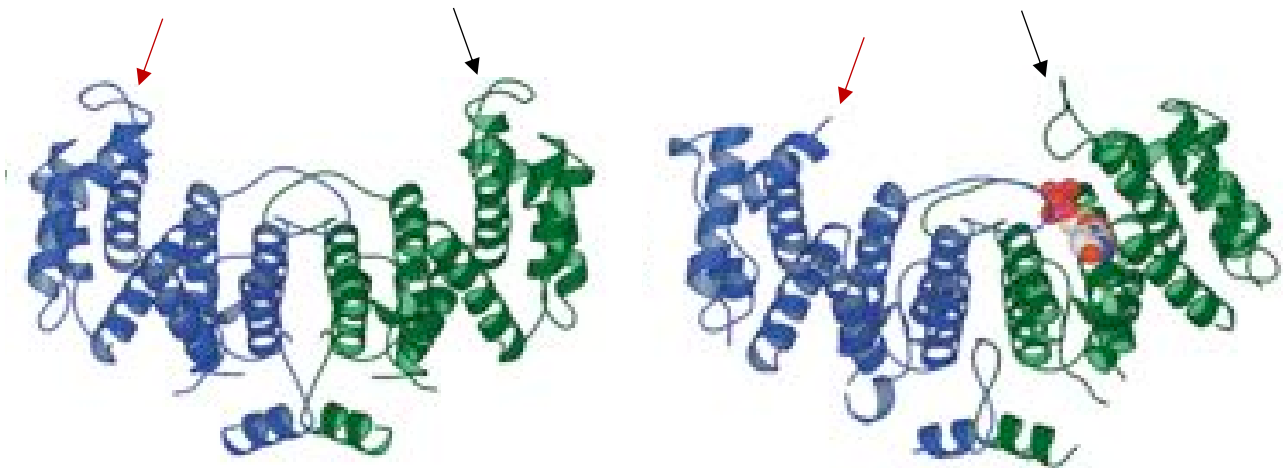


Figura I. 17. Estructuras tridimensionales de Dut de *T. cruzi* en ausencia y presencia de nucleótido. En ausencia de sustrato (*izquierda*), la Dut adopta una conformación abierta en ambas subunidades (azul y verde). En presencia del nucleótido (representado en contorno de esferas de colores por átomo, C en gris, N en azul y O en rojo) en una de las subunidades, el *loop* perteneciente a la misma subunidad se ordena sobre el centro activo (flecha negra), mientras que el *loop* de la otra subunidad (flecha roja) mantiene la misma disposición. Figura adaptada de *Harkiolaki et al. 2004*.

2.3.1. Centro activo.

Los centros activos de las Duts diméricas se ubican en la hendidura formada en la interfase entre los dominios rígido y móvil de cada subunidad (Figura I. 18). Ambos monómeros contribuyen a los dos centros activos aunque el núcleo principal de cada centro activo es aportado por el propio monómero. El segundo monómero aporta residuos provenientes del *loop* que se proyecta sobre el centro activo del compañero. Las Duts diméricas hidrolizan dUTP y dUDP, y no son inhibidas por el dUMP.

La estructura de Dut de *T. cruzi* en complejo con dUDP permitió describir la arquitectura del centro activo y su selectividad por el sustrato. El uracilo es seleccionado por puentes de hidrógeno con las cadenas laterales de residuos localizados en la hélice $\alpha 1$ (N26 y Q22) y el W61 proveniente del lazo de la subunidad complementaria con la que forma el dímero. Tras la unión del nucleótido, éste es bloqueado en el interior del centro activo por la cadena lateral de la metionina 29 de la hélice $\alpha 1$, que establece interacciones de apilamiento sobre el anillo del uracilo. Por su parte, la desoxirribosa es anclada mediante puentes de hidrógeno con la cadena lateral de un residuo de asparragina de la hélice $\alpha 8$ (N201), además de ser flanqueada mediante dos residuos aromáticos de la hélice $\alpha 4$ (H84 y F84). Los fosfatos del nucleótido son posicionados

INTRODUCCIÓN

estableciendo una red de interacciones por puentes de hidrógeno entre residuos de los motivos II a V: E49, E52, W62, E77, K197, R204, Y209, K216, N224 (Figura I. 19). La discriminación del dCTP y dCDP frente al dUTP y dUDP se produciría por la repulsión de cargas entre la amina distal de la citosina y el triptófano 61 (motivo III). La acomodación del dTTP y dTDP sería impedida por la disposición de la cadena lateral del triptófano 62 (motivo III). La discriminación del UTP y UDP se lleva a cabo mediante impedimentos estéricos, ya que la ribosa, con el grupo hidroxilo extra no podría acomodarse entre los residuos H83 y F84 (Hidalgo-Zarco et al. 2001; Harkiolaki et al. 2004).

El dUMP, aunque es producto de reacción, puede unirse débilmente al centro activo dado que esta unión no induce el cierre del centro activo. Las interacciones establecidas entre residuos del dominio móvil y el fosfato β son las que inducen el movimiento de este dominio para que el centro activo adopte la conformación catalíticamente competente (Harkiolaki et al. 2004; Moroz et al. 2004).

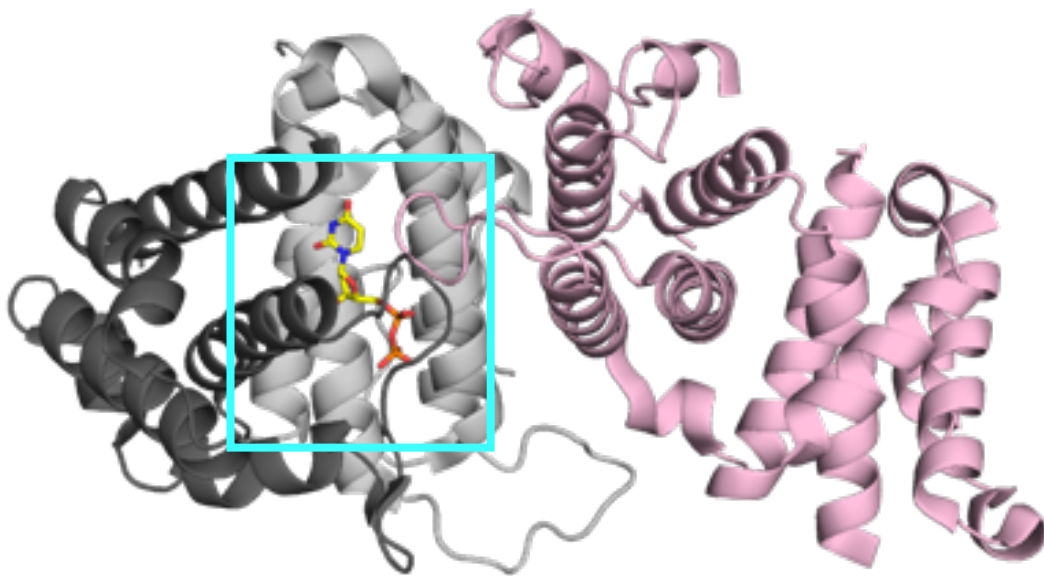


Figura I. 18. Localización del centro activo en Duts diméricas (*T. cruzi*). Se ubica en la hendidura formada en la interfase entre los dominios rígido (gris claro) y móvil (gris oscuro) de cada subunidad (tonos de grises o rosa). El sustrato dUDP se representa en varillas por colores según los átomos (C en amarillo, O en rojo, N en azul, P en naranja). Se encuadra en color cian el centro activo. Código PDB 1OGK.

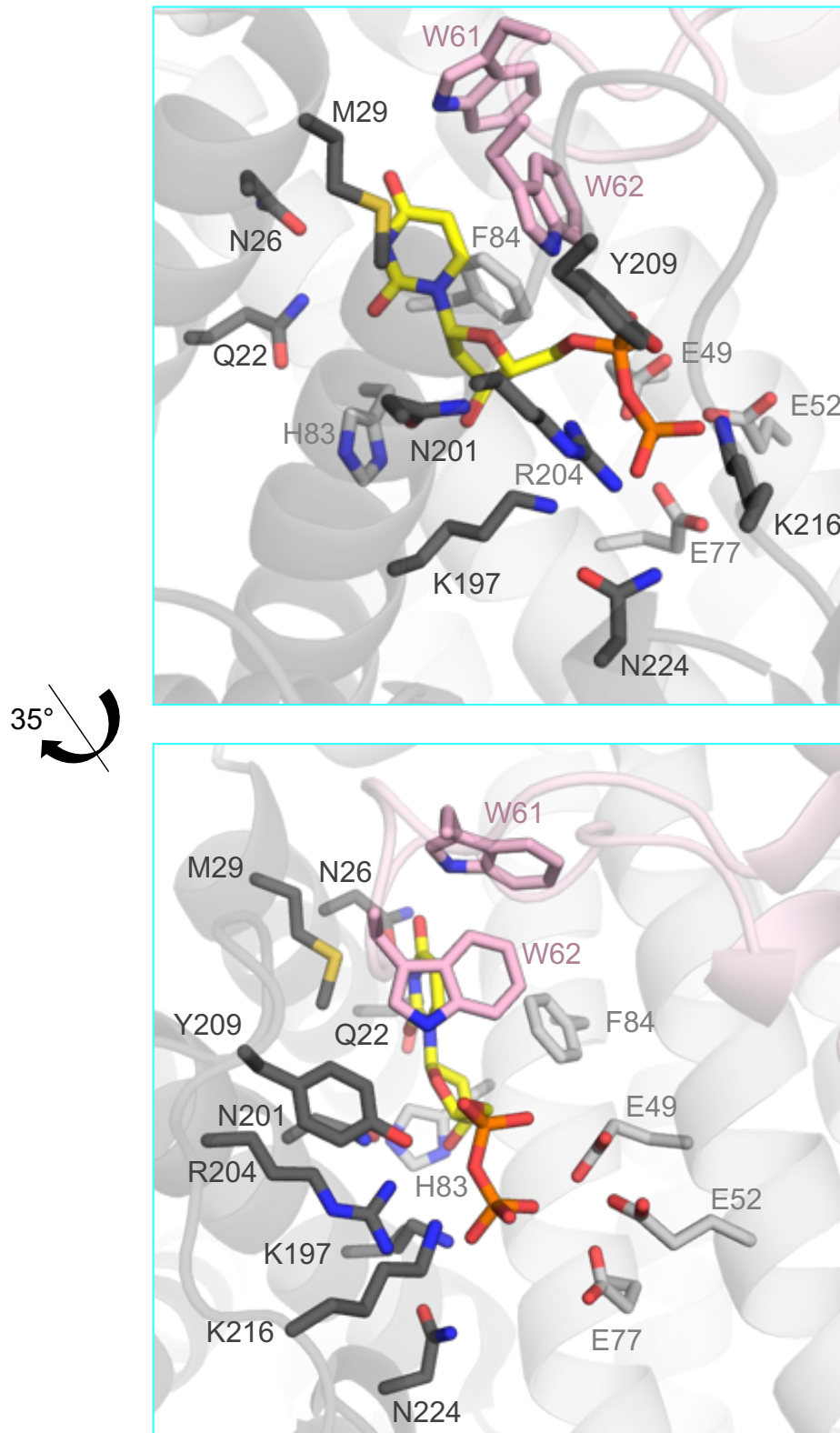


Figura I. 19. Vista detallada del centro activo de Dut de *T. cruzi*. El sustrato se posiciona en la conformación competente para la hidrólisis mediante la interacción con numerosos residuos de la misma subunidad (cadenas laterales identificadas y representadas como varillas en gris claro si pertenecen al dominio rígido, y en gris oscuro si lo son del dominio móvil) y dos residuos de la subunidad adyacente (rosa). El nucleótido se representa en varillas y en color según átomo: C amarillo, O rojo, N azul, P naranja. Código PDB 1OGK.

INTRODUCCIÓN

La estructura de *T. cruzi* ha sido resuelta en presencia de sustrato pero en ausencia de iones. No obstante, sí ha sido resuelta la estructura de la Dut dimérica de *C. jejuni* en complejo con un análogo de sustrato no hidrolizable del dUDP (dUpNHp) y Mg^{2+} (código PDB 1W2Y), la cual ha permitido posicionar los cationes participantes en la catálisis (Moroz et al. 2004). En el centro activo se posicionan 3 cationes Mg^{2+} coordinados por una tétrada acídica de 3 glutámicos y 1 aspártico pertenecientes a los motivos II y IV (E46, E49, E74 y D77). Un primer Mg^{2+} es coordinado por el residuo E46, un oxígeno del fosfato α , un oxígeno del fosfato β , y 3 moléculas de agua. Un segundo Mg^{2+} está coordinado por los residuos E46, E74 y D 77, así como por un oxígeno del fosfato β y tres moléculas de agua. El tercer Mg^{2+} está coordinado por los residuos E49, E74, un oxígeno del fosfato β , y tres moléculas de agua. Estos dos últimos cationes coordinan a una molécula de agua que actuará como nucleófilo (agua catalítica) (Moroz et al. 2004).

No obstante, la estructura de la Dut de *C. jejuni* también ha sido resuelta en presencia de un análogo del dUTP no hidrolizable (dUpNHpp) (código PDB 2CIC), y solo se posicionan dos cationes metálicos identificados también como Mg^{2+} , los cuales ocupan las mismas posiciones que dos de los 3 participantes en el caso del dUDP. Por tanto, se postula que las Duts diméricas utilizan dos cationes Mg^{2+} para la hidrólisis del dUTP, pero tres en el caso del dUDP (Hemsworth et al. 2011).

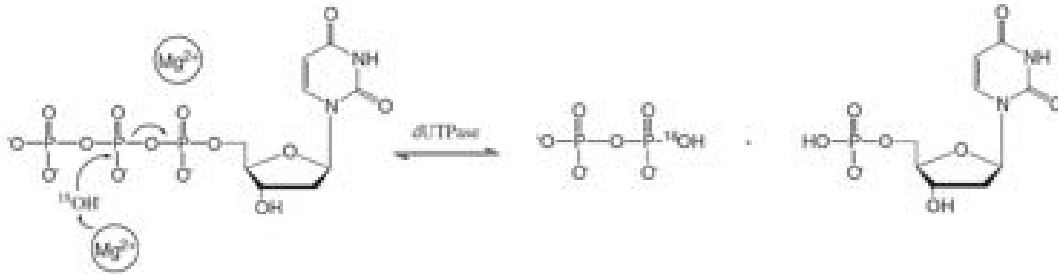
2.3.2. Mecanismo de hidrólisis.

Al igual que en las Duts triméricas, la catálisis en Duts diméricas se inicia por el ataque nucleofílico de la molécula de agua. Sin embargo en estas Duts se realiza sobre el fosfato β , y en sentido $\beta \rightarrow \alpha$. Cuando el sustrato es el dUTP, la molécula de agua es coordinada por un ión Mg^{2+} (Mg2) que la polariza para que inicie el ataque. No obstante, este ión de Mg^{2+} no es el único que participa en la catálisis, sino que un segundo ión Mg^{2+} (Mg1) estabiliza el estado de transición y ayuda a la salida del dUMP, que junto al PPI, se genera como producto de la reacción (Figura I. 20). En el caso de que el sustrato sea el dUDP, la molécula de agua catalítica, coordinada por el Mg2 y un tercer catión Mg^{2+} (Mg3), se dispone en línea con el fosfato β para hacer el ataque nucleofílico y se generan como productos dUMP y Pi.

En las Duts diméricas los metales resultan absolutamente necesarios para la actividad, ya que los dos o tres iones participantes no solo ayudan al posicionamiento del nucleótido, sino que participan directamente en la catálisis al polarizar la molécula de agua catalítica. En el caso de Duts triméricas, aunque también ayudan en el correcto posicionamiento del sustrato y la estabilización del estado de transición, es posible la hidrólisis del dUTP en ausencia de Mg^{2+} ,

aunque a una velocidad mucho menor (Hemsworth, González-Pacanowska, and Wilson 2013).

A)



B)

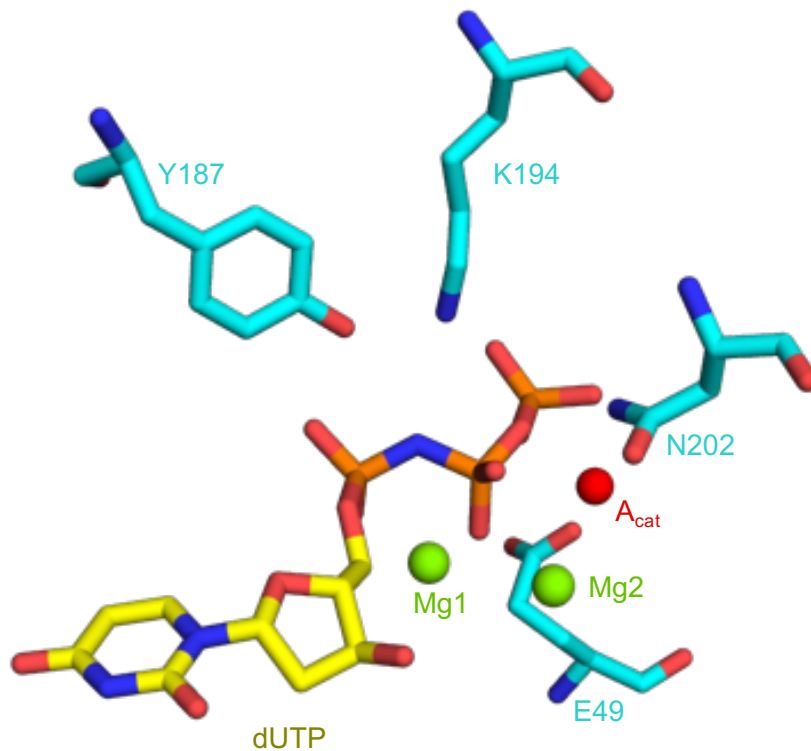


Figura I.20. Mecanismo de hidrólisis del dUTP en Dut dimérica. (A) Un ión Mg^{2+} polariza a una molécula de agua que inicia el ataque nucleofílico sobre el fosfato β , y se produce dUMP y PPI como productos de la reacción. Figura tomada de *Hemsworth, González-Pacanowska, and Wilson 2013*. (B) Representación sencilla del posicionamiento del nucleótido (en varillas y en colores según átomo: átomos de C en amarillo, O en rojo, N en azul y P en naranja), los dos iones Mg^{2+} (esferas verdes) y el agua catalítica (esfera roja) en Dut de *Campylobacter jejuni*. Se representan en varillas y se identifican las cadenas laterales de algunos de los residuos de la Dut participantes (átomos de C en cian, O en rojo, N en azul). Código PDB: 2CIC.

3. Funciones *moonlighting*: ¿Qué son?

En 1941 Beadle y Tatum propusieron la hipótesis de “un gen, una enzima” para referirse a la traducción de una única secuencia génica en una única secuencia aminoacídica. De esta hipótesis se podría derivar un corolario: “una proteína, una función”, pero ambas fueron desestimadas con el descubrimiento del *splicing* alternativo de genes y por el descubrimiento, a finales de los 80, de que una misma proteína era capaz de desarrollar más de una función, siendo éstas totalmente distintas. Fueron denominadas “proteínas *moonlighting*” por Constance Jeffery en 1999, en referencia a su capacidad de tener más de un trabajo, un pluriempleo.

Tradicionalmente se ha asumido que las Duts son proteínas esenciales para la integridad del ADN y la viabilidad tanto en organismos eucariotas como procariotas, y siempre con funciones relacionadas con el papel que la enzima juega previniendo la incorporación del dUTP en el ADN (José R. Penadés et al. 2013). No obstante, existen ejemplos de Duts que realizan, a parte de la función canónica, otras funciones totalmente distintas. A continuación se nombran algunos ejemplos de estas funciones *moonlighting*:

- La Dut de rata interacciona con el factor transcripcional PPAR α , y reprime la activación transcripcional mediada por este factor (Figura I. 21) (Chu et al. 1996).
- Cuando la Dut humana está sobrerregulada es responsable de la desensibilización de tumores tratados con agentes terapéuticos que inhiben la vía de síntesis del timidilato (Tóth et al. 2007).
- La Dut del virus de Epstein-Barr activa la expresión del factor nuclear NF- κ B y produce desregulación del sistema inmune del huésped por unión al receptor celular tipo -Toll2 (Figura I. 21) (Glaser et al. 2006; M.-E. Ariza et al. 2009).
- Este mismo receptor tipo Toll2 es la diana de Dut del retrovirus endógeno K, el cual contribuye a la fisiopatología de la psoriasis (M. E. Ariza and Williams 2011).
- La Dut del sarcoma de Kaposi limita la respuesta inmune al unirse a diversos receptores de citoquinas (Madrid and Ganem 2012).
- Algunas Duts de fagos de *S. aureus* interaccionan con el principal represor de ciertas islas de patogenicidad de *S. aureus* (SaPIs). Esta interacción supone la desrepresión de la correspondiente isla y consecuente empaquetamiento de la SaPI en las cápsulas del fago que porta la Dut

INTRODUCCIÓN

(Tormo-Más et al. 2013; 2010). Más adelante se ahondará en esta fascinante función *moonlighting*.

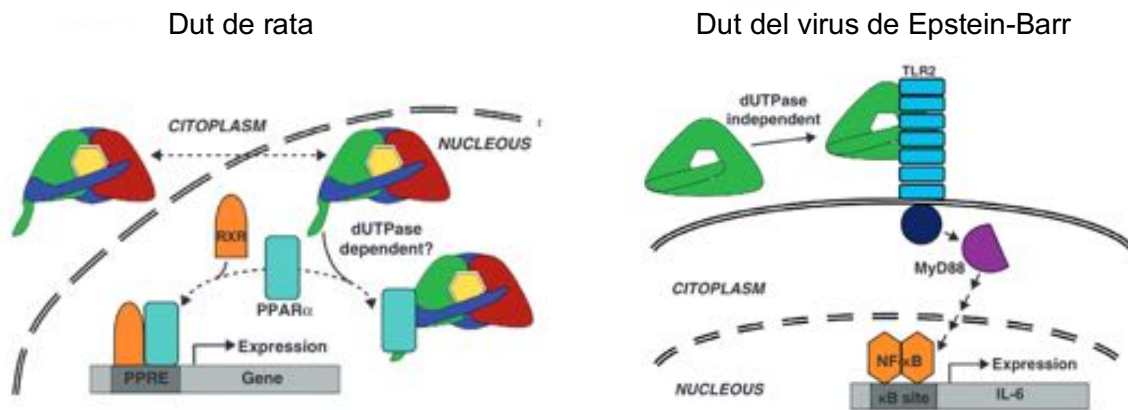


Figura I. 21. Dos ejemplos de estrategias de regulación mediada por Duts. (*Izquierda*) La Dut citoplasmática de rata es transferida al núcleo, y allí interacciona con el receptor PPAR α , impidiendo la heterodimerización con RXR, y reprimiendo la activación transcripcional mediada por PPAR α . (*Derecha*) Dut codificada por el virus de Epstein-Barr interacciona con el receptor tipo Toll2 e induce el reclutamiento de la proteína adaptadora MyD88 que induce la expresión de citoquinas proinflamatorias (IL-6) a través de la activación del factor NF- κ B. Figura adaptada de J.R. Penadés et al. 2013.

4. Islas de patogenicidad de *Staphylococcus aureus* (SaPIs).

Las islas cromosómicas inducibles por fago o PICIs representan una familia única de elementos génicos móviles, y son parásitos extremadamente efectivos, elegantes y sofisticados que llevan incorporado un arsenal de estrategias efectivas para someter a su voluntad a los fagos ayudantes, de los cuales dependen para asegurar su presencia en la naturaleza. Contribuyen sustancialmente a la transferencia génica horizontal, a la adaptación al huésped y a la virulencia de la bacteria que las contiene (José R. Penadés and Christie 2015). Las islas de patogenicidad de *S. aureus* (SaPIs) son las PICIs prototípicas. Las SaPIs portan y diseminan genes de superantígenos y otros genes de virulencia. Están ampliamente distribuidas entre los staphylococcales y son responsables de al menos una enfermedad humana importante, el síndrome de shock tóxico (Tormo-Más et al. 2013). Las SaPIs y otras PICIs tienen un papel importante en la evolución de plásmidos bacterianos y en la transmisión de resistencia antimicrobiana (Humphrey et al. 2021). Ocupan sitios cromosómicos específicos, y para su movilidad y expansión requieren de ciertos fagos, que reciben el nombre de *helpers* o ayudantes. Éstos inducen su ciclo ERE: escisión, replicación y empaquetamiento en partículas infecciosas tipo fago específicas de cabeza reducida (Tormo-Más et al. 2013; José R. Penadés and Christie 2015). Éstas se liberan tras la lisis inducida por el fago y pasan a infectar otras bacterias, donde se integran en el sitio de integración en el cromosoma correspondiente hasta iniciar de nuevo el ciclo ERE (Figura I. 22), perpetuando así su existencia con una elevada frecuencia de transferencia inter e intragénero, y confiriendo toxigenicidad (Úbeda et al. 2008).

Por lo general, las SaPIs contienen de 18 a 22 ORFs, y se agrupan en cuatro módulos principales muy conservados asociados con regulación, integración /escisión, replicación autónoma, y empaquetamiento; y en otros genes accesorios (Úbeda et al. 2008).

- Módulo de control transcripcional: incluye un par de pautas de lectura abierta. Una de ellas es *str* y la otra *stl* (Figura I. 23). Codifican para las proteínas Str y Stl que contienen un motivo HTH, típico de reguladores transcripcionales. Aunque todas las SaPIs que se han descrito codifican para Stl, sus secuencias aminoacídicas no están muy conservadas. La proteína Stl se une a la región divergente entre *str* y *stl* (Figura I. 24) para inhibir la transcripción de ambos y con ello evitar que se inicie el ciclo ERE, siendo por lo tanto el represor maestro de la SaPI. La inducción de dicho ciclo es dependiente del fago ayudante e implica la formación del complejo entre Stl y la proteína antirepresora codificada por el fago. De este modo, la SaPI solo inicia el ciclo ERE cuando hay un fago en

crecimiento activo que asegure su transmisión. En cuanto a la función de la proteína Str, no ha sido aclarada.

- Módulo integración/escisión: las SaPIs residen de manera quiescente en el cromosoma huésped hasta que son desreprimidas. Tras la inducción de la SaPI dependiente del fago ayudante, se produce la escisión de la isla y su circularización. Los genes implicados en la integración (*int*) y escisión (*xis*) se localizan aguas arriba del gen *stI* y aguas abajo de *str*, respectivamente. Todas las SaPIs están flanqueadas por repeticiones directas (*attL* y *attR*), y se insertan en una única orientación en el sitio de inserción específico en el cromosoma (*attC*) (Figura I. 23). En el genoma de *S. aureus* se han identificado 5 sitios *attC*, y son usados por dos o más SaPIs conocidas, pero no son utilizados por otros elementos móviles. La integración de la isla requiere de la actividad de la proteína integrasa Int, mientras que en la escisión están implicadas las proteínas Int y Xis. La mayoría de las SaPIs descritas regulan su propia escisión y transferencia, aunque algunas como SaPI_{bov1} y SaPI_{bov2} muestran niveles detectables de escisión espontánea en ausencia de fagos ayudantes.
- Módulo de replicación: Las SaPIs se replican utilizando su propio replicón. Este módulo está formado por un gen de primasa (*pri*) y un gen *rep* específico, que reconoce el origen de replicación conato y siempre se localiza en 3' del gen *rep* (Figura I. 23). En algunas islas, ambos genes están fusionados. La ADN polimerasa del huésped replica el material genético de la SaPI, y produce diversos centenares de copias del ADN concatenado que serán empaquetadas en partículas fágicas.
- Módulo de empaquetado: hay dos grupos de genes implicados en la modulación del ensamblaje y empaquetamiento de la SaPI. En uno de ellos está el gen *ppi*, que interfiere con el empaquetado del ADN del fago *helper*. En algunas islas, en este primer grupo hay uno o más genes adicionales, de los que no se conoce la función. El segundo grupo de genes contiene al operón I, que incluye los genes de redireccionamiento del tamaño de la cápside (*cpmA* y *cpmB*), los genes de interferencia de la expresión de genes tardíos del fago ayudante, así como los genes implicados en la redirección de la maquinaria del empaquetado del ADN del fago *helper* (TerS) (Figura I. 23). Todos estos genes se transcriben tras la desrepresión de la SaPI desde el promotor *str*. Como fagos satélites que son las SaPIs, éstas se empaquetan en partículas víricas formadas por proteínas estructurales del fago *helper* (Úbeda et al. 2008).

A) Ciclo ERE de la SaPI

B) Infección de la SaPI

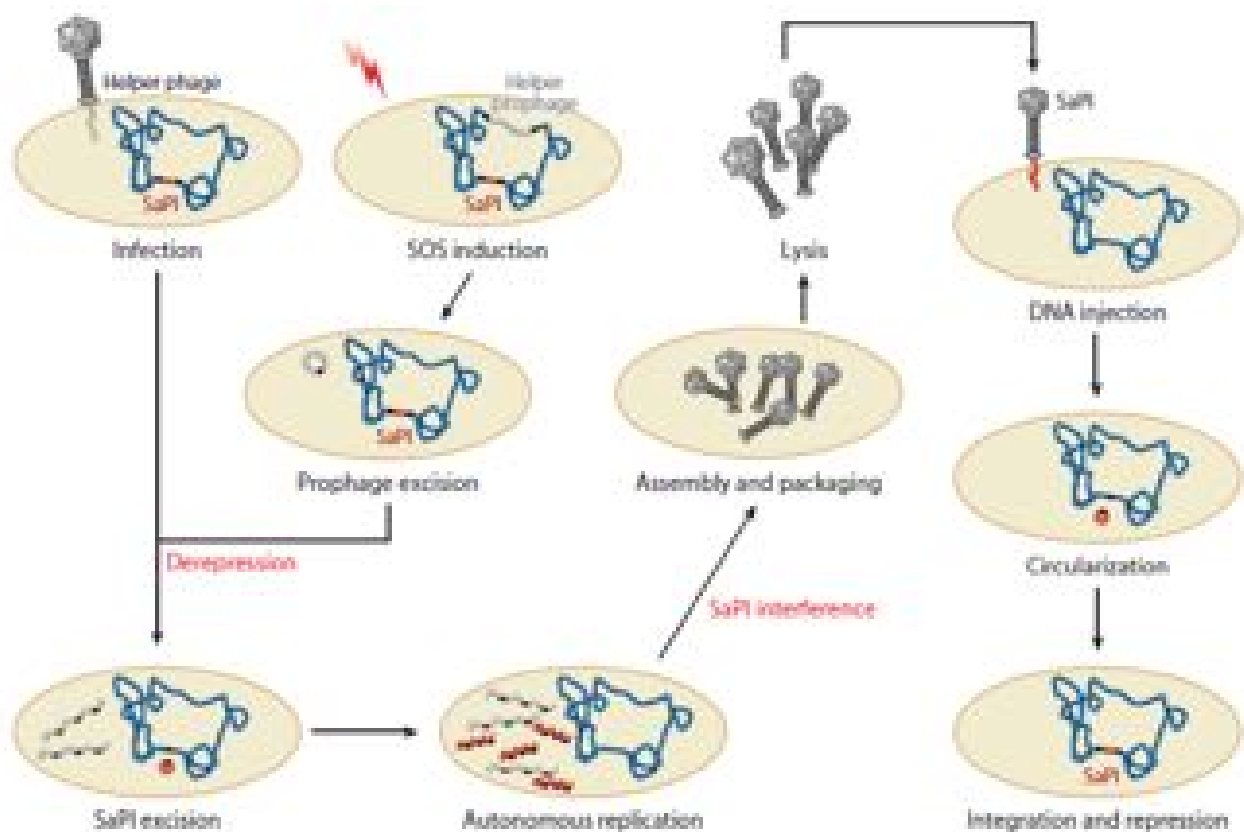


Figura I. 22. Ciclo ERE e infección de las SaPIs. A) Después de la inducción por el sistema SOS o tras la infección por un fago ayudante que codifica una proteína antirrepresora que libera la represión de la SaPI mediada por *Stl*, se inicia el ciclo ERE de las SaPIs. La expresión de las proteínas de escisión (*Xis*) e integrasa (*Int*) codificadas por SaPI promueve la escisión de SaPI. La replicación autónoma posterior de SaPI lleva a la producción de una gran cantidad de copias concatenadas que serán empaquetadas por el heterocomplejo de terminasas del fago y de la SaPI. El ADN de SaPI se empaqueta en partículas constituidas por proteínas de fago redirigidas por proteínas codificadas por la SaPI para formar viriones con cápsidas más pequeñas, que contendrán el genoma de la isla. B) Las partículas víricas con el genoma de SaPI liberadas de las células donantes después de la lisis mediada por el fago ayudante introducen su ADN en una nueva bacteria receptora, donde el ADN de la SaPI entrante se circulariza y se integra en el sitio cromosómico *attC*. Figura tomada de José R. Penadés and Christie 2015.

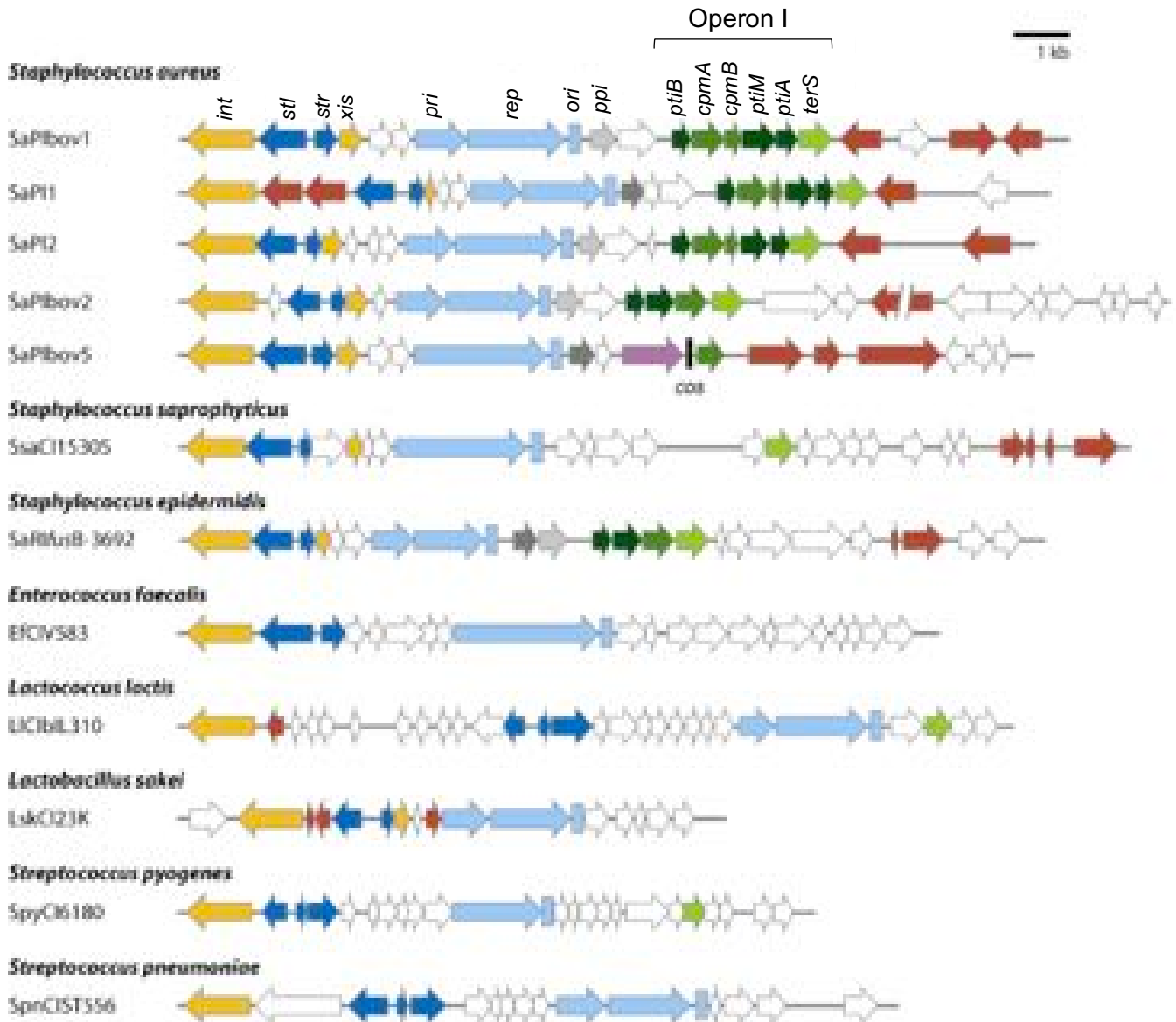


Figura I. 23. Organización génica de PICs encontradas en diferentes especies. Se incluyen PICs de *S. aureus*, *S. saprophyticus*, *S. epidermidis*, *E. faecalis*, *L. lactis*, *L. sakei*, *S. pyogenes*, y *S. pneumoniae*. Se identifica en el extremo izquierdo el gen *int*, mientras que otros genes están coloreados según secuencia y función: *int* y *xis* en amarillo; reguladores de la transcripción (*stl*, *str*) en azul oscuro; genes de replicación (*pri*, *rep* y *ori*) en azul claro; genes implicados en expresión (*pti*) en verde oscuro; genes implicados en ensamblaje de los componentes del virión (*cpm*) en verde de intensidad intermedia. El gen *TerS* en color verde claro; genes de superantígeno y otros genes de viruencia en rojo; genes implicados en la inhibición del empaquetamiento del fago ayudante (*ppi*) se muestran en tonos de gris (distinta intensidad si son variantes distintas). Los genes que codifican proteínas hipotéticas se muestran en blanco. SaPIbov5 lleva un gen similar a la cápside (en púrpura) y un sitio *cos* (en negro). Figura tomada de José R. Penadés and Christie 2015.

INTRODUCCIÓN

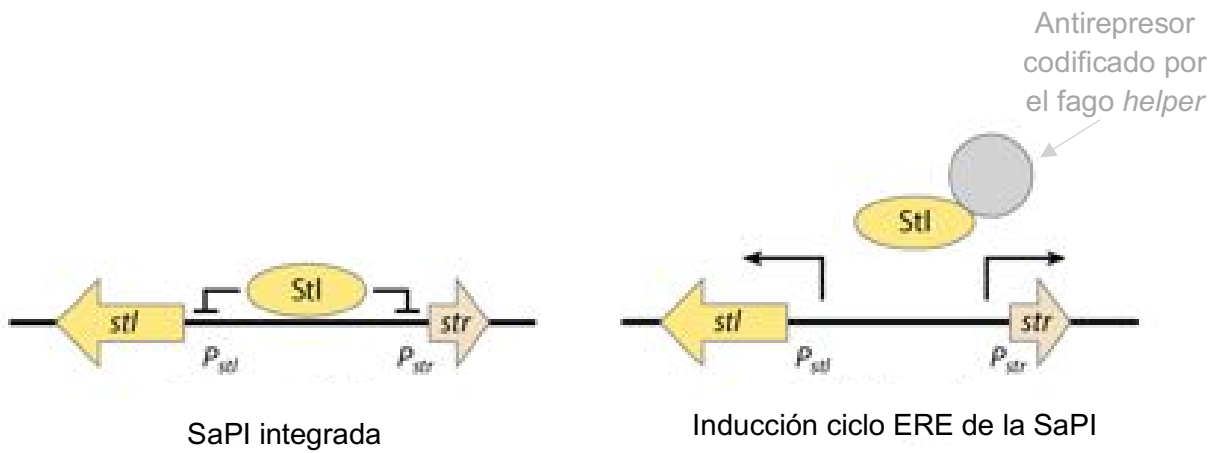


Figura I. 24. Represión y desrepresión de la SaPI. (*Izquierda*) En ausencia del fago *helper*, el represor StI está unido a la región intergénica entre *stl* y *str*, impidiendo la expresión de los genes de la isla. (*Derecha*) Cuando la proteína antirepresora codificada por el fago interactúa con StI, éste se desancla de la isla y se inicia el ciclo ERE. Figura adaptada de José R. Penadés and Christie 2015.

5. Inducción de las SaPIs por Duts.

Las SaPIs se mantienen reprimidas por StI, posiblemente un represor análogo al estándar de fago *cl* (Úbeda et al. 2008). La desrepresión de la isla es llevada a cabo por proteínas de fago específicas con funciones *moonlighting* y que en muchas ocasiones no son esenciales para el fago. Este proceso se realiza por la interacción directa de estas proteínas con el represor StI de la SaPI, induciendo su desanclaje del ADN y permitiendo la expresión de genes que inician el ciclo ERE.

Cada SaPI codifica su propio StI, y cada StI interacciona con una proteína específica del fago *helper* o ayudante. Por ello, los distintos fagos difieren en la habilidad por movilizar las diferentes SaPIs, existiendo casos de fagos que poseen varias proteínas que actúan como antirepresores de diferentes SaPIs (José R. Penadés and Christie 2015). El fago $\phi 80\alpha$ puede inducir las islas SaPI1, SaPI2, SaPIbov1, SaPIbov2 y SaPIbov5, mientras que los fagos ϕ NM1 y ϕ NM2 inducen SaPI1, SaPIbov1 y SaPIbov5. Por su parte, el fago ϕ 11 induce solo SaPIbov1 y SaPIbov5, y el fago ϕ 85 no induce ninguna de estas islas. Por tanto, aquellos fagos que portan un mayor número de proteínas antirepresoras inducen la movilidad de un mayor número de SaPIs. (Figura I. 25) (José R. Penadés and Christie 2015; Tormo-Más et al. 2010).

La variedad de represores StI entre SaPIs y de antirepresores entre fagos o dentro de un mismo fago, ha llevado a hipotetizar que originalmente una sola proteína de fago fue utilizada como antirepresor para un StI ancestral, y tras múltiples ciclos de coevolución de los distintos fagos e islas se ha producido la diversidad existente (José R. Penadés and Christie 2015).

Las proteínas específicas de los fagos ayudantes responsables de la desrepresión no han sido identificadas en todos los casos. Como ejemplos, la SaPI1 es desreprimida por Sri; la SaPIbov2 es desreprimida por una pequeña proteína de función desconocida (gp15); la SaPIbov1 y la SaPIbov5 están desreprimidas por las Duts del fago ϕ 80, ϕ 85, ϕ 71bel y ϕ O11; mientras que las Duts del fago ϕ 11, ϕ NM1 y ϕ DI desreprimen a SaPIbov1 (Figura I. 25) (Szabó et al. 2014; Tormo-Más et al. 2010; 2013; Maiques et al. 2016; Bowring et al. 2017).

INTRODUCCIÓN

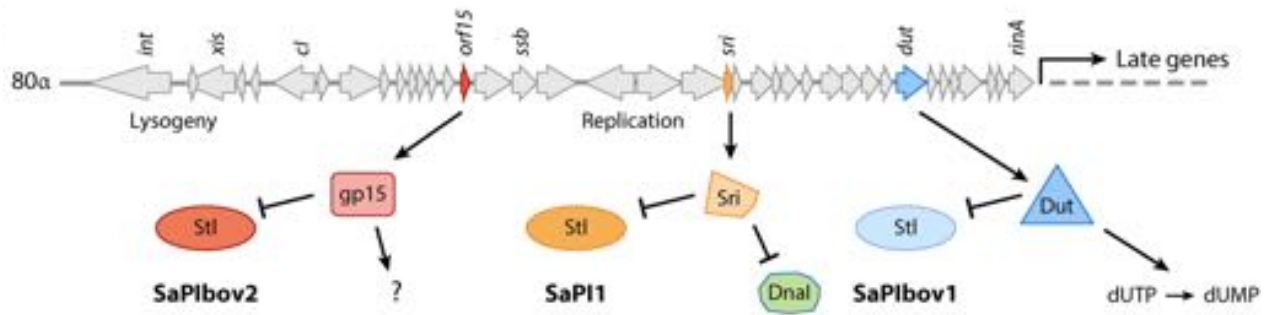


Figura I. 25. El fago 80α es explotado por distintas SaPIs para iniciar su ciclo ERE. Este fago codifica varias proteínas que actúan como antirepresores de distintas SaPIs. La proteína gp15, de función desconocida, interacciona con el Stl de SaPIbov2; la proteína Sri, una proteína de unión al ADN, interacciona con el Stl de SaPI1; la proteína Dut, que hidroliza dUTP, interacciona con el Stl de SaPIbov1. Figura tomada de José R. Penadés and Christie 2015.

5.1. El Stl de SaPIbov1 interacciona con Duts.

El represor Stl de SaPIbov1 (en adelante, Stl) es capaz de unirse a tres regiones de 23 nucleótidos. Dos de éstas constituyen segmentos intergénicos *stl-str*, y el tercero corresponde al segmento de secuencia en la región intergénica *str-xis*. Todos ellos son segmentos palindrómicos (Papp-Kádár et al. 2016).

En el año 2010 se describió que aparte de unirse a estas regiones del ADN, Stl se unía a Duts triméricas de fagos de *S. aureus*, y con ello conseguía la desrepresión de la SaPI de la cual era el represor (Tormo-Más et al. 2010). Desde entonces, mediante estudios funcionales *in vivo* e *in vitro* se ha analizado la participación de los motivos conservados de la Dut, así como el papel del sustrato dUTP en la interacción (Pecsi et al. 2012; Leveles et al. 2013; Tormo-Más et al. 2013; Szabó et al. 2014; Hirmondó et al. 2015; Maiques et al. 2016). Todos estos estudios han llevado a la propuesta de tres modelos de interacción entre esta familia de Duts de fagos de *S. aureus* y este represor Stl. No obstante, aquí solo se explicará el último, siendo éste propuesto por el laboratorio, el más actual y corroborado también con los resultados de esta Tesis doctoral.

Las Duts triméricas de fago, a parte de los 5 motivos conservados descritos anteriormente, poseen una región altamente variable en secuencia y longitud que no poseen otras Duts triméricas (como la humana o la de *M. tuberculosis*). Esta región, en consecución a los cinco motivos catalíticos conservados, se denominó motivo VI (Figura I. 26) (Donderis et al. 2017). Este motivo, no es requerido para la actividad enzimática de Dut pero sí en la interacción con Stl (Maiques et al. 2016; Alite et al. 2017; Leveles et al. 2013).

INTRODUCCIÓN

En el modelo de interacción propuesto, cuando el dUTP se une a la Dut y durante su hidrólisis, el motivo V se pliega de manera muy estable sobre el centro activo. Esta disposición impide que el StI acceda a las regiones de anclaje de los motivos IV y VI de la Dut, totalmente accesibles en conformación *apo* por la elevada flexibilidad del motivo V. En ausencia del sustrato, StI interacciona con los residuos de los motivos IV y VI alrededor del centro activo, y una vez unido, el motivo V interacciona con el StI y estabiliza el complejo. Por tanto, hay una participación sinérgica de los motivos IV, V y VI en la interacción con StI y consecuente inducción de la SaPI (Figura I. 27) (Maiques et al. 2016).

Además del modelado del proceso de interacción, con este planteamiento quedó resuelto el controvertido papel del dUTP en la inducción de las SaPIs: el dUTP no induce la isla, sino más bien mantiene su represión. El sustrato actúa como segundo mensajero que controla los estados de ON/OFF de las proteínas Dut en el proceso de unión al StI (Maiques et al. 2016).

A parte de la interacción de este StI con Duts triméricas de fago de *S. aureus*, también se describió su capacidad de interacción con Duts diméricas de fagos de la misma especie (Dut ϕ NM1 y ϕ NM2 (Dearborn and Dokland 2012)). Desde entonces, en el laboratorio se han añadido Duts diméricas de fagos de *S. aureus* a esta lista: ϕ O46, ϕ DI y ϕ O11 (Donderis et al. 2017). Aunque no se propuso un modelo de interacción como en el caso de Duts triméricas, la resolución de la estructura tridimensional del complejo entre el fragmento C-terminal de StI y la Dut dimérica ϕ O11 ha permitido resolver la estrategia utilizada, la cual ratifica el papel del dUTP como segundo mensajero (Rafael Ciges-Tomas et al. 2019). Esto se abordará en profundidad en el siguiente apartado.

INTRODUCCIÓN

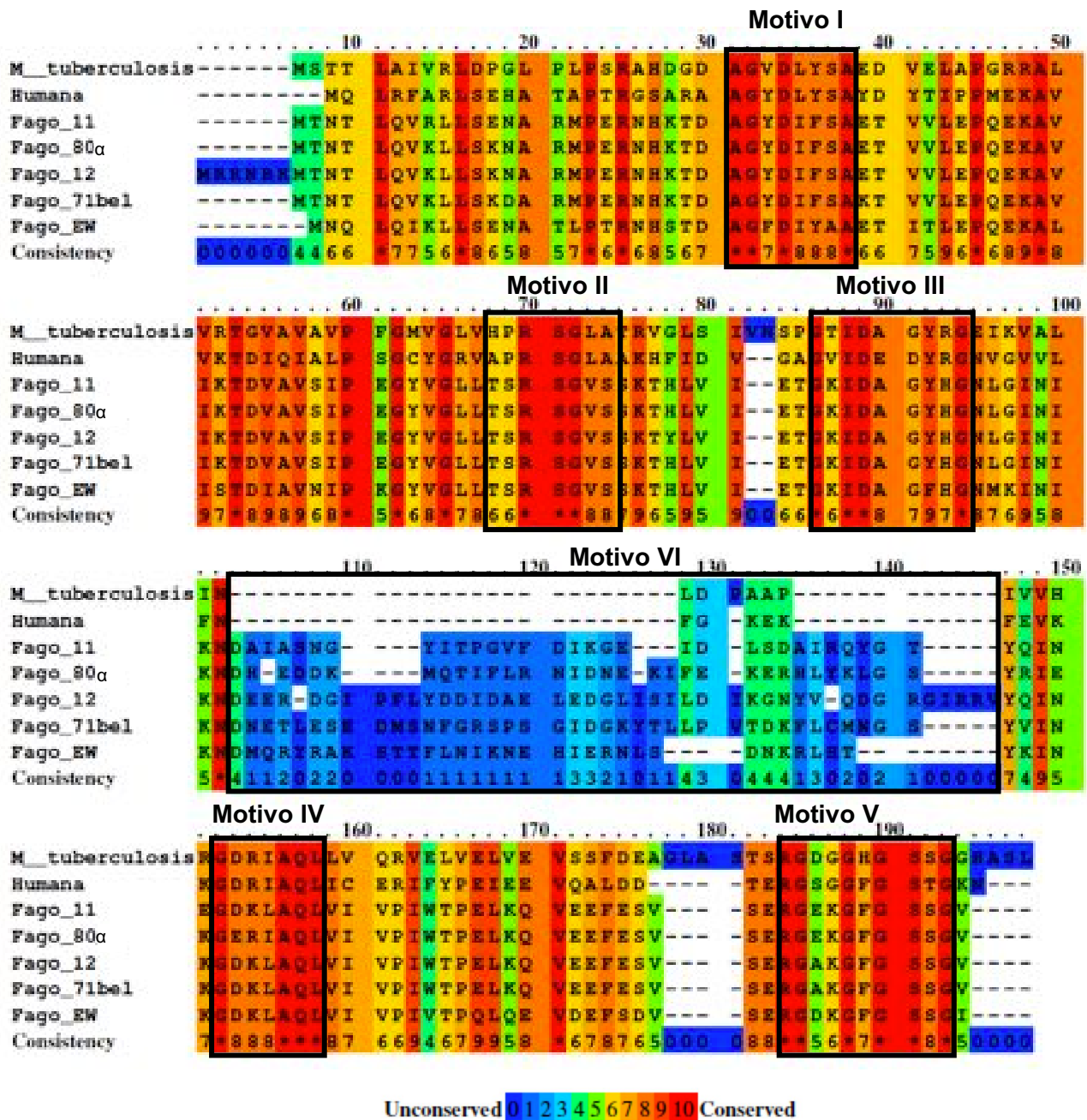


Figura I. 26. Alineamiento de Duts triméricas para la localización del motivo VI específico de Duts de fago. Se incluyen Duts representativas eucariotas (humana) y procariotas (*M. tuberculosis*), así como de fagos de distintos fagos de *S. aureus*. La conservación de residuos se indica con una gradación de colores desde rojo a azul de mayor a menor grado de conservación. Alineamiento realizado en el servidor PRALINE.

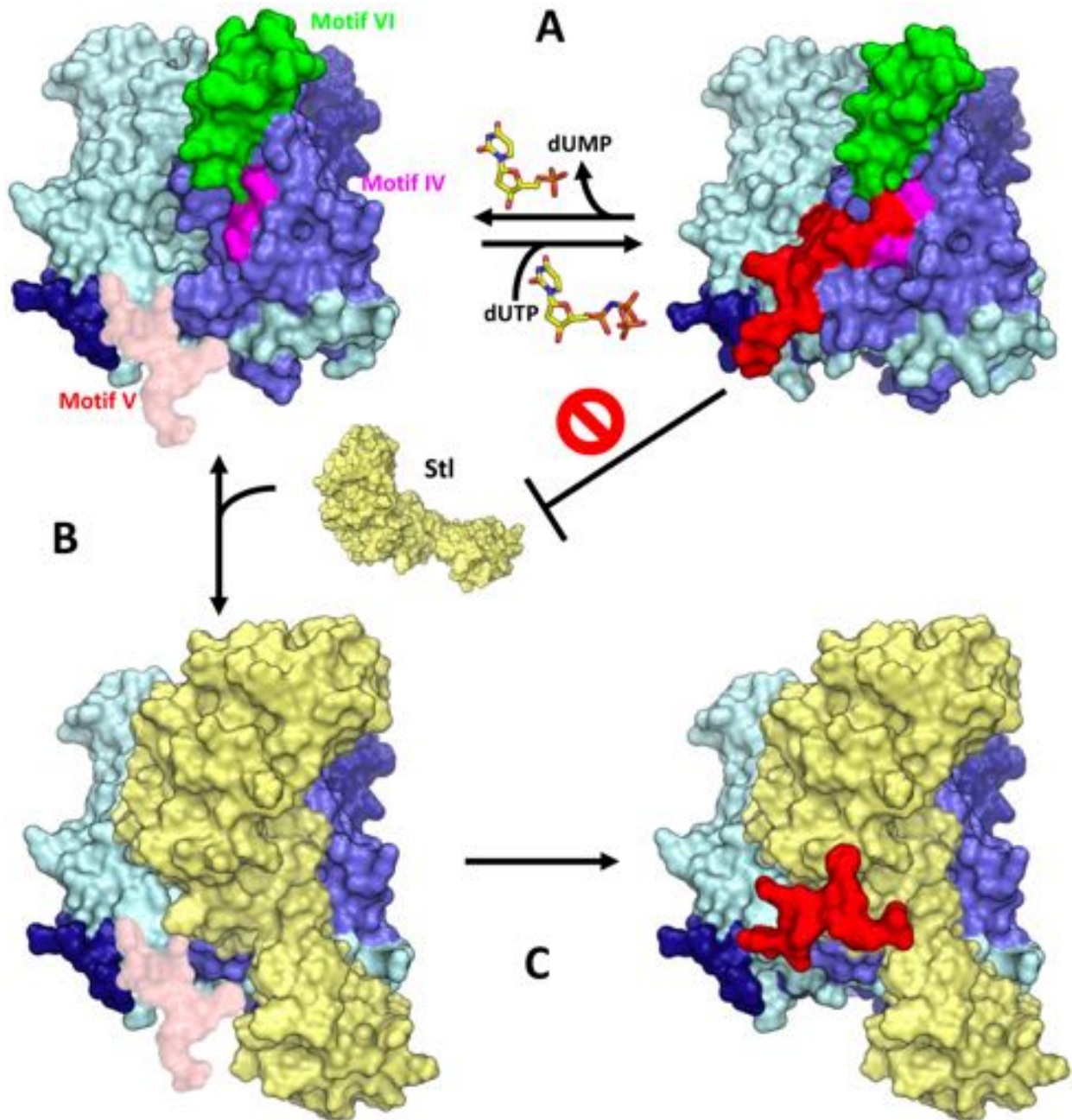


Figura I. 27. Modelo de interacción represor Stl – Dut trimérica de fago de *S. aureus*. (A) La unión y la hidrólisis de dUTP por Dut implican el plegamiento y la estabilización del motivo conservado muy flexible V (en rosa o rojo) sobre el centro activo. Esta disposición del motivo V limita el acceso de Stl a su lugar de anclaje formado por el motivo IV (magenta) y el motivo VI (verde) impidiendo la unión. (B) Por el contrario, en la forma *apo* de Dut, la región de anclaje de Stl es accesible debido a la flexibilidad del motivo V, lo que permite que Stl interactúe con los residuos de los motivos IV y VI en las proximidades del centro activo. (C) Una vez el represor está unido, el motivo V interaccionaría con residuos de Stl, estabilizando el complejo. Las estructuras se muestran como superficies. La Dut del fago $\phi 80\alpha$ se muestra con diferentes colores para los distintos motivos (código PDB 3ZEZ) y el modelado *in-silico* de Stl se muestra en amarillo. La estequiometría de interacción propuesta es 1:1 pero solo se representa un represor para reducir la complejidad de la representación. Figura tomada de *Maiques et al. 2016*.

5. 2. La interacción de StI de SaPI_{bov1} con Duts diméricas y triméricas como fracaso del intento de escape del fago de la inducción de la SaPI.

Debido a que la inducción de la SaPI es un proceso pernicioso para el fago, éstos tratan mediante distintas estrategias complementarias resistir a la intrusión de la SaPI en su proceso de replicación y posterior empaquetamiento. Una de ellas implica la generación de variantes alélicas de los antirepresores que tengan afinidad reducida por los represores StI codificados por las SaPIs; otra implica la expresión reducida de las proteínas antirepresoras; y una tercera estrategia implica la pérdida completa del gen inductor de SaPI codificado por el fago, el cual posiblemente es reemplazado por un gen diferente que proporciona la función necesaria al fago (José R. Penadés and Christie 2015).

Este es el caso de las Duts de fago de *S. aureus*. Estos fagos han intentado escapar de la inducción de la SaPI_{bov1} mediante la utilización de Duts triméricas y diméricas, con distinta afinidad por el represor (Donderis et al. 2017). Éste no es el primer intento de escape realizado por los fagos, ya que la recombinasa del fago $\phi 80\alpha$ que induce a SaPI2 ha sido sustituida en otros fagos de *S. aureus* por genes que codifican recombinasas de hasta 3 familias no relacionadas. No obstante, las diferentes recombinasas son capaces de inducir SaPI2 (Bowring et al. 2017).

Tanto Duts triméricas como diméricas catalizan la misma reacción enzimática, aunque son estructural y funcionalmente diferentes. El mecanismo de interacción con StI es conceptualmente igual en ambos, donde el dUTP impide la interacción entre Dut y el represor, actuando como interruptor de apagado (Donderis et al. 2017). Por tanto, las SaPIs han ido un paso más allá en la carrera armamentista con los fagos y ya no apuntan a una proteína específica, sino a un proceso biológico básico para los fagos *helpers*, imponiendo la transmisión de la isla en sus cápsidas para la propia supervivencia.

Cómo StI ha conseguido reconocer un proceso biológico del fago y no una proteína específica es una cuestión estudiada en profundidad en el laboratorio y la parte experimental que llevó a su explicación ha sido recogida en dos tesis doctorales anteriores (Ciges-Tomas 2019; Alite Hernández 2021; Rafael Ciges-Tomas et al. 2019; Tormo-Más et al. 2013; Maiques et al. 2016; Donderis et al. 2017).

Los estudios estructurales en el laboratorio mostraron que la estructura de StI está formada completamente por hélices α y que presenta tres dominios diferenciados (Rafael Ciges-Tomas et al. 2019). Por una parte, un dominio N-terminal con un plegamiento HTH que media la unión al ADN (hélices $\alpha 1$ a $\alpha 4$); por otra, un dominio intermedio o medio (hélices $\alpha 6$ a $\alpha 10$); y finalmente, uno C-terminal (hélices $\alpha 11$ a $\alpha 13$), que media la dimerización del receptor necesaria

INTRODUCCIÓN

para la unión al operador de ADN palindrómico de la SaPI (Figura I. 28). La unión entre el dominio intermedio y el C-terminal se realiza por una conexión desestructurada de 20 residuos que da flexibilidad entre estos dos dominios. Ensayos *in vitro* e *in vivo* mostraron que el dominio intermedio es el encargado de reconocer Duts triméricas, mientras que el dominio C-terminal reconoce Duts diméricas. La resolución de las estructuras tridimensionales de los complejos de los dominios N-terminal e intermedio de StI con la Dut trimérica del fago $\phi 11$ de *S. aureus* y del dominio C-terminal de StI con la Dut dimérica del fago $\phi O11$ de *S. aureus*, permitieron elucidar la base molecular de la interacción entre el represor y ambas Duts, dianas biológicas del represor.

La estructura de Dut trimérica del fago $\phi 11$ ($\phi 11$ Dut) en complejo con un fragmento N-terminal de StI (StI^{N-ter}) que incluía los dominios HTH e intermedio mostró que Dut mantiene su estado trimérico uniéndose a tres monómeros StI independientes (Figura I. 29).

Cada monómero de StI se une a uno de los tres centros activos del trímero de Dut e interactúa con residuos catalíticos de 4 de los 5 motivos conservados (Figura I. 30). Las interacciones que realiza StI con esos residuos catalíticos emulan a las interacciones con el sustrato dUTP y el Mg^{2+} , por tanto, StI sigue un mecanismo de mimetismo molecular para reconocer Duts triméricas (Figura I. 31).

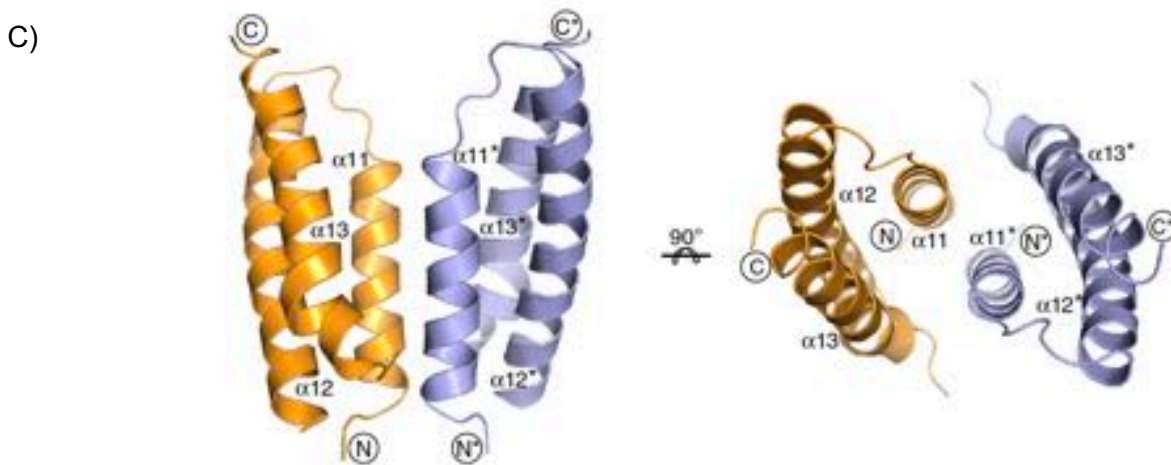
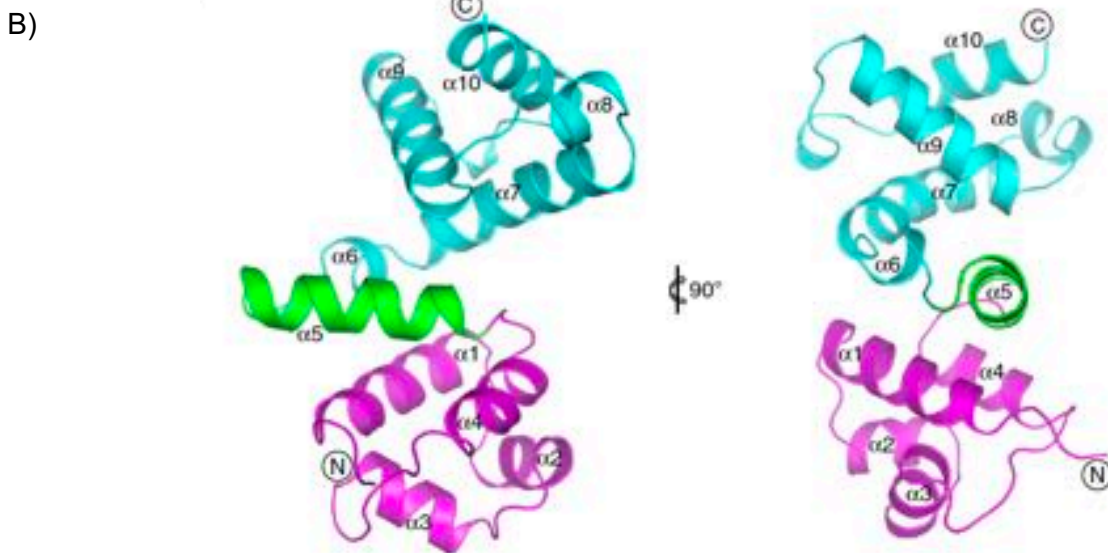
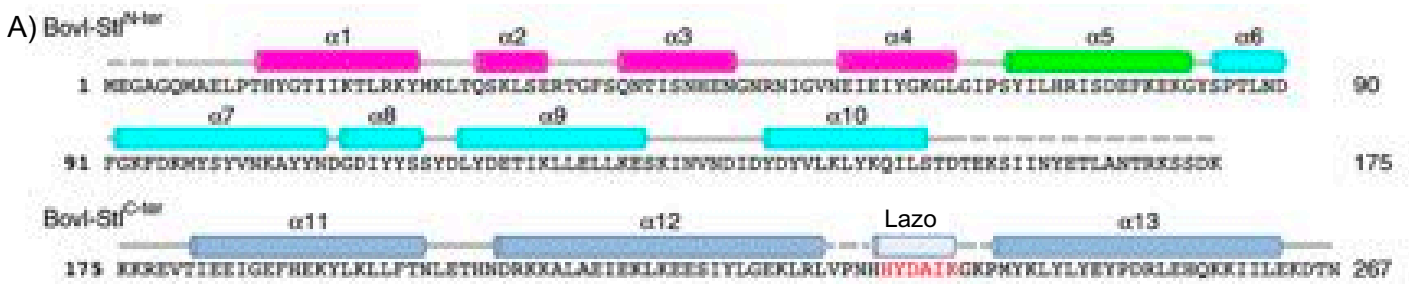


Figura I. 28. Estructura secundaria y terciaria del represor Stl dividida en los dos fragmentos utilizados para la resolución estructural. (A) Secuencia aminoacídica de Stl (segmentos N y C terminal) con las estructuras secundarias especificadas y numeradas en la parte superior. Cada estructura secundaria aparece coloreada en magenta, cian o azul marino según corresponden al dominio N-terminal (HTH), dominio intermedio o dominio C-terminal, respectivamente. En verde se representa la hélice $\alpha 5$, que sirve de conexión entre los dominios N-terminal e intermedio. El fragmento de la secuencia en color rojo corresponde con el lazo (no aparece en la estructura tridimensional). (B) Estructura tridimensional de los dominios N-terminal e intermedio con el mismo código de colores que en A. (C) Estructura tridimensional del dímero del dominio C-terminal con un protómero en azul marino y otro en naranja. Figura tomada de *Rafael Ciges-Tomas et al. 2019*.

INTRODUCCIÓN

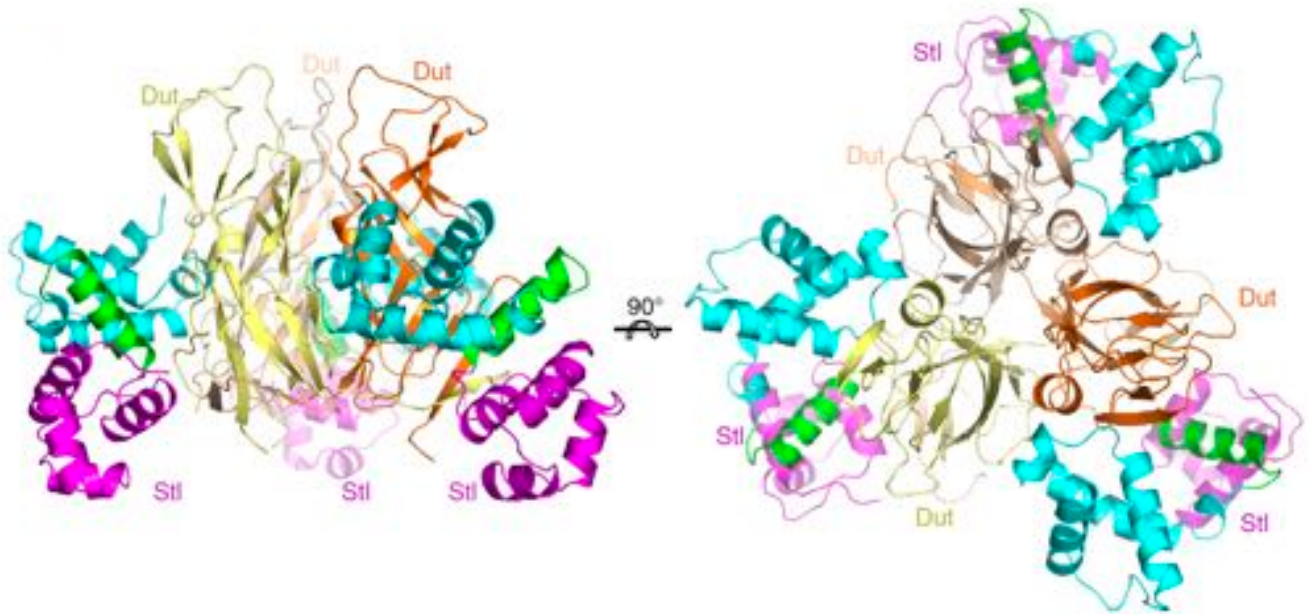


Figura I. 29. Estructura tridimensional del complejo de Dut trimérica del fago $\phi 11$ de *S. aureus* con Stl^{N-ter}. Un trímero de Dut, donde cada protómero está coloreado con tonos amarillo-naranja, tiene unidos tres monómeros de Stl^{N-ter} coloreados según el dominio al que pertenecen (magenta, cian, azul marino o verde según corresponden al dominio N-terminal (HTH), dominio intermedio, dominio C-terminal o hélice $\alpha 5$, respectivamente). Se muestran dos vistas ortogonales. Figura tomada de *Rafael Ciges-Tomas et al. 2019*.

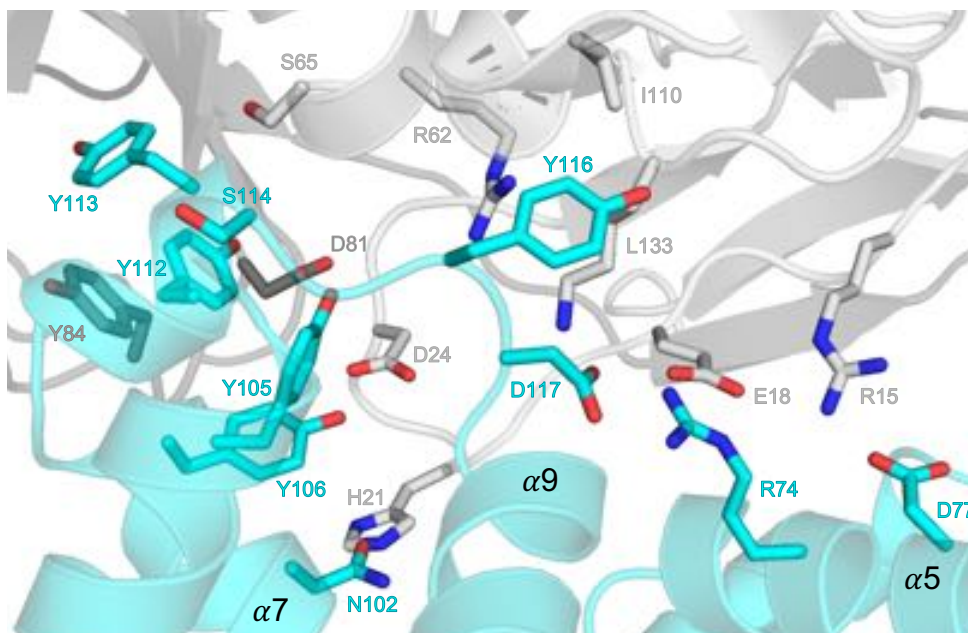


Figura I. 30. Vista detallada del centro activo del complejo $\phi 11$ Dut-Stl^{N-ter}. Se representan en varillas y se etiquetan los residuos implicados en la interacción (Stl^{N-ter} en cian y Dut en gris). Código PDB: 6H4C

INTRODUCCIÓN

Por su parte, la estructura del complejo entre el dominio C-terminal de StI ($\text{StI}^{\text{C-ter}}$) con la Dut dimérica del fago ϕO11 (ϕO11Dut) es el único ejemplo estructural disponible para observar las interacciones de $\text{StI}^{\text{C-ter}}$ con Duts diméricas. Esta estructura también muestra un mecanismo molecular similar, donde residuos de $\text{StI}^{\text{C-ter}}$ interactúan con residuos catalíticos conservados de Duts diméricas (Rafael Ciges-Tomas et al. 2019).

Sin embargo, para la interacción entre Dut dimérica y el represor, este último induce la disociación del dímero de Dut, formando un heterocomplejo $\phi\text{O11Dut-StI}^{\text{C-ter}}$. Además, en este complejo StI media contactos con las hélices implicadas en la dimerización de Dut (Figura I.32 e I. 33). Por tanto, el represor se une a la Dut dimérica por un doble mecanismo de mimetismo, emulando tanto al sustrato como al monómero complementario en la oligomerización.

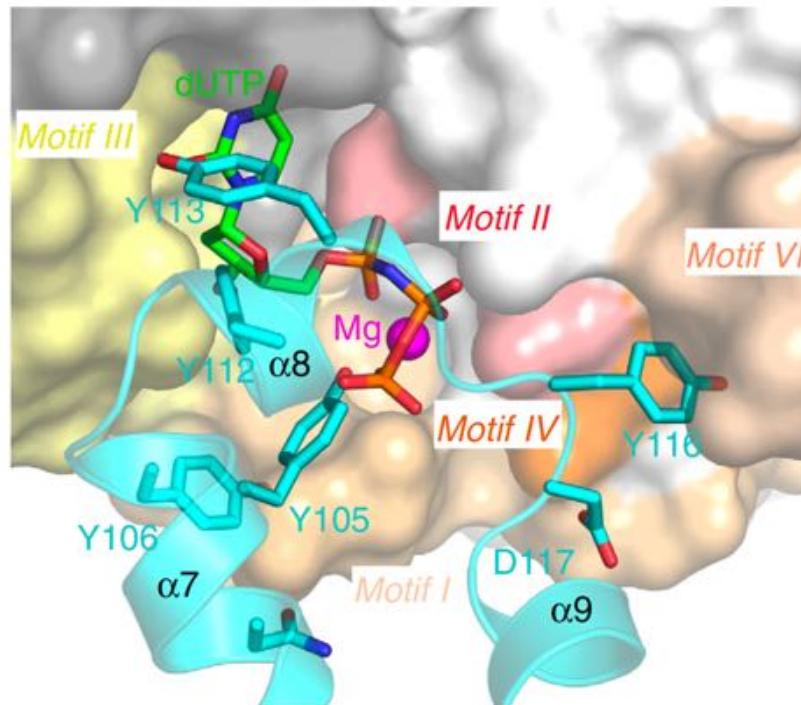


Figura I. 31. StI mimetiza las interacciones del sustrato dUTP con ϕO11Dut . La superposición de la Dut unida al dUTP (PDB 4GV8) con la Dut unida al StI (PDB 6H4C) muestra la clara estrategia de mimetismo llevada a cabo por el represor para unirse a la Dut. Se representa en varillas el sustrato (átomos de C en verde, O en rojo, N en azul y P en naranja) y las cadenas laterales de los residuos de StI implicados (en cian y etiquetados). El ión Mg^{2+} se representa como una esfera en color magenta. La Dut se representa en superficie y en distintos colores según el motivo: I en color trigo, II en rojo, III en amarillo, IV en naranja y VI en color crema. Figura tomada de *Rafael Ciges-Tomas et al. 2019*.

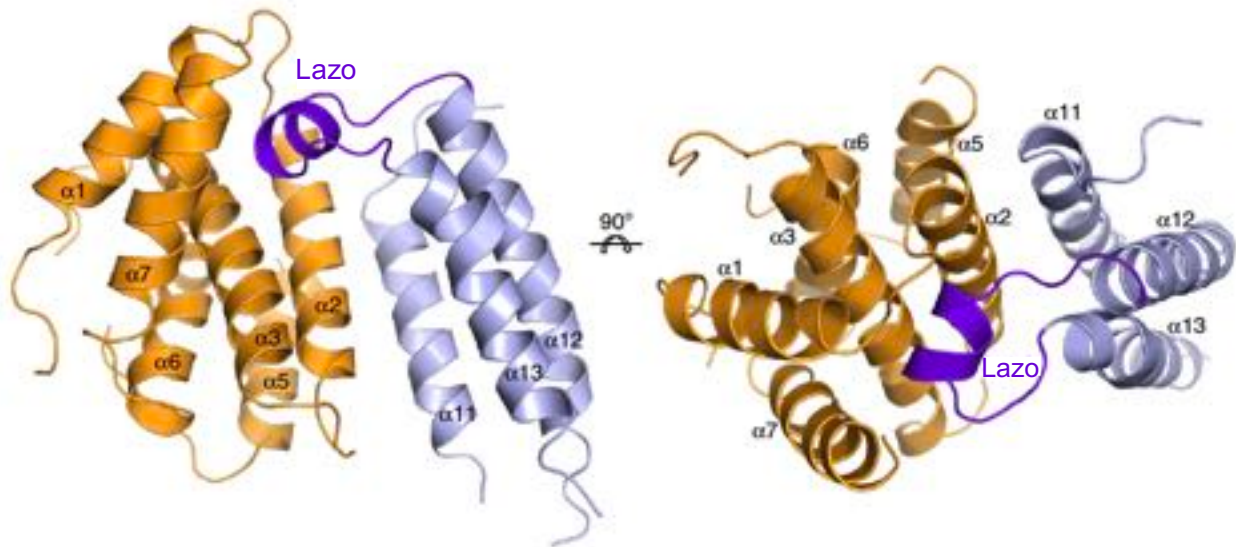


Figura I. 32. Estructura tridimensional del complejo de Stl^{C-ter} con la Dut dimérica del fago φO11 de *S. aureus*. El lazo del represor (azul-gris) que es proyectado sobre el centro activo de la Dut (naranja) está destacado en morado. Se identifican las estructuras secundarias de ambas proteínas y se muestran dos vistas ortogonales. Figura tomada de *Rafael Ciges-Tomas et al. 2019*.

La resolución de las estructuras de ambos complejos ratificó, por una parte, que la entrada del dUTP o del Stl en el centro activo de la Dut es puramente excluyente; por otra, que el complejo Dut : dUTP es inaccesible para el represor; y finalmente, que no es la actividad enzimática de Dut la necesaria para desreprimir el ciclo SaPI (Tormo-Más et al. 2010), sino que el papel de la Dut en el ciclo de la SaPI se restringe a su inducción (Maiques et al. 2016).

La estrategia de mimetización llevada a cabo por Stl se dirige a un proceso del fago y no a un dominio o secuencia aminoacídica específica. De este modo, la posibilidad de escape por variaciones en la secuencia es muy baja dado que los residuos catalíticos están altamente conservados dentro de cada familia. Además, la sustitución de una Dut por otra de diferente familia tampoco permite el escape ya que Stl reconoce a ambas familias. Dado que el reconocimiento del dUTP y su mecanismo de hidrólisis es diferente en Duts diméricas y triméricas, Stl ha evolucionado hasta reclutar dos dominios independientes para reconocer ambos tipos de enzimas e impedir la evasión de los fagos. La precisión mostrada en el mimetismo del sustrato y la alta conservación de los residuos catalíticos en Duts de diferentes organismos podría indicar que Stl tiene un amplio espectro de interacción con Duts, y se ha propuesto como un inhibidor proteico universal de Duts (Benedek et al. 2018; Hirmondó et al. 2015; Nyíri et al. 2018).

INTRODUCCIÓN

Diferentes estudios han confirmado este escenario mostrando que Stl de SaPI_{bov1} es capaz de interactuar e inhibir las Duts triméricas de *Mycobacterium tuberculosis*, *Drosophila melanogaster* y humano, cuya actividad enzimática se reduce por la naturaleza de la interacción (hasta el 40 % en la Dut de *D. melanogaster*, del 80 % en Dut de *M. tuberculosis* y del 100 % en Dut del fago $\phi 11$ de *S. aureus* (Benedek et al. 2018)). Aunque la caracterización *in vitro* de la interacción de Stl con Duts triméricas procariontas y eucariotas propone un mecanismo molecular de reconocimiento similar al descrito en Duts de fago de *S. aureus*, se han observado algunas discrepancias en cuanto a ratios molares, proponiéndose ratios equimolares (un trímero de Dut con tres monómeros de Stl), pero también estequiometrías alternativas (un trímero de Dut interactuando con un dímero de Stl o 2 monómeros del represor) (Benedek et al. 2018).

A) Zona dimerización

B) Centro activo

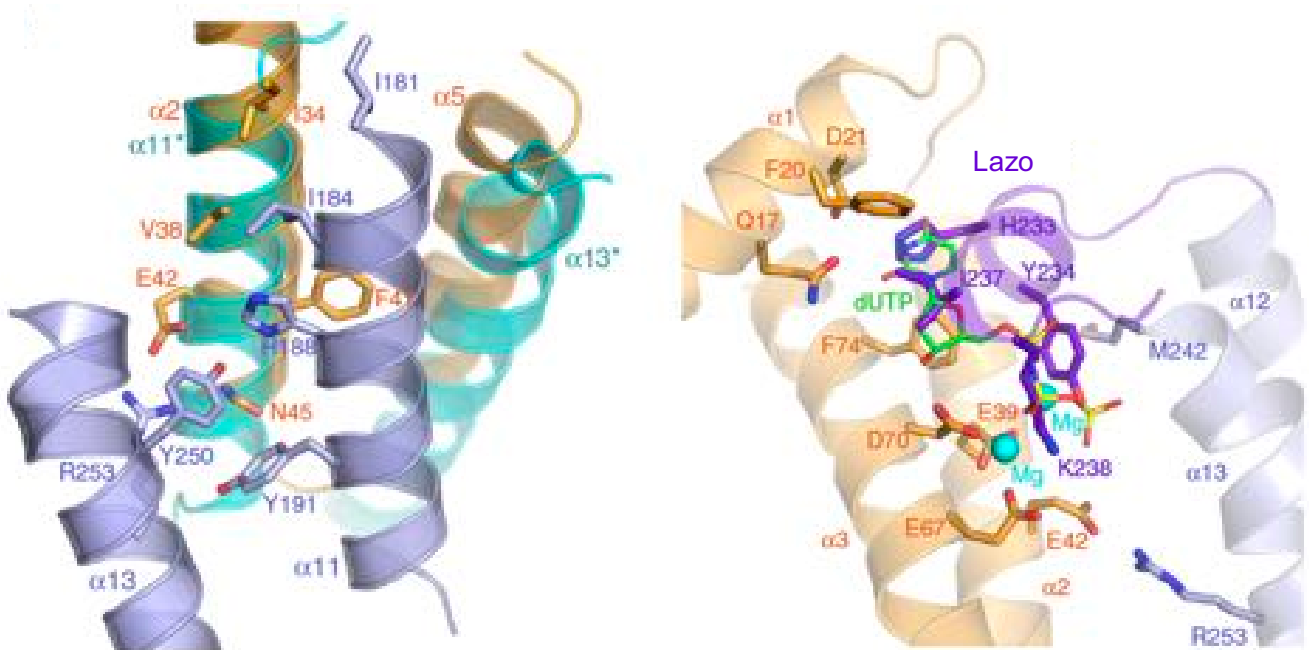


Figura I. 33. Estrategia de mimetización por parte del Stl para unirse a Dut dimérica del fago de *S. aureus* $\phi O11$. (A) Tanto Dut como Stl establecen interacciones mimetizándose las respectivas zonas de dimerización. El protómero de Stl se colorea en azul y el protómero de Dut en naranja, con las cadenas laterales de los residuos implicados etiquetados y representados en varillas y coloreados según átomo (O en rojo, N en azul, C en el respectivo). Las hélices semitransparentes y numeradas en color cian son las correspondientes al otro protómero de Stl^{C-ter}. (B) La superposición de la Dut unida al dUTP (PDB 5MIL) con la Dut unida al Stl (PDB 6H4B) muestra la misma estrategia de mimetismo llevada a cabo por el represor para unirse a Duts diméricas. Se representa en varillas el sustrato (átomos de C en verde, O en rojo, N en azul y P en amarillo) y las cadenas laterales de los residuos de Stl y Dut implicados (etiquetados y coloreados según átomo: O en rojo, N en azul, C en azul o morado si pertenecen al represor y en naranja si pertenecen a Dut). Los iones Mg²⁺ se representan como esferas en color cian. Figura tomada de *Rafael Ciges-Tomas et al. 2019*.

6. Proteínas MazG.

6.1. ¿Qué son las proteínas MazG?

Las MazG son proteínas de la superfamilia todo- α NTP pirofosfohidrolasas no específicas, generalmente capaces de hidrolizar oxi y desoxinucleótidos canónicos y no canónicos a sus respectivos monofostatos y PPI (Moroz et al. 2005). Esta familia está formada a su vez por las siguientes subfamilias: dCTPasas, fosforibosil-ATP pirofosfatasa HisE, NTP pirofosfatasa MazG, diversas familias de proteínas como la proteína humana XTP3TPA y la proteína DR2231 de *Deinococcus radiodurans* (Moroz et al. 2005). Esta superfamilia se incluye a su vez en la superfamilia de NTP-hidrolasas, que engloba además a las hidrolasas Nudix, las α/β dITPasas, y las Duts (Mota, Gonçalves, and de Sanctis 2016).

Estas proteínas realizan funciones de *house-cleaning* al hidrolizar NTPs no canónicos y son homólogas funcionalmente a las hidrolasas de la familia Nudix, con las que no están relacionadas estructuralmente. Con sus actividades enzimáticas previenen la incorporación de los nucleótidos di- y trifosfatos aberrantes en el ADN o el ARN naciente (Moroz et al. 2005).

Las proteínas MazG están ampliamente distribuidas en los tres dominios del árbol de la vida (Moroz et al. 2005). Podrían sustituir a las Duts en ciertos organismos que son de vida libre, como *Thermotoga maritima* o *Deinococcus radiodurans*, que no codifican para Duts conocidas ni para dCTP desaminasa (principal enzima que produce dUTP). También hay microorganismos que no codifican para ninguna Dut, pero esto podría ser explicado por un tipo de vida parasítica en la que utilicen la maquinaria del huésped para su replicación (Moroz et al. 2004).

6.1.1. Organización de dominios MazG.

Las MazG se clasifican en cuatro subgrupos de acuerdo con la presencia de dominios adicionales al de MazG y su organización (Tabla I. 1) (Lu et al. 2010):

- Tipo I: representado por MazG de especies de *Vibrio*. Contienen un único dominio MazG.
- Tipo II: es el predominante en Actinobacteria, y se compone de dos posibles organizaciones (subgrupos) en las que siempre está presente un dominio MazG flanqueado por uno o dos dominios de función desconocida. En el caso de estar flanqueado por un único dominio, éste se localiza en el extremo C-terminal.
- Tipo III: presente en *Proteobacteria*, y conteniente de dos dominios MazG en tándem.
- Tipo IV: se organizan en tándem dos dominios MazG fusionados a un dominio metilasa en el extremo N-terminal.

Tabla T. I. 1. Clasificación de las proteínas MazG en función de la organización de sus dominios. Se dividen en 4 subgrupos, el primero presenta el dominio MazG de manera independiente, el segundo presenta el dominio MazG fusionado a uno o dos dominios sin función conocida, el tercero lo presenta en dos dominios fusionados en tándem, y el cuarto lo presenta en dos dominios en tándem junto a un dominio con actividad metilasa en el extremo N-terminal. Tabla adaptada de *Lu et al. 2010*.

Tipo	Organización de los dominios	Filo que lo contiene
I		Ampliamente distribuido
II		Actinobacteria (45 %)
		Actinobacteria (35 %)
III		Bacteroidetes (86 %)
		Proteobacteria (80 %)
		Thermotogae (100 %)
IV		Firmicutes (50 %)

La proteína MazG de *T. maritima*, *E. coli* y muchas otras bacterias contienen duplicaciones en tándem del dominio principal (tipo III), pero solo uno de ellos parece ser requerido para la actividad NTP-pirofosfatasa (Moroz et al. 2005).

6.2. Estructura tridimensional de las MazG.

Las MazG son proteínas formadas totalmente por hélices α . El módulo estructural básico característico y que da nombre a la familia de las MazG está formado por 4 hélices α . De éstas, 2 hélices α conforman un dominio rígido, y las otras dos hélices α conforman un dominio móvil (Gonçalves, de Sanctis, and McSweeney 2011). Habitualmente, este módulo o dominio MazG oligomeriza interaccionando en disposición cabeza-cola. Para ello, las hélices del módulo rígido de un monómero se apilan de manera antiparalela con las hélices homólogas del otro formando un dímero. A su vez, en muchas ocasiones, el dímero oligomeriza con otro homodímero, asumiendo una estructura tetramérica de dímero de dímeros (Gonçalves, de Sanctis, and McSweeney 2011). En la Figura I. 34 A, puede visualizarse la estructura prototípica ejemplificada en la MazG de *Bacillus cereus* (Kim and Hong 2016).

INTRODUCCIÓN

Esta organización se ha observado tanto en MazG procariontas como eucariotas. En la Figura I. 34 B puede visualizarse la MazG de *Mus musculus*, en la cual queda patente la organización similar de las hélices α . En la misma figura puede observarse la MazG de *D. radiodurans*, formada por 2 monómeros que se ensamblan en una estructura dimérica a través de las hélices $\alpha 2$ y $\alpha 3$ (dominio rígido) de cada monómero (Gonçalves, de Sanctis, and McSweeney 2011) (Figura I. 34 C)

Aunque la estructura tridimensional es similar, no siempre está formada por 4 monómeros. Éste es el caso de MazG de *E. coli*, en la cual un monómero está formado por dos dominios que incluyen 5 hélices α cada uno, y se encuentran entrelazados mediante un conector. Ambos dominios interaccionan con los dos dominios de otro monómero a través de las hélices $\alpha 2$ y $\alpha 3$ de cada dominio (Lee et al. 2008), las mismas dos hélices α que en las otras estructuras. Es decir, este motivo estructural de las dos hélices es el que dirige la oligomerización (Figura I.34 D).

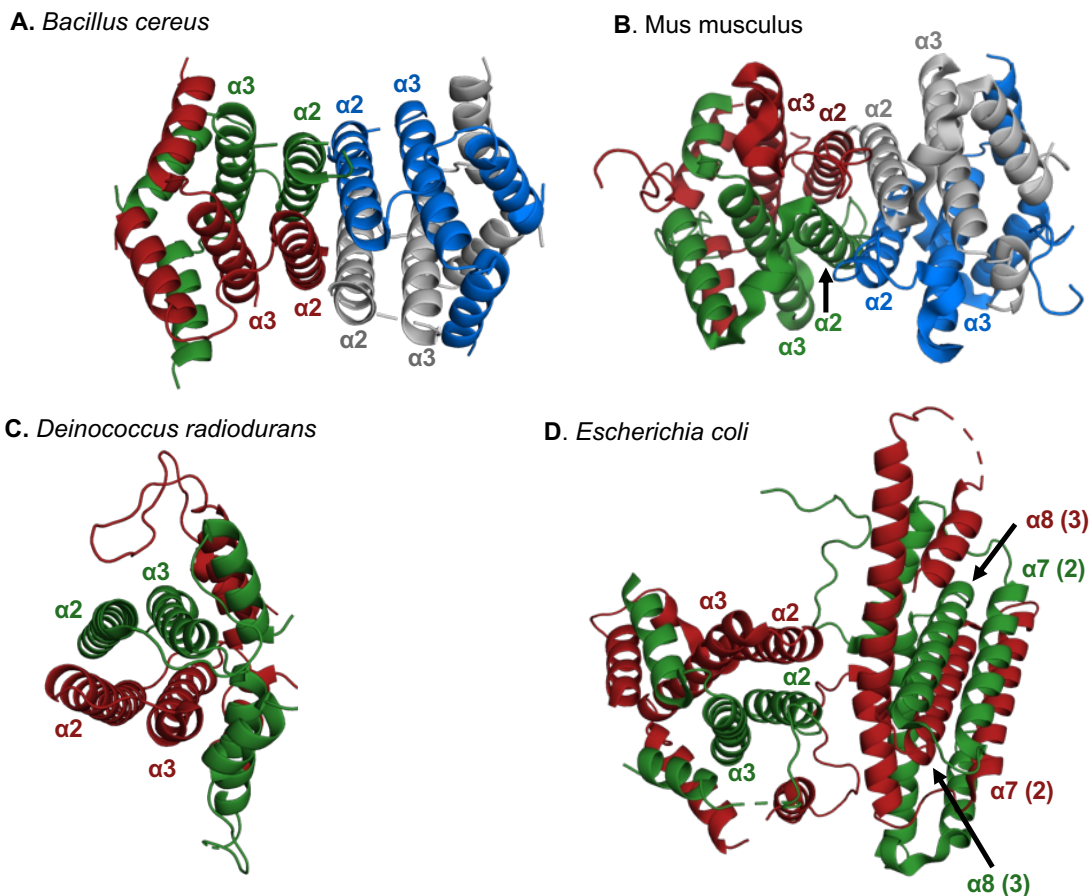


Figura I. 34. Estructuras tridimensionales de proteínas MazG de diferentes organismos. A) *Bacillus cereus* (código PDB 5IE9), B) *Mus musculus* (código PDB 2A3Q), C) *D. radiodurans* (código PDB 2YF4), D) *E. coli* (código PDB 3CRC). Las MazG de *D. radiodurans* y *E. coli* no tetramerizan. Se identifican las hélices α participantes en la dimerización y tetramerización ($\alpha 2$ y $\alpha 3$, y $\alpha 7$ y $\alpha 8$ también en el caso de dimerización de Dut de *E. coli*).

6.2.1. Estructura tridimensional de MazG de fago de *Enterococcus faecalis*.

En una tesis anterior realizada en el laboratorio, se resolvieron las estructuras tridimensionales de las MazG del fago ϕ FL2A de *Enterococcus faecalis* (EnMazG) y del fago ϕ IPLA7 *Staphylococcus epidermidis* (SeMazG), pertenecientes al subgrupo tipo I (Ciges-Tomas 2019). La identidad de secuencia entre ambas MazG es casi del 62 %, lo que se refleja en una alta homología estructural, formando ambas tetrámeros muy similares (RMSD de 0,452 Å (99 C α) y 0,513 Å (389 C α) para la superposición de los monómeros y tetrámeros), respectivamente. Por ello, en esta introducción solo se detallará la estructura tridimensional tetramérica de una de ellas, EnMazG. La estructura del monómero de EnMazG incluye 4 hélices α (α 1 a α 4) consecutivas seguida por 2 hebras β (β 1- β 2) que forman una horquilla (Figura I. 35. A).

Las hélices α 2 y α 3 forman una horquilla que genera el *core* del centro activo donde se posiciona el sustrato dUTP y Mg²⁺. Por su parte, las hélices α 1 y α 4 se extienden oblicuamente y en direcciones contrarias con respecto a la horquilla formada por α 2 y α 3. Al formarse el dímero, las hélices α 1 y α 4 quedan en paralelo con la hélice opuesta (α 1 con α 4) del monómero complementario, mientras que las hélices α 2 y α 3 de cada monómero quedan orientadas antiparalelamente generando el núcleo central del dímero. Por su parte, la hoja β se proyecta sobre el centro activo del monómero con el que forma el dímero, posicionándose a modo de tapa sobre su centro activo (Figura I. 35.B).

Para formar el tetrámero, los homodímeros enfrentan sus hélices α 2 de forma paralela, formándose un haz de cuatro hélices que constituye el núcleo del tetrámero (Figura I. 35.C). La estructura tetramérica presenta dos estados, abierto y cerrado, que corresponden a la forma *apo* y unido a sustrato, respectivamente.

La entrada del sustrato en el centro activo supone la transición del estado abierto al cerrado, en la cual se produce una rotación en cuerpo rígido del dímero (a partir de una región fija formada por los residuos C-terminales de la hélice α 2 y su conexión con la hélice α 3) que provoca la reducción de la superficie de tetramerización. El tetrámero resultante es más compacto, lo cual es clave para la correcta organización de los centros activos de la enzima. Finalmente, el extremo C-terminal se ordena sobre el centro activo (Ciges-Tomas 2019).

INTRODUCCIÓN

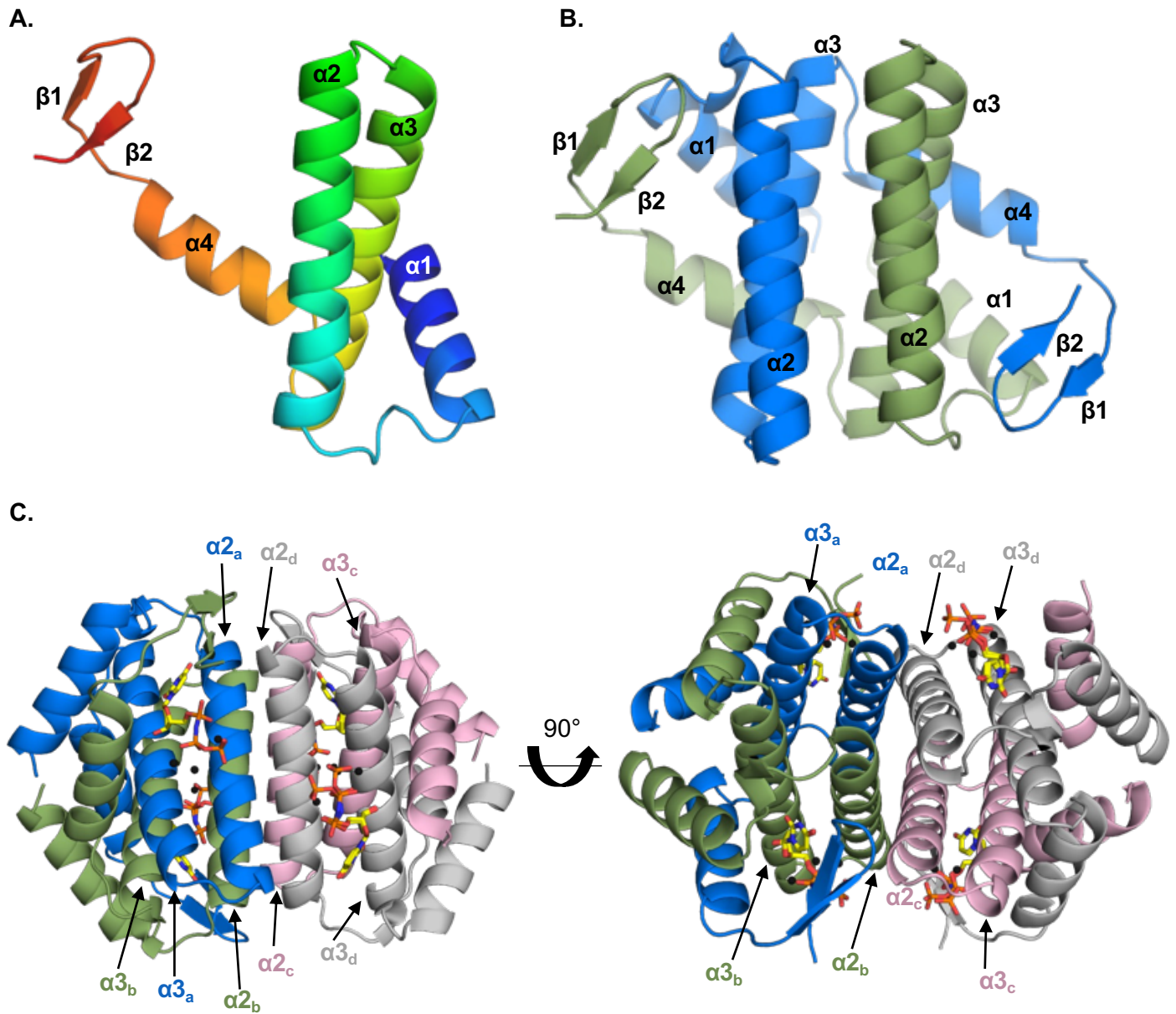


Figura I. 35. Estructuras tridimensionales de la proteína EnMazG. A) Monómero con las 4 hélices α y la hoja β identificadas. B) Dímero con los elementos estructurales identificados. Un monómero está coloreado en verde y el otro en azul. C) Dos vistas ortogonales del tetramero, con las hélices $\alpha 2$ y $\alpha 3$ identificadas en cada monómero. Ambas hélices son las implicadas en la dimerización entre monómeros, mientras que las hélices $\alpha 2$ son las participantes en la tetramerización. Cada monómero se identifica con un color y una letra: azul-a, verde-b, rosa-c, gris-d. También se representan los análogos de nucleótido no hidrolizables (dUpNHpp) coloreados en función del átomo (C amarillo, P en naranja, O en rojo y N en azul). Los cationes de Mg^{2+} se representan como esferas negras. PDB tomado de Ciges-Tomas 2019.

6.2.2. Organización del centro activo de MazG del fago ϕ FL2A de *Enterococcus faecalis* (EnMazG).

Cada tetrámero posee 4 centros activos, y en cada uno de ellos participan residuos de tres de los cuatro monómeros que lo conforman. Esta participación sinérgica de monómeros explica la cooperatividad observada previamente en EnMazG (Ciges-Tomas 2019). Esta enzima es específica de nucleótido, degradando solo el dUTP a dUMP y PPi.

El monómero que nuclea cada centro activo aporta desde las hélices $\alpha 2$ y $\alpha 3$ la mayor parte de residuos en el reconocimiento y selección del sustrato. El anillo de uracilo establece puentes de hidrógeno con la cadena lateral de la glutamina (Q25, hélice $\alpha 2$), además de contactos hidrofóbicos con las cadenas principales del W11, L17, K28, V57, T58, e I61 (todos de provenientes de la hélice $\alpha 3$). Tras el reconocimiento, el sustrato es fijado en esa posición mediante la cadena lateral de la K28 (hélice $\alpha 2$), la cual se apila sobre el anillo del uracilo. Por su parte, el anillo de desoxirribosa establece interacciones de tipo hidrofóbico con los residuos V57 y T58 al introducir el C2' entre ellos (Figura I. 36). Por su parte, el segundo monómero del dímero aporta la hoja β que actúa como una tapa sobre el centro activo, y establece interacciones hidrofóbicas a través de la fenilalanina 93 con el nucleósido (Figura I.36). Esta organización del centro activo descarta la entrada de otros nucleótidos por impedimentos estéricos o repulsión de cargas. En el caso de la adenina y la guanina no podría acomodarse el anillo de purina, y tampoco podría establecerse el puente de hidrógeno entre el grupo carboxilo en la posición C2' del uracilo con el residuo Q26 por la ausencia de éste en ambas bases nitrogenadas. En cuanto a las pirimidinas, la timina sufriría impedimentos estéricos entre su grupo metilo en C5' y el residuo F93, mientras que la citosina lo haría por repulsión de cargas positivas entre el grupo amina (N4) y el nitrógeno de la cadena principal del mismo residuo (F93). Por su parte, los oxinucleótidos quedarían descartados porque los residuos que flanquean el anillo de la desoxirribosa (V57 y T58) supondrían un impedimento estérico y de repulsión de cargas por el grupo hidroxilo de la ribosa (Ciges-Tomas 2019).

La unión y correcta disposición de los fosfatos y los tres cationes Mg^{2+} también requiere de la participación de las hélices $\alpha 2$ y $\alpha 3$. El fosfato α , a través del oxígeno 1, interacciona con la K28 de la hélice $\alpha 2$. Los residuos R84 y K95 (del monómero complementario del mismo homodímero) establecen puentes salinos con el oxígeno 2 del fosfato β y el oxígeno 2 del fosfato γ , respectivamente. La lisina 95 también establece contactos con el enlace fosfoanhidro entre los fosfatos β y γ (Figura I.37). El tercer monómero participante en el centro activo aporta el residuo arginina 42, y éste establece contacto con los fosfatos α y γ , ayudando a la estabilización del sustrato (Figura I. 37) (Ciges-Tomas 2019).

INTRODUCCIÓN

Los cationes Mg^{2+} se posicionan mediante un motivo de unión a metal similar al descrito previamente en la propia superfamilia de las todo- α NTPasas, y que corresponde con una secuencia consenso EExx(E/D) – EExx(E/D), localizada en dos hélices consecutivas. Se ha descrito en las MazG de especies del género *Vibrio*, *E. coli*, *D. radiodurans*, *B. cereus*, *S. solfataricus*, *B. subtilis* y *Mus musculus* (Robinson et al. 2007; Lee et al. 2008; Gonçalves, de Sanctis, and McSweeney 2011; Kim and Hong 2016). En el caso de EnMazG el primer motivo se conserva, siendo $E_{31}E_{32}xxE_{35}$, mientras que el segundo es $D_{50}xxxD_{54}$. Esta péntada acídica se localiza en las hélices $\alpha 2$ y $\alpha 3$. Los tres cationes establecen una geometría octaédrica, con una hexacoordinación en la que participan, además de la péntada acídica, oxígenos de los tres fosfatos del sustrato y moléculas de agua. Cada uno de ellos ocupa una posición que determina su participación en la catálisis. Mientras que el Mg^{2+} que ocupa la posición 1 interacciona con los oxígenos de los 3 fosfatos, los ocupantes de las posiciones 2 y 3 se proponen como los responsables de activar a la molécula de agua catalítica (Figura I.37) (Ciges-Tomas 2019).

La numeración de los cationes está relacionada con la numeración en Duts diméricas. Éstas utilizan tres cationes para la hidrólisis de dUDP, mientras que utilizan dos para la de dUTP. Las posiciones numeradas como 1 y 2 son las coincidentes en ambos sustratos, mientras que la posición 3 es asociada por correlación. Es cierto que en el caso de MazG de *Deinococcus radiodurans* se ha encontrado una referencia en que el catión 1 mantiene la nomenclatura, pero el catión 2 la intercambia con el 3. No obstante, dado que la posición del tercer catión no es la misma que en EnMazG, se ha decidido continuar con la nomenclatura asociada con las Duts diméricas (Moroz et al. 2004).

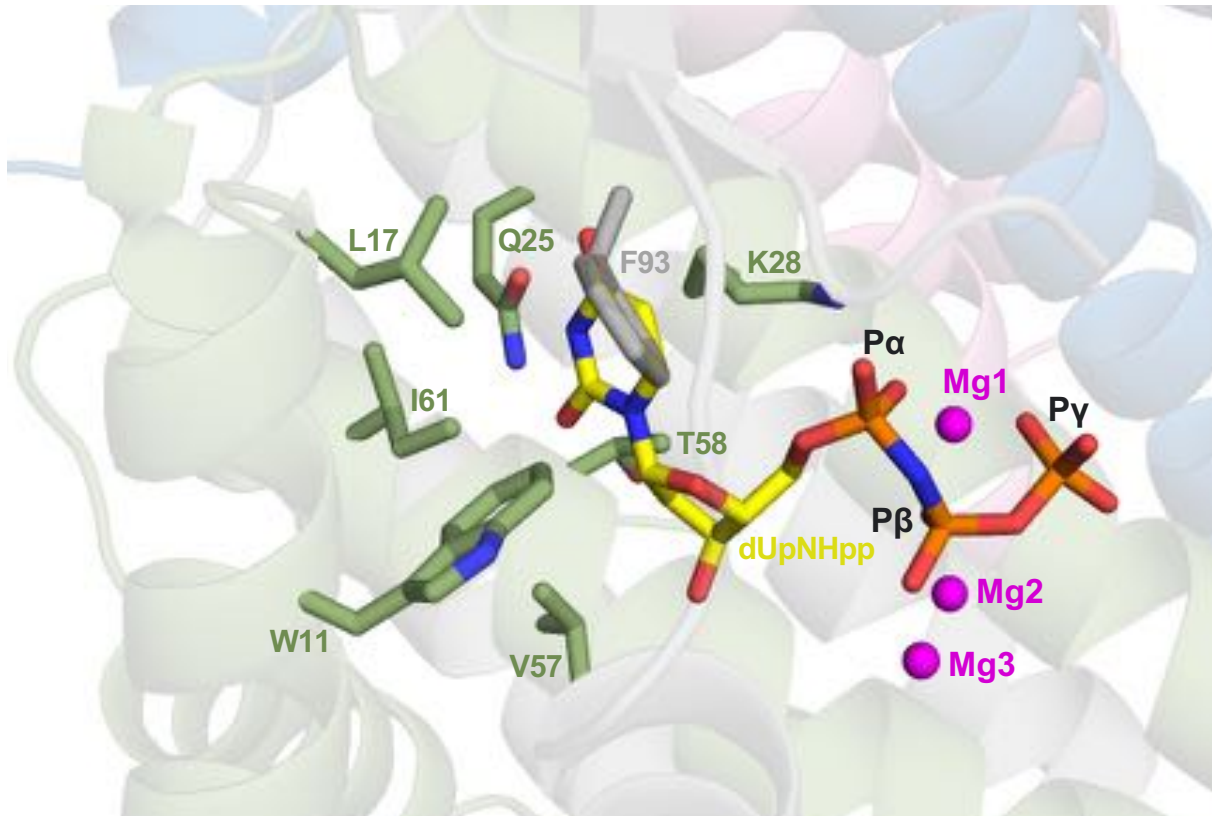


Figura I. 36. Selectividad y acomodación de la base nitrogenada uracilo y la desoxirribosa en el centro activo de EnMazG. Se representan en 4 colores los monómeros: azul, rosa, verde y gris. El sustrato no hidrolizable dUpNHpp se representa en varillas y en colores en función del átomo (C amarillo, O rojo, N nitrógeno, P naranja). Las cadenas laterales responsables de la selectividad del sustrato se identifican y representan en varillas y en colores en función del átomo (C según el monómero al que corresponden, O en rojo, N en azul). Los cationes de Mg^{2+} se representan como esferas en color magenta. El archivo PDB utilizado fue tomado de *Ciges-Tomas 2019*.

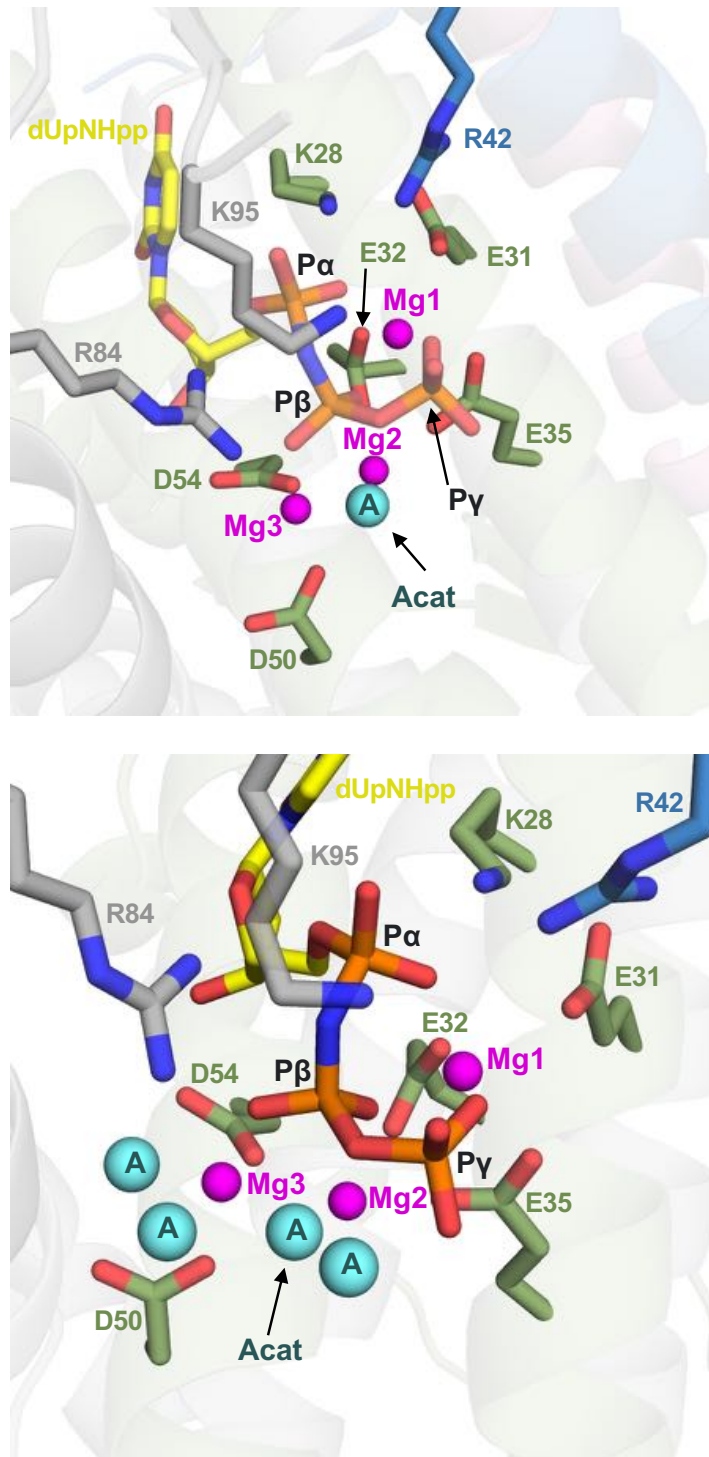


Figura I. 37. Posicionamiento de los tres fosfatos del nucleótido y los tres cationes Mg^{2+} en el centro activo de EnMazG. Se representan en 4 colores los monómeros: azul, rosa, verde y gris. El sustrato no hidrolizable dUpNHpp se representa en varillas y en colores en función del átomo (C amarillo, O rojo, N nitrógeno, P naranja). Las cadenas laterales responsables del posicionamiento del trifosfato y de los cationes se identifican y representan en varillas y en colores en función del átomo (C según el monómero, O en rojo, N en azul). Los cationes de Mg^{2+} se representan como esferas en color magenta y las aguas como esferas azules. El agua catalítica se identifica como Acat. Se muestran dos ángulos del centro activo para una mejor visualización. El archivo PDB utilizado fue tomado de *Ciges-Tomas 2019*.

6.2.3. El nucleótido dUTP puede adoptar dos conformaciones en el centro activo de EnMazG y SeMazG.

Cuando las estructuras de SeMazG y EnMazG en presencia del sustrato no hidrolizable dUpNHpp fueron resueltas, se observó que el nucleótido es capaz de asumir una doble conformación en el centro activo (Ciges-Tomas 2019). Ambas conformaciones difieren fundamentalmente en el fosfato γ y del nucleótido. Mientras que el fosfato α y parte del β mantienen su posición, el oxígeno 1 del fosfato β y el enlace fosfoanhidro entre los fosfatos β y γ rotan unos 25° respecto a su posición en la otra conformación (Figura I. 38). La conformación del nucleótido vista en otras Duts y MazG fue denominada “ $\gamma P - a$ ”, y la alternativa “ $\gamma P - b$ ” (Ciges-Tomas 2019).

En las dos conformaciones participan los 3 cationes de Mg^{2+} . No obstante, en la conformación $\gamma P - a$ es posible posicionar el agua catalítica, que es activada por los cationes 2 y 3 y que realizará el ataque nucleofílico sobre el fosfato β en dirección $\beta \rightarrow \alpha$; mientras que en la conformación $\gamma P - b$ esta molécula de agua no puede posicionarse y podría imposibilitar la hidrólisis. Por tanto, la conformación del nucleótido $\gamma P - a$ sería la catalíticamente activa, y la $\gamma P - b$ sería la catalíticamente inactiva (Figura I. 38) (Ciges-Tomas 2019).

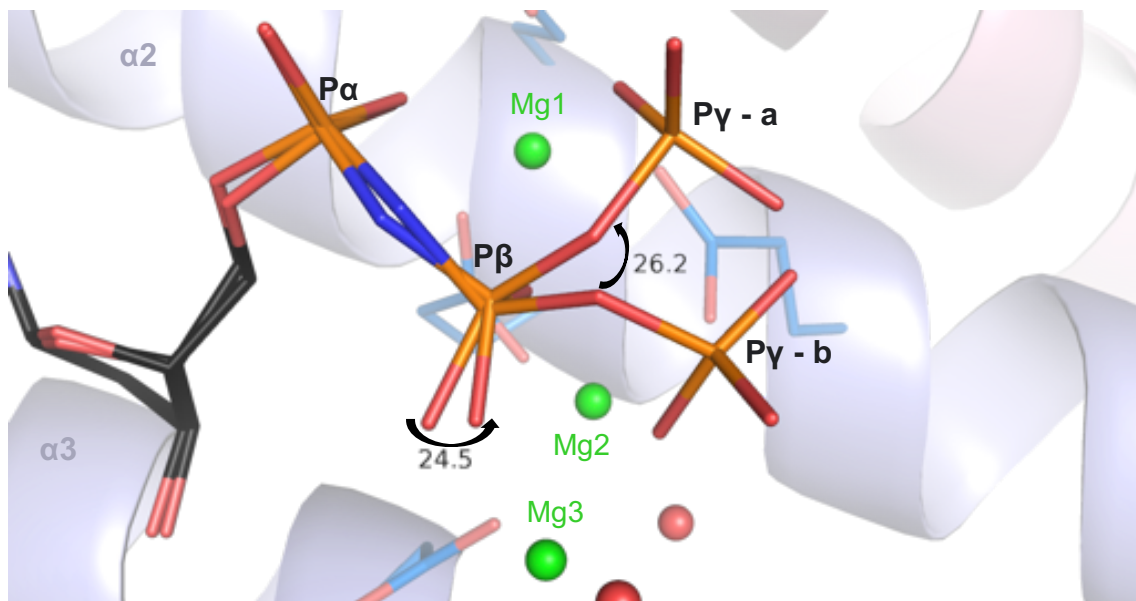


Figura I. 38. Diferencias entre las dos conformaciones adoptadas por el nucleótido en el centro activo de EnMazG y SeMazG. Las dos conformaciones se diferencian fundamentalmente por la rotación llevada a cabo por el O1 del fosfato β y por el enlace fosfoanhidro entre los fosfatos β y γ . Los cationes Mg^{2+} se representan en esferas verdes, y el nucleótido en varillas con los átomos de C en negro, N en azul, O en rojo y P en naranja. Figura tomada de Ciges-Tomas 2019.

INTRODUCCIÓN

Esta misma conformación alternativa inactiva había sido descrita previamente para MazG de *Deinococcus radiodurans* (DR2231) en presencia del nucleótido no hidrolizable dUMpNHpp (código PDB 5HWU) (Mota, Gonçalves, and de Sanctis 2016). En este mismo artículo también se describió otra doble conformación (código PDB 5HVA), siendo una de las conformaciones catalíticamente inactiva por imposibilitar el posicionamiento de la molécula de agua catalítica, y en la cual participan 2 cationes Mg^{2+} . Esta nueva conformación inactiva no coincide con la vista en SeMazG y EnMazG (Figura I. 39).

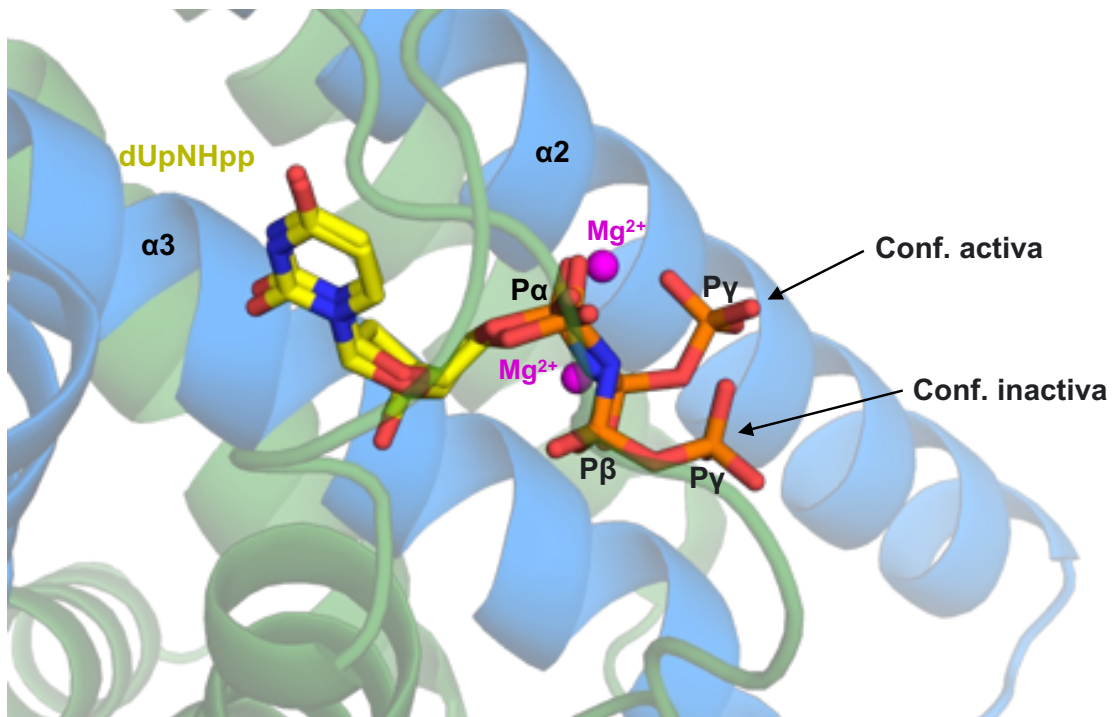


Figura I. 39. Doble conformación del nucleótido dUpNHpp en el centro activo de MazG de *D. radiodurans* (DR2231). Se representa en dos colores los monómeros de MazG, uno en azul y otro en verde. El nucleótido se representa en varillas y en colores según el átomo (C en amarillo, O en rojo, N en azul y P en naranja). Los cationes Mg^{2+} se representan en esferas en color magenta. Se indican las dos conformaciones del nucleótido. Código PDB 5HVA.

6.3. Las proteínas MazG degradan más de un nucleótido, y utilizan cationes distintos al Mg^{2+} .

Aunque las MazG de fago de *E. faecalis*, *S. epidermidis* y *D. radiodurans* (DR2231) solo degraden dUTP (Ciges-Tomas 2019; Gonçalves, de Sanctis, and McSweeney 2011), se ha descrito que ciertas MazG son capaces de hidrolizar sustratos alternativos. La MazG presente en *Vibrio* DAT722 es capaz de degradar el dCTP y el dATP (Mustafi et al. 2003). La MazG de *M. tuberculosis* degrada todos los dNTPs canónicos, además de los no canónicos dUTP, 5-OH-dCTP y 8-oxo-7,8-dihidro-2' dGTP (Lyu et al. 2013). La MazG de *B. cereus* degrada el ATP entre los NTPs canónicos (Kim and Hong 2016). Las MazG de *E. coli* y *T. maritima* degradan dATP, dGTP, dCTP y dTTP y también los correspondientes oxinucleótidos (Zhang and Inouye 2002; Zhang, Zhang, and Inouye 2003). La MazG de un fago del género *Synechococcus* degradada dGTP y dCTP. Este último caso se relaciona con la diferente composición genómica del fago y el hospedador ya que el genoma del primero es rico en adenina y timina mientras que el del segundo lo es en guanina y citosina, de modo que la actividad de la MazG del virus le facilitaría la propia replicación (Rihtman et al. 2019).

Además de la variabilidad de sustrato encontrada entre las distintas MazG, se ha descrito que ciertas MazG pueden utilizar cationes divalentes alternativos al Mg, como por ejemplo MazG de *D. radiodurans* (DR2231), capaz de degradar el dUTP tanto en presencia del Mg^{2+} como del Mn^{2+} con una eficiencia similar, y con actividad residual en presencia de Ca^{2+} (Gonçalves, de Sanctis, and McSweeney 2011). En las estructuras de DR2231 el centro activo está ocupado por el sustrato y 2 cationes divalentes, aunque no se descarta el uso de un tercer catión, al igual que ocurre en las MazG de fagos de *E. faecalis* y *S. epidermidis*.

6.4. Las proteínas MazG del tipo I están relacionadas con las Duts diméricas.

La homología estructural entre las Duts diméricas y las MazG del tipo I permite una buena superposición de un monómero de Dut dimérica con un dímero de MazG (Figura I. 40). Las secuencias de MazG contienen el conjunto de hélices α mínimo requerido para la formación del centro activo, el cual conserva una elevada similitud con el de las Duts diméricas. La región variable de las Duts diméricas, formada por entre 4 y 9 hélices α , se reduce a unos pocos residuos en MazG; mientras que las Duts poseen otras hélices en su cuerpo rígido requeridas para la estabilización (Moroz et al. 2005).

El motivo de unión a metal conservado en el centro activo de proteínas MazG también se conserva en Duts diméricas (los residuos del motivo de unión a metal provienen de los motivos conservados II y IV) (Figura I. 40), y constituye la firma de identidad de ambas proteínas junto a proteínas dCTPasas y HisE. Estos motivos constituyen el mínimo grupo de residuos requeridos para la hidrólisis de NTPs, y a excepción de ellos, la homología de secuencia es baja incluso entre proteínas MazG, aunque la homología estructural es elevada (Figura I. 41) (Gonçalves, de Sanctis, and McSweeney 2011; Moroz et al. 2005).

Ha sido propuesto que las Duts diméricas han evolucionado por duplicación génica y posterior pérdida de uno de los dos centros activos de una MazG tetramérica ancestral. Además, las Duts diméricas habrían añadido a lo largo de la evolución elementos estructurales al andamio de 8 hélices α ancestral (Moroz et al. 2005; Gonçalves, de Sanctis, and McSweeney 2011).

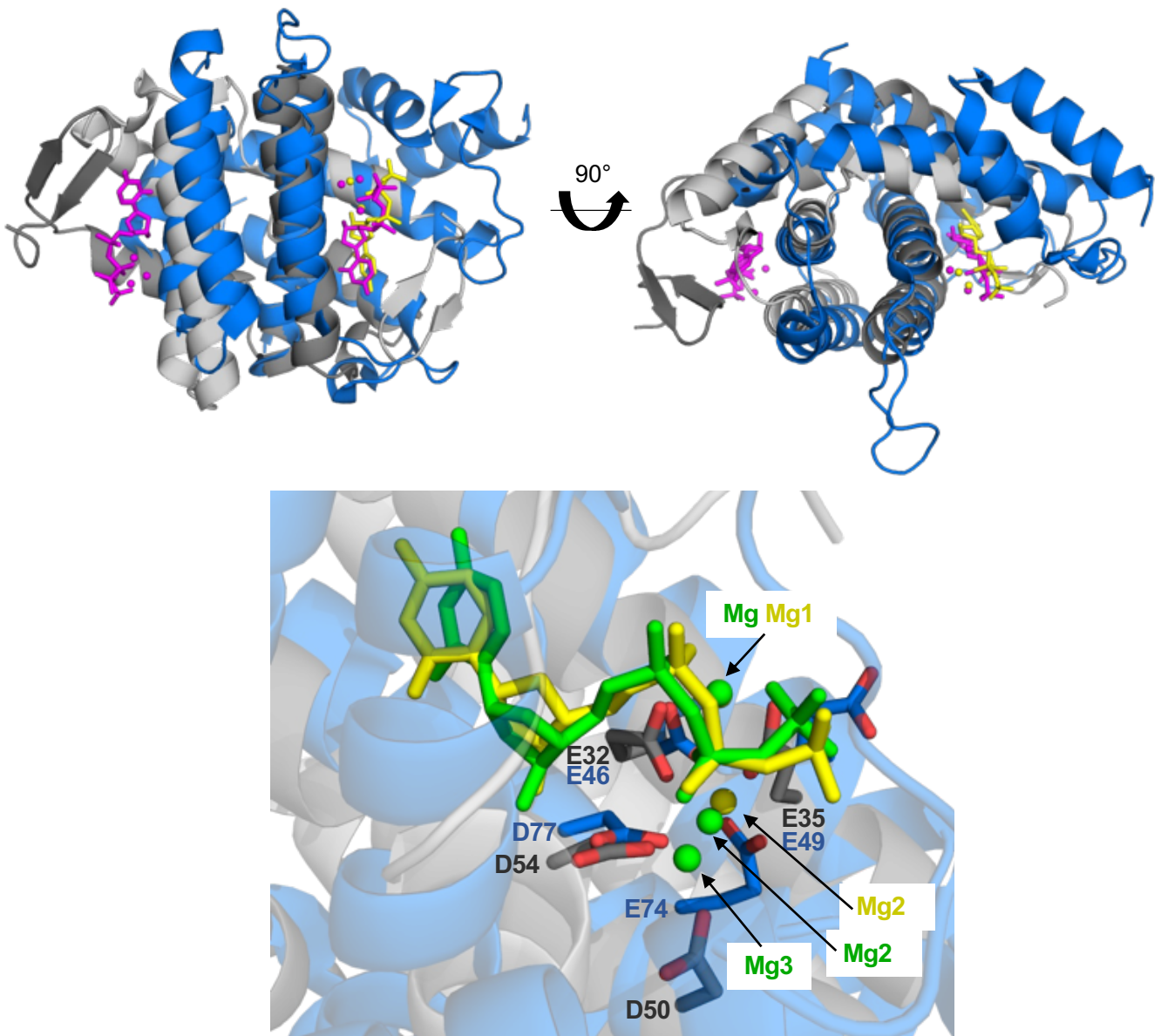


Figura I. 40. Superposición de las estructuras tridimensionales de EnMazG y Dut dimérica de *C. jejuni*. (Arriba) Se muestran dos vistas ortogonales de la superposición del monómero de la Dut dimérica (en azul) y un dímero de EnMazG (cada monómero de EnMazG se representa en una tonalidad grisácea distinta). Se representa en varillas el nucleótido y en esferas los cationes de Mg^{2+} . Tanto los sustratos como los iones se representan en color magenta en la estructura de la MazG y en amarillo en el caso de la Dut. El código PDB de la Dut dimérica es 2CIC. (Abajo) Superposición de los centros activos, donde se representa en varillas el sustrato y las cadenas laterales de los residuos del motivo de unión del metal (MazG, en gris: E32, E35, D50 y D54; Dut, en azul: E46, E49, E74 y D77). Los iones Mg^{2+} se representan en esferas. El sustrato y los cationes de MazG se representan verde, y los de la Dut dimérica en amarillo.

INTRODUCCIÓN

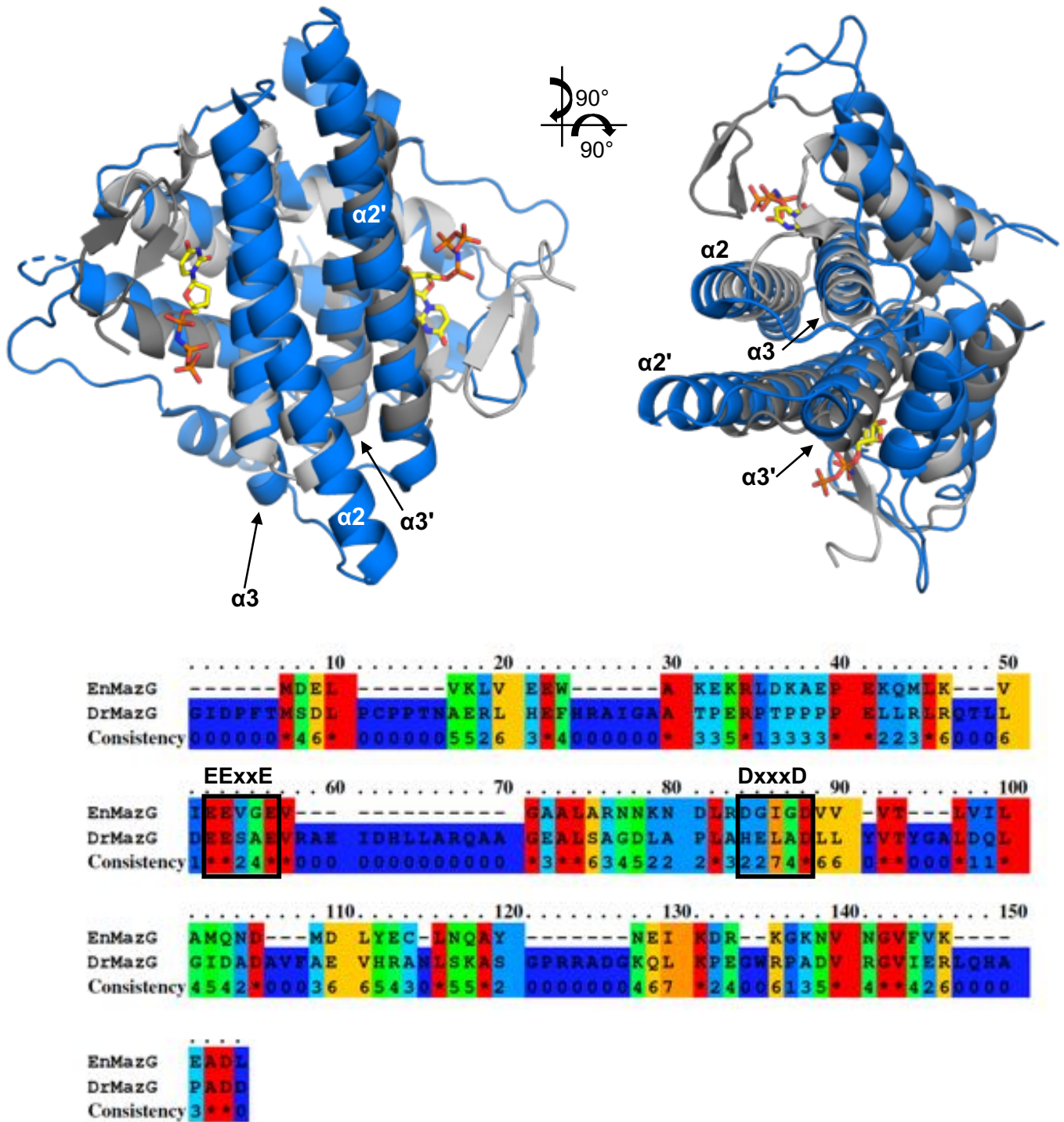


Figura I. 41. Alineamiento y superposición de las secuencias aminoacídicas y las estructuras de las MazG de *D. radiodurans* y de fago de *E. faecalis*. (Arriba) Se muestran dos vistas de la superposición de un dímico de ambas MazG (EnMazG en gris, DrMazG en azul). Se representan en varillas los nucleótidos (átomos de C en amarillo, O en rojo, N en azul y P en naranja). (Abajo) Alineamiento de secuencia de ambas MazG. Mientras que la homología estructural es elevada entre ambas proteínas, la conservación de residuos es baja y se restringe básicamente al primer motivo de unión a metal (EExxE). El segundo motivo de unión a metal solo es conservado en EnMazG, y en DrMazG solo se mantiene el segundo residuo de arpártico (DxxxD).

OBJETIVOS

OBJETIVOS

El objetivo general de la Tesis es ampliar la conocida capacidad de señalización de las Duts o variantes íntimamente relacionadas, arrojar luz sobre las bases moleculares de este proceso señalizador y caracterizar nuevas dianas.

Para alcanzar este objetivo general se establecen tres objetivos específicos que se detallan a continuación:

- Objetivo específico 1: Estudio biofísico y estructural de la interacción de Duts representativas procariotas y eucariotas con el represor StI de la isla de patogenicidad de *Staphylococcus aureus* bovina 1 (SaPI_{bov1}).
- Objetivo específico 2: Análisis del mecanismo catalítico de proteínas MazG de los fagos ϕ IPLA7 de *Staphylococcus epidermidis* y ϕ FL2A de *Enterococcus faecalis* a través de su estudio enzimático y estructural.
- Objetivo específico 3: Caracterización bioquímica y estructural de proteínas candidatas a interaccionar con la Dut de *Mycobacterium tuberculosis*.

MATERIAL Y MÉTODOS

1. TÉCNICAS EN BIOLOGÍA MOLECULAR

1.1. Construcción de vectores.

La obtención de todas las proteínas usadas en esta Tesis se inició con su expresión heteróloga en *Escherichia coli*, para lo cual se prepararon los correspondientes plásmidos de expresión.

1.1.1. Sistema de expresión génica.

Los plásmidos o vectores de expresión utilizados fueron pLIC-SGC1, pETNKI 1.1, 1.5, y 1.10, pNIC28-Bsa4 y pET28a. En todos ellos, los genes de interés se clonaron aguas abajo del promotor T7. Todas las construcciones utilizadas se basan en el operón *lac* de *E.coli*, y de ellas destacaremos los siguientes elementos (Fig. M.1):

- Promotor T7: secuencia de ADN a la cual se unirá la ARN polimerasa T7.
- Operador *lac*: secuencia de ADN a la cual se une el represor *lac*.
- Cola His: secuencia de ADN correspondiente a 6 o 10 residuos de histidinas para la posterior purificación por afinidad de la proteína de interés.
- Gen de interés, clonado aguas abajo del promotor.
- Terminador T7: secuencia de ADN que le indica a la T7 ARN polimerasa el fin de la transcripción.
- Gen *lacI*: codifica para el represor del operador *lac*.
- Promotor *lacI*: controla la expresión del gen *lacI*.

Las células de *E. coli* BL21 (DE3) contienen la secuencia codificante de la ARN polimerasa T7 en el profago λ DE3 bajo el control de un promotor *lac*, inducible por lactosa o su derivado no hidrolizable Isopropil- β -D-1-tiogalactopiranosido (IPTG). En ausencia de lactosa o IPTG, el plásmido con la construcción de interés expresa de manera constitutiva el represor *lacI*, que se une al operador *lac*. Cuando el inductor es añadido al medio, por una parte éste secuestra al represor *lacI* y por otra, induce la transcripción y expresión de la ARN polimerasa T7. En consecuencia, la ARN polimerasa producida se posiciona libremente sobre el promotor T7 e inicia la transcripción de los genes de interés (Figura M.1).



Figura M.1. Estructura común simplificada de las construcciones génicas contenidas en los vectores de expresión usados en esta Tesis. Se destacan en distintos colores los elementos clave para la transcripción génica.

1.1.2. Sistemas de clonación

Las construcciones definidas en el apartado anterior fueron obtenidas por distintos mecanismos de clonación: por enzimas de restricción, por el sistema independiente de ligación (*LIC: ligation independent cloning*, de sus siglas en inglés), por el método Gibson o por la vía sintética.

- Clonación mediante enzimas de restricción.

Para llevar a cabo la clonación por este mecanismo, se analizaron los sitios de clonación múltiple (*MCS, multiple cloning site*) de cada vector con el fin de elegir aquellas enzimas de restricción de tipo II de corte único que permitieran el posicionamiento del gen en el sitio y orientación deseada en la construcción, y que al mismo tiempo su sitio de restricción no se encontrara en las secuencias de los genes de interés a clonar. Una vez seleccionadas las enzimas de restricción (dos para cada gen y vector), las secuencias de restricción fueron añadidas a los genes de interés mediante reacción en cadena de la polimerasa (PCR) con los cebadores correspondientes y el preparado de la ADN polimerasa de alta fidelidad NZYProof 2x Colourless (NZYTech), que ya contiene los nucleótidos, el tampón y la enzima. Los genes de interés clonados por este sistema se habían obtenido previamente por PCR desde ADN genómico de la bacteria (*M. tuberculosis*). Una vez obtenidos los productos de PCR, se comprobó por gel de agarosa al 1% la presencia de ADN inespecífico y las bandas correspondientes a los genes de interés fueron recortadas del gel y purificadas por el kit GeneJET® Purification Kit (ThermoScientific). A continuación, los productos de PCR puros y los correspondientes plásmidos fueron sometidos a las enzimas de restricción según el protocolo comercial de ambas enzimas (ThermoScientific), por el cual, en reacciones de un volumen total de 20 µL, se empleó una unidad de enzima por microgramo de ADN, ya fuese el producto de PCR o el plásmido. Las mezclas, que incluían también el tampón óptimo para las enzimas, se incubaron a 37 ° C durante 2 h.

Tras la incubación, la reacción con el gen de interés se pasó por las columnas del kit nombrado previamente para aislar el ADN y éste se eluyó con el mínimo volumen de agua milliQ (20 µL) para obtener la máxima concentración. La reacción con el plásmido fue sometida a electroforesis en gel de agarosa al 1% para separar las bandas del plásmido digerido y sin digerir. La banda correspondiente al plásmido digerido fue recortada del gel de agarosa y purificada utilizando el mismo kit. De nuevo, se eluyó con el mínimo volumen de agua (20 µL).

Para realizar la reacción de ligación una vez cuantificados por NanoDrop™ el plásmido digerido y el gen, se preparó una mezcla de un volumen total de 20 µL incluyendo la ADN ligasa T4, una ratio molar gen : plásmido 5 : 1, y el buffer correspondiente. En esta reacción se empleó 1 unidad de ligasa por 30 ng de vector linealizado, y se incubó durante 1h

a 22 ° C, seguido de la inactivación de la ligasa a 65 ° C durante 10 min. Después de la inactivación, se transformaron por electroporación células de *E. coli* DH5 α competentes (producidas en el laboratorio) con 10 μ L de la mezcla de ligación (Figura M. 2).

- Clonación independiente de ligación: Este mecanismo se basa en la capacidad dual polimerasa 5' \rightarrow 3' y exonucleasa 3' \rightarrow 5' de la ADN polimerasa T4 y en la complementariedad entre hebras del ADN. Una de sus ventajas es que no deja los sitios de restricción en el producto final, es decir, no hay cicatrices de la clonación. En nuestro caso se utilizaron dos plásmidos distintos, pETNKI 1.1 y pLIC-SGC1. El primero requiere ser sometido a digestión por la enzima de restricción *KpnI* y el segundo por la *Bse RI*. Para ello, siguiendo el protocolo de la casa comercial, se prepararon dos reacciones de un volumen total de 150 μ L que incluían 20 unidades de la enzima correspondiente por cada microgramo de ADN, 3 μ g del ADN a digerir, el tampón óptimo para las enzimas y agua milli-Q hasta el volumen final. Ambas reacciones se incubaron 1 h a 37 ° C y luego se purificaron por gel de agarosa las bandas correspondientes a los vectores digeridos. En paralelo, se añadieron mediante PCR los correspondientes *overhangs* a los insertos de ADN, que se obtuvieron previamente por la misma técnica a partir de ADN genómico de *M. tuberculosis*. Una vez purificados los plásmidos por el kit nombrado previamente, éstos y los insertos con los *overhangs* fueron sometidos a la ADN polimerasa T4. Para ello, en reacciones separadas para cada inserto y para cada plásmido, se mezcló 1 U de polimerasa por cada microgramo de ADN, el ADN a tratar, 25 mM de dTTP/dGTP o dATP/dCTP dependiendo de si la reacción era del plásmido (pETNKI 1.1/pLIC-SGC1) o del gen, 50 mM de DTT, 1 μ g de BSA (suero albumina bovina), y el tampón óptimo para la enzima. Las mezclas se incubaron durante 25 min a 22 ° C , tras lo cual se inactivó térmicamente la polimerasa a 75 ° C durante 20 minutos.

Tras la digestión del vector con la enzima *Kpn I* o *Bse RI*, los extremos producidos por la enzima de restricción son cohesivos, quedando dos extremos 3' \rightarrow 5' de cadena simple que la ADN polimerasa T4 atacará e eliminará los nucleótidos uno a uno por su actividad exonucleasa, así como intentará tras cada eliminación colocar de nuevo el nucleótido complementario a la otra hebra del ADN. No obstante, en el medio de reacción solo se añade dTTP o dGTP (según el plásmido), con lo que eliminará todos los nucleótidos desde el extremo 3' hasta que encuentre en la cadena complementaria una adenina o citosina, para la cual sí podrá posicionar una timina o guanina. De esta manera, quedarán de nuevo extremos cohesivos, pero la parte de cadena monocatenaria será de mayor longitud (10 a 15 pb). Por su parte, cuando sometemos a la misma

reacción a los fragmentos de ADN pero en presencia de dATP o dCTP (según el plásmido receptor), aunque los extremos son romos, la polimerasa actuará del mismo modo sobre los *overhangs* en dirección 3' → 5'. La clave de este sistema es que las secuencias nucleotídicas (*overhangs*) que se añaden a ambos lados del inserto de ADN que se desea clonar, poseen la secuencia complementaria a la que se encuentra a ambos lados del sitio de corte de la enzima de restricción con la que se digiere el plásmido (en 5' o en 3' según corresponda). De este modo, tras el tratamiento con la polimerasa T4 se generan extremos cohesivos de 10 a 15 nucleótidos de longitud complementarios entre plásmido e inserto. Tras el tratamiento con la polimerasa T4, se prepararon las mezclas correspondientes de vector e inserto con una ratio molar 1 : 3 vector : inserto y se incubaron a temperatura ambiente 5 minutos. En este tiempo, inserto y vector se unen por complementariedad sin necesidad de enzima ligasa, y 5 µL de la mezcla se transforman por electroporación en células de *E. coli* DH5α competentes. Una vez internalizado el ADN, la bacteria sella la unión de ambos fragmentos (Figura M. 2).

- Clonación por el método Gibson (Gibson et al. 2009): este sistema de clonación se basa de nuevo en la complementariedad entre hebras del ADN y en la actividad exonucleasa 5' → 3' de la exonucleasa del bacteriófago T5, y permite la clonación de varios fragmentos de ADN en el mismo vector al mismo tiempo. No obstante, en esta Tesis solo se ha utilizado para clonar un único gen de interés. Este mecanismo tiene la ventaja de que no depende de sitios de restricción, con lo que puede utilizarse en cualquier vector. Conociendo la posición en el vector donde se quiera insertar el gen, se diseñan dos pares de cebadores, siendo uno de ellos para linealizar el plásmido desde la posición de inserción, y el otro/s para colocar ambas secuencias de ADN del plásmido a un lado y otro de la inserción en 5' o 3' del inserto (*overhangs*), según se quiera situar en una orientación u otra. Cuando se dispuso del plásmido linealizado (purificado por el kit nombrado anteriormente) y del inserto con los *overhangs* (ya sea mediante PCR o por síntesis), se preparó una mezcla de 20 µL de volumen total compuesta por inserto : vector de ratio 2 : 1 junto con la mezcla de exonucleasa T5, ADN polimerasa Phusion y la ADN ligasa Taq. Esta mezcla se incubó durante 15 minutos a 50 ° C y se utilizaron 2 µL de esta mezcla para transformar células de *E. coli* quimiocompetentes. En los 15 minutos de incubación, la exonucleasa T5 empieza a eliminar nucleótidos de ambos fragmentos de ADN (plásmido e inserto) desde los extremos 5' y crea salientes monocatenarios complementarios entre ellos. Estas secuencias monocatenarias de vector y fragmento se hibridan y entonces empieza la actividad polimerasa de la Phusion, que es mucho más rápida que la exonucleasa, y que rellena

rápidamente los huecos. Finalmente, la ADN ligasa Taq sella las cadenas y acaba de fusionar totalmente la construcción vector e inserto. Este mecanismo presenta una elevada eficacia de clonación, incluso en insertos de gran tamaño, que podría deberse a que las tres actividades sobre el ADN se realizan en el mismo volumen de reacción y de manera secuencial sin competencia entre ellas, y al mayor tamaño de las regiones que hibridan. Debido a que este mecanismo ha sido optimizado en un kit por la compañía *New England Biolabs* (NEBuilder® HiFi DNA Assembly Cloning Kit), la composición exacta de la mezcla de las tres enzimas no es especificada y no puede detallarse en esta Tesis (Figura M. 2).

A lo largo de esta Tesis se clonaron 6 genes con dos orígenes distintos y con distintas construcciones y plásmidos, todos ellos para ser expresados en *E. coli*. En Tabla M. 1 se recoge toda la información asociada a este proceso de clonación para cada una de estas construcciones: plásmido, tamaño del inserto, sistema de clonación, cebadores para la clonación, *overhangs*, y cebadores para la secuenciación. No obstante, más adelante se explicarán con detalle estas construcciones.

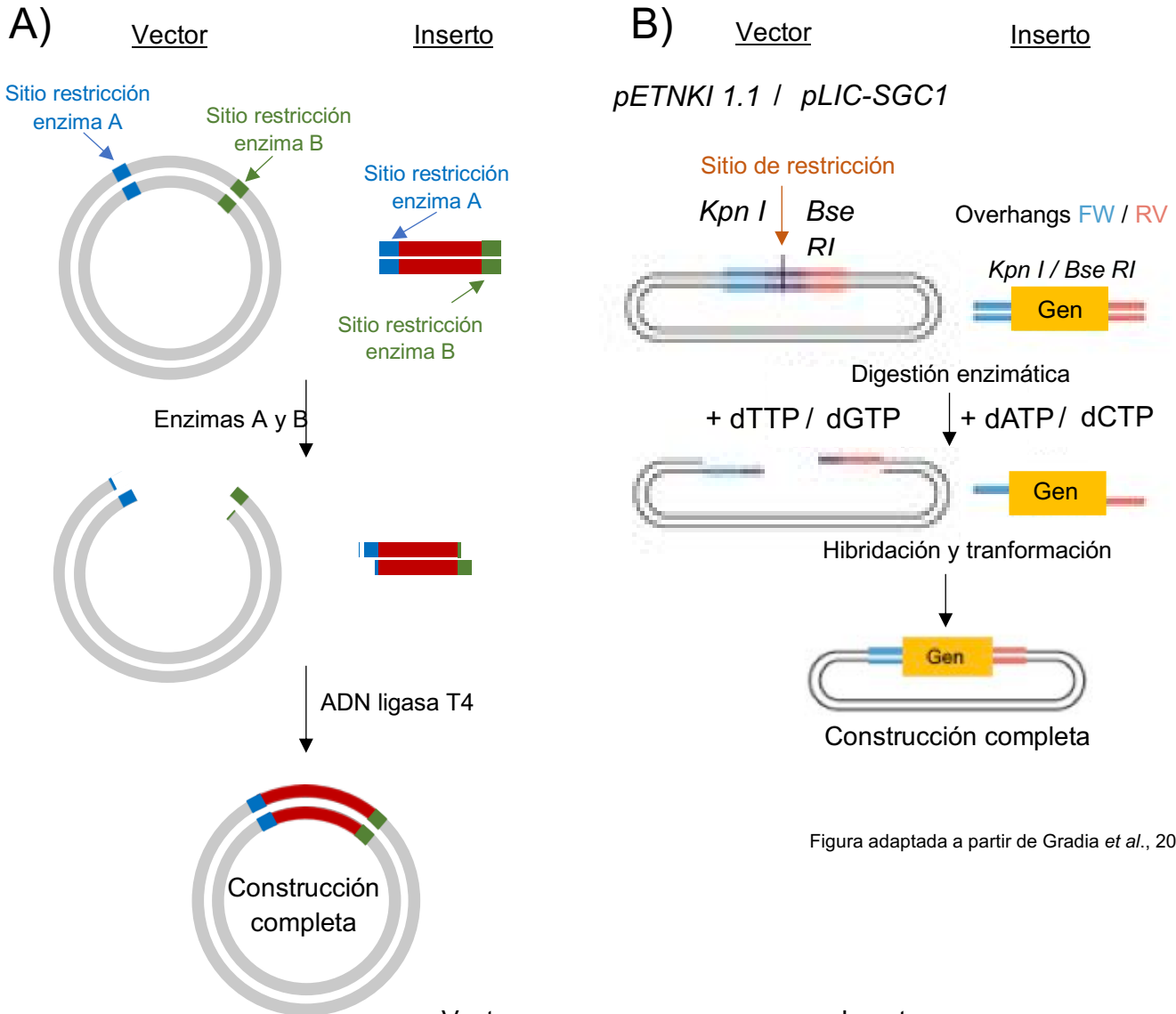


Figura adaptada a partir de Gradia *et al.*, 2017

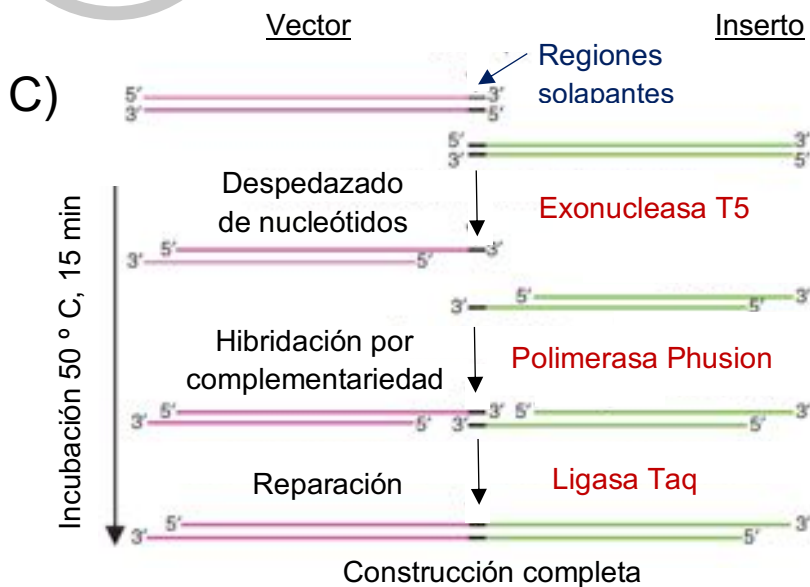


Figura adaptada a partir de Gibson *et al.*, 2009

Figura M. 2. Diagrama de los tres sistemas de clonación utilizados en esta Tesis: por enzimas de restricción (A), por el sistema LIC (B) y por el método Gibson (C).

Tabla M. 1. Tabla detallada de las clonaciones exitosas preparadas en esta Tesis. Las construcciones se explican más adelante en el texto. (Continúa)

Gen (NCBI GeneID)	Proteína (Acrónimo) - UniProt	Tamaño gen (pb)	Vector	Sistema de clonación	Gen de resistencia	Células de expresión
1854	dUTPasa humana (hDut) - P33316	487	pET28a	Gibson	Kanamicina	BL21 (DE3)
887290	dUTPasa de <i>M. tuberculosis</i> (mDut) - P9WNS5	465	pETNKI 1.1	LIC	Kanamicina	BL21 (DE3)
			pETNKI 1.5		Ampicilina	
887574	ADN-helicasa dependiente de ATP de <i>M. tuberculosis</i> (Rv3202c) - O53348	885	pETNKI 1.1 y 1.5	LIC (construcción 1)	Kan y Amp	BL21 (DE3)
		2343	pET28a	Sintético (construcción 2)	Kanamicina	
		3168	pET28a	Sintético (construcción 3)		
888250	Proteína de membrana de <i>M. tuberculosis</i> (Rv3524) - I6X7J6	798	pETNKI 1.1	LIC (construcción 1)	Kanamicina	BL21 (DE3)
		1032	pLIC-SGC1	LIC (construcción 2)	Ampicilina	
887643	Proteína sin función conocida (DUF) de <i>M. tuberculosis</i> (Rv3354) O50383	294	pETNKI 1.1	LIC (construcción 1)	Kanamicina	Rosseta 2 (DE3)
		390	pLIC-SGC1	LIC (construcción 2)	Ampicilina	
				LIC (construcción 3)		
886658	Regulador transcripcional de la familia PucR de <i>M. tuberculosis</i> (Rv1429) - O06829	1269	pETNKI 1.5	Enzimas de restricción <i>Kpn I</i> y <i>Xho I</i>	Ampicilina	BL21 (DE3)

Tabla M. 1 (Continuación). Tabla detallada de las clonaciones preparadas en esta Tesis. Las construcciones se explican más adelante en el texto. En negrita se identifican los **overhangs** de los cebadores necesarios en el inserto para su clonación.

Proteína (Acrónimo)	Construcción	Cebadores clonación 5' → 3'	Cebadores secuenciación 5' → 3'
hDut	Gen completo	CTAGCATAACCCCTTGGG (<i>foward</i> linealización vector) GTGATGATGATGATGATGG (<i>reverse</i> linealización vector) GCCATCATCATCATCAC (overhang en inserto en 5') CTAGCATAACCCCTTGGGGC (overhang en inserto en 3')	TAATACGACTCACTATAGGGGAATTG (Cebador T7, <i>forward</i> (FW)) CTAGTTATTGCTCAGCGGTGGCAG (Cebador T7, <i>reverse</i> (RV))
mDut	Gen completo	CAGGGACCCGGT GTGTCGACCACTCTGGCGATC (FW) CGAGGAGAAGCCCGTT ATCACAACTCGCATGTCCGCC (RV)	Cebadores T7 <i>forward</i> y <i>reverse</i>
Rv3202c	1 dominio entero (E295)	CAGGGACCCGGT ATGTCACATATCTGGGGTGTGCG (FW) CGAGGAGAAGCCCGTT ACTCGCCGCCGCGGAACCC (RV)	Cebadores T7 <i>forward</i> y <i>reverse</i>
	2 dominios enteros (D781)	Sintético	
	Gen entero	Sintético	Cebadores T7 <i>forward</i> y <i>reverse</i>
Rv3524	LIC (G79 a final)	CAGGGACCCGGT GGGTCGTCCACGTCCAGGC (FW) CGAGGAGAAGCCCGTT ATTAAGCCGTCGGCAGTTTTAAC (RV)	Cebadores T7 <i>forward</i> y <i>reverse</i>
	LIC (gen completo)	TACTTCCAATCC ATGGTGAAGTTCACACCCGATA (FW) TATCCACCTTTACTG TTAAGCCGTCGGCAGTTTTAAC (RV)	
Rv3354	LIC (G47 a final)	CAGGGACCCGGT GGCGATCCGGTCGACGCC (FW) CGAGGAGAAGCCCGTT ACGAGGGGCAGTACATCGAAATC (RV)	Cebadores T7 <i>forward</i> y <i>reverse</i>
	LIC (N33 a final)	TACTTCCAATCC ATGAACCCCGTCGACGACGCGTT (FW) TATCCACCTTTACTG CTAGGACCCGGGCAGCCC (RV)	
	LIC (gen completo)	TACTTCCAATCC ATGAACCTACGGCGCCATCAG (FW) TATCCACCTTTACTG CTAGGACCCGGGCAGCCC (RV)	
Rv1429	Gen completo	CGGGGTACC TTGTGGCTGAAGCTGGTGGC (FW-Kpn I) CCGCTCGAG CTATTGTTTGGCGCGGAGC (RV-Xho I)	Cebadores T7 <i>forward</i> y <i>reverse</i>

1.2. Transformación de células competentes y selección positiva de clones.

Todos los procesos de clonación descritos anteriormente concluyen en un mismo paso, la transformación del constructo en células competentes. En nuestro caso siempre se trata de células DH5 α de *E. coli*, pero dependiendo del proceso de clonación se han transformado células electrocompetentes o quimiocompetentes. No obstante, hay que resaltar que el proceso de clonación no determina el tipo de transformación celular, sino que es la forma en la que las células son preparadas para ser competentes la que lo determina.

En la electroporación, de 1 a 5 ng de ADN se añaden a 50 μ L de células electrocompetentes conservadas en agua y 10% de glicerol a -80 ° C. Tras descongelarlas en hielo para no perder competencia de transformación, la mezcla se transfiere a una cubeta de electroporación fría, y se introduce en el electroporador (*Electroporator 2510* de eppendorf), se especifica la descarga que se administrará a las células (en nuestro caso, 1800 W), y se inicia el proceso. Tras la finalización de la descarga, se añaden a las células 950 μ L de LB libre de antibiótico y estéril previamente enfriado para que se recuperen del choque eléctrico y se dejan crecer durante 1 h a 37 ° C en agitación (200 rpm).

En el choque térmico, de 1 a 2 ng de ADN se añaden a 25 μ L de células quimiocompetentes (comerciales, en nuestro caso New England Biolabs) conservadas a -80 ° C. Tras la descongelación en hielo durante 30 min, se las somete a un cambio brusco de temperatura (42 ° C) durante 30 segundos y de nuevo en hielo durante 2 minutos más. Después, de igual modo que en la electroporación, se añaden a las células 975 μ L de LB libre de antibiótico y estéril previamente enfriado para que se recuperen del choque térmico y se dejan crecer durante 1 h a 37 ° C en agitación (200 rpm).

En ambos casos, la preparación previa de las células permite que el choque eléctrico o térmico cause la formación de poros en la membrana de las bacterias por los que incorporarán el vector de interés.

Tras la hora de agitación, se plaquean 100 μ L del cultivo de células en una placa de LB - agar suplementada con el antibiótico correspondiente. El resto de cultivo (900 μ L) se centrifuga durante 2 minutos a 8000 g y el pellet celular se resuspende en 200 μ L de LB y se plaquea en otra placa de LB - agar suplementada con el antibiótico correspondiente. Las placas se dejan toda la noche a 37 ° C para el crecimiento de las bacterias.

El antibiótico nos permitirá hacer una selección positiva inicial de las células que contengan la construcción de interés, aunque posteriormente se comprueba por PCR y su exactitud en la secuencia se analiza por secuenciación. Cada vector contiene un gen de resistencia que le aporta a la bacteria la capacidad de

sobrevivir en presencia de dicho antibiótico. La resistencia de los plásmidos utilizados en esta Tesis puede consultarse en la Tabla M.1.

Habría que resaltar el caso de la clonación en el vector pLIC-SGC1 por el mecanismo LIC. Cuando las construcciones hechas con este mecanismo son introducidas en las bacterias, éstas son plaqueadas tras la transformación en placas de LB suplementadas con el antibiótico Ampicilina más un 5% de sacarosa. La digestión del plásmido con *Kpn I* para introducir el inserto elimina el gen *SacB*, un enzima que metaboliza la sacarosa en levan, un carbohidrato que se acumula en el periplasma y es tóxico para *E. coli*. Por tanto, aquellas bacterias que hayan sido transformadas por un remanente de vector completo vacío no crecerán, funcionando como sistema de selección.

1.3. Comprobación de células positivas.

Tras dejar toda la noche a 37 ° C las células bacterianas de *E.coli* plaqueadas en las correspondientes placas de LB suplementadas con antibiótico (y cuando se precisa, en sacarosa), se comprueban las colonias que han crecido. En principio son colonias que poseen el inserto dentro del vector, pero puede tratarse de células que contienen el vector sin el inserto o resistentes naturales al antibiótico. En el caso de las colonias que aparecen en placas suplementadas con antibiótico y sacarosa, no podrían ser colonias que contuvieran el vector vacío.

Para confirmar la presencia del inserto se seleccionan colonias individuales y se las somete a una PCR, la cual puede plantearse de dos maneras. Una posibilidad sería usar cebadores que se unan al ADN antes y después del MCS del plásmido. De ese modo, tras realizar la PCR y correr un gel de agarosa las colonias con insertos presentarían fragmentos de mayor tamaño que aquellas que porten vectores sin inserto. La otra posibilidad sería usar los cebadores que se usaron para amplificar los insertos en el proceso de clonación. En este caso, en el gel de agarosa veríamos una banda del tamaño correspondiente al inserto en las colonias positivas, y en las negativas veríamos ausencia de banda de ADN.

No obstante, cuando se usa la primera opción se adelanta un paso, porque de las colonias que resultan ser positivas, se debe comprobar por secuenciación si el inserto no presenta mutaciones y si se ha integrado correctamente. Y para ello el producto de reacción de PCR puede mandarse directamente a un servicio de secuenciación tras un rápido tratamiento con el kit ExoProStar (illustra®), que contiene la exonucleasa I de *E. coli* y una fosfatasa alcalina. La actividad exonucleasa digiere el ADN monocatenario (cebadores no incorporados) en dirección 3' → 5' liberando desoxirribonucleótidos 5' fosfatos, los cuales son

desfosforilados junto a los nucleótidos no incorporados por la actividad de la fosfatasa alcalina. Los cebadores específicos de cada construcción utilizados en su secuenciación pueden consultarse en la tabla M.1.

1.4. Preparación de geles de agarosa.

Los geles de agarosa fueron utilizados principalmente para ver el resultado de las PCR por electroforesis. La agarosa es un polisacárido que al polimerizar forma canales con diámetros de tamaño variable según el porcentaje al que se prepare el gel, siendo una relación inversamente proporcional. Los poros suelen oscilar entre los 50 y los 200 nm.

Los geles de agarosa utilizados en esta Tesis fueron preparados al 1 % de agarosa mayoritariamente, aunque ocasionalmente se prepararon al 0,8 % para moléculas de ADN de mayor tamaño. Para ello se pesó la cantidad correspondiente de agarosa y se añadió el volumen final de tampón TBE (Tris, borato y EDTA), variando éste en función del tamaño del gel. Para poder disolver la agarosa en el tampón, la mezcla se calentó hasta la completa desaparición de la agarosa sólida. Una vez disuelta, se dejó enfriar la mezcla para añadir la solución de tinte *Red Safe*® (ThermoFisher) or *Green Safe*® (NZYTech) y después se distribuyó en el molde para dejarlo polimerizar completamente antes de su uso. Ambas soluciones de tinte se añaden en una proporción volumen : volumen 1 : 25, y sustituyen al bromuro de etidio para la visualización del ADN por luz ultravioleta.

Una vez preparado el gel, las muestras se cargaron en volúmenes de 5 μ L del producto de PCR y 1 μ L del tampón de carga 6X (Anexo II) y se añadió un marcador de pesos moleculares acorde a las muestras. Las electroforesis se realizaron a temperatura ambiente y a 100 V hasta que el azul de bromofenol alcanzó el límite inferior del gel.

1.5. Sistemas de expresión bacteriana

Hasta el momento nos hemos referido a células de almacenamiento y replicación de vectores, pero la misma técnica es utilizada para transferir construcciones completas y comprobadas por secuenciación en células de expresión, que permitan la inducción de la expresión génica y posterior traducción proteica.

Para poder transformar dichas células se extrajo previamente el plásmido a partir de una colonia ya secuenciada que se creció durante toda la noche a 37 ° C en un cultivo de 5 mL de LB suplementado con el antibiótico correspondiente. Este cultivo se centrifugó a 11.000 g durante 2 minutos y se extrajo el plásmido con la construcción de interés a partir del *pellet* celular siguiendo el protocolo del kit *GeneJET Plasmid MiniPrep Kit*® (Thermo Scientific).

Tras el aislamiento y purificación, se cuantificó la concentración de la extracción mediante NanoDrop y se transformaron las células con de 1 a 5 ng de ADN.

Como sistemas de expresión bacteriano en este trabajo se han utilizado las cepas de *E. coli* BL21 (DE3), Rosseta 2 (DE3), y ArticCells (DE3). Las primeras son usadas en una gran mayoría de casos, mientras que las células Rosseta 2 (DE3) se utilizan cuando las proteínas a expresar están codificadas con codones raramente usados por *E. coli*. Esta cepa contiene un plásmido con resistencia a Cloranfenicol que les suple con los tRNAs deficitarios en *E. coli*. Las células Artic Cells (DE3) son una alternativa cuando las proteínas expresadas de forma heteróloga presentan problemas de insolubilidad por mal plegamiento, ya que esta cepa coexpresa chaperonas, de otro organismo, adaptadas al frío con una identidad de secuencia aminoacídica de entre el 54 y 74% a las GroEL y GroES de *E. coli*, mejorando el procesado de proteínas a bajas temperaturas (4 a 12 ° C) y con ello, aumentando el rendimiento de proteína soluble. El vector que porta estas chaperonas también incluye el gen de resistencia a la Gentamicina, que es utilizado para su selección.

Los tres tipos de células usadas en esta Tesis presentan el mismo sistema de expresión génica (DE3) es decir, la expresión proteica se induce con IPTG, la cual ha sido abordada anteriormente.

Tanto las células de almacenamiento como las de expresión proteica fueron conservadas en forma de glicerinado y a - 80 ° C. Los glicerinados se prepararon en tubos criogénicos y con un volumen de 800 µL, siendo la mitad de éste cultivo crecido en LB suplementado con los correspondientes antibióticos de selección, y la otra mitad glicerol al 50%.

1.6. Mutagénesis dirigida.

La mutagénesis dirigida permite sustituciones, inserciones o deleciones de nucleótidos en posiciones específicas en un fragmento de ADN o plásmido completo. En este trabajo se presentan un total de 7 mutantes, para los cuales se diseñaron cebadores específicos (Tabla M. 2) que se utilizaron en el protocolo de mutagénesis del kit Q5® *Site-Directed Mutagenesis kit* de NEB. Brevemente, la mutagénesis se realiza mediante PCR en un solo paso a partir del vector que contiene la construcción con el gen salvaje a mutar, y el producto de PCR es tratado con la mezcla KLD, formado por quinasa, ligasa y *DpnI*, para fosforilar los extremos 5' de los vectores linealizados creados, ligarlos para recircularizar el producto y eliminar el ADN utilizado como molde, respectivamente. Parte de esta mezcla es utilizada para transformar por choque térmico a células DH5α que son seleccionadas en placas de LB suplementadas con el antibiótico correspondiente. Tras crecer toda la noche a 37 ° C, se seleccionan varias colonias de las cuales se purifican los correspondientes vectores mediante el kit *GeneJET Plasmid MiniPrep Kit*® (Thermo Scientific) y los insertos de interés son

secuenciados confirmando la presencia de la mutación deseada. Una construcción comprobada de cada mutagénesis se utilizó para la transformación de células de expresión.

Tabla M. 2. Tabla recopilatoria de los cebadores utilizados en la mutagénesis dirigida. En **negrita** se resaltan los codones mutados.

Proteína	Vector	Cebadores	T _m (T ^a de hibridación)
EnMazG R42A	pET28a	(FW) TGCATTAGCAG CAAACA CAAAAATGATTTAAGGG	64 ° C
		(RV) GCGCCAACCTTCTCCGACT	
EnMazG I30Y		(FW) GCTAAAAGTGT ATG GAGGAAGTCG	56 ° C
		(RV) ATTTGCTTTTCAGGCTCTG	
EnMazG I30A		(FW) GCTAAAAGT GGCTG GAGGAAGTCGG	58 ° C
		(RV) ATTTGCTTTTCAGGCTCTG	
SeMazG R48A		(FW) AGCATTATCAG GCT GGTAATTTAGAGGCTTTAAAGACGGTATAG	64 ° C
		(RV) GCACCAACTTCTCCTGCC	
SeMazG Y36I		(FW) GTTGAAGTTT ATC GAAGAGGCAG	57 ° C
		(RV) GCTTGTCTATCCGGATTG	
SeMazG Y36A		(FW) GTTGAAGTTT GCC GAAGAGGCAG	57 ° C
		(RV) GCTTGTCTATCCGGATTG	
hDut H135F	pET28a	(FW) CGACGGGGCTTTGGTCCCCGGCGG	70 ° C
		(RV) CCGCGGGAGCGAGGCC	

2. EXPRESIÓN Y PURIFICACIÓN DE PROTEÍNAS.

2.1. Pruebas de expresión proteica.

Una vez obtenidas las cepas de expresión, se analizaron 4 condiciones distintas con el objetivo de seleccionar aquella óptima para obtener la mayor cantidad de proteína. Para ello se prepararon precultivos de 5 mL suplementado con el antibiótico correspondiente (Tabla M. 3) y se dejaron crecer a 37 ° C y en agitación durante toda la noche. Al día siguiente se prepararon cultivos de 25 mL de LB suplementado con el antibiótico y con 625 µL del precultivo, y se dejaron crecer a 37 ° C y en agitación hasta una densidad óptica a 600 nm de 0,5 O.D que corresponde a un estado en fase exponencial de crecimiento del cultivo, y entonces se inducía la expresión proteica.

Tabla M. 3. Concentraciones de antibiótico usadas en los cultivos bacterianos.

Antibiótico	Concentración en el cultivo
Kanamicina	33 µg / mL
Ampicilina	100 µg / mL
Cloranfenicol	33 µg / mL
Gentamicina	20 µg / mL

Las diferencias entre las 4 condiciones probadas implican a la concentración de inductor (IPTG), la temperatura y el tiempo de expresión tras la inducción:

- Condición 1: 0,1 mM IPTG, 37 ° C durante 3 horas.
- Condición 2: 1 mM IPTG, 37 ° C durante 3 horas.
- Condición 3: 0,1 mM IPTG, 20 ° C durante 16 horas.
- Condición 4: 1 mM IPTG, 20 ° C durante 16 horas.

Se tomaron muestras del cultivo antes y después de la inducción y se sometieron a electroforesis en gel de poliacrilamida en condiciones desnaturalizantes.

No obstante, estas cuatro condiciones no siempre fueron directamente válidas para la expresión de todas las proteínas, sino que en el caso de algunas proteínas fue necesario recurrir a la transformación de células Rosseta 2 por presencia de codones raros. Tras la transformación, se probaron con las nuevas células estas mismas 4 condiciones.

En alguna proteína cuya expresión había sido ínfima y el porcentaje de codones raros no era un limitante, se decidió probar la expresión en otros 4 medios de cultivo comerciales (Molecular Dimensions): *Turbo*, *Hyper*, *Superior* y *Terrific*.

Todos ellos incrementan el rendimiento de la expresión, obteniendo una mayor cantidad de proteína expresada que en el medio LB.

Una vez realizadas las pruebas de expresión, se llevaron a cabo las pruebas de solubilidad.

2.2. Pruebas de solubilidad proteica.

El *pellet* celular obtenido en las pruebas de expresión fue utilizado para las pruebas de solubilidad. Para ello se tuvo en cuenta principalmente el punto isoeléctrico de la proteína (pI) y se eligieron 4 tampones de prueba para cada proteína. El pH del tampón se elige siempre siendo un punto o un punto y medio mayor o menor al pI de la proteína a producir. A parte del agente tamponador, se elige la concentración de sal y si hay algún cofactor que podría necesitar la proteína. En nuestro caso, algunas de las proteínas han necesitado $MgCl_2$ que se ha añadido en una concentración final del 5 o 10 mM. Por tanto, para cada proteína, generalmente se probaron soluciones tamponantes con 100 mM de agente tamponador, dos concentraciones de NaCl y en ausencia o presencia de $MgCl_2$.

El pH elegido determinó el uso de un agente tamponador u otro, haciendo uso del Tris cuando se requería un pH de 8-8.5 y Hepes para pH 7-7,5. A modo de resumen las soluciones tamponadas probadas sin éxito pueden consultarse en la Tabla M. 4, aunque en el apartado de resultados se especifican para cada proteína por separado.

Para llevar a cabo las pruebas de solubilidad, se distribuyó en 4 partes iguales el *pellet* celular resultante de la prueba de expresión y se resuspendieron cada una de ellas en un tampón distinto. Tras la resuspensión se lisaron las células y el lisado se centrifugó a 10.000 g durante 30 min. Se tomaron muestras de sobrenadante y *pellet* (fracción soluble e insoluble), y se sometieron a electroforesis en gel de poliacrilamida en condiciones desnaturizantes.

El lisado celular se realizó por dos métodos distintos: por ultrasonidos o por homogenización con microesferas de vidrio. Para el primero se utilizó un sonicador Ultrasonic processor (Sonics) de 130 vatios, con un programa de sonicación de 3 ciclos de 2 minutos compuestos por 10 segundos de sonicación (amplitud 40 %) seguidos de 30 segundos de descanso. La rotura con las microesferas de vidrio se realizó con un homogenizador Precellys® (Bertin instruments), usando los tubos adecuados y añadiendo un volumen de esferas correspondiente a un tercio del volumen del cultivo a lisar. Las muestras se sometieron a 3 ciclos de 15 segundos de agitación de las microesferas seguidos de 30 segundos de descanso. En los descansos, las muestras se refrigeraban en hielo.

En aquellos casos que los resultados de la electroforesis con las muestras de fracción soluble y no soluble no eran del todo concluyentes, se hizo uso de bolitas magnéticas y resinas para detectar si la proteína de interés estaba en la fracción soluble, aunque lo estuviera en pequeñas cantidades. Las bolitas magnéticas son microesferas de metal que en su superficie tiene iones de Ni divalentes unidos irreversiblemente. Simplemente, la muestra a probar se mezcla con un pequeño volumen de estas bolitas magnéticas y luego se recupera lo unido por afinidad a ellas. Las resinas son las mismas usadas a gran escala (His-Pur™ Ni-NTA Resin, ThermoScientific), pero para las pruebas de solubilidad se preparan columnas de muy pequeño volumen (Columnas de microcentrífuga Pierce™, ThermoScientific) y se pasa la muestra por ellas. Al igual que en el mecanismo anterior, tras el paso de la muestra, se recupera lo unido a la resina. En ambos casos, las proteínas poseían una etiqueta de histidinas.

Tras las pruebas de expresión y solubilidad se compararon los resultados y se seleccionaron aquellas condiciones que maximizaban la extracción de proteína soluble. Las condiciones seleccionadas para la extracción de las proteínas utilizadas en esta Tesis se resumen en la Tabla M. 5, aunque se especifican posteriormente para cada proteína en el correspondiente apartado de resultados.

Tabla M. 4. Tabla recopilatoria de todos los tampones probados sin éxito para solubilizar las proteínas de esta Tesis. (Continúa)

Proteína	Tamaño (KDa)	Punto isoeléctrico (pI)	Tampones probados en prueba de solubilidad sin éxito
Rv3202c (1 dominio)	30,82	8,76	PBS 1X pH 7,4
			PBS 1X pH 7,4 + 5 mM MgCl ₂
			PBS 1X pH 7,4 + 5 mM MgCl ₂ + Urea (1 – 4M)
Rv3202c (2 dominios)	81,14	9,40	200 mM Tris-HCl pH 8 + 200 mM NaCl
			10 mM Tris-HCl pH 8 + 1 mM EDTA + 200 mM NaCl + 50% glicerol + 2mM DTT + 5mM MgCl ₂
			10 mM Tris-HCl pH 8 + 200 mM NaCl + 20% glicerol + 1 mM β-mercaptoetanol + 5 mM MgCl ₂
			50 mM Tris-HCl pH 8 + 1 M KCl + 1 mM EDTA + 2 mM DTT
			200 mM Tris-HCl pH 8 + 200 mM NaCl + 5 mM MgCl ₂
Rv3202c (entero)	111,50	8,72	200 mM Hepes pH 7 + 400 mM NaCl + 5 mM MgCl ₂
			200 mM Hepes pH 7 + 400 mM NaCl + 5 mM MgCl ₂ + 1 mM β-mercaptoetanol + 20 % glicerol
			200 mM Hepes pH 7 + 400 mM NaCl + 5 mM MgCl ₂ + 20 % glicerol
Rv3354 (G47 a final)	8,27	3,90	200 mM Tris-HCl pH 8 + 200 mM NaCl
			200 mM Hepes pH 7 + 400 mM NaCl
			200 mM Tris-HCl pH 8 + 400 mM NaCl
			200 mM Tris-HCl pH 8,5 + 400 mM NaCl
			200 mM Tris-HCl pH 8 + 200 mM NaCl + 4 M urea

Tabla M. 4 (Continuación). Tabla recopilatoria de todos los tampones probados sin éxito para solubilizar las proteínas de esta Tesis.

Proteína	Tamaño (KDa)	Punto isoeléctrico (pI)	Tampones probados en prueba de solubilidad sin éxito
Rv1429 (entero)	47,70	6,08	150 mM Tris pH 8,5 + 300 mM NaCl
			75 mM Hepes pH 7,5 + 400 mM NaCl + 5 mM MgCl ₂
			200 mM Tris pH 8 + 200 mM NaCl + 5 mM MgCl ₂
			200 mM Tris pH 8 + 500 mM NaCl + 5 mM MgCl ₂
			200 mM Hepes pH 7 + 200 mM NaCl + 5 mM MgCl ₂
			200 mM Hepes pH 7 + 500 mM NaCl + 5 mM MgCl ₂

Tabla M. 5. Tabla recopilatoria de las condiciones de expresión y solubilidad seleccionadas para cada proteína utilizada en esta Tesis.

Proteína	Tamaño (KDa)	Punto isoelectrico (pI)	Medio de expresión	Condiciones de expresión seleccionadas	Tampón seleccionado en prueba de solubilidad
hDut	13,40	6,12	LB	0,5 mM IPTG, 20 °C, 16 h	100 mM Tris-HCl pH 8 + 300 mM NaCl
mDut	16,60	6,23	LB	0,5 mM IPTG, 37 °C, 3 h	200 mM Tris-HCl pH 8 + 400 mM NaCl
Rv3202c (entero)	111,50	8,72	LB	0,1 mM IPTG, 20 °C, 16 h	200 mM Hepes pH 7 + 200 mM NaCl + 5 mM MgCl ₂ (medio LB)
					200 mM Hepes pH 7 + 200 mM NaCl + 5 mM MgCl ₂ + 1 mM EDTA + 50 % glicerol + 2 mM DTT (medio LB)
Rv3354 (N33 a final)	13,80	6,10	LB	0,1 mM IPTG, 37 °C, 3 h	200 mM Tris-HCl pH 8 + 200 mM NaCl
Rv3524 (G47 a final)	27,68	4,44	LB	0,1 mM IPTG, 20 °C, 16h	200 mM Tris-HCl pH 8 + 200 mM NaCl
Rv1429	47,70	6,08	<i>Hyper</i> ®	1 mM IPTG, 20 °C, 16 h	200 mM Tris pH 8 + 500 mM NaCl + 5mM MgCl ₂ + 2 M urea
EnMazG	14,80	5,95	LB	1 mM IPTG, 20 °C, 16 h	150 mM Tris-HCl pH 8,5 + 250 mM NaCl
SeMazG	15,60	6,28	LB	1 mM IPTG, 20 °C, 16 h	150 mM Tris-HCl pH 8,5 + 250 mM NaCl
EBVDut	33,20	9,43	LB	*0,5 mM IPTG, 37 °C, 6 h	Condiciones y tampones extraídos de la bibliografía (Tarbouriech et al. 2005)*
Stl ^{N-ter}	18,00	5,63	LB	** 0,1 mM IPTG, 20 °C, 16 h	Condiciones y tampones extraídos de la bibliografía (Rafael Ciges-Tomas et al. 2019)**
Stl ^{full-length}	31,40	6,48			

2.3. Purificación a gran escala.

El proceso de purificación a gran escala se explica siguiendo como hilo argumental el orden en el que se realizan los distintos experimentos, y no agrupados por técnicas. Se introduce brevemente cada una de las técnicas.

2.3.1. Expresión proteica y lisis bacteriana.

Una vez seleccionadas las condiciones de expresión y solubilidad para cada proteína, el siguiente paso fue obtener proteína en cantidad y pureza suficiente para los ensayos bioquímicos, biofísicos y estructurales. Para ello se procedió a la expresión a gran escala, utilizando volúmenes de cultivo de 4 L de media por cada purificación y proteína analizada, aunque en algunos casos podríamos alcanzar los 8 L o realizar múltiples purificaciones para aquellas que se requería más cantidad de proteína. Cada litro de cultivo suplementado con el antibiótico o antibióticos necesarios fue inoculado con 25 mL de precultivo de la cepa de *E. coli* portando el vector de expresión de la proteína a producir y se creció 37 ° C en agitación (180 rpm) hasta alcanzar la densidad óptica a 600 nm de 0,5 para la inducción, acorde con las condiciones seleccionadas. Una vez realizada la expresión, los cultivos fueron centrifugados a 4000 g durante 30 minutos a 4 ° C. El *pellet* celular se lavó con PBS 1X y se resuspendió en la solución tamponada de lisis correspondiente (TL, Tabla M. 6) para su posterior sonicación en presencia de 1 mM del inhibidor de proteasas PMSF (fluoruro de fenilmetilsulfonilo). La sonicación se realizó con los mismos parámetros usados en la prueba de solubilidad.

Los lisados celulares fueron centrifugados a 10.000 g durante 35 min a 4 ° C y la fracción soluble fue inicialmente purificada por cromatografía de afinidad.

Tabla M. 6. Tabla recopilatoria de los tampones usados en la purificación de las proteínas utilizadas en esta Tesis.

Proteína	Tamaño (KDa)	Punto isoelectrico (pI)	Tampones usados en la purificación
hDut	13,40	6,12	(TL) 100 mM Tris-HCl pH 8 + 300 mM NaCl
			(TE) 40 mM Tris-HCl pH 8 + 200 mM NaCl + 500 mM imidazol
			(TF) 20 mM Tris pH 8 + 125 mM NaCl + 1 mM MgCl ₂
mDut mDut H145F	16,60	6,23	(TL) 200 mM Tris-HCl pH 8 + 400 mM NaCl
			(TE) 150 mM Tris-HCl pH 8 + 400 mM NaCl + 250 mM imidazol
			(TF) 100 mM Tris-HCl pH 8 + 400 mM NaCl
Rv3524 (fragmento)	37,10	4,44	(TL) 200 mM Tris-HCl pH 8 + 200 mM NaCl
			(TE) 200 mM Tris-HCl pH 8 + 200 mM NaCl + 500 mM imidazol
			(TF) 100 mM Tris-HCl pH 8 + 150 mM NaCl
Rv1429	47,70	6,08	(TL) 200 mM Tris-HCl pH 8 + 500 mM NaCl + 2 M urea
			(TE) 200 mM Tris-HCl pH 8 + 500 mM NaCl + 2 M urea + 2,5 mM destiobiotina
EnMazG	14,80	5,95	(TL) 150 mM Tris-HCl pH 8,5 + 250 mM NaCl
			(TE) 150 mM Tris-HCl pH 8,5 + 200 mM NaCl + 500 mM imidazol
			(TF) 50 mM Tris-HCl 8,5 + 150 mM NaCl
SeMazG	15,60	6,28	(TL) 150 mM Tris-HCl pH 8,5 + 250 mM NaCl
			(TE) 150 mM Tris-HCl pH 8,5 + 200 mM NaCl + 500 mM imidazol
			(TF) 50 mM Tris-HCl 8,5 + 150 mM NaCl
EBV ^{Dut}	33,20	9,43	(TL) 50 mM Hepes 7,5 + 1 M NaCl + 10 mM MgCl ₂ + 10 mM imidazol
			(TE) 25 mM Hepes 7,5 + 500 mM NaCl + 10 mM MgCl ₂ + 500 mM imidazol
			(TF) 20 mM Hepes 7,5 + 250 mM NaCl + 10 mM MgCl ₂ + 200 mM imidazol
Sti ^{N-ter}	18,00	5,63	(TL) 75 mM Hepes pH 7,5 + 400 mM NaCl + 5 mM MgCl ₂
			(TE) 75 mM Hepes pH 7,5 + 400 mM NaCl + 5 mM MgCl ₂ + 500 mM imidazol
			(TF) 75 mM Hepes pH 7,5 + 250 mM NaCl + 5 mM MgCl ₂
Sti ^{full-length}	31,40	6,48	(TL) 75 mM Hepes pH 7,5 + 400 mM NaCl + 5 mM MgCl ₂
			(TE) 75 mM Hepes pH 7,5 + 400 mM NaCl + 5 mM MgCl ₂ + 500 mM imidazol
			(TF) 75 mM Hepes pH 7,5 + 250 mM NaCl + 5 mM MgCl ₂

TL: solución tamponante de lisis; TE: solución tamponante de elución; TF: solución tamponante final.

2.3.2. Cromatografía de afinidad.

Esta técnica permite la separación de proteínas en base a una interacción reversible entre una proteína y un ligando que está unido covalentemente a una matriz. En nuestro caso, el ligando fueron iones de níquel inmovilizados o puntos de unión a biotina en la sefarosa StrepTactin (Cytiva), según si se purificaba proteínas con una etiqueta de histidinas o cola *StrepII-tag*, respectivamente.

En el caso de las proteínas con etiqueta de histidinas, una columna His-Trap HP (GE Healthcare) fue equilibrada con la misma solución de lisis y a continuación se cargó con la fracción soluble tras la centrifugación del lisado celular. Una vez la columna estuvo cargada con el sobrenadante, ésta fue lavada con la misma solución tamponada de lisis pero con 10 mM de imidazol para eliminar los posibles contaminantes que se hubieran pegado por afinidad a la fase estacionaria. Tras el lavado, las proteínas fueron eluidas de las columnas con la solución tamponada de elución (TE, Tabla M. 6) en un sistema de FPLC ÄKTA pure (GE-Healthcare), el cual permite el seguimiento de la elución proteica por medida de la absorbancia a 280 nm. Las soluciones tamponadas para la elución contienen imidazol que compite con las histidinas por unirse al níquel, de modo que el imidazol se une al níquel y desplaza a la proteína, que sale de la columna.

En el caso de las proteínas con cola *StrepII-tag*, la solución tamponada de elución contenía 2,5 mM de D-destiobiotina, una forma modificada de la biotina que se une a los puntos de unión de ésta en la resina y compite con la *StrepII-tag*, desplazando a la proteína y provocando su salida de la columna.

Tras cada uso, ambos tipos de columnas fueron regeneradas para su uso futuro y se conservaron en 20 % de etanol a 4 ° C. De todos estos pasos de purificación se tomaron muestras: fracción soluble e insoluble tras la sonicación, fracción no retenida en la columna de afinidad, lavado, y las diferentes fracciones en las que se dividió la elución. Todas ellas se sometieron a una electroforesis SDS-PAGE, y se seleccionaron aquellas fracciones de la cromatografía de afinidad de mayor pureza.

2.3.3. Electroforesis en gel de poliacrilamida en condiciones desnaturalizantes (SDS-PAGE).

La electroforesis es una técnica ampliamente usada en la cual se separan las moléculas en función de su movilidad en un campo eléctrico.

En el caso de la electroforesis en gel de poliacrilamida en condiciones desnaturalizantes (SDS-PAGE de sus siglas en inglés, *sodium dodecyl sulfate polyacrylamide gel electrophoresis*), las muestras proteicas llevan una solución tamponada con SDS, β -mercaptoetanol y glicerol. El primero le aporta carga negativa a la proteína, el segundo rompe los posibles enlaces disulfuro que

podrían formarse en la proteína y el tercero le aporta densidad a la muestra para que quede en el fondo del pocillo hasta que se aplique el campo eléctrico y no difunda por la cubeta. Además, se incuban durante 5 minutos a 95 ° C antes de cargarse en el gel para desnaturalizar completamente las proteínas. Con ello se consigue que las proteínas avancen por el gel en función de su peso molecular y no en función de su forma ni de su carga, ya que la acrilamida forma una red de poros por los cuales las proteínas más pequeñas podrán avanzar más rápidamente y las más grandes se ralentizarán. El gel está formado por dos partes, la parte resolutive y la parte de apilamiento. Esta última es en la que se encuentran los pocillos y posee un bajo % de acrilamida, y su función es que todas las moléculas de todas las muestras se apilen en el final de esta parte y entren al mismo tiempo en la parte resolutive. La parte resolutive posee una mayor concentración (en %) de acrilamida, la cual se decide en función del tamaño de las proteínas que van a atravesar el gel. Cada parte del gel posee, aparte de un porcentaje de acrilamida distinto, un distinto pH (Anexo II), y SDS para mantener la carga negativa en las muestras.

Una vez cargadas todas las muestras del procesado anterior, la electroforesis se lleva a cabo a 100 V durante unos 40-50 minutos. El tiempo dependerá del tamaño de la proteína, del porcentaje de acrilamida y del tamaño del gel usado. El tampón utilizado en la cubeta de electroforesis está compuesto por 25 mM Tris-HCl pH 8,3 + 192 mM de glicina y 0,1 % de SDS.

Para visualizar las proteínas tras la electroforesis pueden utilizarse diferentes técnicas, pero en nuestro caso utilizamos la tinción con Azul Coomassie que contiene ácido clorhídrico (0,11 %) y azul brillante Coomassie (93,7 µM), actuando el primero de fijador y el segundo de teñidor. Tras la electroforesis el gel se sumerge en 40-50 mL de esta solución de teñido que se ha precalentado a 50-60 ° C y se deja en agitación durante 10 minutos, tras lo que se retira la solución de teñido y se lava múltiples veces con agua desionizada.

Los porcentajes de acrilamida usados para los geles de SDS-PAGE en las proteínas de esta Tesis se especifican en la tabla M. 7.

Visto el resultado del gel, las fracciones de mayor pureza fueron mayormente sometidas a digestión para la eliminación de la etiqueta utilizada para la cromatografía de afinidad, y parte de la proteína purificada se conservó con la correspondiente etiqueta. En este último caso, las fracciones fueron concentradas y sometidas a cromatografía de exclusión molecular.

Tabla M.7. Relación de los porcentajes de acrilamida usados en los geles para SDS-PAGE de las proteínas utilizadas en esta Tesis.

Proteína	% de acrilamida
hDut	17
mDut	17
Rv3524	15
Rv1429	15
Rv3202c	12
Rv3354	17
EnMazG	15
SeMazG	15
EBVDut	15
Stl ^{full-length}	17
Stl ^{N-terminal}	15

2.3.4. Digestión de la etiqueta unida a la proteína.

Todas las proteínas en las que se ha trabajado en esta Tesis, a excepción de EnMazG y SeMazG, fueron clonadas en vectores que permitían la eliminación de la etiqueta utilizada en la purificación por afinidad. Normalmente, la etiqueta no imposibilita la función de la proteína, pero podría dificultar su uso en algunas técnicas. Además, es posible que no se ordene espacialmente de una manera constante y dificulte la cristalización de la proteína, algo clave para los estudios estructurales.

Para ello, entre la etiqueta y el inicio de la proteína se encuentra una secuencia proteica (de unos 6 aminoácidos) que es el sitio de reconocimiento y restricción de proteasas específicas. En la tabla M. 8 se encuentra la relación de proteínas cuya etiqueta fue digerida y la proteasa que se utilizó para tal fin. Se ha hecho uso de las proteasas TEV, PreScission y SENP2. Para la reacción de digestión de la TEV se requiere, para el buen funcionamiento de la proteasa, añadir EDTA y β -mercaptoetanol en concentraciones de 0,5 mM y 5 mM, respectivamente. En el caso de la SENP2 se requiere añadir 1 mM de DTT, y en el caso de la PreScission, no se requiere añadir ningún otro componente a la reacción.

Tabla M. 8. Relación de proteasas utilizadas en el corte de la etiqueta de las proteínas utilizadas en esta Tesis.

Proteína	Proteasa usada en el corte de la cola
hDut	TEV
mDut	PreScission
Rv3524	PreScission
Rv1429	PreScission
Rv3202c	PreScission (pETNKI 1.1) / TEV (pET28a)
Rv3354	PreScission (pETNKI 1.1) / TEV (pLIC-SGC1)
EnMazG	Sin sitio de corte
SeMazG	Sin sitio de corte
EBV _{Dut}	TEV
Stl ^{full-length}	SENP2
Stl ^{N-terminal}	SENP2

Las reacciones de proteólisis se llevaron a cabo a 4 ° C durante 16 horas y en agitación para asegurar el acceso de la proteasa a todas las moléculas de proteína, haciendo uso de una ratio molar 1 : 50 proteasa : proteína. Tras las reacciones de proteólisis, que habitualmente se utilizaron al mismo tiempo para el cambio de solución tampón, las moléculas de proteasa fueron separadas de las de proteína sin etiqueta mediante cromatografía de afinidad, ya que las proteasas TEV y SENP2 poseen una etiqueta de 6-histidinas, y la proteasa PreScission una etiqueta GST (glutación S- transferasa). En el caso de las digestiones con TEV y SENP2 esta cromatografía de afinidad también sirve para separar aquella fracción de la proteína en la que la etiqueta de histidina no ha sido eliminada.

La proteína sin etiqueta y con la solución tamponante final fue sometida de nuevo a SDS-PAGE, y si la pureza era elevada se concentró, alicuotó y vitrificó a - 80 ° C para su conservación hasta ser utilizada. En caso de no tener la suficiente pureza, se concentró y se sometió a cromatografía de exclusión molecular.

Antes de finalizar esta técnica, hay que destacar que también fue utilizada tras la cromatografía de afinidad en aquellas proteínas cuya concentración de contaminantes era muy elevada. Si los contaminantes se habían unido inespecíficamente a la columna de afinidad, lo harían de nuevo al igual que la etiqueta separada de la proteína, con lo cual la proteína de interés eluiría de la columna sin contaminantes.

2.3.5. Cromatografía de exclusión molecular.

La cromatografía de exclusión molecular, conocida también como cromatografía de filtración en gel, separa las moléculas en función de su tamaño molecular. En ella, la fase estacionaria está formada por una matriz compuesta por esferas con un tamaño de poro determinado empaquetada en una columna y con las partículas esféricas rodeadas de solvente. La fase móvil está constituida por la muestra, que estará formada por moléculas de distinto tamaño, que en caso de ser más pequeñas o más grandes que el tamaño del poro, difundirán o no dentro de éstos, respectivamente. Por tanto, aquellas moléculas de menor tamaño pasarán por los poros ralentizando su movimiento en la columna, mientras que las de mayor tamaño, son “excluidas” de las esferas y se mueven a mayor velocidad por la columna, saliendo las moléculas en orden decreciente de tamaño de la columna. No obstante, la forma de las moléculas también influirá en su comportamiento, teniendo problemas para entrar por los poros las moléculas no globulares, por lo que muestran volúmenes de elución no acordes con sus pesos moleculares en este tipo de cromatografía.

Tal y como se ha avanzado anteriormente, las muestras con elevada presencia de contaminantes tras la cromatografía de afinidad se concentraron y sometieron a cromatografía de exclusión molecular, tanto para descartar proteínas no deseadas como para cambiar la solución tampón de la proteína para futuros experimentos. Tras el paso de la muestra por la columna de exclusión molecular, las muestras de las fracciones recogidas se sometieron a SDS-PAGE para seleccionar aquellas de mayor pureza. De nuevo, las de mayor pureza fueron concentradas, alicuotadas, vitrificadas y guardadas a -80°C hasta su utilización.

Finalmente, habría que puntualizar que a lo largo de esta Tesis se han usado dos tipos de columnas de exclusión molecular con distinta capacidad resolutive, las de menor resolución para separar contaminantes de las proteínas de interés y las de mayor resolución para detectar la formación de complejos entre proteínas, la formación de oligómeros de una misma proteína, o la estimación de su peso molecular.

Para la separación de contaminantes se ha usado la columna Superdex S75 (GE Healthcare) y el FPLC ÄKTA Pure para las proteínas mDut, mDut H145F, hDut y Rv3524; mientras que se ha utilizado una columna HiLoad™ 16/60 Superdex™ 75 PrepGrade (GE Healthcare) y el FPLC ÄKTA Purifier para las proteínas EnMazG y SeMazG. El resto de las proteínas no se han sometido a exclusión molecular con esta finalidad.

Tanto la cromatografía de afinidad como la de exclusión molecular fueron realizadas a 4°C , con la excepción de la Dut del virus de Epstein Barr, cuyo cultivo se centrifugó a 15°C y el resto del proceso se realizó a temperatura

ambiente por la reducción radical de solubilidad sufrida por la proteína a bajas temperaturas.

2.3.6. Diálisis

La diálisis es una técnica basada en la ósmosis, donde el solvente se mueve a través de una membrana semipermeable desde una solución de menor concentración de soluto a una de mayor concentración hasta lograr el equilibrio osmótico, es decir, la misma concentración de soluto a ambos lados de la membrana.

A lo largo de esta Tesis se ha empleado esta técnica en los últimos pasos de aquellas purificaciones que tras la cromatografía de afinidad la proteína presentaba una gran pureza y no había contaminantes presentes. Mediante la diálisis se eliminaba el imidazol o el agente competidor utilizado en la elución de la columna de afinidad. Además, se modificaba la composición del tampón en el cual se encontraba la proteína, con una menor concentración de agente tamponante y de NaCl que podían interferir en los ensayos posteriores.

Para llevar a cabo esta técnica, se enfrentó la proteína proveniente de la cromatografía de afinidad a la solución tamponada final (TL, Tabla M. 6). Fueron empleadas membranas deshidratadas con un tamaño de poro de 6 a 8 kDa (Spectrum™), las cuales fueron hidratadas con la solución tampón final. El tamaño de la bolsa de diálisis preparada con la membrana difería en función del volumen de proteína, el cual siempre se sobreestimó para evitar la rotura de la bolsa al incrementarse el volumen por la entrada del tampón durante el intercambio osmótico. El proceso de diálisis se realizó a 4 ° C durante 16 horas en agitación.

2.3.7. Concentración centrífuga de proteínas.

Como consecuencia de la cromatografía de afinidad y exclusión molecular o diálisis, la proteína se diluye, de modo que el siguiente paso es su concentración.

Para ello se recurrió a la concentración centrífuga usando concentradores cuyos filtros poseían distintos tamaños de poro (3, 10 o 30 kDa), que permitían pasar a través de ellos las macromoléculas de menor tamaño (concentradores centrífugos AMICON™). Por tanto, nuestra proteína de interés no podía atravesarlo y se concentraba a medida que la solución tamponada sí lo hacía. Se realizaron múltiples ciclos de centrifugación a distintos tiempos y velocidades, éstas últimas especificadas por el tamaño del concentrador. El tiempo fue decreciendo a medida que la proteína estaba más concentrada para evitar un exceso de concentración y consecuente precipitación.

Tras la máxima concentración, se cuantificó la concentración de proteína por el método Bradford, se alicuotó en pequeños volúmenes que se vitrificaron en nitrógeno líquido, y se guardaron las alícuotas a - 80 ° C hasta su posterior uso.

El uso de concentradores también permite el cambio de solución tamponada en la que se encuentra la proteína, mediante repeticiones de concentración y dilución en la solución tamponada final hasta haberla intercambiado totalmente.

2.3.8. Cuantificación proteica por el método Bradford.

La cuantificación de proteínas puede hacerse por diversos métodos pero en esta Tesis se ha utilizado el método de Bradford.

El método Bradford es una técnica colorimétrica que se basa en el viraje de color del reactivo Bradford proporcionalmente a la concentración de proteína, y se cuantifica por medida de la absorbancia a 595 nm de la muestra respecto a un blanco de reactivo. Dicho reactivo contiene el Azul Coomassie Brillante G-250, que se fija a las proteínas por sus residuos básicos y aromáticos. Para ello, se prepara una curva patrón con soluciones de BSA a concentraciones conocidas que sirve para hacer la extrapolación y conocer la concentración de la proteína. Las muestras se midieron añadiendo 2 µL de la muestra a medir en 1 mL del reactivo de Bradford cuya curva patrón se calculaba previamente.

2.4. Ensayo de termoestabilidad (termofluor)

Este ensayo permite conocer el estado de plegamiento de una proteína, y en esta Tesis se ha utilizado para comprobarlo en una construcción proteica diseñada. Se realiza en un sistema de PCR cuantitativa a tiempo real, en el cual se incrementa gradualmente la temperatura y se registra el aumento de la fluorescencia emitida por un fluoróforo unido a los residuos hidrofóbicos de la proteína. A medida que aumenta la temperatura se produce la desnaturalización de la proteína, de modo que los residuos hidrofóbicos quedan expuestos al solvente y la señal de fluorescencia aumenta.

En esta Tesis se ha utilizado el sistema de PCR cuantitativa 7500 Fast (Applied Biosystems), y el fluoróforo Sypro Orange (Sigma). La temperatura se incrementó desde los 20 ° C hasta los 85 ° C con un incremento de 1 ° C por ciclo. La intensidad de fluorescencia se normalizó y analizó con el *software* Graphpad Prism 9.0.1.

2.5. Western-Blot.

A lo largo de esta Tesis se ha recurrido al uso de esta técnica para la identificación de las proteínas de interés.

Esta técnica se basa en el reconocimiento específico antígeno-anticuerpo. Las muestras a analizar se separan por SDS-PAGE, junto a un marcador de pesos moleculares ya teñido, y son transferidas a una membrana de nitrocelulosa. Aunque la transferencia puede ser húmeda, semi-seca y seca, y en esta Tesis se ha llevado a cabo la húmeda y la semi-seca. En ambos tipos de transferencia el gel de acrilamida y la membrana se colocan uno encima del otro y son tapados en ambas caras por papel Whatman® y una esponja. En el caso de la transferencia húmeda todo se sumerge en la solución tamponada de transferencia (20 mM Tris-HCl + 192 mM glicina + 20 % metanol, pH 8,3) y se introduce en un campo eléctrico en el cual la corriente va desde el gel en dirección a la membrana y las proteínas se transfieren. Para esta transferencia se aplica un voltaje 100 V durante 1 hora y se mantiene a 4 ° C. En el caso de la transferencia semi-seca, el montaje de gel y membrana es el mismo, pero en lugar de realizarse sumergido en el tampón de transferencia, solo las esponjas y el papel Whatman® están humedecidos con el tampón de transferencia. Esta transferencia se hace fijando el amperaje a 2,5 A durante tan solo 7 min a temperatura ambiente.

Una vez transferidas las proteínas a la membrana, ésta se bloquea con una solución de BSA y TBS-T durante 45 minutos a temperatura ambiente para evitar uniones inespecíficas del anticuerpo a la membrana. Tras el bloqueo, la membrana se incuba con el anticuerpo específico, y después con un anticuerpo secundario que reconoce al específico y tiene conjugado una peroxidasa o una molécula fluorescente que permitirá su detección. No obstante, en nuestro caso se disponía de anticuerpo específico contra la etiqueta de histidinas que ya tiene conjugado peroxidasa de rábano picante.

Una vez el anticuerpo de la membrana se lava con tampón TBS-T, se procede al revelado de la membrana, el cual variará en función de la naturaleza del anticuerpo o anticuerpos utilizados. Particularmente, en nuestro caso se utilizó del kit de revelado Pierce® ECL Western Blotting Substrate, que contiene el luminol y un potenciador de la señal de éste. Cuando la peroxidasa entra en contacto con el luminol, éste se transforma en un producto excitado que emite quimioluminiscencia a 428 nm, con lo que aquellas proteínas que posean la etiqueta de histidinas (las proteínas de interés), emitirán señal en la membrana. Con el potenciador del luminol se incrementa la señal que emite y también la duración de la señal, ya que de manera natural es muy breve. La señal se recogió con el aparato Image Quant® y se superpuso la imagen de la señal del anticuerpo con la imagen de la luz visible, donde se pueden visualizar los

patrones moleculares que permiten determinar si la señal del anticuerpo se sitúa en el tamaño de proteína que se espera.

2.6. MALS y DLS

Tal y como se ha comentado con anterioridad, la cromatografía de exclusión molecular no se utilizó con la única finalidad de separar proteínas contaminantes de las de interés, sino que también se recurrió a ella para estimar el peso molecular de proteínas individuales y/o complejos.

En este último caso se requiere de la preparación de una recta patrón con estándares moleculares a partir de la cual se realiza la extrapolación del tamaño de la proteína. No obstante, se asume que las proteínas de las muestras poseen volúmenes hidrodinámicos similares a los estándares, aunque esta asunción suele ser acertada para proteínas globulares. En el caso de aquellas que son más fibrilares no se cumple, pudiéndose obtener cálculos distorsionados de la masa molecular.

Con el objetivo de obtener información libre de este problema respecto al peso molecular, se recurrió a asociar (acoplar) esta técnica con una segunda: el MALS.

El MALS, de sus siglas en inglés *Multi-Angle Light Scattering*, proporciona el peso molecular absoluto de una macromolécula en solución. En esta técnica, un láser atraviesa la muestra que contiene la macromolécula a analizar, la cual lo dispersa. La luz dispersada es recogida y medida en detectores situados en distintos ángulos fijos. La intensidad de la luz dispersada en cada ángulo es proporcional a la masa molecular de la macromolécula, la concentración, y el cociente $(dn/dc)^2$ (el incremento del índice de refracción del analito, es decir, la diferencia entre el índice de refracción del analito y el tampón). Con estos equipos se pueden medir todos estos parámetros y se calcula la masa molar de la macromolécula (Folta-Stogniew and Weeks). El MALS está asociado a una cromatografía de exclusión molecular previa que separará las macromoléculas de la muestra en función del tamaño, de forma que podamos calcular las masas moleculares de los especímenes individuales.

Además de conocer la masa molecular absoluta de una proteína, en la caracterización de la muestra interesa conocer el estado de dispersión de la muestra, el tamaño hidrodinámico o el estado de agregación, y para ello se puede recurrir al DLS o *Dynamic Light Scattering*.

El DLS, conocido también como dispersión QELS (*Quasi Elastic Light scattering*) o la espectroscopía de correlación de fotón, no mide la dispersión de luz, sino que mide fluctuaciones en la intensidad de la luz al atravesar una muestra cuyas moléculas en solución poseen el movimiento Browniano, el cual dependerá del tamaño de las moléculas, la temperatura, la viscosidad del solvente, el

coeficiente de difusión y el radio hidrodinámico. En esta técnica, un láser atraviesa la muestra y a un cierto ángulo se sitúa el contador de fotones, que registra la fluctuación de éstos, y la información llega a un procesador digital que la transforma. Por tanto, tras un experimento DLS podremos conocer cómo de polidispersa es la muestra, como de agregada está y cuáles son los diámetros hidrodinámicos presentes en ella (Folta-Stogniew and Weeks).

En el centro donde se ha desarrollado esta Tesis se dispone de ambas técnicas, conectadas en tándem (Wyatt TREOS II, Wyatt Optilab T-rEX, Wyatt DynaPro NanoStar), de modo que se utilizaron para conocer la masa molecular y el estado de agregación de las proteínas a estudio. Para utilizar la técnica utilizamos la columna KW-803-4F para la cromatografía de exclusión molecular previa al análisis por el detector MALS con un sistema cromatográfico FPLC ÄKTA Pure (GE Healthcare). El tampón utilizado fue 50 mM Hepes pH 7,5 + 150 mM NaCl, y se inyectaron al menos 100 µg de cada proteína que se eluyeron a un flujo de 0,35 mL/min. El *software* de análisis utilizado fue Astra 7.1.2.

2.7. ENSAYOS DE INTERACCIÓN DE PROTEÍNAS.

2.7.1. Geles de acrilamida en condiciones nativas (Native-PAGE).

Esta técnica utiliza una aproximación electroforética en geles de acrilamida (PAGE) pero manteniendo las proteínas en sus condiciones nativas. Para ello se elimina en todo el proceso el detergente SDS, y el reductor β-mercaptoetanol. Las muestras tampoco son previamente calentadas a 95 ° C, si no que se mantienen en las condiciones consideradas más apropiadas para su estabilidad. Los geles son preparados de la misma manera que los utilizados en el SDS-PAGE pero sin añadir el SDS y preparando la parte resolutive del gel al porcentaje de acrilamida de interés según el tamaño de las proteínas de las muestras (en nuestro caso, del 8 %) (Anexo II).

En este trabajo de Tesis se ha utilizado esta técnica para analizar la interacción entre diferentes proteínas y observar la formación de complejos. Para ello, se incubaron de forma conjunta las proteínas implicadas antes de ser sometidas a la electroforesis a 100 V durante varias horas y a 4 ° C para evitar que el sobrecalentamiento del gel provocara la alteración de la interacción proteica. El tiempo de electroforesis requiere de optimización para cada interacción analizada. Dado que en estos geles las proteínas migran en función de su relación carga-masa, no podemos utilizar marcadores de peso molecular, y se analiza la formación de complejo por comparación de la movilidad en las muestras de los complejos frente a carreras donde solo se encuentran de forma

individual cada una de las proteínas integrantes del complejo. Una vez terminada la electroforesis, los geles se tiñen de la misma manera que en la técnica SDS-PAGE.

2.7.2. Interferometría de biocapa.

La interferometría de biocapa es una técnica analítica que permite medir interacciones entre macromoléculas en tiempo real. En este trabajo de Tesis se utilizó específicamente para medidas de interacción proteína-proteína. En esta metodología una de las proteínas, denominada ligando, es capturada en un biosensor a través del cual se emite una luz blanca cuya reflexión se recoge y se analiza el patrón de interferencia. La segunda proteína, denominada analito, se prepara en soluciones a distintas concentraciones en las que el biosensor es sumergido. Cuando el analito se une al ligando la reflexión de la luz blanca cambia, y con ello, el patrón de interferencia.

No se requiere el marcaje de las moléculas participantes, no obstante, tiene la limitación de que la inmovilización de la proteína capturada en el sensor podría limitar o impedir la interacción con la otra proteína. Para evitar conclusiones erróneas, es conveniente hacer el experimento en la otra dirección, es decir, intercambiar ligando y analito.

En esta Tesis se han utilizado dos dispositivos distintos para realizar esta técnica, BLITZ y OCTET. La diferencia entre ellos es la sensibilidad de detección y la cantidad de interacciones que se pueden medir simultáneamente. Con BLITZ se puede medir una única interacción mientras que en el OCTET se pueden medir hasta 8 simultáneamente y la sensibilidad es mayor, pudiendo medir la interacción con moléculas de menor tamaño.

Esta técnica presenta 5 etapas tras la hidratación del sensor en el tampón de ensayo durante al menos 10 minutos (Figura M. 3):

1. Captura de la proteína en el sensor: el sensor es sumergido en una solución a una concentración de ligando suficiente para que todos los sitios disponibles del sensor sean ocupados por moléculas del ligando. La concentración de ligando utilizada es aquella que permita una unión gradual de moléculas al biosensor en una señal de entre 1 y 3 nm en 5 minutos. Hay diferentes tipos de sensores, y su elección dependerá de la proteína que se quiera capturar a él o del tamaño de la molécula libre que se unirá a la capturada. En nuestro caso, siempre se han usado sensores Ni-NTA para la captura de proteínas con etiqueta de histidinas.
2. Eliminación de las moléculas excedentes: una vez el ligando es capturado en el sensor, éste se sumerge en tampón de ensayo para que las moléculas que no se han unido al sensor se desprendan de él. El tiempo de este paso es determinado por la señal de interferencia, que tiene que ser estable tras el descenso.

3. Línea base: el sensor con el ligando unido y sin uniones inespecíficas se sumerge en el tampón para establecer las condiciones basales de interferencia. Suele durar entre 30 y 60 segundos, el tiempo requerido para que la señal de interferencia se mantenga estable.
4. Asociación: el sensor se sumerge en una solución que contiene el analito, susceptible de unirse a ligando. A medida que el analito se une al ligando, el grosor de la biocapa del sensor aumenta y con ello cambia el patrón de interferencia. El tiempo de asociación se determina en función de la afinidad de la interacción. La asociación dura hasta que la señal de interferencia se estabilice al llegar a *plateau*, siendo el tiempo de asociación determinado en función de la afinidad de la interacción
5. Disociación: el sensor con las dos proteínas unidas se sumerge en tampón para analizar cómo el analito se disocia del ligando, siguiéndose esta disociación por el cambio en el patrón de interferencia. La disociación se alargará hasta que la señal de interacción sea similar a la del inicio del paso de asociación.

A partir de los patrones de asociación y disociación el *software* Octet Data Analysis HT calcula las cinéticas de unión en función de la información biológica que se disponga sobre la interacción, pudiendo seleccionar entre un modelo de interacción 1:1 o 1:2 (ligando:analito). De este análisis se obtiene la constante de asociación (K_a), la constante de disociación (K_d), así como la constante de afinidad (K_D). Siempre en el experimento se incluyen los controles negativos de la unión inespecífica de moléculas del tampón al sensor, así como la posible unión inespecífica del analito al sensor. También se incluyen al menos 4 concentraciones de la proteína analito para un cálculo fiable y preciso de la K_D , siendo dos de ellas 10 veces más y 10 veces menos la K_D .

En la Tabla M. 9 se incluyen los tampones usados en los distintos experimentos realizados con esta técnica, las concentraciones de los correspondientes ligandos y analitos, y los tiempos de cada una de las 5 etapas para cada uno de los ensayos realizados.

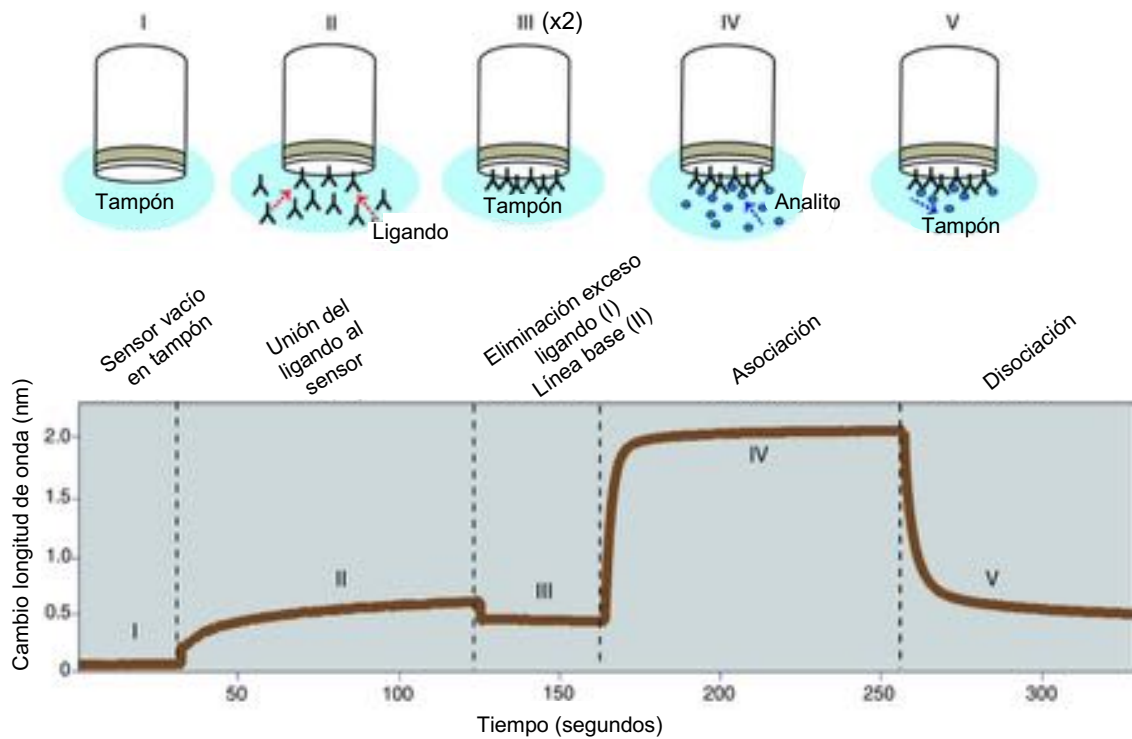


Figura M. 3. Diagrama de la técnica interferométrica de biocapa. En la parte superior se muestra un diagrama del proceso, y en la parte inferior el sensograma obtenido: I. Inmersión del sensor vacío en tampón para su equilibrado; II. Unión de ligando al sensor; III (I). Eliminación del exceso de ligando; III (II). Línea base; IV. Asociación; V. Disociación. Figura adaptada de *Sultana and Lee 2015*.

Tabla M. 9. Tabla recopilatoria de las condiciones utilizadas en las medidas de interferometría de biocapa. Se especifican las concentraciones finales utilizadas de la proteína ligando y analito, así como los tiempos de cada etapa del proceso y el tampón de reacción.

Medida probada ligando – analito (aparato utilizado)	Concentración		Tiempo (segundos)					Composición
	Proteína ligando	Proteína analito	Equilibrado inicial	Captura del ligando	Eliminación exceso ligando /Línea base	Asociación	Disociación	Tampón
mDut - Stl ^{N-ter} (OCTET)	218,7 nM	218,7 to 3,4 nM (diluciones seriuadas ½)	600	280	200 / 60	120	120	75 mM Hepes pH 7,5 + 250 mM NaCl + 5 mM MgCl ₂ + 1 % BSA + 0,005 % Tween – 20 + 10 mM imidazol
mDut ^{H135F} - Stl ^{N-ter} (OCTET)								
hDut - Stl ^{N-ter} (OCTET)								
Φ11Dut - Stl ^{N-ter} (OCTET)								
EBV ^{Dut} - Stl ^{N-ter} (BLITz)	410 nM	4100 nM		300	30 / 30	300	300	20 mM Hepes pH 7,5 + 250 mM NaCl + 10 mM MgCl ₂
mDut - Rv3524 (BLITz)	300 nM	3000 nM		120	30 / 30	300	300	50 mM Tris pH 8 + 400 mM NaCl
Rv3524 – mDut (BLITz)	600 nM	6000 nM						

2.8. ENSAYOS DE INTERACCIÓN PROTEÍNA - ADN.

En esta Tesis solo se ha probado la interacción entre proteína y ADN mediante ensayos electroforéticos de retraso en gel o EMSA (de sus siglas en inglés *electrophoretic mobility shift assay*).

Es un método rápido y sensible para detectar este tipo de interacciones, y se basa en que la movilidad electroforética de un complejo proteína – ADN es normalmente menor a la del ADN libre (Hellman and Fried 2007). Aunque habitualmente se utiliza con fines cualitativos también es posible para hacer aproximaciones cuantitativas, como se ha hecho en esta Tesis.

Las muestras son sometidas a electroforesis en gel de acrilamida al porcentaje de acrilamida acorde al tamaño de las proteínas participantes (en nuestro caso al 8 %). La composición es la misma que la del PAGE, pero con la excepción de que no posee parte de apilamiento, y el tampón Tris es sustituido por tampón TBE (Anexo II). Tras la electroforesis, que se realiza a 4 ° C para evitar el sobrecalentamiento del gel, el ADN puede ser visualizado por diferentes aproximaciones como el Red Safe o sondas fluorescentes incorporados al ADN durante su síntesis.

Para los ensayos realizados en esta Tesis, las muestras se incubaron a temperatura ambiente durante 30 minutos antes de ser cargadas en el tampón de reacción (50 mM Hepes pH 7,5 + 75 mM NaCl + 5 mM MgCl₂ + 0,5 mM EDTA), y posteriormente se añadió el tampón de carga de ADN. Antes de cargar las muestras en el gel, éste fue pre-corrido a 100 V durante 1 hora. Una vez incubada la proteína y el ADN, las muestras fueron cargadas en el gel, y la electroforesis se realizó a 100 V durante 150 minutos.

Dado que el ADN utilizado estaba marcado fluorescentemente en ambos extremos con el fluoróforo IR700 (Eurofins Genomics), una vez terminada la electroforesis, el resultado en el gel fue visualizado en el sistema de imagen Odyssey (Li-Cor).

2.9. MEDIDA DE ACTIVIDAD ENZIMÁTICA POR REACTIVO VERDE MALAQUITA.

Con el objetivo de medir la actividad enzimática de Duts y MazG purificadas en este trabajo, se recurrió al ensayo basado en el reactivo Verde Malaquita, en el cual se cuantifica la producción de fosfato inorgánico mediante una reacción colorimétrica (Pegan et al. 2010; Vardakou et al. 2014).

Los ensayos se realizan con una solución de Verde Malaquita compuesta por: 1.22 g/L del reactivo Verde Malaquita (Acros), responsable del viraje de color, 16.67% de ácido sulfúrico, que frena la reacción por desnaturalización de la enzima, y agua milliQ. Esta solución de color anaranjado se prepara con anterioridad a la realización del experimento y se almacena a 4 ° C. Solo antes de realizar los ensayos de actividad, se mezcla la solución de Verde Malaquita con amonio de molibdeno al 7.5 % y el Tween-20 al 0,02 %, para generar la mezcla de verde malaquita, la cual tiene una viabilidad de 4 horas y es la que se utilizará en el ensayo.

La interacción de fosfato con la mezcla del Verde Malaquita da lugar a la formación del complejo Verde Malaquita – fosfomolibdato, de color verde con un máximo de absorbancia a 623 nm. Por tanto, el viraje del color naranja a verde es proporcional a la concentración de Pi presente. Por tanto, a partir de la absorbancia puede calcularse la concentración de Pi producida por la enzima mediante la interpolación a una recta patrón, la cual se prepara a partir de concentraciones conocidas de PO_4^{4-} .

Las reacciones enzimáticas son iniciadas con la adición del sustrato a una mezcla de la proteína y el tampón de reacción, y se toman muestras de la reacción a distintos tiempos en función de la enzima y/o factor a ensayar. La reacción es frenada a dichos tiempos por la adición de la solución Verde Malaquita en proporción 1:4 (muestra de reacción: solución Verde Malaquita). Tras 20 minutos de incubación a temperatura ambiente, la lectura colorimétrica es realizada con un espectrofotómetro a 630 nm (en nuestro caso Victor² 1420 Multilabel counter (Wallac)). Toda la reacción enzimática se realiza a temperatura ambiente.

Los ensayos realizados en esta Tesis fueron llevados a cabo en 300 μL de mezcla de reacción, que contiene el tampón de reacción, la Dut correspondiente, y 0,01 unidades de pirofosfatasa inorgánica (Thermo Scientific). Se tomaron muestras cada 1 o 2 minutos durante 5 o 10 minutos respectivamente, según la proteína a testar. La utilización de la pirofosfatasa fue requerida por la naturaleza de la hidrólisis llevada a cabo por estas enzimas, que producen PPi en lugar de Pi. Por tanto, se requiere de la reacción acoplada de la pirofosfatasa para la hidrólisis del PPi a 2 Pi.

Estos ensayos han tenido como finalidad la detección y/o cuantificación de la inhibición de la proteína Dut de Epstein Barr por presencia de otra proteína o distintas concentraciones de un reactivo en el caso de las MazG, así como la alteración de la actividad enzimática consecuencia de mutaciones puntuales (también en MazG). En el caso de la Dut, ésta ha sido incubada junto al inhibidor durante 15 minutos a temperatura ambiente antes de llevar a cabo el ensayo. En el caso de las MazG no ha habido incubación previa pero se han testado distintas concentraciones de distintos reactivos para ver sus efectos sobre la enzima.

En la Tabla M. 10 se recopilan los detalles específicos de concentraciones de enzima, sustrato, tiempos de toma de muestras y tampones de reacción. En el correspondiente apartado de resultados se explicarán con mayor detalle los ensayos llevados a cabo.

Tabla M. 10. Tabla recopilatoria de las concentraciones de Dut/MazG, sustrato, solución tampón y tiempos de toma de muestra de los experimentos de actividad enzimática realizados en esta Tesis.

Proteína testada	[Proteína] (Dut)	[sustrato] (dUTP)	Tampón	Toma de muestras
EnMazG	2 ng/ μ L	300 μ M	50 mM Tris-HCl 8,5 + 250 mM NaCl + 5 mM MgCl ₂	Cada 2 min durante 10 min
SeMazG	1 ng/ μ L			
EBVDut	6 ng/ μ L	15 μ M	75 mM Hepes pH 7,5 + 250 mM NaCl + 5 mM MgCl ₂	Cada 30 seg o 1 min durante 5 min

Para el análisis de los datos en esta Tesis, se restó la contribución del sustrato a la absorbancia final restando el valor del control reacción sin enzima iniciada con sustrato, y también a cada reacción el correspondiente valor a tiempo 0. El análisis se llevó a cabo con el software Graphpad Prism 9. 0. 1.

3. CRISTALOGÉNESIS, CONGELACIÓN y DIFRACCIÓN DE CRISTALES DE PROTEÍNA.

3.1. Puesta de placas de *screening*.

A lo largo de esta Tesis doctoral se ha tenido como objetivo determinar la estructura tridimensional de diferentes proteínas en complejo con otras proteínas o en diferentes estados enzimáticos. Para ello se ha utilizado la técnica de la difracción de rayos X de cristales de proteína.

Un cristal de proteína es un sólido ordenado en el que se repite una unidad básica. Para poder conocer la estructura tridimensional de la proteína se requiere que el orden interno sea suficiente para alcanzar una resolución de 3 Å o mejor, lo cual puede dificultarse en parte por las elevadas cantidades de agua que contienen (sobre un 50 % del volumen de cristal) que los hace especialmente frágiles.

Para generar los cristales se recurrió a la técnica de difusión de vapor. En ella, una gota formada de proteína y de solución de cristalización se mantiene en una

cámara bien sellada que además incluye un reservorio con un mayor volumen (2 o 3 órdenes de magnitud) de la solución de cristalización. La gota y el reservorio no están en contacto directo. La presión de vapor de agua es mayor en la gota que en el reservorio, de modo que la gota poco a poco disminuye su contenido en agua hasta que las concentraciones de los precipitantes se equiparan en gota y reservorio.

De forma habitual se utilizan dos modalidades de la difusión de vapor conocidas como gota colgante y gota sentada, que se diferencia, como su nombre refleja, por la disposición de la gota de cristalización en el ensayo. En nuestro caso utilizamos gota sentada, donde la gota de cristalización queda sobre una elevación cóncava que se mantiene sobre el nivel del reservorio donde se coloca la solución cristalizante o madre (*Mother liquor* en inglés), una mezcla de tampones, aditivos y soluciones precipitantes que ayudarán a precipitar la proteína. El uso de las placas y un robot que dispensa las gotas y el reservorio, ha permitido que existan placas donde se pueden colocar de 1 a 3 gotas distintas, y con ello, simultáneamente se puede poner a cristalizar para cada condición hasta tres muestras de proteínas. Habitualmente se usaron placas con dos gotas, de modo que se pusieron las proteínas solas (conformación *apo*) y en presencia del ligando / iones / otra proteína (co-cristalización), que por lo general, siempre estabilizan la organización proteica y con ello, favorecen la cristalización. Además de la co-cristalización, se han intentado otras aproximaciones, pero debido a su especificidad se abordarán en el apartado de resultados.

Una vez rellena la placa con las 96 condiciones, ésta se sella con una lámina de plástico para dejar que la difusión de vapor tenga lugar. Las placas ya selladas se dejan a temperatura constante para que alcancen el equilibrio, habitualmente a 21 ° C o 4 ° C. La cristalización proteica depende de diversos factores, tanto de las soluciones de cristalización (aditivos, precipitantes, pH...) como factores de la propia muestra (concentración, pureza, homogeneidad, estabilidad...incluso presencia o ausencia de una etiqueta en el extremo N o C-terminal de la proteína), con lo que se puede hacer cribados con todas estas condiciones.

En esta Tesis, se han realizado tanto cribados de condiciones comerciales como JBS I y JBS II (Jena Bioscience), JCSG+ y Aditivos ANGSTROM (Molecular Dimensions), como ampliaciones de las mejores condiciones, o utilizando un 85 % de la solución madre. Las ampliaciones eran diseñadas alrededor de la condición de cristalización que mayor resolución nos había proporcionado, con lo que manteniendo el pH y la concentración del tampón, se variaba la concentración de los aditivos y precipitantes en ambas direcciones de magnitud. En cuanto a la solución madre, esta aproximación se basó en utilizar un *screening* de condiciones conocido y que suponía el 15 % de la solución de cristalización comercial, mientras que el otro 85 % era la solución de

cristalización que mayor resolución nos había proporcionado (ésta, la solución madre, era preparada un 15 % más concentrada para que en la gota quedara a la concentración exitosa).

A lo largo de esta Tesis se han cristalizado varias proteínas, aunque no todos los cristales obtenidos fueron lo suficientemente válidos para poder obtener la estructura tridimensional de la proteína. Las condiciones específicas de cada proteína o complejo del que se ha conseguido la estructura tridimensional sobre concentración utilizada, ratio molar en la cristalogénesis, y solución de cristalización pueden consultarse en la Tabla M.11, aunque se han especificado igualmente en el correspondiente apartado de resultados.

Tabla M. 11. Tabla recopilatoria de las condiciones de cristalización y crioprotección de los cristales de proteína cuya estructura tridimensional se ha obtenido en este trabajo de Tesis. (PEG: polietilenglicol)

Estructura resuelta	[Proteína]	Solución cristalización	Solución crioprotectora
Complejo mOut-SI TM (sin etiquetas de purificación)	6,2 mg / mL – 7,7 mg / mL	20 % PEG 10.000 + 100 mM Na-Hepes pH7,5	Solución cristalización suplementada con 35 % PEG 10.000 + 10 % etilenglicol
Complejo nOut-SI TM (nOut con etiqueta His)	7,56 mg / mL – 5,46 mg / mL	1,6 M (NH ₄) ₂ SO ₄ + 1 M Li ₂ SO ₄	Solución cristalización suplementada con 5 % glicerol + 10 % etilenglicol
EnMazG + MnCl ₂	15 mg / mL	30 % propanol + 100 mM Tris pH 8,5 + 200 mM acetato de amonio	Solución cristalización suplementada con 10 % glicerol + 10 % etilenglicol + 10 % PEG 300
EnMazG + MnCl ₂ : MgCl ₂ 1: 2,5	10 mg / mL	15 % Etanol + 100 mM Tris pH 8,5 + 20 % PEG 300	Solución cristalización suplementada con 10 % PEG 400
EnMazG + MnCl ₂ : MgCl ₂ 1: 4	10 mg / mL	150 mM Li ₂ SO ₄ + 100 mM Tris pH 8,5 + 30 % PEG 300	-
EnMazG + MnCl ₂ : MgCl ₂ 1: 8	10 mg / mL	170 mM Li ₂ SO ₄ + 100 mM Tris pH 8,5 + 25,5 % PEG 4000 + 3 % Etanol	Solución cristalización suplementada con 10 % PEG 4000
EnMazG + CaCl ₂ : MgCl ₂ 1: 6	10 mg / mL	15 % Etanol + 100 mM Tris pH 8,5 + 20 % PEG 300	Solución cristalización suplementada con 15 % PEG 400

3.2. Congelación de cristales.

De los cristales obtenidos se seleccionaron los de tamaño suficiente (por encima de 10 -20 µm) para ser difractados en el sincrotrón.

La difracción de rayos X de los cristales de proteína requiere de su protección para evitar la formación de hielo en el solvente del cristal. Las soluciones crioprotectoras suelen ser alcoholes de bajo peso molecular como el glicerol, etilenglicol o polietilenglicol, aunque a veces las mismas condiciones de cristalización ya son crioprotectoras. Cuando la propia solución madre del cristal

no lo es, se prepara la solución crioprotectora de manera gradual. Es decir, se añade una pequeña concentración de la solución crioprotectora en función de la composición de la solución madre y se sumerge un cristal en ella. Si el cristal no sufre daño macroscópico, se vitrifica en menos de 2 segundos en nitrógeno líquido. En el caso de sufrir daño, se busca una alternativa de solución crioprotectora y se repite el proceso.

Tras la vitrificación, son difractados en sincrotrones. En esta Tesis se han llevado a difractar al sincrotrón ALBA (Barcelona, España) y DLS (Didcot, Inglaterra).

Las soluciones crioprotectoras utilizadas específicamente en cada cristal a partir del cual se ha obtenido la estructura tridimensional de la proteína o complejo pueden consultarse en la Tabla M. 11, aunque también se han especificado más adelante en el apartado correspondiente de resultados.

3.3. Difracción de rayos X de cristales de proteínas.

La difracción de cristales y recogida de datos presentados en este trabajo de Tesis se llevó a cabo a 100 K en los sincrotrones ALBA (Barcelona, España) o DLS (Didcot, Inglaterra).

Para poder resolver las estructuras tridimensionales a partir de los juegos de datos recogidos, es necesario que éstos provengan del impacto del rayo X en diversos puntos del cristal. Por ello, es necesario establecer una estrategia de difracción. Ésta es calculada automáticamente por programas del sincrotrón cuando se testa inicialmente el cristal. El testado inicial del cristal se lleva a cabo haciendo impactar el rayo X en tres puntos del cristal en tres orientaciones distintas (0, 45 y 90 °). De la difracción del cristal en estos tres puntos se diseña la estrategia de difracción, en la cual se especifica el porcentaje del rayo, el tiempo de exposición, los grados de rotación, los grados totales del cristal a difractar y la resolución máxima que puede obtenerse de dicho cristal. En función de la simetría que el cristal presente, la cantidad de información contenida en él que necesitemos recoger será inversamente proporcional, de manera que a mayor simetría, menos grados del cristal se necesitarán recoger (de los 360 ° totales).

A largo de esta Tesis se testaron cientos de cristales (más de 400), recogiendo juegos de datos completos para más de 70 cristales.

4. RESOLUCIÓN DE LA ESTRUCTURA TRIDIMENSIONAL.

Una vez recogido el espectro completo de difracción, el cual estará compuesto por cientos de imágenes (Figura M. 4), o en su defecto unas pocas películas (según el detector utilizado en la difracción), éste se digitaliza y se almacena hasta su posterior procesamiento.

La transformación de la información contenida en el espectro total de difracción en la estructura tridimensional se divide en diversos procesos: indexado, refinado de la celda, integración, traslación, escalado y mezcla, cálculo de la fase, trazado y refinado.

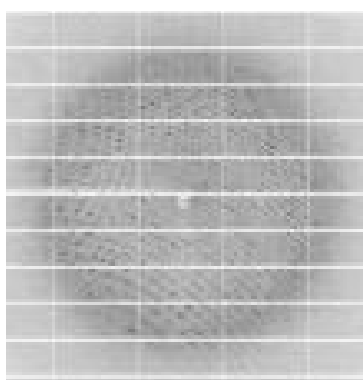


Figura M. 4. Imagen de difracción. Ejemplo de imagen recogida por el detector tras el impacto del rayo X sobre un cristal de proteína. Los puntos (*spots*) son las reflexiones, producto del impacto de los rayos X dispersados. Fuente: imagen de difracción obtenida en esta Tesis.

4.1. Indexado y refinado de la celda.

Ambos procesos son llevados a cabo de manera automática por los programas iMOSFLM y DIALS (en los datos más antiguos se utilizó el primero, y para los más recientes el segundo). En el indexado se seleccionan unas pocas imágenes de difracción correspondientes a distintas orientaciones del cristal, y el programa mapea las posiciones de las reflexiones en dichas imágenes e identifica la orientación del cristal, además de estimar de manera inicial las dimensiones de la celda unidad, la simetría cristalográfica y se atribuye un índice a cada reflexión (Powell 2017).

A continuación, se realiza el refinado de la celda, en el cual el mismo programa compara el ajuste de la celda inicial y la simetría con los posibles grupos espaciales, para asociarle el más óptimo. El objetivo de este proceso es ajustar al máximo el tamaño de la celda que se repite en el cristal y establecer la

correlación de reflexiones entre imágenes, de modo que se identifique una misma reflexión en todas las imágenes en las que esté presente (Powell 2017).

4.2. Integración.

En la integración, la intensidad de todas las reflexiones de todas las imágenes es calculada y se integran en un todo, y las fracciones de una misma reflexión son también asociadas entre sí.

4.3. Traslación, escalado y mezcla de datos.

El integrado de las reflexiones no escala las intensidades, simplemente las calcula. Esto implica que reflexiones equivalentes tienen distinta intensidad, lo cual es debido al daño de la radiación causada en el cristal y la detección no uniforme de los rayos X por parte del detector, entre otros motivos (Powell 2017). Por tanto, el siguiente paso es el escalado de los datos, o mejor dicho, el escalado del integrado.

El escalado fue llevado a cabo por el programa Aimless (Evans and Murshudov 2013) dentro del paquete de programas CCP4 (Winn et al. 2011), que incluye a su vez a tres programas que pueden ser utilizados de manera independiente: Pointless, Scala y C-truncate (Evans 2011).

Antes del propio escalado, el programa Pointless analiza el grupo espacial asignado al cristal para añadirle los *screw axes*. Estos ejes son ejes de traslación de la molécula alrededor del eje de simetría. El grupo espacial asignado previamente no tiene en cuenta la traslación de la molécula, solo la simetría, con lo que al pasar por Pointless, en caso de ser necesario, se le añaden. Son representados con subíndices junto al eje de simetría correspondiente.

Scala llevará a cabo el escalado: pondrá las intensidades que se han calculado en el integrado en una misma escala y unificará (mezclará) reflexiones equivalentes de todo el espectro de difracción. Aquí se establecerán las reflexiones únicas.

Finalmente, una vez escalados los datos, el programa C-truncate tendrá dos funciones principales. Por una parte reservará un 5 % de las reflexiones únicas que se etiquetarán como *Free flags*, que no se tendrán en cuenta para el refinado posterior de la estructura tridimensional de la proteína; y por otra calculará las amplitudes a partir de las intensidades ya escaladas.

Al finalizar el escalado, se analiza la calidad de los datos, para decidir a qué resolución se utilizan los datos. Para hacer este análisis, en esta Tesis se recurrió a los valores de los siguientes parámetros (Powell 2017):

- Mean I/σ : es la medida de la señal sobre el ruido. Se consideró una buena resolución cuando este parámetro tuvo un valor igual a superior a 1,5.
- R_{merge} : es una medida del error del valor medio de intensidad asociado a todas las reflexiones. Incluye aquellas equivalentes por simetría, con lo que el parámetro no es del todo preciso porque aumenta con la multiplicidad. Se consideró un valor aceptable cuando era igual o inferior a 0,6.
- R_{meas} : es el mismo parámetro que R_{merge} , pero teniendo en cuenta la multiplicidad, con lo que es más preciso. También se consideró un valor aceptable 0,6 o inferior. Ambos parámetros se refieren a las reflexiones no unificadas.
- R_{pim} : es una medida del error de la intensidad media asociada a todas las reflexiones ya estando unificadas las equivalentes. Se consideró un valor aceptable igual o inferior a 0,4.
- $CC_{1/2}$: es una medida de la similitud entre las dos mitades del mismo juego de datos dividido al azar. A partir de 0,5 significa que hay correlación significativa.
- Completitud: es un indicativo de cómo de completos son los datos. Cuanto más cercano a 100 sea su valor, mejor.
- Multiplicidad: es un indicativo de cuántas veces se ha recogido la misma información, de cuantas reflexiones equivalentes hay. Con una multiplicidad de 3 ya se consideró una buena resolución.

Todos los valores se analizaron en la capa de mayor resolución.

4.4. Cálculo de la fase.

En el paso desde el espectro de difracción que se obtiene en el sincrotrón, al mapa de densidad electrónica que esclarecerá la estructura tridimensional, es necesaria la resolución de la función Transformada de Fourier, cuya ecuación es la siguiente:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl}^{+\infty} |F(hkl)| \cos 2\pi (hx + ky + lz - \varphi(hkl))$$

En esta ecuación, $\rho(xyz)$ define la densidad electrónica, y la otra parte de la ecuación define el espacio recíproco (espectro de difracción total). En el espacio recíproco interviene la magnitud F , conocida como factor de estructura y que define cada punto de dicho espacio como una onda, que tiene su módulo ($|F(hkl)|$) y su fase ($\varphi(hkl)$). Los módulos representan las amplitudes de los haces difractados y se obtienen a partir de las intensidades recogidas en el detector. La fase no puede obtenerse experimentalmente y se recurre a otros métodos, ya que constituye el último factor a calcular de la transformada de Fourier que limita

el paso a la densidad electrónica (Hermoso Domínguez and Martínez Ripoll 2003).

Tres métodos pueden ser utilizados para calcular la fase: Reemplazo Isomórfico Múltiple (MIR, de sus siglas en inglés *Multiple Isomorphous Replacement*), Dispersión Anómala Múltiple (MAD, de sus siglas en inglés *Multi-wavelength Anomalous Diffraction*), y Reemplazo Molecular (MR, de sus siglas en inglés *Molecular Replacement*) (Hermoso Domínguez and Martínez Ripoll 2003).

- Reemplazo Isomórfico Múltiple: consiste en introducir en el cristal átomos que sean grandes dispersores de los rayos X, es decir, átomos con alto número atómico. Éstos penetran por los canales de los cristales por difusión, y son coordinados por las cadenas laterales que son capaces de hacerlo. Tiene la dificultad de que dichos átomos no deben distorsionar la estructura nativa de la proteína, de ahí que sea más fácil aplicar esta técnica en metaloproteínas, en las que se puede sustituir sus metales endógenos por otros más pesados.

Estos átomos pesados con mayor capacidad de dispersión modifican ligeramente la intensidad del espectro de difracción del cristal que los contiene (cristal derivatizado) respecto al que no los contiene (cristal nativo). Esas diferencias de intensidad se utilizan para calcular un mapa de vectores para situar los átomos pesados (mapa de Patterson) dentro de la celda unidad. A partir de este mapa, se calculan los factores de estructura (F), con sus módulos y fases, de modo que se completa la *Transformada de Fourier* y se obtiene el mapa de densidad electrónica.

- Dispersión Anómala Múltiple: consiste en provocar cambios en la intensidad de la difracción de un determinado átomo, haciendo incidir el rayo X con una energía próxima a la frecuencia natural de oscilación de los electrones de ese átomo, que está presente en el cristal y que producirá la dispersión anómala. La dispersión anómala es el nombre que recibe el cambio de intensidad entre las llamadas reflexiones de *Friedel*, que son parejas de reflexiones que en condiciones normales (no dispersión anómala) poseen la misma amplitud y fase pero con símbolos opuestos. En esta técnica se recoge un mismo cristal que contenga el átomo dispersor a tres longitudes de onda: una que maximiza f' (pico), otra que minimice f'' (inflexión), y otra energía alejada de las dos anteriores (remoto); siendo f' y f'' dos componentes de la difracción anómala (real e imaginaria, respectivamente). Cuando se analizan las diferencias entre los datos recogidos a las tres longitudes de onda, se calculan las amplitudes y las fases generadas por el átomo dispersor, y a partir de ellas se calcula el mapa de densidad electrónica para toda la proteína.

- Reemplazo molecular: consiste en utilizar la estructura de una proteína de secuencia de aminoácidos homóloga a la proteína cuya estructura va a resolverse, y servirá como primer modelo que será refinado a posteriori. Para posicionar el modelo en la celda unidad de los nuevos datos se requiere la correcta orientación y posicionamiento preciso de la molécula en ella. Una vez orientada y trasladada convenientemente, se utilizan los factores de estructura experimentales del modelo (módulo y fase) para calcular el mapa de densidad electrónica.

En esta Tesis, el cálculo de la fase en todas las estructuras resueltas se llevó a cabo mediante reemplazo molecular con el programa Phaser MR (McCoy et al. 2007), también del paquete de programas CCP4 (Winn et al. 2011).

4.5. Trazado y refinado de la estructura.

Una vez conocida la fase por uno u otro método, se podrá completar la Transformada de Fourier, y con ello se obtendrá el mapa de densidad electrónica. La calidad de los mapas dependerá de la resolución de los datos de difracción, la cual se define como el menor espaciado obtenido en el experimento de difracción (Hermoso Domínguez and Martínez Ripoll 2003). A mayor resolución, mayor detalle veremos en el mapa y podremos inferir con mayor precisión parámetros como las distancias o ángulos entre átomos. Deberá tener al menos la resolución suficiente para poder corregir posibles errores en la asignación de residuos de la secuencia primaria de la proteína.

Una vez realizado el reemplazo molecular, se revisó el modelo y se trazó manualmente en algunos extremos donde el modelo de la cadena polipeptídica usado terminaba pero el mapa de densidad electrónica indicaba que podían posicionarse más residuos. Esto se llevó a cabo con el programa Coot (Emsley et al. 2010). Tras el trazado, se refinó el modelo con el programa Refmac5 (Murshudov et al. 2011).

El modelo de una estructura está definido para cada átomo por las coordenadas atómicas y un factor de temperatura (β -factor), que define la movilidad térmica del átomo alrededor de su posición de equilibrio (Hermoso Domínguez and Martínez Ripoll 2003). En el refinado, se ajustan ambos parámetros para optimizarse entre las observaciones recogidas en el sincrotrón y la predicción de la estructura (el modelo).

El modelo final se obtuvo con repeticiones del ciclo trazado-refinado. En este punto hay que destacar el papel de las reflexiones *Free Flags*, porque éstas no entran en el proceso de refinado, sino que solo están para compararlas con el mapa que se construye con el modelo. De esta comparación se calcula el parámetro R_{free} , el cual junto al R_{factor} , que mide el ajuste entre el mapa construido con el modelo y el mapa del escalado, sirve como indicador del grado de

optimización del modelo respecto a los datos experimentales recogidos en el sincrotrón.

Para que el modelo de la estructura sea válido, debe ser lo suficientemente ajustado a las observaciones experimentales y no contener choques entre cadenas laterales, enlaces con distancias o ángulos no permitidos, y los factores β deben ser razonables. Además, se le incluirán todas las moléculas de ligandos, iones o moléculas de agua, sobre todo aquellas que formen parte de la capa de solvatación, que sean estructurales o que sean esenciales para la compresión de la función de la proteína.

Los modelos de los complejos resueltos en este trabajo de Tesis que han sido publicados fueron depositados en la base de datos Protein Data Bank para su consulta pública.

4.6. Cálculo de picos de dispersión anómala.

En uno de los proyectos desarrollados en esta Tesis, se ahondó en la entrada de los iones alternativos al Mg^{2+} en el centro activo de EnMazG para conocer cómo afectaba a la catálisis del sustrato. Para ello, los iones fueron posicionados en la estructura tridimensional haciendo uso de la difracción anómala, pero difractando los cristales que contenían los iones de interés en el pico máximo de las correspondientes energías de oscilación de los átomos a posicionar, es decir, la energía suficiente para provocar una transición electrónica desde un nivel energético inferior a uno superior.

El cálculo de la fase se hizo por reemplazo molecular, y la localización de los iones se hizo con la utilidad *Finding and analysing difference map peaks*, del paquete de programas Phenix (Liebschner et al. 2019). Ésta permite calcular un mapa de diferencias anómalas, cuyo resultado se presenta en formato .pdb, como un modelo, que al superponerse con el modelo de la estructura en Coot (Emsley et al. 2010), permite la localización de los picos de anómalo (el catión de interés). Además, el mismo archivo abierto como un texto permite conocer la señal de anómalo correspondiente a cada pico.

RESULTADOS

CAPÍTULO I

1.1. El fragmento N-terminal de Stl es capaz de unirse a las Duts humana y de *Mycobacterium tuberculosis*.

La capacidad inhibidora de Stl de SaPIbov1 (en adelante, Stl) sobre las Duts triméricas ha sido confirmada bioquímicamente tanto en enzimas procariotas como eucariotas (Benedek et al. 2018; Nyíri et al. 2018; Maiques et al. 2016). Con el objetivo de conocer la base molecular de esta inhibición, se abordó la caracterización estructural por difracción de rayos X de los complejos del represor Stl con Duts representativas de diferentes reinos, una procariota y una eucariota, además del análisis comparativo con la estructura del complejo con su diana natural, la Dut del fago $\phi 11$ de *S. aureus*.

Para ello se seleccionaron la Dut de *M. tuberculosis* como representante procariota (mDut), y la Dut humana (hDut) como representante eucariota, dado que su interacción con el represor Stl había sido caracterizada *in vitro* previamente en otros laboratorios (Nyíri et al. 2018; Hirmondó et al. 2015; Pecsí et al. 2012). Además, ambas Duts habían sido caracterizadas estructural y funcionalmente en profundidad (códigos PDB hDut 1Q5U y 2HQU, códigos PDB 1MQ7 y 1SIX para mDut) (Mol, Harris, Mcintosh, et al. 1996; Varga et al. 2007; Chan et al. 2004; Hirmondó et al. 2015).

Mediante análisis funcionales y estructurales realizados previamente en el laboratorio donde se ha realizado esta Tesis, fue demostrado que Stl es una proteína modular con una región altamente flexible entre su dominio intermedio y el dominio C-terminal, dominios implicados en la interacción con las Duts triméricas y diméricas, respectivamente (Bowring et al. 2017). Esta flexibilidad ha impedido obtener su estructura tridimensional completa, aunque se intentase cristalizar solo o en complejo con sus Duts diana. Dado que para obtener la estructura tridimensional del Stl en complejo con una de sus dianas naturales, la Dut trimérica del fago $\phi 11$ de *S. aureus* ($\phi 11$ Dut) (Rafael Ciges-Tomas et al. 2019), se utilizó una construcción que incluye los residuos 1 a 156 de Stl, se decidió utilizar esta misma construcción para nuestros ensayos. Esta construcción se denominó Stl^{N-ter} e incluye los dominios de unión al ADN (DBD, del inglés *DNA Binding Domain*) y el dominio intermedio de Stl.

Aunque se preveía que el mecanismo de reconocimiento y unión a estas Duts sería similar al utilizado con las Duts de fago, se decidió confirmar que este fragmento de Stl también era el responsable de la unión a estas Duts eucariota y procariota antes de ahondar en los ensayos estructurales.

Con esta finalidad se recurrió a la técnica de electroforesis en condiciones nativas (Native-PAGE). Se realizó un ensayo de titración, en el cual la concentración de Stl^{N-ter} se mantenía constante y la de Dut se incrementaba en ratios molares Stl^{N-ter}:Dut crecientes desde 1:0,5 hasta 1:4. Estos ensayos se realizaron por triplicado para cada Dut y se introdujo como control positivo la

proteína $\phi 11\text{Dut}$. En las tres repeticiones, las bandas correspondientes a las proteínas solas (tanto las Duts como $\text{Stl}^{\text{N-ter}}$) desaparecían de manera simultánea a la aparición de la banda correspondiente al complejo $\text{Dut-Stl}^{\text{N-ter}}$ (Figura R.1.1). Estos ensayos de titración mostraron que para una ratio molar 1:1 la formación del complejo $\text{Dut-Stl}^{\text{N-ter}}$, tanto en mDut como hDut , es máxima con una mínima cantidad residual de cada una de las proteínas participantes, al igual que podía verse en el control con $\phi 11\text{Dut}$.

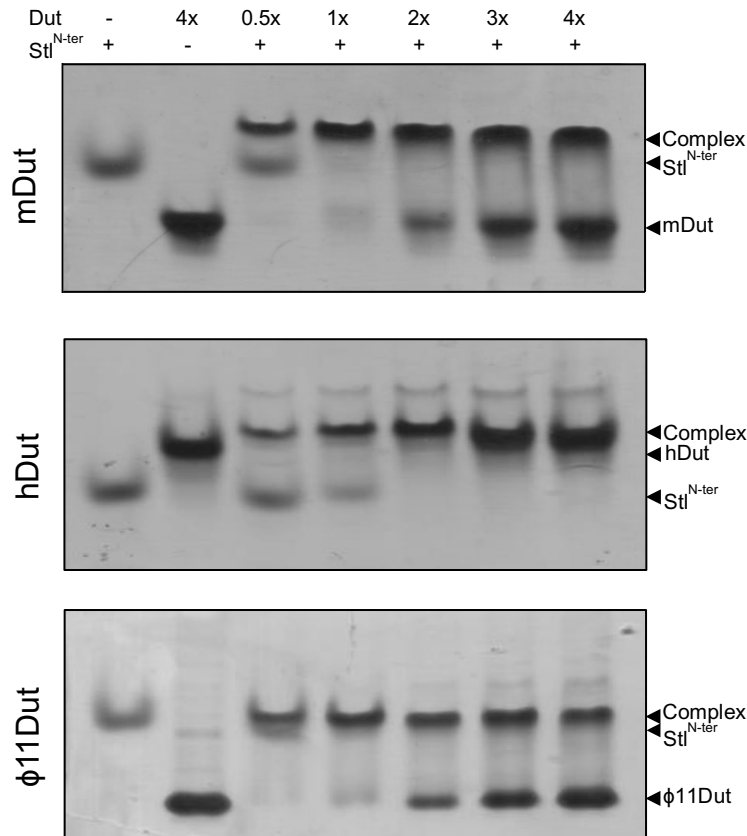


Figura R.1.1. Geles al 8 % de acrilamida en condiciones nativas para la visualización del complejo de $\text{Stl}^{\text{N-ter}}$ con mDut , hDut y $\phi 11\text{Dut}$. Ratios molares crecientes de Dut fueron añadidos a una concentración fija de $\text{Stl}^{\text{N-ter}}$ ($11 \mu\text{M}$). Figura tomada de *Sanz-Frasquet et al. 2023*.

Estos resultados apoyan el modelo de interacción 1:1 observado en la estructura del complejo $\phi 11\text{Dut-Stl}^{\text{N-ter}}$, en el cual un trímero de $\phi 11\text{Dut}$ une tres monómeros independientes de $\text{Stl}^{\text{N-ter}}$ (Rafael Ciges-Tomas et al. 2019); la misma estequiometría propuesta anteriormente por la misma técnica utilizando el represor en su longitud total (Stl^{FL}) (Hirmondó et al. 2015; Nyíri et al. 2018). No obstante, también había sido propuesta la estequiometría $\text{Dut-Stl}^{\text{N-ter}}$ 3:2 (Nyíri et al. 2018).

Una vez confirmada cualitativamente la interacción, se procedió a la medida cuantitativa de ésta para comparar la afinidad de interacción del represor por estas Duts con la que presenta por su diana natural, $\phi 11\text{Dut-Stl}^{\text{N-ter}}$. Para ello se hizo uso de la interferometría de biocapa (BLI), siendo el ligando inmovilizado al sensor la Dut mediante una etiqueta de histidinas situada en el extremo N-terminal, y el analito el represor $\text{Stl}^{\text{N-ter}}$. Al contrario que en los geles nativos, el ligando se mantuvo a una concentración molar fija (la Dut, 218,7 μM), y se midieron las afinidades en presencia de diluciones seriadas $\frac{1}{2}$ del analito ($\text{Stl}^{\text{N-ter}}$). Se incluyeron 7 concentraciones del analito que abarcaron ratios molares desde 1:1 a 1:64 $\text{Stl}^{\text{N-ter}} : \text{Dut}$. Los resultados, que se recogen en la Tabla R. 1. 1, se analizaron utilizando un modelo de interacción 1:1 considerando las 7 medidas de cada ensayo.

Tabla R. 1.1. Afinidades de unión entre $\text{Stl}^{\text{N-ter}}$ y Duts de *M. tuberculosis* (salvaje y mutante, mDut y mDut H145F, respectivamente), humana (hDut) y de fago $\phi 11$ de *S. aureus* ($\phi 11\text{Dut}$). Se incluye la constante de afinidad (K_D), la constante de asociación (K_{on}), la constante de disociación (K_{off}), y la correlación entre el ajuste y datos experimentales (R^2). A partir de un valor de 0,95 para R^2 se considera un buen ajuste. Tabla tomada de *Sanz-Frasquet et al. 2023*.

Dut in complex with $\text{Stl}^{\text{N-ter}}$	K_D (M)(10^{-9})	K_{on} ($\text{M}^{-1}\text{s}^{-1}$)(10^5)	K_{off} (s^{-1})(10^{-3})	R^2
mDut WT	39,25	2,56	10,00	0,990
mDut H145F	12,46	3,88	4,83	0,992
hDut WT	34,44	5,83	20,20	0,974
$\phi 11$ Dut WT	7,94	2,84	2,25	0,990

Estos ensayos de interacción mostraron una K_D de 34,44 y 39,25 nM para hDut y mDut, respectivamente, es decir, una afinidad casi idéntica de ambas Duts por $\text{Stl}^{\text{N-ter}}$. Ésta es solo 5 veces más baja que la mostrada por $\phi 11\text{Dut}$ (7,94 nM), pero muy similar a la mostrada por la Dut del fago $\phi 80\alpha$ (40 nM) (Alite et al. 2017), ambas dianas biológicas del represor. Es decir, hDut y mDut presentan una afinidad de interacción de la misma magnitud que las dianas naturales del represor. Previamente, ya se habían reportado diferencias en la misma dirección en afinidad de la $\phi 11\text{Dut}$ y hDut por el Stl completo con la técnica ITC: 0,23 and 0,10 μM para hDut and $\phi 11\text{Dut}$, respectivamente (Nyíri et al. 2018; Szabó et al. 2014). Estas diferencias de afinidad respecto al control $\phi 11\text{Dut-Stl}^{\text{N-ter}}$ son debidas principalmente a una mayor constante de disociación en el caso de mDut y hDut, llegando a una K_{off} entre 5 y 10 veces mayor, respectivamente (Tabla R.

1.1 y Figura R. 1.2). En el caso de la constante de asociación (K_{ON}), mDut presentaba un valor similar y la hDut incluso más elevado que la $\phi 11$ Dut (Tabla R. 1.1 y Figura R. 1.2). Por tanto, estos resultados apoyan que este fragmento N-terminal de Stl ha evolucionado para reconocer Duts triméricas.

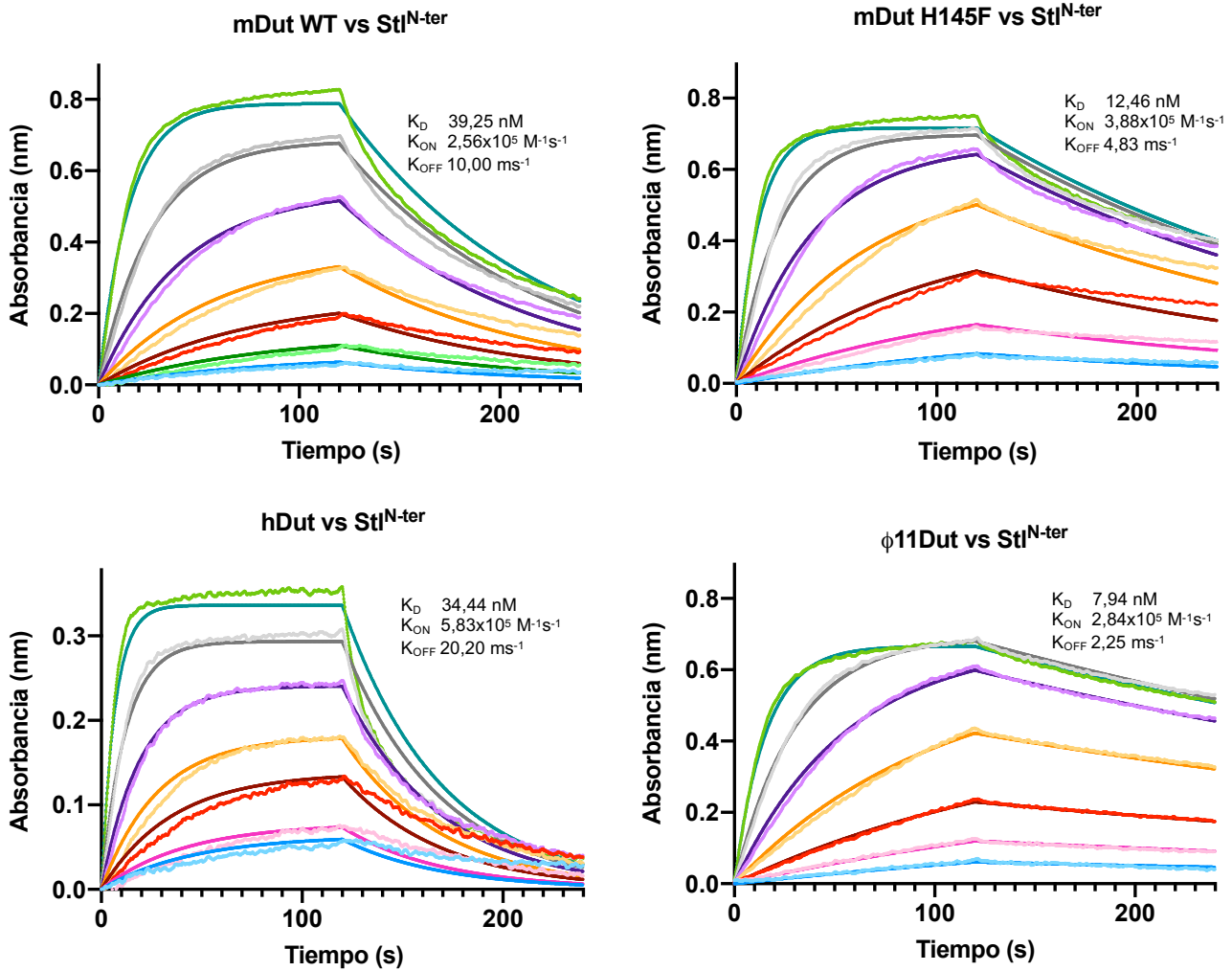


Figura R. 1. 2. Determinación de las constantes de afinidad en el equilibrio de las interacciones de Stl^{N-ter} con mDut, mDut H145F, hDut y $\phi 11$ Dut por interferometría de biocapa (BLI). Se muestran los sensogramas de los ensayos con 7 concentraciones crecientes de Stl^{N-ter} en los pasos de asociación y disociación. Los colores claros representan los datos experimentales y los colores oscuros representan el ajuste a un modelo de interacción 1:1. Se indican los valores de afinidad para el experimento representado. Figura tomada de *Sanz-Frasquet et al. 2023*.

La unión de Dut de fago al Stl interfiere con la capacidad del represor de reconocer a su ADN diana. Por ello, la interacción Dut-Stl desencadena la expresión de los genes de la SaPI reprimidos. Antes de ahondar en los ensayos estructurales con hDut y mDut, se llevaron a cabo ensayos EMSA con la proteína Stl completa (Stl^{FL}) con la finalidad de confirmar si la interacción con ambas Duts también provocaba la disrupción de la unión de Stl a su ADN.

La unión de Stl a su ADN está mediada por la fracción N-terminal del represor, pero para ello necesita estar formando un dímero. La dimerización está mediada por la fracción C-terminal, de modo que resulta obligatoria la utilización de la proteína completa. Junto al Stl y las Duts, se empleó el ADN diana del represor. Este ADN, marcado fluorescentemente, incluye la región del genoma de SaPI-bov1 13730 - 13949, el cual contiene la región intergénica *str* y *stl*, con dos sitios de unión del Stl para su regulación (Papp-Kádár et al. 2016).

Al igual que en los geles nativos, se realizaron ensayos de titración con la cantidad de ADN y la concentración de Stl^{FL} constantes, y ratios molares crecientes de Dut desde 1:0,5 a 1:32 (Stl^{FL}: Dut). Se realizaron tres réplicas y como control positivo se incluyó la ϕ 11Dut. El resultado fue muy similar independientemente de la Dut testada. El incremento de la ratio molar de la Dut coincide con la desaparición de las bandas correspondientes a los dos complejos Stl-ADN y con la aparición de la banda correspondiente al ADN libre (figura R. 1. 3). Del mismo modo que se había podido observar en los geles nativos, con una ratio molar 1:1 (Dut:Stl^{FL}) ya se provoca la liberación mayoritaria del ADN por parte del Stl, lo cual está de acuerdo con la relación equimolar Dut:Stl propuesta anteriormente.

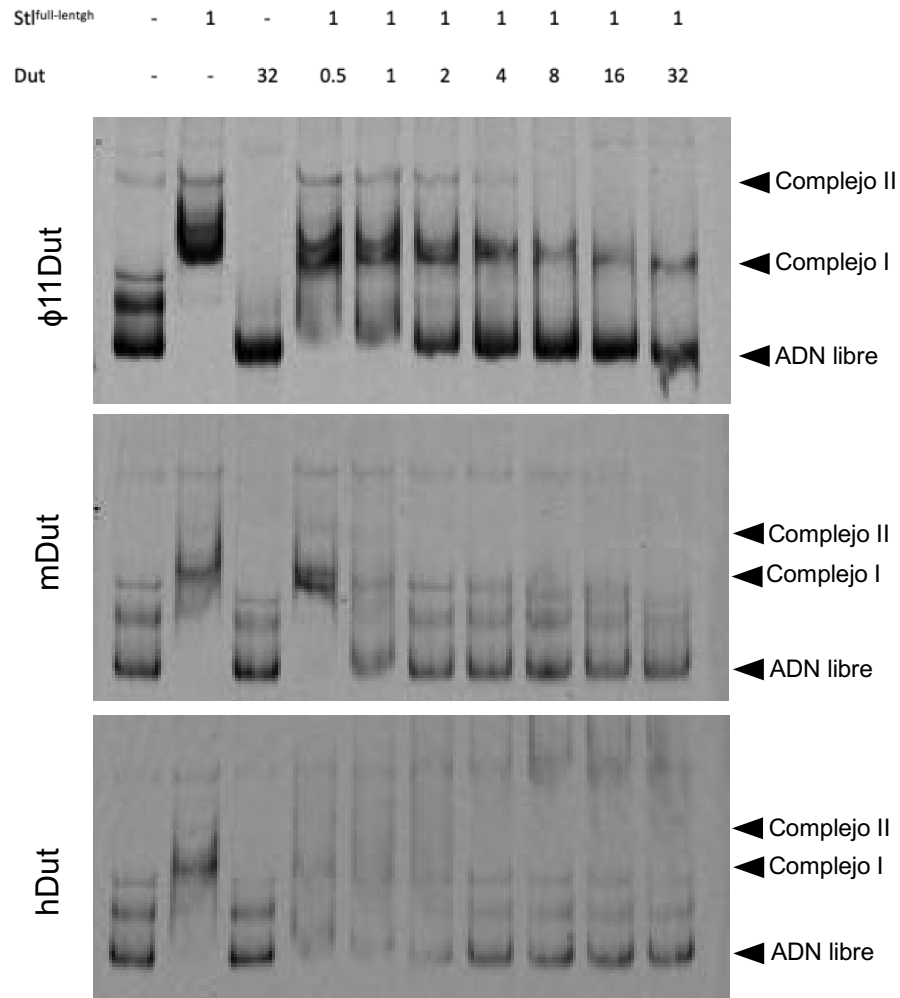


Figura R. 1. 3. Disrupción del complejo ADN-Stl^{full-length} en presencia de ϕ 11Dut (gel superior), mDut (gel intermedio) o hDut (gel inferior) por ensayos EMSA. Ratios molares crecientes de Dut fueron añadidos a una concentración constante de Stl^{FL} (4,8 μ M) y a 10 ng de ADN. En la parte superior se indica la ratio molar del represor y/o Dut presente en cada carrera. En la parte lateral derecha se indican las bandas correspondientes a los dos complejos Stl^{FL}- ADN, así como la banda de ADN libre. Figura tomada de *Sanz-Frasquet et al. 2023*.

1.2. Determinación de las estructuras tridimensionales de Stl^{N-ter} unido a mDut y a hDut.

Una vez confirmadas las altas afinidades de las Duts humana y de *M. tuberculosis* por Stl^{N-ter}, abordamos la resolución de las estructuras de sus respectivos complejos por difracción de rayos X.

1.2.1. Estructura tridimensional del complejo mDut-Stl^{N-ter}.

El complejo mDut-Stl^{N-ter} cristalizó a una concentración de 17 mg/mL en una solución al 20 % de PEG 10.000 y 100 mM de HEPES sódico a pH 7,5, con una mezcla de ambas proteínas en ratio molar 1:1 (calculada para el monómero). Los cristales se formaron a los 4 días de la puesta de las placas de cristalización, a una temperatura constante de 21 ° C.

Los cristales del complejo, con grupo espacial P2₁, fueron crioprotectados en una solución con la misma composición que la solución de cristalización suplementada con un 35 % de PEG 10.000 y 10 % de etilenglicol, y posteriormente congelados en nitrógeno líquido hasta su difracción. Ésta se llevó a cabo en el sincrotrón ALBA (Barcelona, España) en la línea XALOC. El juego de datos completo recogido permitió obtener la estructura a una resolución de 2,75 Å. La fase requerida para la resolución de la estructura tridimensional fue obtenida por reemplazo molecular utilizando como modelos de búsqueda las estructuras de los componentes individuales: Stl^{N-ter}, PDB 6H49 (Rafael Ciges-Tomas et al. 2019) y mDut, PDB 1MQ7 (Chan et al. 2004). Los parámetros cristalográficos se recogen en la Tabla R. 1. 2.

La estructura tridimensional del complejo muestra en su unidad asimétrica 2 trímeros de mDut y 6 monómeros de Stl^{N-ter}, formando dos complejos completos mDut-Stl^{N-ter} con una estequiometría 1:1 en la cual cada trímero de mDut interacciona con 3 monómeros independientes de Stl^{N-ter} (Figura R. 1. 4). Con el objetivo de conocer si había diferencia entre los monómeros, todos los monómeros de Dut y de Stl fueron superpuestos entre sí, y mostraron valores de RMSD inferiores a 0,8 Å para la superposición de los C α de los 130 a 136 residuos del monómero de Dut, y valores inferiores a 0,7 Å para la superposición de los C α de los 141 a 146 residuos del monómero de Stl^{N-ter} (Anexo III).

Tanto en el caso de la mDut como en el del Stl^{N-ter}, las estructuras tridimensionales en complejo son muy similares a las estructuras de las proteínas por separado. Es decir, la Dut mantiene su oligomerización trimérica formada mayoritariamente por hojas β , mientras que el fragmento del represor está formado totalmente por hélices α distribuidas en dos dominios (dominio de unión al ADN, hélices α 1 a α 4, dominio intermedio, hélices α 6 a α 10, y hélice conectora entre ambos, hélice α 5).

Tabla R. 1. 2 . Colección de datos cristalográficos y estadísticos de refinado de la estructura del complejo mDut-Sti^{N-ter}.

	mDUT-Sti ^{N-ter}
Datos de recogida	
Línea del haz	ALBA-XALOC
Longitud de onda (Å)	0,979
Grupo espacial	P2 ₁
Dimensiones de la celda (Å)	a=79,41 b=150,43 c=107,37 α=90 β=98,85 γ=90
Resolución (Å)	75,21 - 2,75 (2,85 - 2,75) ^a
Reflexiones totales	272538 (19164) ^a
Reflexiones únicas	61491 (6168) ^a
Compleitud (%)	95,04 (95,91) ^a
Multiplicidad	4,40 (4,20) ^a
Media I/σ (I)	8,30 (1,70) ^a
Rpim	0,05 (0,42) ^a
CC 1/2	0,98 (0,59) ^a
Refinado	
Factor R	0,200
Factor Rfree	0,250
Número de átomos	12305
Proteína	12105
Agua	108
Otras moléculas	16 ^b
Rmsd, enlaces (Å)	0,014
Rmsd, ángulos (°)	1,86
Probabilidad molecular de choque	5,67
Zonas de Ramachandran	
Preferidas (%)	94,70
Permitidas (%)	5,24
Atípicas (%)	0,06
Código PDB	7PWX

^aLos números entre paréntesis indican los valores para la capa de mayor resolución.

^bLos átomos corresponden con 8 moléculas de etilenglicol, 4 de dietilenglicol y 4 de Tris.

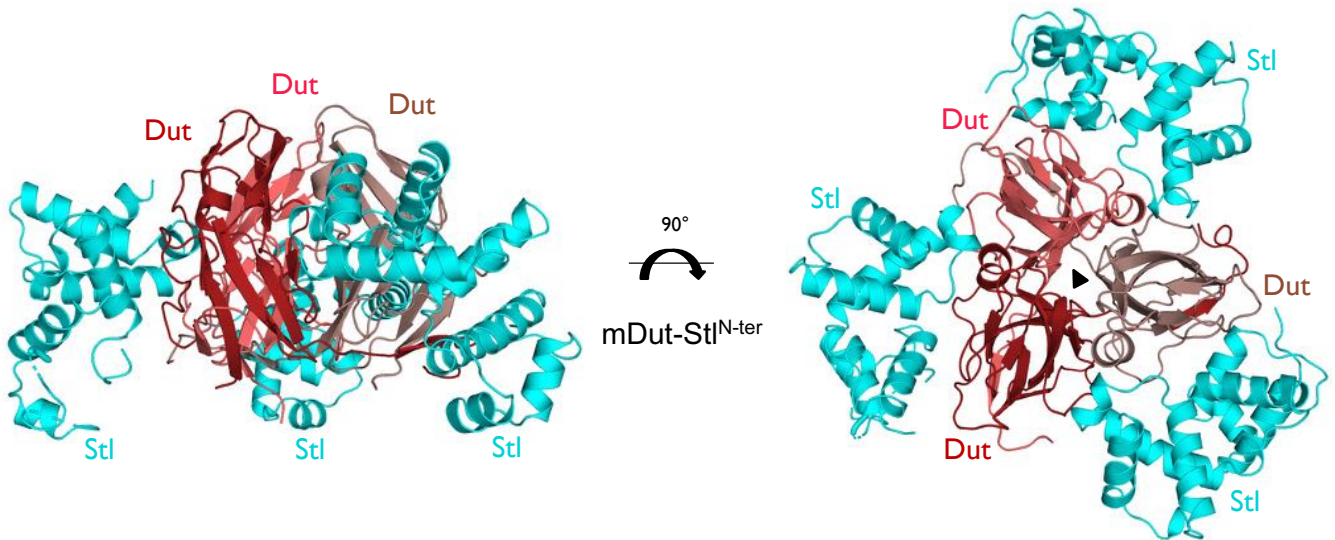


Figura R. 1. 4. Representación del complejo Dut de *M. tuberculosis* con el represor Stl^{N-ter}. Se muestran dos vistas ortogonales. Cada protómero de mDut está coloreado en tonos de rojo y los protómeros de Stl^{N-ter} en cian. Figura tomada de *Sanz-Frasquet et al. 2023*.

Los dos complejos independientes presentes en la unidad asimétrica son esencialmente idénticos. La superposición de los 642 átomos C α correspondientes a los trímeros de Dut y los tres monómeros de Stl^{N-ter} de cada complejo (Figura R. 1. 5), mostró tan solo 1,1 Å de desviación de la raíz cuadrada media (RMSD).

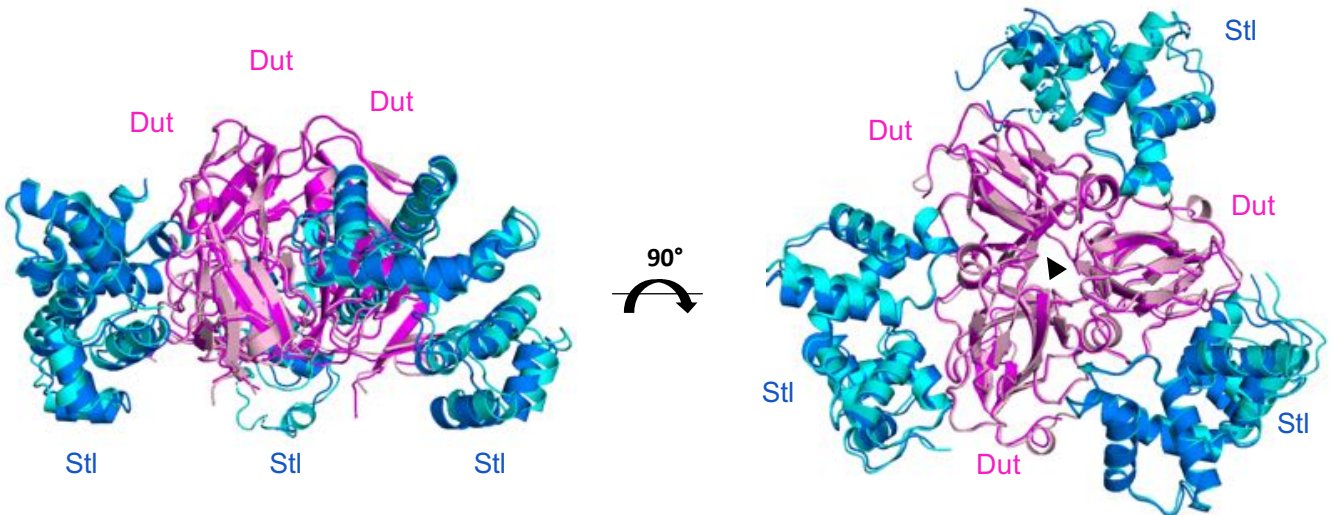


Figura R. 1. 5. Superposición de los dos complejos mDut-Stl^{N-ter} presentes en la unidad asimétrica de la celda unidad. Se muestran dos vistas ortogonales, uno de los complejos está coloreado en magenta (Dut) y cian (Stl^{N-ter}), y el otro en rosa claro (Dut) y azul marino (Stl^{N-ter}). Figura tomada de *Sanz-Frasquet et al. 2023*.

En ambos complejos mDut-Stl^{N-ter}, la densidad electrónica del mapa permitió trazar los protómeros de Dut casi enteros, con la excepción de 8 residuos en el extremo C-terminal que corresponden con el motivo V catalítico (*P-loop*), el cual cubre el centro activo de la enzima una vez el nucleótido ha sido posicionado. Esto es debido a que en el complejo, los centros activos de la Dut son ocupados por las moléculas de Stl^{N-ter}, de modo que este motivo es proyectado hacia el solvente, donde probablemente disposiciones alternativas en los distintos monómeros que impiden su visualización y trazado.

En cuanto a las moléculas de Stl^{N-ter}, los primeros 8 residuos del extremo N-terminal y los 3 últimos del C-terminal no son visibles, lo cual da soporte a una elevada flexibilidad derivada de la no participación en la interacción con la Dut. Además, dos de los 6 monómeros presentes en la estructura, uno de cada complejo, mostraron regiones del dominio DBD (residuos 33-53 y 59-62 en uno de los protómeros, y residuos 38-39 y 47-61 en el otro) cuya densidad electrónica era tan débil que imposibilitó el trazado de la cadena (Figura R. 1. 6). De nuevo, esto indica una mayor libertad de movimiento atribuible a una no participación en la interacción con la mDut.

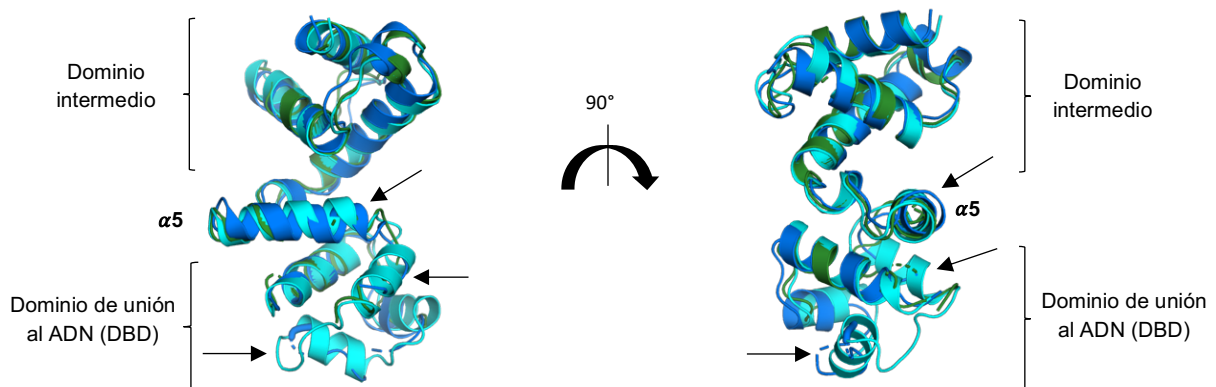


Figura R. 1. 6. Superposición de las estructuras de tres de los monómeros Stl^{N-ter} presentes en el complejo con mDut. Se muestran dos vistas ortogonales de la superposición de los dos monómeros de Stl^{N-ter} incompletos (cian y verde) sobre uno completo (azul marino). Las regiones de elevada flexibilidad que impidieron trazar la estructura por completo se indican en flechas. Figura tomada de *Sanz-Frasquet et al. 2023*.

1.2.2. Estructura tridimensional del complejo hDut-Stl^{N-ter}.

El complejo hDut-Stl^{N-ter} cristalizó a 3 mg/mL en una solución de 1.6 M (NH₄)₂SO₄ y 1 M Li₂SO₄, con una mezcla de ambas proteínas en ratio molar 1:1,25 calculada para el monómero (represor:Dut). Los cristales aparecieron a los 7 días de la puesta de placas de cristalización, a una temperatura constante de 21 ° C.

Los cristales del complejo, con grupo espacial P2₁2₁2₁, fueron crioprotegidos en una solución con la misma composición que la solución de cristalización suplementada con un 5 % de glicerol y 10 % de etilenglicol, y congelados en nitrógeno líquido hasta su difracción. Ésta se llevó a cabo en el sincrotrón DLS (Didcot, Reino Unido) en la línea I24. El juego de datos completo recogido permitió obtener la estructura a una resolución de 1,94 Å. La fase requerida para la resolución de la estructura tridimensional fue obtenida por reemplazo molecular utilizando como modelos de búsqueda las estructuras de los componentes individuales: Stl^{N-ter}, PDB 6H49 (Rafael Ciges-Tomas et al. 2019) y hDut, PDB 1Q5U (Mol et al. 1996). Los parámetros cristalográficos se recogen en la Tabla R. 1. 3.

La estructura tridimensional contiene en la unidad asimétrica un trímero de hDut y tres monómeros de Stl^{N-ter}, con un ensamblaje 1:1 idéntico al observado en el complejo mDut-Stl^{N-ter} y al publicado previamente en el complejo ϕ 11Dut-Stl^{N-ter} (Rafael Ciges-Tomas et al. 2019) (Figura R. 1. 7). De nuevo, con el objetivo de conocer si había diferencia entre los monómeros, los 3 monómeros de Duts y los 3 de Stl fueron superpuestos entre sí, y mostraron valores de RMSD inferiores e 0,4 Å para la superposición de los C α de los 134 a 140 residuos en el caso de Dut, y valores inferiores a 0,9 Å para la superposición de los C α de los 141 a 143 residuos en el caso de Stl (Anexo IV).

Los tres monómeros del represor presentes en este complejo pudieron ser trazados con excepción de los primeros 8-10 residuos del extremo N-terminal, porque del mismo modo que sucede con el complejo mDut-Stl^{N-ter}, no se observaba densidad electrónica.

Al igual que en la estructura del complejo mDut-Stl^{N-ter}, esta estructura ratifica la estequiometría equimolar Dut:Stl propuesta por los ensayos bioquímicos, además del mecanismo de reconocimiento de las Duts triméricas por parte del represor Stl observado previamente en el complejo ϕ 11Dut-Stl^{N-ter}, su diana biológica (Rafael Ciges-Tomas et al. 2019). La Dut mantiene su estado oligomérico trimérico y une tres monómeros del represor, cada uno de los cuales utiliza fundamentalmente su dominio intermedio para interactuar con un centro activo de la Dut.

Tabla R. 1. 3 . Colección de datos cristalográficos y estadísticos de refinado de la estructura del complejo hDut-Stl^{N-ter}.

hDUT-Stl^{N-ter}	
Datos de recogida	
Línea del haz	DLS-I24
Longitud de onda (Å)	0,992
Grupo espacial	P2 ₁ 2 ₁ 2 ₁
Dimensiones de la celda (Å)	a=77,72 b=81,80 c=198,40 α=β=γ=90
Resolución (Å)	63,11 - 1,94 (2,01 – 1,94) ^a
Reflexiones totales	701503 (33067) ^a
Reflexiones únicas	92354 (7967) ^a
Complejitud (%)	98,48 (86,22) ^a
Multiplicidad	12,90 (10,00) ^a
Media $I/\sigma(I)$	18,90 (2,50) ^a
Rpim	0,03 (0,24) ^a
CC 1/2	1,00 (0,90) ^a
Refinado	
Factor R	0,215
Factor Rfree	0,255
Número de átomos	7045
Proteína	6585
Agua	318
Otras moléculas	31 ^b
Rmsd, enlaces (Å)	0,015
Rmsd, ángulos (°)	1,92
Probabilidad molecular de choque	5,11
Zonas de Ramachandran	
Preferidas (%)	96,89
Permitidas (%)	2,99
Atípicas (%)	0,12
Código PDB	7PWJ

^aLos números entre paréntesis indican los valores para la capa de mayor resolución.

^bLos átomos corresponden con 19 moléculas de etilenglicol, 6 de sulfato y 6 de glicerol.

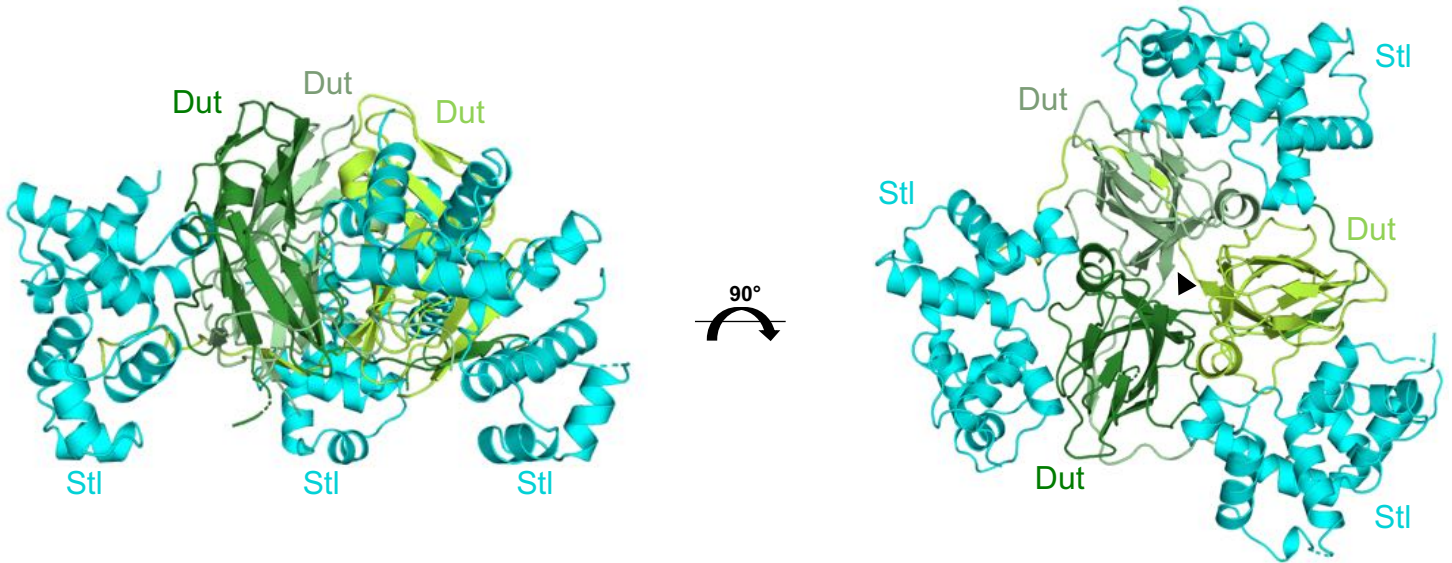


Figura R. 1. 7. Representación del complejo de Dut humana con el represor $\text{Stl}^{\text{N-ter}}$. Dos vistas ortogonales son mostradas, cada protómero de hDut está coloreado en tonos de verde y los protómeros de $\text{Stl}^{\text{N-ter}}$ en cian. Figura tomada de *Sanz-Frasquet et al. 2023*.

Sorprendentemente, en el caso de este complejo, tanto el cuerpo central del trímero de la hDut como el motivo V pudieron ser trazados. Este *P-loop*, incapaz de posicionarse sobre el centro activo que está ocupado por el represor, se proyecta hacia el Stl e introduce el fragmento C-terminal entre el dominio intermedio y el dominio DBD del represor.

Por tanto, esta estructura permitió ampliar la visión molecular respecto a la contribución del motivo catalítico V en la interacción con Stl, el cual había sido propuesto previamente mediante ensayos bioquímicos como participante en la interacción con Duts de fago de *S. aureus* (Tormo-Más et al. 2013; Rafael Ciges-Tomas et al. 2019; Maiques et al. 2016), y recientemente observado en la estructura de la Dut del crustáceo *Litopenaeus vannamei* en complejo con el represor (lvDut- $\text{Stl}^{\text{N-ter}}$) (Wang et al. 2021).

1.3. La diana de StI es el trímero de Dut.

La unión de StI^{N-ter} al trímero de ϕ 11Dut no afecta a la oligomerización trimérica de ésta última (Rafael Ciges-Tomas et al. 2019). Dado que tampoco es alterada en el caso del complejo con mDut o hDut, se planteó la posibilidad de que el trímero fuese la diana del represor.

Con el objetivo de comprobar esta hipótesis, se compararon por superposición con el programa Gesamt (Krissinel 2012) las estructuras de ambas Duts en complejo con StI^{N-ter} con las respectivas estructuras de Dut en conformación *apo* y unida al nucleótido sustrato (Figura R. 1. 8). La comparación estructural mostró valores de RMSD inferiores a 0,8 Å (superposición de los C α de 386 a 390 residuos) y a 0,5 Å (superposición de los C α de 384 residuos) para mDut y hDut, respectivamente. Es decir, la Dut presenta una estructura tridimensional muy similar en conformación *apo* y en complejo con el sustrato o con el represor. Por tanto, la diana del StI es el trímero de Dut.

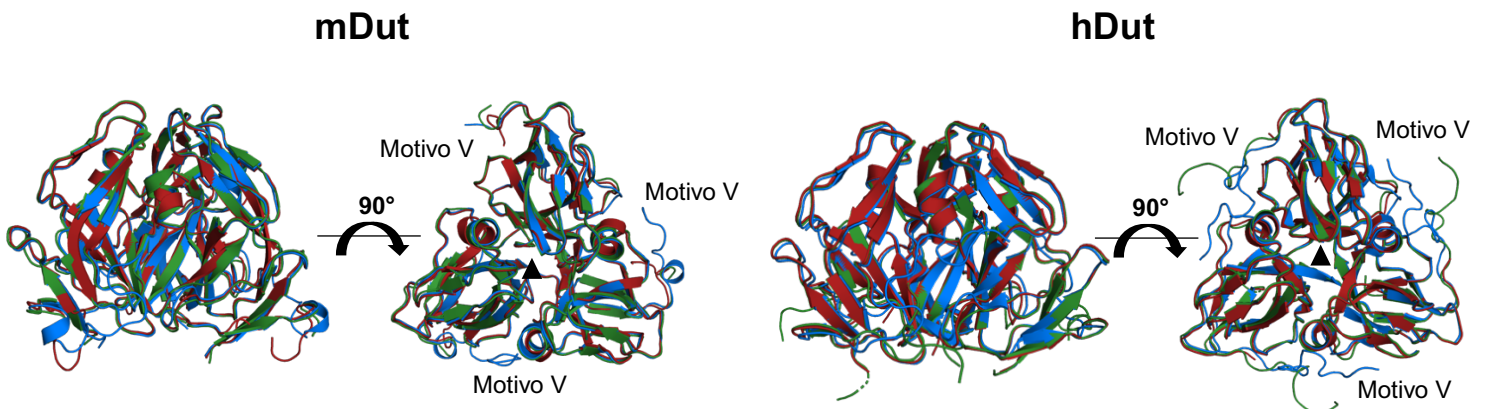


Figura R. 1. 8. Las Duts triméricas presentan una conformación competente para la unión de StI^{N-ter}. Se muestran dos vistas ortogonales de las estructuras superpuestas de mDut (izquierda) y hDut (derecha) en su conformación *apo* (rojo), unida a dUTP (azul) y en complejo con StI^{N-ter} (verde). El motivo V está desordenado en la conformación *apo* mientras que se ordena sobre el centro activo en la conformación unida al sustrato. Figura tomada de Sanz-Frasquet et al. 2023.

Además, la superposición del monómero de mDut en las distintas conformaciones (*apo*, en complejo con el sustrato, o en complejo con el represor), mostró una ligera mayor similitud entre la estructura en complejo con el sustrato y en complejo con el represor que con la conformación *apo*, con valores de RMSD de 0,7 y 0,9 Å para la superposición de los C α de 128 a 130 residuos, respectivamente). Las diferencias están principalmente concentradas en la única hélice α corta del motivo catalítico III de mDut (residuos 65 a 70). En el caso de la conformación *apo*, la hélice α está distorsionada, alterando la

topología respecto a la presente en las estructuras de mDut unida al nucleótido o unida al represor, en las cuáles es casi idéntica en ambos casos (Figura R. 1. 9). Esta hélice α está compuesta por los residuos 65 a 70, y la distorsión de la topología se produce en dos sentidos. Es decir, entre los residuos de S65 a A68 la hélice α se expande, mientras que de los residuos A68 a R70 se estrecha, comparando la hélice α en la estructura en *apo* respecto a la estructura unida a Stl o unida al dUTP. Los C α de un mismo residuo llegan a estar desplazados entre 2,8 y 4 Å.

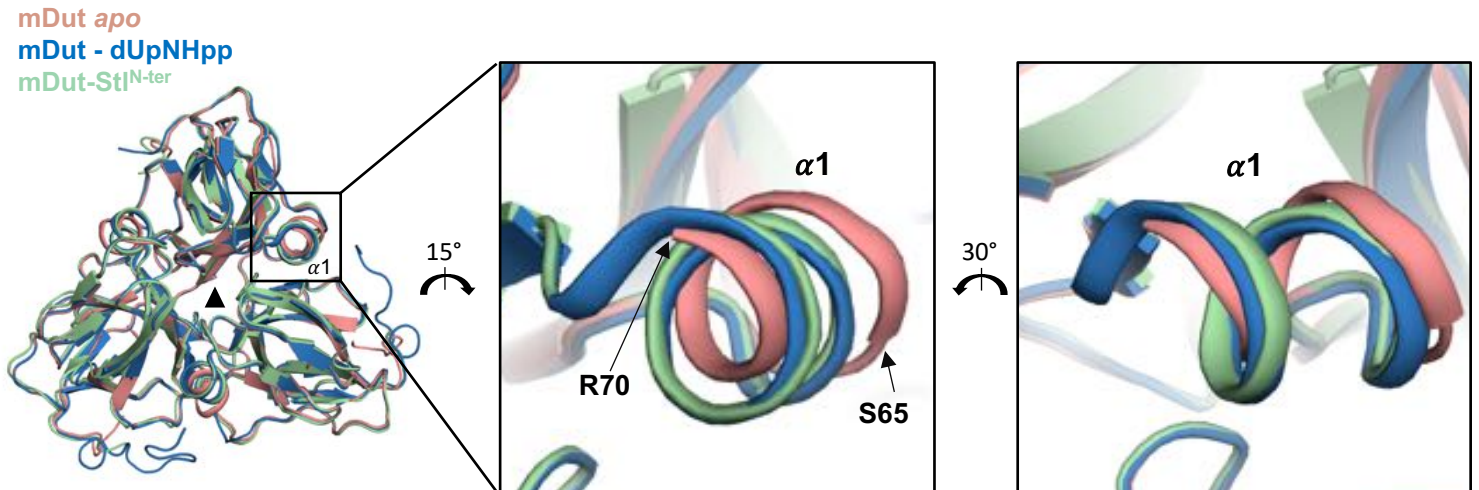


Figura R. 1. 9. La estructura del trímero de mDut en complejo con el Stl es más similar a la forma unida al sustrato que a la conformación *apo*. Vista detallada de la superposición de la hélice α de mDut (residuos 65 a 70) en las estructuras en conformación Apo (salmón), unida al nucleótido (azul) y en complejo con Stl (verde). Hay una pérdida de la topología de hélice en la conformación *apo* que se mantiene en el caso de la mDut en complejo con el represor o el sustrato. Figura tomada de *Sanz-Frasquet et al. 2023*.

Por tanto, Stl reconoce el centro activo de la Dut en su conformación competente para la unión del sustrato, lo cual concuerda con la capacidad inhibitoria del represor sobre la actividad catalítica de la enzima (Bowring et al. 2017; Hirmondó et al. 2015), derivada del mecanismo de interacción.

Una vez confirmado que el Stl reconoce a la Dut en la conformación de unión al sustrato, se analizó si el Stl realiza cambios importantes para unirse a las Duts. Con esta finalidad, se superpusieron los monómeros de Stl^{N-ter} en conformación *apo* y unido a ϕ 11Dut, mDut y hDut (Tabla R. 1. 4, Figura R. 1. 10). Las superposiciones de los C α mostraron diferencias con valores de RMSDs superiores a 1,6 Å. Es decir, el fragmento N-terminal del monómero Stl presenta la flexibilidad necesaria para adquirir una conformación competente para reconocer los centros activos de las Duts triméricas, aunque la transición de la conformación *apo* a la unida a Dut no incluya cambios drásticos (Figura R. 1.10).

Los centros activos de las Duts triméricas se conservan en un elevado grado desde los fagos hasta los humanos. No obstante, se analizaron los cambios conformacionales inducidos en el Stl^{N-ter} debido al reconocimiento y unión a las Duts de humano y de *M. tuberculosis*, respecto a los cambios inducidos por la unión a la ϕ 11Dut, una de sus dianas naturales. Con la excepción de los dos monómeros de Stl en complejo con mDut que presentan una elevada flexibilidad en el dominio DBD que impidieron su trazado completo (Anexo V, monómeros D y F'), el resto de los monómeros de Stl mostraron una conformación muy similar independientemente de si estaban unidos a una Dut fágica, bacteriana o humana (Figura R. 1. 11), con diferencias de RMSDs en la superposición entre 0,3 y 1,1 Å (superposición de los C α de 141-147 residuos) (Anexo V). En algunos casos hay diferencias más grandes entre protómeros de Stl^{N-ter} pertenecientes al mismo complejo (monómeros F y E de hDut) que entre complejos de diferentes especies (monómeros G de ϕ 11Dut y F de mDut) (Anexo V).

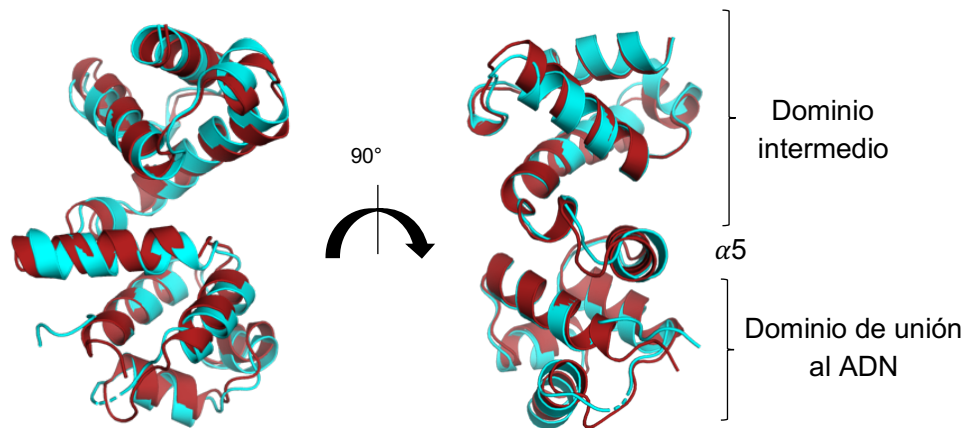


Figura R. 1. 10. Superposición de las estructuras de Stl^{N-ter} en su conformación libre (*apo*) y unido a mDut (rojo y cian, respectivamente). Se muestran dos vistas ortogonales. Figura tomada de *Sanz-Frasquet et al. 2023*.

Tabla R. 1. 4. Tabla de valores RMSD en Å derivados de la superposición de los monómeros de Stl^{N-ter} unidos a ϕ 11Dut, mDut y hDut con el monómero Stl^{N-ter} en conformación *apo*. Entre paréntesis se indica el número de residuos superpuestos.

Stl ^{N-ter} en complejo con...		Stl ^{N-ter} <i>apo</i>
mDut	Monómero D	1.835 (123)
	Monómero E	1.002 (146)
	Monómero F	1.344 (146)
	Monómero D'	0.995 (146)
	Monómero E'	1.139 (145)
	Monómero F'	1.355 (114)
hDut	Monómero D	1.275 (142)
	Monómero E	1.459 (142)
	Monómero F	1.333 (141)
ϕ 11Dut	Monómero D	1.453 (145)
	Monómero E	1.65 (146)
	Monómero F	1.461 (146)
	Monómero G	1.426 (144)

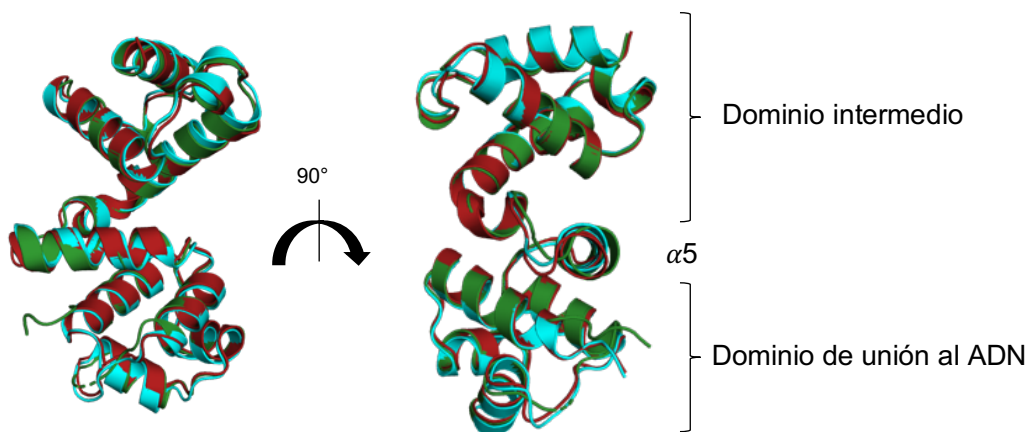
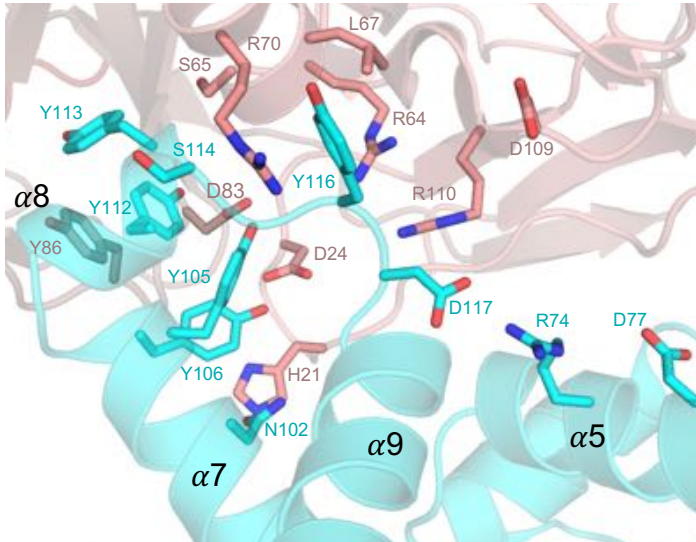


Figura R. 1. 11. Superposición de las estructuras de Stl^{N-ter} unidas a mDut, hDut y ϕ 11Dut (rojo, verde y cian, respectivamente). Se muestran dos vistas ortogonales. Figura tomada de Sanz-Frasquet *et al.* 2023.

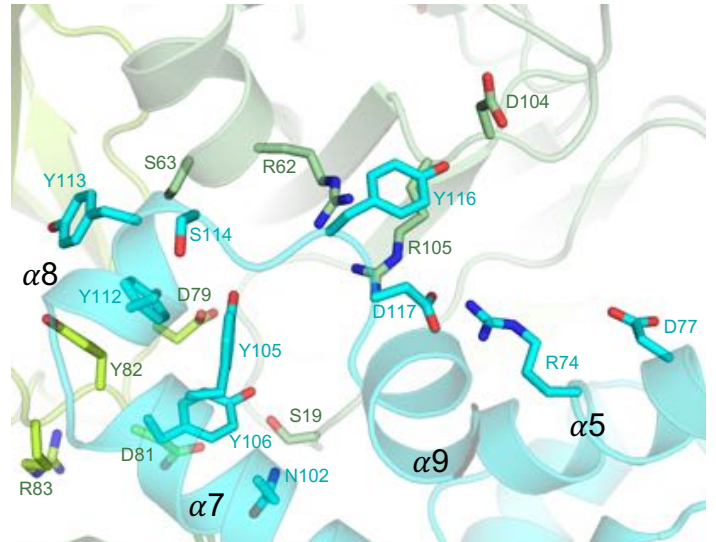
1.4. Stl mimetiza al nucleótido para interactuar con las Duts triméricas.

El análisis de las estructuras de los complejos mDut-Stl^{N-ter} y hDut-Stl^{N-ter} muestra que Stl utiliza una estrategia similar para interactuar con las Duts triméricas eucariotas y procariotas. En ambos complejos, el mecanismo principal de interacción del represor con el trímero de Dut es la inserción del dominio intermedio en el centro activo de la Dut, siendo los elementos estructurales principales en la interacción la hélice $\alpha 8$ y los *loops* de unión de esta hélice con las hélices $\alpha 7$ y $\alpha 9$. Este mecanismo de unión es idéntico al observado en el complejo de Stl con $\phi 11$ Dut, una de sus dianas naturales (figura R. 1. 12) (Rafael Ciges-Tomas et al. 2019). Específicamente, los residuos Y112 y Y113 de la hélice $\alpha 8$ del represor interactúan con los residuos conservados tirosina y aspartato del motivo III de las Duts triméricas (D83 y Y86 en mDut, D79 y Y82 en hDut, D81 y Y84 en $\phi 11$ Dut), situados en el *loop* de conexión entre las hojas $\beta 7$ y $\beta 8$. Ambos residuos de tirosina del Stl (Y112 y Y113) mimetizan al sustrato de la Dut ocupando el mismo sitio que ocuparía la ribosa y el anillo de pirimidina del nucleótido, y el agua catalítica (figura R. 1. 13). Además, el residuo de Stl Y113 junto con la Y105 y la Y116 también mimetizan el sustrato al interactuar con los residuos catalíticos conservados de las Duts que interactúan con los fosfatos del nucleótido, como la arginina y la serina del motivo II (R64 y S65 en el caso de mDut, R62 y S63 en hDut y R64 y S65 en $\phi 11$ Dut) y la tirosina del motivo III (Y86, Y82 y Y84 en el caso de mDut, hDut y $\phi 11$ Dut, respectivamente) (Figura R. 1.12, R. 1. 13, R. 1. 14, Anexos VI y VII).

A



B



C

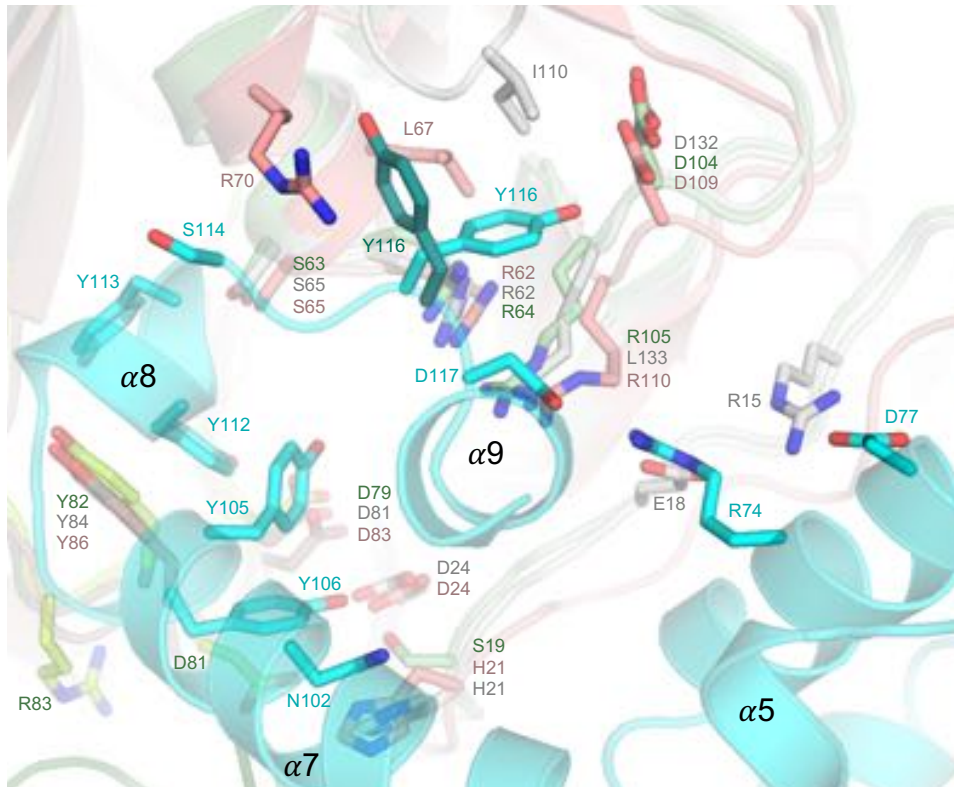


Figura R. 1. 12. El represor StI mimetiza las interacciones con el sustrato para interactuar con las Duts triméricas. Vista detallada de un centro activo de la mDut (A) y uno de hDut (B) en complejo con un monómero de StI^{N-ter}. mDut está coloreada en tonos rojos (A), hDut en tonos verdes (B), y StI^{N-ter} en cian. Las cadenas laterales de los residuos implicados en la interacción se han etiquetado y coloreado en función del tipo de átomo (los carbonos mantienen el mismo color que las cadenas principales correspondientes, O en rojo, N en azul). (C) Vista detallada de la superposición de los centros activos de mDut, hDut y ϕ 11Dut en complejo con StI^{N-ter} (rojo, verde, gris y cian, respectivamente). Figura tomada de Sanz-Frasquet et al. 2023.

En el mecanismo de mimetización del nucleótido llevado a cabo por el represor, la hélice $\alpha 8$ representa el principal punto de anclaje a la Dut, el cual queda reforzado con interacciones adicionales con los residuos conservados de las Duts. Estos son los casos de los residuos de Stl Y106 (hélice $\alpha 7$), que interacciona con los residuos conservados o parcialmente conservados de los motivos I y III (H21, D24, R87 y G88 en el caso de mDut; D81, R83 y G84 en el caso de hDut), o los residuos del represor S114 y D117 (hélices $\alpha 8$ y $\alpha 9$) que interaccionan con residuos cargados positivamente (arginina o lisina) localizados en los motivos II y IV de las Duts (R70 y R110 en el caso de mDut, K68 y R105 en el caso de hDut) (Figura R. 1.12, R. 1. 13, R. 1. 14, Anexos VI y VII).

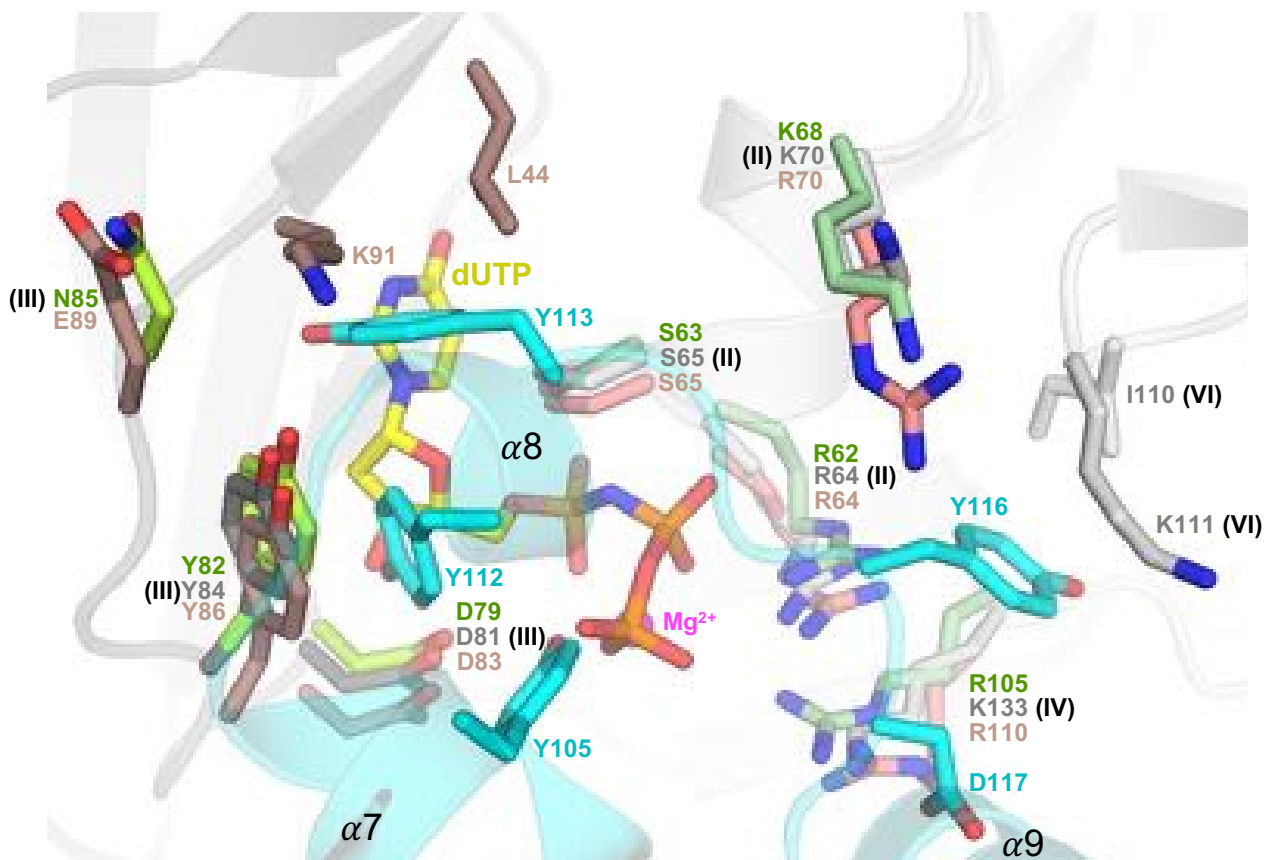


Figura R. 1. 13. El represor Stl mimetiza las interacciones del sustrato dUTP con las Duts triméricas. Vista detallada de la región de interacción Dut-Stl de los tres complejos superpuestos: mDut en marrón-salmón, hDut en verdes, $\phi 11$ Dut en grises y Stl^{N-ter} en cian. Una molécula del análogo de sustrato (representado en varillas y los átomos de C en amarillo) fue posicionado en el centro activo mediante la superposición de la estructura de $\phi 11$ Dut con el sustrato no hidrolizable α, β -imino dUTP (PDB 4GV8) sobre el complejo $\phi 11$ Dut-Stl^{N-ter}. El ion quelado de Mg²⁺ se representa en una esfera color magenta. Las cadenas laterales de los residuos participantes en la interacción se muestran en varillas etiquetadas, y en color según átomo (C según molécula, O en rojo, N en azul y P en naranja). Figura tomada de *Sanz-Frasquet et al. 2023*.

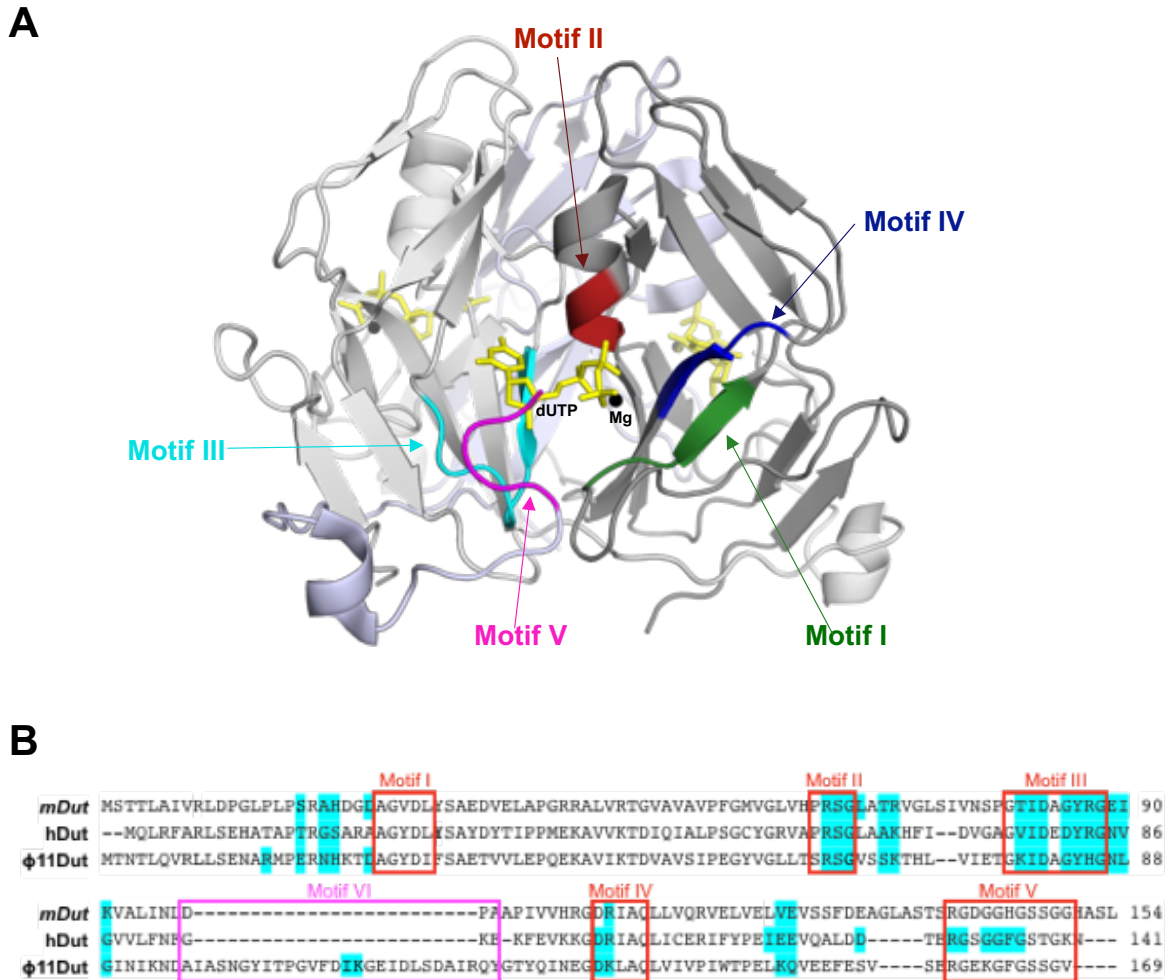


Figura R. 1. 14. Motivos catalíticos de las Duts triméricas. (A) Los cinco motivos catalíticos de las Duts triméricas se destacan en uno de los tres centros activos con distintos colores en la estructura del complejo mDut-dUpNHpp. Cada monómero está coloreado en un tono de gris. Las moléculas de sustrato se representan en varillas (amarillo) y los iones Mg^{2+} en esferas (negro). (B) Alineamiento de mDut, hDut y ϕ 11Dut. Los residuos de cada Dut que interactúan con Stl^{N-ter} están resaltados en cian, los motivos catalíticos I a V están encuadrados en cajas rojas, y el motivo VI, específico de Duts de fagos, en una caja magenta. Figura tomada de *Sanz-Frasquet et al. 2023*.

Un trabajo previo del laboratorio con las Duts de fago de *S. aureus* ϕ 80 α y ϕ 11 había mostrado que la variabilidad en el motivo IV de ambas Duts tenía un fuerte efecto en la afinidad por Stl y la inducción de la SaPIbov1 (Maiques et al. 2016). Este motivo, localizado en la convergencia de los motivos V y VI, está formado por 5 residuos, y mostraba variabilidad en tres de ellos, los cuales fueron intercambiados entre ambas proteínas (E133, R134 y I135 en el caso de ϕ 80 α Dut, y D132, K133 y L134 en el caso de ϕ 11Dut). En el caso de hDut y mDut, este motivo es idéntico al de la Dut de ϕ 80 α , con excepción del primer

residuo (glutámico en el caso de $\phi 80\alpha$ Dut, y aspártico en el caso de mDut y hDut). De los tres residuos que mostraron variabilidad con la $\phi 11$ Dut, el único con carga positiva es la arginina (R134, R110 y R105 en $\phi 80\alpha$ Dut, mDut y hDut, respectivamente) o lisina en el caso de la $\phi 11$ Dut (K133). En las estructuras de mDut y hDut en complejo con el represor, este residuo básico de la Dut es el que mayor participación de los 3 tiene, interaccionando con 2 residuos de StI. Por una parte, establece un puente salino con el aspartato 117 (D117), y por otra, establece interacciones hidrofóbicas y polares con el residuo Y116.

No obstante, las interacciones con el residuo Y116 del represor difieren según la Dut reconocida. En los complejos $\phi 11$ Dut-StI^{N-ter} y $\phi 80\alpha$ Dut-StI^{N-ter}, el residuo Y116 del represor se inserta completamente en el cuerpo de la Dut e interacciona con residuos de los motivos catalíticos II (R64 y R64 en $\phi 11$ Dut y $\phi 80\alpha$ Dut respectivamente), IV (K133 y R134), y con el motivo VI, específico de fago (I110 e I109) (Rafael Ciges-Tomas et al. 2019; Alite Hernández 2021). En el caso del complejo del represor con hDut, la conformación de la cadena lateral del residuo Y116 es idéntica a la que presenta en complejo con $\phi 11$ Dut y $\phi 80\alpha$ Dut, aunque hDut no posee el motivo VI, e interacciona con la R105 (motivo IV). En contraste, en el caso de la Dut de *M. tuberculosis* sí explota la ausencia de este motivo VI específico, ya que la cadena lateral del residuo Y116 presenta, en las diferentes subunidades de StI, conformaciones rotaméricas alternativas a la observada en el caso de los complejos $\phi 11$ Dut-StI^{N-ter}, $\phi 80\alpha$ Dut-StI^{N-ter} y hDut-StI^{N-ter}. Cuando asume la conformación adoptada en complejo con las Duts de fagos y humana, interacciona con los residuos D109 y R110 de la Dut, siendo la interacción mediante la R110 la conservada con las otras Duts (R105, K133 y R134 en hDut, $\phi 11$ Dut y $\phi 80\alpha$ Dut, respectivamente). Sin embargo, cuando la Y116 de StI adopta otras conformaciones rotaméricas pierde las interacciones a través de estos residuos, pero establece nuevas interacciones mediante los residuos del motivo conservado II L67 y R70 (figura R. 1. 12). Hay que destacar que no solo se observan las dos conformaciones “opuestas” del residuo Y116, sino que los distintos protómeros de StI presentan conformaciones rotaméricas intermedias, cuyas interacciones con la mDut se establecen con los residuos D109 y R110 o con L67 y R70 según la proximidad a la cadena lateral del residuo Y116 de StI.

Por tanto, las diferencias vistas en afinidad entre estas Duts por el represor StI (7,94 nM en el caso de $\phi 11$ Dut y alrededor de 40 nM en $\phi 80\alpha$ Dut, mDut y hDut) podrían explicarse con la versatilidad del StI para mimetizar las interacciones de Dut con el nucleótido, junto a las interacciones periféricas proporcionadas por los residuos de StI del dominio de unión al ADN (DBD) y la hélice $\alpha 5$, las cuales difieren en función de la Dut con la que forma el complejo (Anexos VI y VII para los complejos mDut-StI^{N-ter} y hDut-StI^{N-ter} respectivamente, las estructuras resueltas en esta Tesis).

1.5. El motivo V de las Duts triméricas es un elemento activo en la unión con StI.

La participación del motivo V de las Duts triméricas en el reconocimiento e interacción con el represor StI ha sido controvertida (M. Ángeles Tormo-Más et al. 2013; Maiques et al. 2016; Takács et al. 2010). Un trabajo previo realizado en el laboratorio con Duts triméricas de fagos de *S. aureus*, mostró que el nivel de participación del motivo V variaba entre las Duts (Maiques et al. 2016). Aunque se describió por primera vez en el laboratorio la estructura del complejo de una Dut trimérica (la $\phi 11$ Dut) con el represor (StI^{N-ter}), no pudo visualizarse el extremo C-terminal de la Dut, donde se sitúa el motivo V (Rafael Ciges-Tomas et al. 2019). No obstante, el complejo lvDut-StI^{N-ter} fue posteriormente descrito y aportó una primera visión de como este extremo C-terminal de la Dut reconocería al represor StI (Wang et al. 2021).

La estructura del complejo hDut-StI^{N-ter} mostró un mecanismo de reconocimiento del represor y de estabilización del complejo similar al observado en el complejo lvDut-StI^{N-ter}. El residuo conservado fenilalanina del motivo V de Dut se introduce en un bolsillo hidrofóbico formado por las hélices $\alpha 5$, $\alpha 6$ y $\alpha 7$ del StI. Además del residuo de fenilalanina, las cadenas principales o laterales de los residuos del motivo V de la hDut (D127, R130, G133, G134 y G136) establecen interacciones con los residuos del *loop* entre las hélices $\alpha 4$ - $\alpha 5$ (G66, I67, P68) y la hélice $\alpha 7$ (Y98, S99, N102, K103 y N107) (Figura R.1.15 y Anexo VII). Para unirse a StI, el extremo C-terminal de la Dut se dispone de una manera totalmente distinta a la observada para el reconocimiento del nucleótido. Este cambio drástico en la dirección del motivo V es permitido por la flexibilidad que le aportan los múltiples residuos de glicina presentes en este motivo, así como la labor del residuo R130, situado al principio de la cola C-terminal, que actúa a modo de bisagra e interacciona con el residuo N102 del represor. Por tanto, Dut explota residuos del motivo V para el reconocimiento de StI anclando la cola C-terminal por la fenilalanina, proporcionándole flexibilidad por los residuos de glicina, y actuando la arginina como una bisagra (figura R. 1. 15).

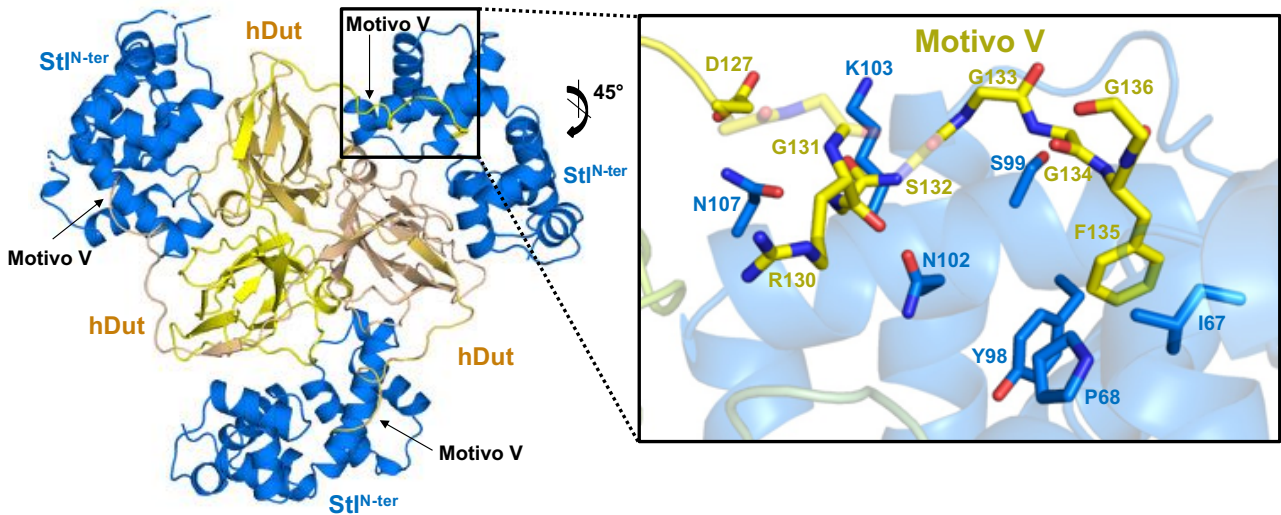


Figura R. 1. 15. El motivo V de hDut está implicado en el reconocimiento y unión de Stl. *Izquierda.* Representación del complejo hDut-Stl^{N-ter} (tonos amarillos y azul, respectivamente) con la inserción del motivo V de la Dut en Stl para asegurar la formación del complejo. *Derecha.* Vista detallada de la interacción del motivo V de hDut con Stl^{N-ter}. Las cadenas laterales de los residuos participantes se muestran en varillas y etiquetados (colores según átomo: C en amarillo o azul, O en rojo, N en azul). Figura tomada de *Sanz-Frasquet et al. 2023*.

En el caso de la estructura de mDut-Stl^{N-ter} no se pudo atribuir densidad electrónica a la porción C-terminal de la Dut. La principal diferencia entre mDut y hDut en este dominio corresponde al punto de anclaje, justo la fenilalanina presente en hDut (y muchas otras Duts) es sustituida por una histidina (H145) en mDut. Se trata de la sustitución de un residuo hidrofóbico por uno polar, por lo que su acomodación en el bolsillo hidrofóbico de Stl es (o sería) energéticamente desfavorable. Esto podría suponer la eliminación por parte de mDut del principal punto de anclaje de su porción C-terminal a Stl, impidiendo o dificultando la participación de este elemento estructural en la unión al represor. Con el objetivo de confirmar esta hipótesis se generó un mutante puntual de mDut en el que se sustituyó la histidina 145 por una fenilalanina (mDut^{H145F}) y se midió su afinidad por el represor. Este residuo en las Duts monoméricas y triméricas es un aminoácido aromático, principalmente una fenilalanina, la cual se apila sobre el anillo de uracilo del sustrato. En estudios enzimáticos previos había sido mutado esta histidina de mDut por triptófano, un residuo aromático no polar, sin mostrar alteraciones en la actividad enzimática del enzima (Hirmondó et al. 2015).

La afinidad de mDut^{H145F} por Stl^{N-ter} fue medida por la técnica BLI, mostrando un valor de K_D de 12,45 nM. Dado que la afinidad de la mDut salvaje por Stl^{N-ter}

medida por esta misma técnica es de 39,25 nM, la sustitución por un residuo hidrofóbico en esta posición supone un incremento de tres veces en la afinidad (tabla R. 1. 1 y figura R. 1. 2). Por tanto, esta ganancia de función apoya la participación del motivo V en la unión de la Dut al StI, y la labor de anclaje que realiza el residuo hidrofóbico de este motivo. Con estos resultados podemos concluir, finalmente, la controversia sobre la participación del motivo V en el reconocimiento y unión al StI.

El aumento de la afinidad de este mutante de mDut por StI podría permitir la visualización del motivo V en la estructura tridimensional del complejo. Con esta finalidad se realizaron los primeros intentos de cristalización de este mutante con StI^{N-ter}, pero los cristales difractados hasta el momento no poseen la suficiente resolución, y por falta de tiempo no se ha podido completar este ensayo estructural.

1.6. El represor StI tiene capacidad universal para unirse a las Duts.

De las tres familias de Duts que existen, las monoméricas se originaron a partir de la duplicación y fusión génica en tándem de las triméricas, de modo que ambas exhiben una elevada conservación de los centros activos y del mecanismo catalítico (Tarbouriech et al. 2005; Baldo and McClure 1999). Debido a esta elevada similitud, se planteó la hipótesis de que StI fuera también capaz de reconocer a las Duts monoméricas, convirtiéndose así en un inhibidor universal de Duts.

Con el objetivo de comprobar dicha hipótesis, se llevó a cabo una aproximación *in-silico* para producir un modelo estructural del complejo Dut monomérica con StI^{N-ter}, utilizando para ello la información estructural propia y también la disponible en el Protein Data Bank. Como Dut monomérica representativa se utilizó la del virus de Epstein-Barr, haciendo uso de su estructura en complejo con un análogo de sustrato no hidrolizable del dUTP (EBVDut, PDB 2BT1 (Tarbouriech et al. 2005)). Como Duts triméricas se utilizaron las de *M. tuberculosis* (PDB 1SIX (Chan et al. 2004)) y la humana (PDB 2HQU (Varga et al. 2007)), ya que son representativas de Duts procarionota y eucariota, y se disponía tanto de las estructuras del complejo con un análogo de sustrato no hidrolizable del dUTP, como de las estructuras de sus complejos con el represor (resueltos en esta Tesis). El primer paso para generar el modelo fue la superposición de los nucleótidos de las correspondientes estructuras de Duts trimérica y monomérica, la cual confirmó la relación estructural de los centros catalíticos de ambas familias de enzimas (Figura R. 1. 16).

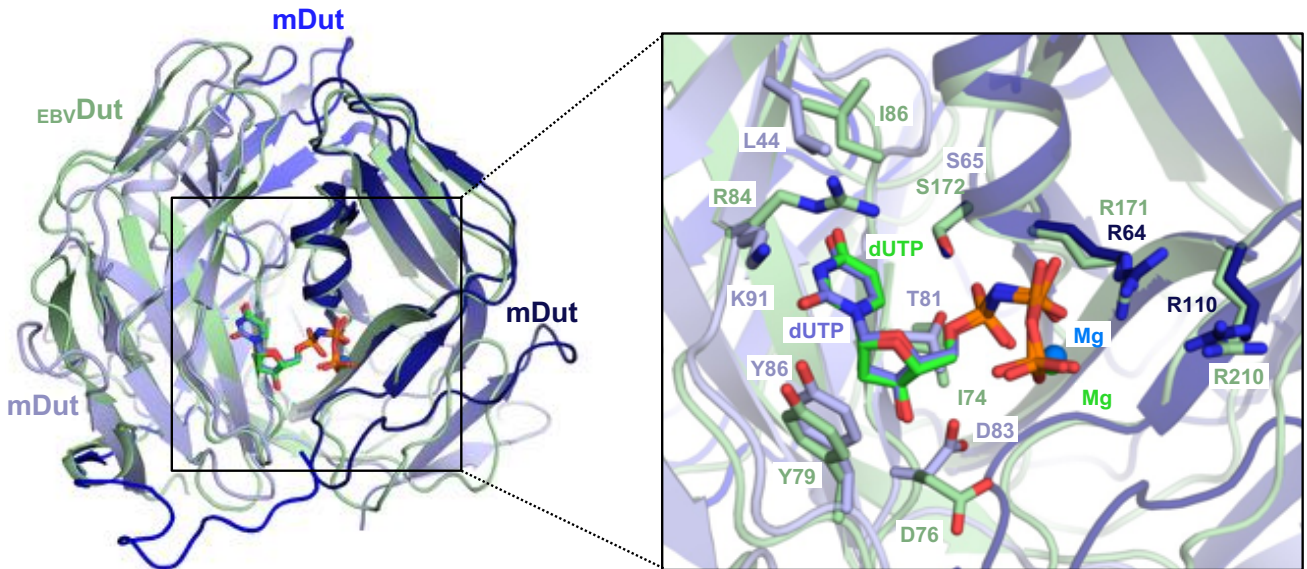


Figura R. 1. 16. Comparación estructural de las Duts trimérica y monomérica. La superposición de los nucleótidos en los centros activos de mDut (trimérica) y $_{EBV}Dut$ (monomérica) destaca la conservación estructural de los centros activos de ambas familias de Duts. (Izquierda) La representación de ambas Duts superpuestas, con los monómeros de mDut en tonalidad azul y $_{EBV}Dut$ en verde, muestra el alineamiento de las estructuras secundarias después de la superposición. (Derecha) Vista detallada de los centros activos superpuestos, donde los iones Mg^{2+} se representan en esferas, y los nucleótidos y las cadenas laterales de los residuos implicados en varillas y etiquetados. Cada elemento conserva los átomos de C del color de la estructura de la Dut de la que proviene, mientras que los átomos de O están en rojo, los de N en azul, y los de P en naranja. Figura tomada de Sanz-Frasquet et al. 2023.

A continuación, Stl^{N-ter} fue posicionado junto a la $_{EBV}Dut$ por superposición de las Duts en las estructuras en complejo con el represor. Es decir, sobre la estructura de $_{EBV}Dut-dUTP$, se superpuso inicialmente la de $hDut-dUTP$ o $mDut-dUTP$ utilizando como punto de referencia el nucleótido. Tras esto, los complejos $hDut-Stl^{N-ter}$ y $mDut-Stl^{N-ter}$ se superpusieron sobre la correspondiente superposición previa ($_{EBV}Dut-dUTP$ con $hDut-dUTP$ o $_{EBV}Dut-dUTP$ con $mDut-dUTP$) utilizando como punto de referencia $hDut$ o $mDut$, respectivamente. Aunque se generaron dos modelos, uno utilizando $hDut$ y otro utilizando $mDut$, el posicionamiento resultante de Stl^{N-ter} en la Dut monomérica fue muy similar en ambos casos, no mostrando impedimentos estéricos en ninguno de los modelos generados.

Ambos modelos mostraron que Stl insertaría su hélice $\alpha 8$ en el centro activo de $_{EBV}Dut$ y utilizaría sus tirosinas (Y105, Y112, Y113 y Y116) para reconocer los residuos catalíticos conservados de la Dut monomérica. Las tirosinas 112 y 113 del represor interaccionarían con el aspártico catalítico (D76) y la Y73 del $_{EBV}Dut$, ésta última implicada en el reconocimiento del anillo de uracilo (Figura R.1.16, Figura R.1.17, Tabla R. 1. 5). Además de estas interacciones, los modelos

señalaban otras con residuos conservados de los motivos catalíticos II y IV de las Duts triméricas que se replicaban en la Dut monomérica (Anexo VIII). Este es el caso del D117 de StI, que interacciona mediante un puente salino con R280 de $EBVDut$, y cuyo residuo equivalente en las Duts triméricas es una lisina o una arginina del motivo IV; y también de los residuos de $EBVDut$ R171 y S172 que replican las interacciones con el StI con los mismos residuos del motivo II que en las otras Duts triméricas: R64/S65 en mDut y R62/S63 en hDut. Aunque algunas interacciones StI^{N-ter}-Dut trimérica no se ven replicadas en el modelo del complejo $EBVDut$ -StI^{N-ter}, como es el caso del residuo de StI S114, que en el caso de Duts triméricas interacciona con un residuo cargado positivamente del motivo II, la mayoría lo hace (Anexo VIII) y da soporte a la hipótesis de que StI podría unir Duts monoméricas. Más aún, el posicionamiento del StI sería apropiado para que el P-loop del dominio C-terminal de la $EBVDut$ (cuya elevada flexibilidad se deduce por la ausencia de densidad electrónica en las estructuras disponibles en el PDB), que también posee un residuo de fenilalanina, pueda reconocer a StI de una manera similar al observado en el complejo hDut-StI^{N-ter}, insertando este residuo en el bolsillo hidrofóbico.

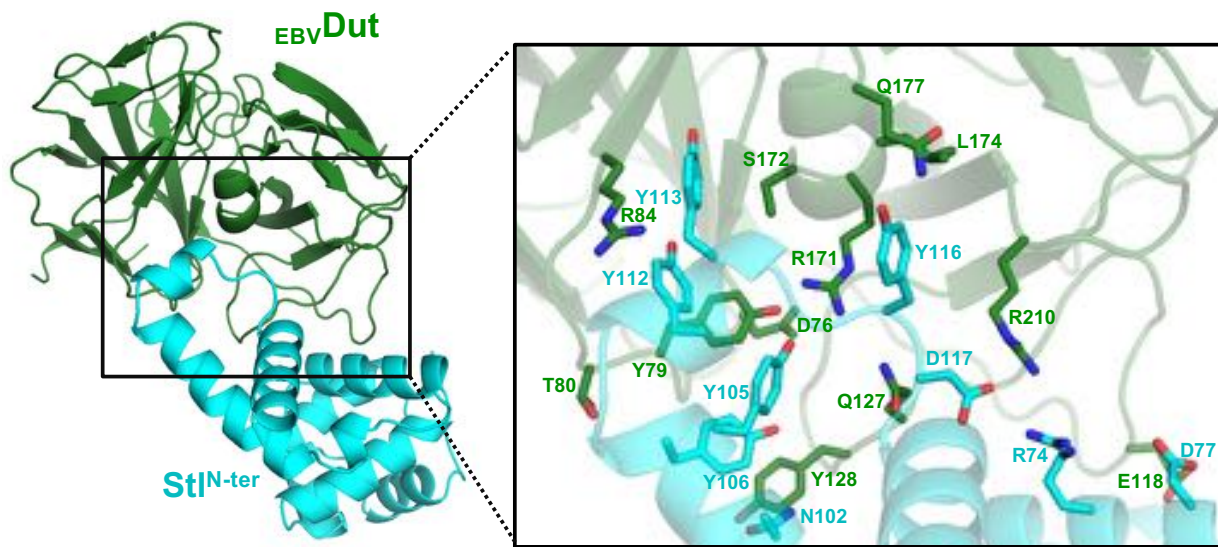


Figura R. 1. 17. Modelo del complejo Dut monomérica - StI^{N-ter}. (Izquierda) Representación del modelo $EBVDut$ -StI^{N-ter} con la Dut y el StI en verde y cian, respectivamente. (Derecha) Vista detallada del centro activo de la Dut donde el StI se acomodaría. Las cadenas laterales de los residuos que participarían en la interacción se representan en varillas y coloreados según átomo (C en cian o verde según si pertenecen a StI o Dut, O en rojo y N en azul). Figura tomada de *Sanz-Frasquet et al. 2023*.

Con el objetivo de probar la interacción entre Duts monoméricas y StI, se expresó y purificó la $EBVDut$. No obstante, la proteína presentaba problemas de solubilidad que limitaron su obtención, y en consecuencia, los análisis posteriores.

Tabla R. 1. 5. Interacciones residuo-residuo en el modelo del complejo SaPIBov1-Stl^{N-ter} - EBVDut.

SaPIBov1-Stl ^{N-ter}		EBVDut	
Elemento estructural	Residuo	Elemento estructural	Residuo
α4	55 Val	Lβ12 - β13	118 Glu
	56 Asn		117 Glu
	59 Glu		115 Gln
102 Asn	128 Tyr		
α7	106 Tyr	L β9 - β10	78 Gly
	110 Asp		80 Thr
	112 Tyr		84 Arg
		β9	79 Tyr
α9	116 Tyr	L β16 - α1	171 Arg
		α1	173 Gly
	117 Asp	L β19 - β20	210 Arg

Dado que Stl se une al centro activo de las Duts triméricas impidiendo la unión del sustrato, actúa como inhibidor competitivo de estas enzimas (Hirmondó et al. 2015; Nyíri et al. 2018). Los modelos del complejo EBVDut-Stl^{N-ter} muestran que el mecanismo de unión a Duts monoméricas sería similar al de Duts triméricas, por lo que Stl debería tener capacidad inhibidora sobre Duts monoméricas. Se decidió por ello evaluar la actividad enzimática de EBVDut en presencia y ausencia de Stl^{N-ter} haciendo uso de una ratio molar 1: 20 Dut: represor. Los ensayos enzimáticos basados en Verde Malaquita mostraron una disminución de la actividad dUTPasa sobre el 60 % en presencia de Stl^{N-ter} (figura R. 1.18), la cual es comparable a la inhibición de entre el 40 y el 80 % de la actividad dUTPasa inducida por el represor en mDut, hDut o la Dut de *Drosophila melanogaster* (Benedek et al. 2018; Hirmondó et al. 2015; Nyíri et al. 2018), apoyando un mecanismo de unión común de Stl a Duts monoméricas y triméricas. Tabla tomada de Sanz-Frasquet et al. 2023.

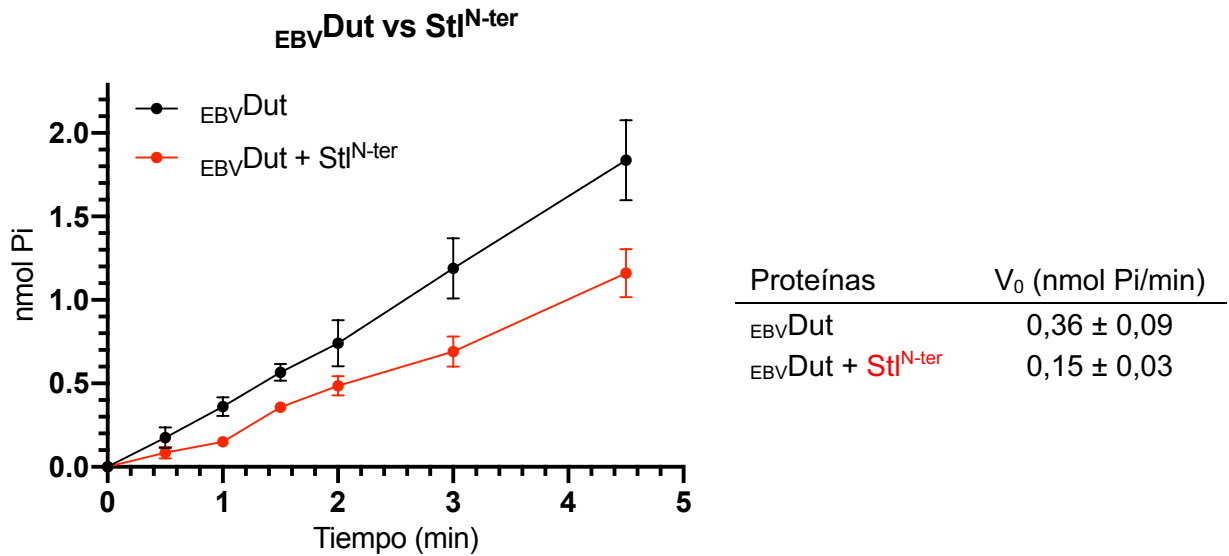


Figura R. 1. 18. Stl inhibe la actividad enzimática de EBVDut. La actividad enzimática de EBVDut se ha medido como producción de Pi en el tiempo en ausencia (negro) o presencia (rojo) del represor. La gráfica muestra los resultados y la desviación estándar de cuatro ensayos independientes. En la tabla de la derecha se indica la velocidad inicial de la Dut en ausencia o presencia del represor, junto con los errores estándar correspondientes entre paréntesis. Figura tomada de *Sanz-Frasquet et al. 2023*.

Finalmente, con el objetivo de evaluar la posible capacidad de Stl para unirse de manera universal a Duts realizamos un análisis de secuencias. Para ello se tomaron 10.616 secuencias de Duts de la base de datos PFAM bajo la ID PF00692, las cuales fueron alineadas computacionalmente y agrupadas por tamaño. Aunque bajo esta ID de PFAM se incluyen Duts de las 3 familias, el proceso computacional incluyó las Duts mayoritarias y las similares a éstas, siendo las monoméricas y las triméricas. Debido a la reorganizada secuencia que presentan las Duts monoméricas respecto a las triméricas (Tarbouriech et al. 2005), su inclusión en el análisis limita a que solo una parte de los motivos catalíticos puedan ser correctamente alineados, reduciendo el grado de conservación. Tras la obtención del LOGO introduciendo el alineamiento en el servidor Skylign, se observó que la conservación es especialmente alta en residuos de los motivos II, III y V, y muestra una mayor variabilidad en los residuos del motivo IV, los cuales modularían la afinidad entre Duts y Stl según estudios *in vitro* e *in vivo* previos (Rafael Ciges-Tomas et al. 2019; Maiques et al. 2016; Tormo-Más et al. 2013). Los residuos del motivo I, que contribuyen en menor medida en la unión del represor, también muestran una mayor variabilidad (Figura R. 1. 19).

Esta elevada conservación soporta la idea que Stl pueda reconocer e interactuar con un amplio rango de Duts triméricas y monoméricas,

proponiendo el alcance de esta capacidad como universal. Además, esta capacidad de unión podría extenderse a otros StIs de otras SaPIs que tengan una homología de secuencia elevada con la del StI^{BovI} en sus dominios intermedio y N-terminal (Bowring et al. 2017).

Dada la naturaleza inhibidora de su interacción con las Duts, los resultados mostrados en esta Tesis apoyan el uso de StI^{BovI} como un inhibidor proteico de Duts, como había sido propuesto previamente (Nyíri et al. 2018).

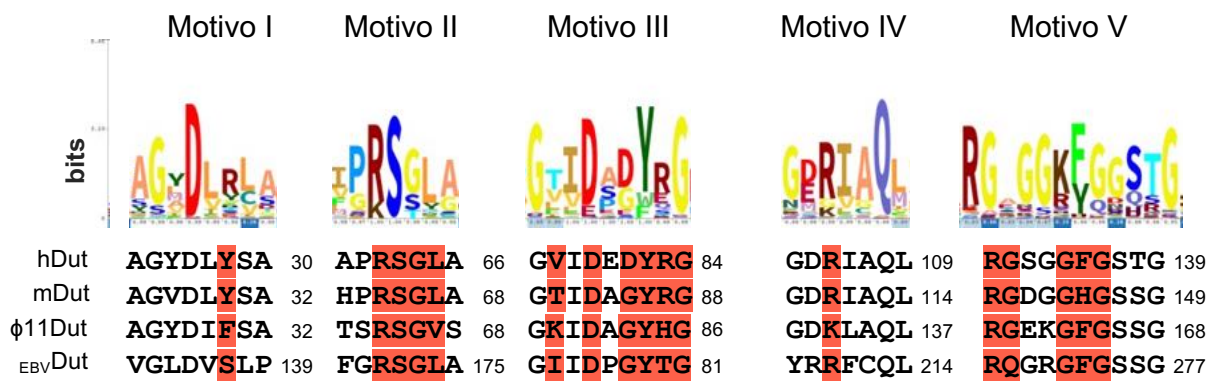


Figura R. 1. 19. Análisis de secuencia de los residuos de los motivos catalíticos de las Duts triméricas y monoméricas implicados en el reconocimiento y unión de StI^{N-ter}. (Arriba) El LOGO de los 5 motivos catalíticos. La altura del símbolo del aminoácido es proporcional a su frecuencia en esa posición específica. (Abajo) Las secuencias de hDut, mDut, φ11Dut and EBVDut se alinean con el LOGO y las posiciones participantes en la interacción con el represor se resaltan en color rojo (estructuras resueltas o modelo *in silico* para Duts triméricas o monomérica, respectivamente). Figura tomada de *Sanz-Frasquet et al. 2023*.

CAPÍTULO II

2.1. Análisis molecular de las diferencias cinéticas entre SeMazG Y EnMazG.

En una Tesis anterior se caracterizaron enzimática y estructuralmente las proteínas MazG del fago ϕ IPLA7 de *Staphylococcus epidermidis* (SeMazG) y del fago ϕ FL2A de *Enterococcus faecalis* (EnMazG), con la hipótesis de que estas proteínas constituían una alternativa a las Duts en dichos fagos, y por tanto, estarían realizando funciones homólogas. Ambas MazG solo son capaces de hidrolizar el dUTP, al contrario de las MazG de otros organismos que presentan la capacidad de degradar una amplia gama de desoxinucleótidos (Mustafi et al. 2003; Lyu et al. 2013; Kim and Hong 2016; Zhang and Inouye 2002; Zhang, Zhang, and Inouye 2003; Rihtman et al. 2019).

Las proteínas MazG descritas con anterioridad a éstas, como la MazG de *M. tuberculosis* y de *Bacillus* (Chan et al. 2004; Kim and Hong 2016), presentan una cinética característica de Michaelis-Menten; mientras que un ortólogo humano de MazG (CTPasa DCTPP1) (Requena et al. 2014) muestra una cinética sigmoideal, al igual que las dos MazG utilizadas en esta Tesis. Ambas MazG muestran cooperatividad, con valores alrededor de 2,2 para el coeficiente de Hill. Otros parámetros cinéticos como la actividad específica, la constante de recambio y la $K_{0,5}$ (homóloga a la K_m en cinética michaeliana) difieren de manera significativa, posicionando a SeMazG con una eficiencia superior a EnMazG entre cuatro y cinco veces (constante de especificidad de 0,1 y 0,023 $\mu\text{M}^{-1}\text{s}^{-1}$, respectivamente) (Tabla R. 2.1).

Tabla R. 2.1. Parámetros cinéticos de EnMazG y SeMazG. Para cada enzima se muestran los valores de 5 variables cinéticas con las unidades y los errores estándar correspondientes entre paréntesis. Tabla adaptada de *Ciges-Tomas 2019*.

	EnMazG	SeMazG
Coeficiente de Hill (n)	2,28 (\pm 0,4)	2,19 (\pm 0,4)
$K_{0,5}$ (μM)	206 (\pm 29,38)	23,79 (\pm 1,5)
Actividad específica ($\mu\text{mol PPI/mg}\cdot\text{min}$)	18,87 (\pm 2,35)	9,26 (\pm 0,22)
K_{cat} (S^{-1})	4,7 (\pm 0,58)	2,4 (\pm 0,06)
Cte. de especificidad ($K_{cat}/K_{0,5} - \mu\text{M}^{-1}\text{s}^{-1}$)	0,023	0,1

2.1.1. La participación del tercer monómero en el centro activo es necesario para una actividad catalítica óptima.

EnMazG y SeMazG son tetrámeros formados por un dímero de dímeros. Esta organización cuaternaria explica la cooperatividad positiva observada en ambas, quedando reflejada en un valor mayor a 1 del coeficiente de Hill, y en la participación de residuos de más de una subunidad en cada uno de los cuatro centros activos que existen en los tetrámeros. En concreto, residuos de 3 subunidades participan en cada uno de estos centros activos, siendo dos de ellas, las formadoras de un dímero, las que aportan la mayor parte de los residuos, mientras que la subunidad que proviene del segundo dímero aporta un residuo de arginina (Figura I.37). Este residuo básico, R42 en el caso de EnMazG y R48 en SeMazG, interacciona con los fosfatos α y γ del dUTP, ayudando a la estabilización del sustrato.

Con el objetivo de conocer cuál era el papel de este residuo en la catálisis de ambas Duts, se diseñaron y produjeron dos mutantes puntuales donde la arginina es sustituida por una alanina (mutante R42A en el caso de EnMazG y R48A en el caso de SeMazG). La actividad enzimática de ambos mutantes fue analizada frente a la actividad desarrollada por las correspondientes enzimas salvajes mediante ensayos colorimétricos basados en el reactivo Verde Malaquita (Figura R. 2.1). En la figura R. 2.1 se muestra la actividad catalítica como la producción de nmol de Pi en el tiempo para EnMazG (*arriba*), y SeMazG (*abajo*) tanto es su forma mutante como salvaje.

La producción de fosfato inorgánico en 10 minutos en el caso de EnMazG WT fue de 15,90 nmol de Pi, mientras que el mutante R42A produjo solo 4,66 nmol de Pi. En el caso de SeMazG WT la producción fue 15,40 nmol de Pi, mientras que la producción en el caso del mutante R48A fue de 7,70 nmol de Pi. Por tanto, en ambos casos la actividad enzimática se ve afectada por la sustitución de este residuo básico por alanina, llegando a una reducción de entre 2 y 3,5 veces en SeMazG y EnMazG, respectivamente.

Junto al análisis de la actividad enzimática, se comprobó si el estado oligomérico de los mutantes era distinto al de las MazG no mutantes. Para ello, se recurrió a la cromatografía de exclusión molecular acoplada a la dispersión de luz en múltiples ángulos (SEC-MALS), con el objetivo de conocer el peso absoluto de las proteínas mutantes y con ello saber si la estructura cuaternaria de éstas era distinta.

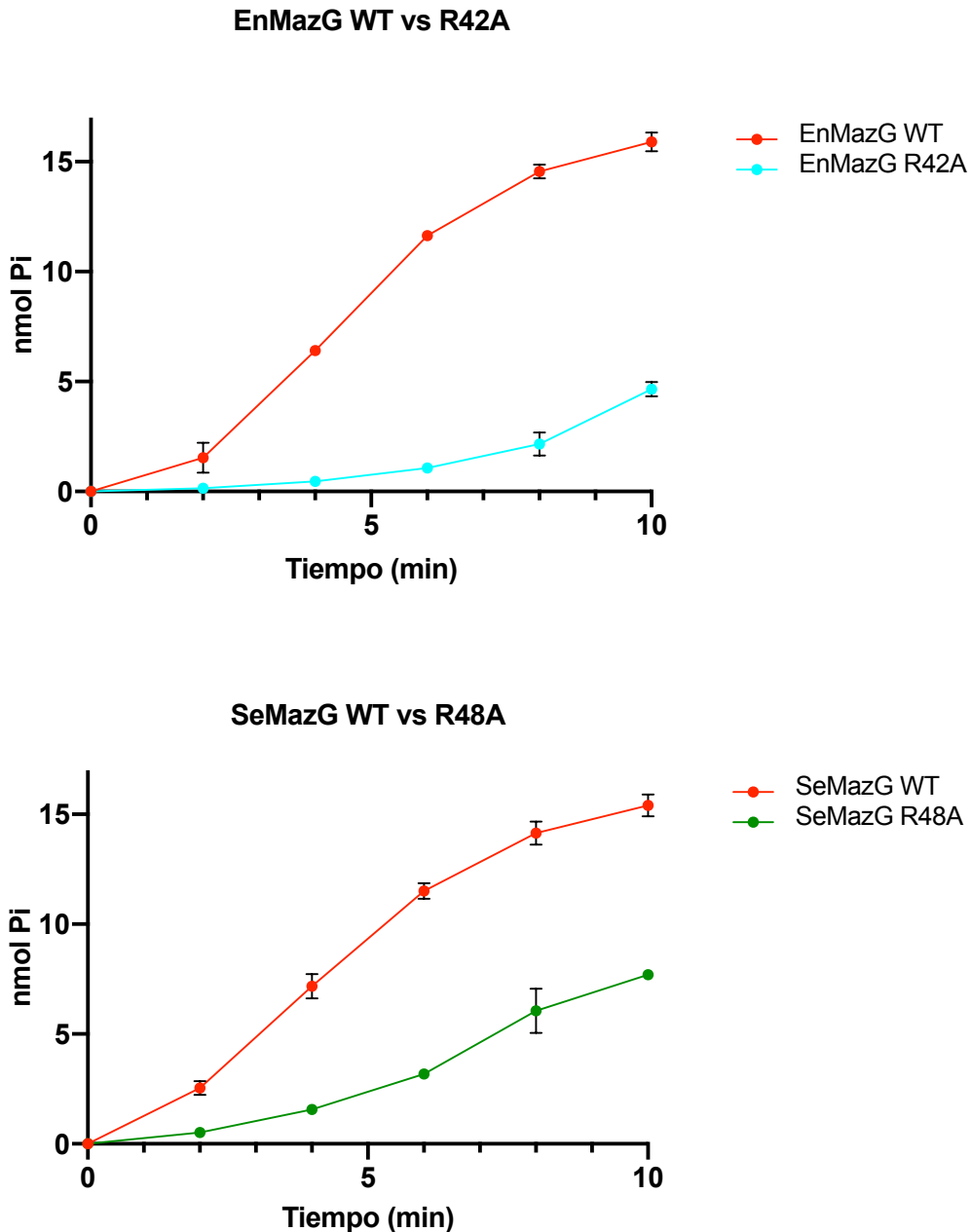


Figura R. 2.1. Perturbación de la actividad enzimática de los mutantes puntuales EnMazG R42A (*arriba*) y SeMazG R48A (*abajo*). Se muestra en rojo la producción de Pi en el tiempo tras la hidrólisis del dUTP de ambas MazG no mutantes, en cian la de EnMazG R42A y en verde SeMazG R48A. La concentración de enzima en los ensayos era distinta (EnMazG 2 ng/ μ L, SeMazG 1 ng/ μ L). Análisis realizado con *GraphPad* y dos réplicas.

Los resultados se presentan en la Tabla R. 2. 2. La muestra de EnMazG no mutante es monodispersa, con un único pico de UV cuya masa molecular corresponde con el cálculo teórico de un tetrámero (57,32 kDa experimentales frente a 59,2 kDa teóricos). La muestra EnMazG R42A también es monodispersa y la masa molecular experimental es 55,4 kDa, con lo cual el estado oligomérico es el mismo, un tetrámero. La muestra de SeMazG no mutante es polidispersa,

con dos picos de UV cuyas masas moleculares corresponderían con los estados oligoméricos dímero y tetramero (39,6 y 67,7 kDa experimentales frente a 31,2 y 62,4 kDa teóricos). Por su parte, la muestra del mutante SeMazG R48A es monodispersa con una masa molecular experimental de 71,04 kDa, la cual se podría asociar con una oligomerización tetramérica.

Tabla R. 2.2. Pesos moleculares de EnMazG y SeMazG no mutantes y mutantes de arginina calculados por SEC-MALS. Se muestran los pesos moleculares teóricos calculados mediante Protparam y los experimentales con sus correspondientes errores entre paréntesis (en kDa).

Masa Molecular (kDa)	Teórica	Experimental (SEC-MALS)			
	No mutante (tetramero)	No mutante		Mutante Arginina (R42/48A)	
		Tetramero	Dímero	Tetramero	Dímero
EnMazG	59,2	57,32 (± 1,5 %)	-	55,4 (± 1,2 %)	-
SeMazG	62,4	67,7 (± 1,2 %)	39,6 (± 8,7%)	71,04 (± 0,86%)	-

Por tanto, los mutantes puntuales de este residuo de arginina no alteran la organización oligomérica de ninguna de las dos proteínas. No obstante, la disminución del 50 % en el caso de SeMazG y del 70 % en el caso de EnMazG en la producción del Pi destacan la participación clave de este residuo en la actividad catalítica, dado su posicionamiento sobre el sustrato y consecuente estabilización. Para una actividad enzimática óptima es necesaria la participación del único residuo procedente del tercer monómero en cada centro activo.

2.1.2. La tetramerización es necesaria para la actividad catalítica.

La superposición de los tetrameros de ambas MazG muestra que la organización cuaternaria es muy similar (RMSD 0,506 - 0,553 Å para la superposición de los C α de 386 - 394 residuos), pero más aún sus centros activos, que mantienen dos dominios de unión de metal formados por una péntada acídica de aspárticos y glutámicos (E31/32/35 y D50/54 en EnMazG, y E37/38/41 y D56/60 en SeMazG). En presencia del sustrato, el tetramero de ambas MazG muestra un movimiento respecto a la conformación *apo* (conformación abierta) mediante el

cual se estrecha la zona de tetramerización y la estructura cuaternaria se vuelve más compacta (conformación cerrada) (Ciges-Tomas 2019). Esta transición a la conformación cerrada implica movimientos de las hélices $\alpha 2$ y $\alpha 3$, donde la primera se aproxima al dímero complementario, con el que establece más contactos de tetramerización. En estas hélices se posicionan las péntadas acídicas (Figura R. 2. 2)



Figura R. 2. 2. Cambios conformacionales de las hélices $\alpha 2$ y $\alpha 3$ del dímero de EnMazG en el cambio de conformación abierta a cerrada. Superposición de un dímero de EnMazG en conformación *apo* (abierto, en rosa) sobre un dímero en conformación cerrada (con dUpNHpp, en azul). Cada monómero del dímero está coloreado con una tonalidad distinta del mismo color. Se identifican las estructuras secundarias (hélices α) y el desplazamiento realizado entre ambas conformaciones (distancias indicadas en Å). Se identifica con un apóstrofo el segundo monómero formador del dímero. Figura tomada de Ciges-Tomas 2019.

El análisis estructural profundo de ambas estructuras no llevó a plantear la hipótesis de que las diferencias entre ambas enzimas fueran debidas a un distinto grado de plasticidad entre ambas; no obstante, dado el carácter

cooperativo, un distinto grado de plasticidad podría permitir una organización más óptima de los centros activos.

El análisis mostró que la superficie de tetramerización está altamente conservada, con tan solo un cambio que llamó nuestra atención (Tabla R. 2.3). Este residuo distintivo es una isoleucina en EnMazG (I30) y una tirosina (Y36) en SeMazG, situados en la hélice $\alpha 2$ (Ciges-Tomas 2019) (Figura R. 2.3 y R. 2.4). Aunque es cierto que no podría considerarse un cambio conservativo, ambos residuos son los únicos que interaccionan con su homólogo en la hélice $\alpha 2$ del otro dímero, estableciendo dos puntos de anclaje de la zona de tetramerización mediante interacciones hidrofóbicas. Además, en el caso de la tirosina de SeMazG, el grupo hidroxilo interacciona con dos residuos de la hélice $\alpha 2$ del monómero con el que forma el dímero (mediante el nitrógeno de la G42 y el oxígeno de la A39). En el caso de la isoleucina de la EnMazG, varios átomos de C de la cadena lateral interaccionan con el residuo V33 (con un C de la cadena lateral y el O de la cadena principal) y del residuo G34 (con el O de la cadena principal de la I30 y el C asimétrico de la glicina), ambos del monómero complementario con el que forma el dímero. Por tanto, ambos residuos parecen ser cruciales en la nucleación de los respectivos tetrámeros, y por ello resulta llamativo el cambio de EnMazG a SeMazG de un residuo con una cadena lateral tan voluminosa.

Tabla R. 2. 3. Residuos equivalentes en EnMazG y SeMazG que median la tetramerización. Tabla adaptada de *Ciges-Tomas 2019*.

Residuos de contacto en la tetramerización

EnMazG	SeMazG
23 Glu	29 Asp
24 Lys	30 Arg
26 Met	32 Ala
27 Leu	33 Leu
30 Ile	36 Tyr
31 Glu	37 Glu
34 Gly	40 Gly
38 Ala	44 Ala
40 Leu	46 Leu
41 Ala	
42 Arg	48 Arg

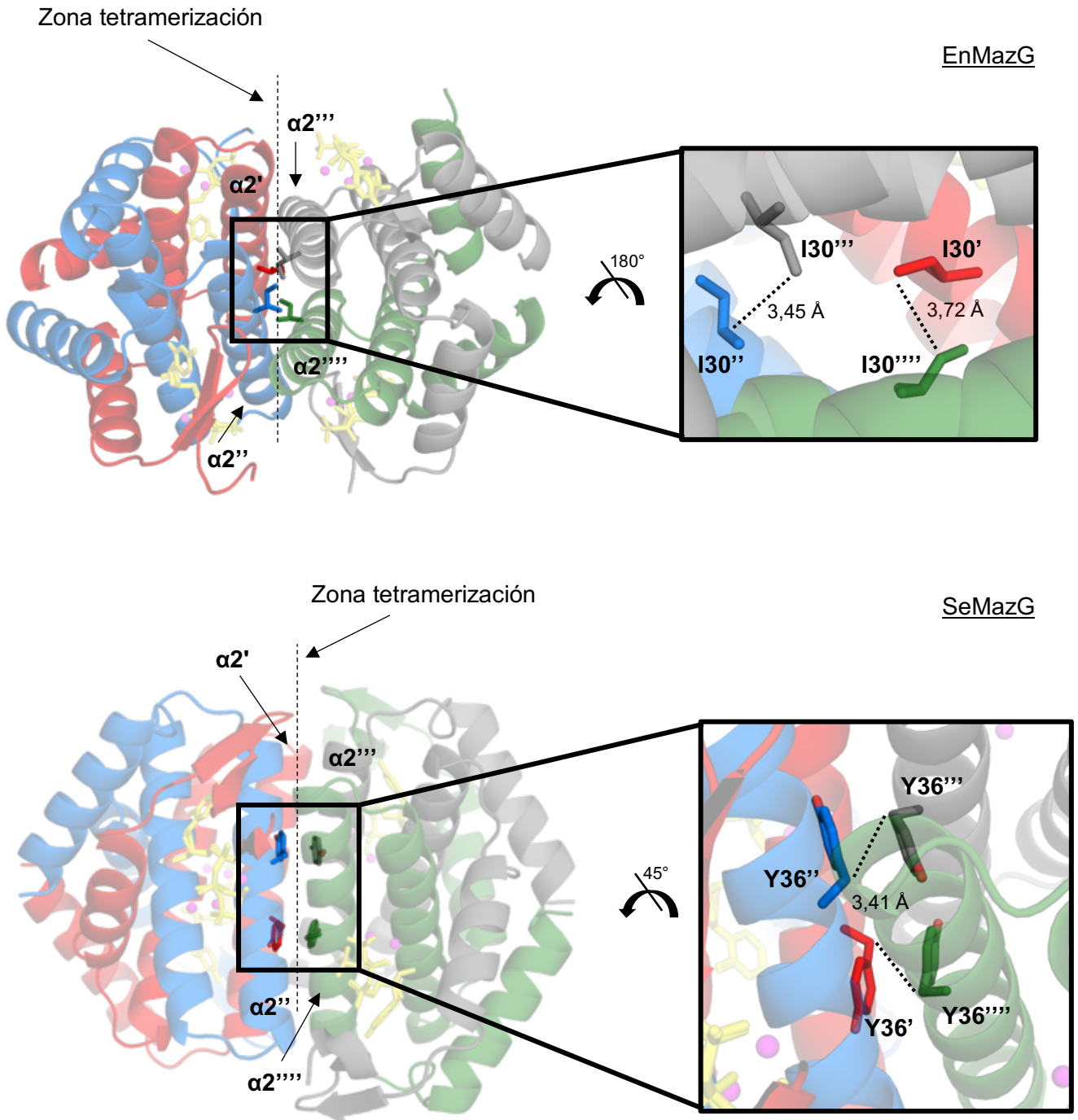


Figura R. 2.3. Localización estructural de los residuos I30 y Y36 en la zona de tetramerización de EnMazG (*arriba*) y SeMazG (*abajo*), respectivamente. A una distancia de unos 3,4 Å las cadenas laterales de isoleucina o tirosina establecen interacciones hidrofóbicas entre las hélices $\alpha 2$ de monómeros contiguos de homodímeros distintos. Se representan en distintos colores los 4 monómeros formadores de cada enzima (azul, rojo, verde y gris, *izquierda*), y se destacan las cadenas laterales del residuo de interés (*derecha*) con el color correspondiente al monómero del que provienen. También se identifican las hélices $\alpha 2$ de cada monómero, así como la zona de tetramerización de cada enzima. PDBs tomados de Ciges-Tomas 2019.

EnMazG

SeMazG

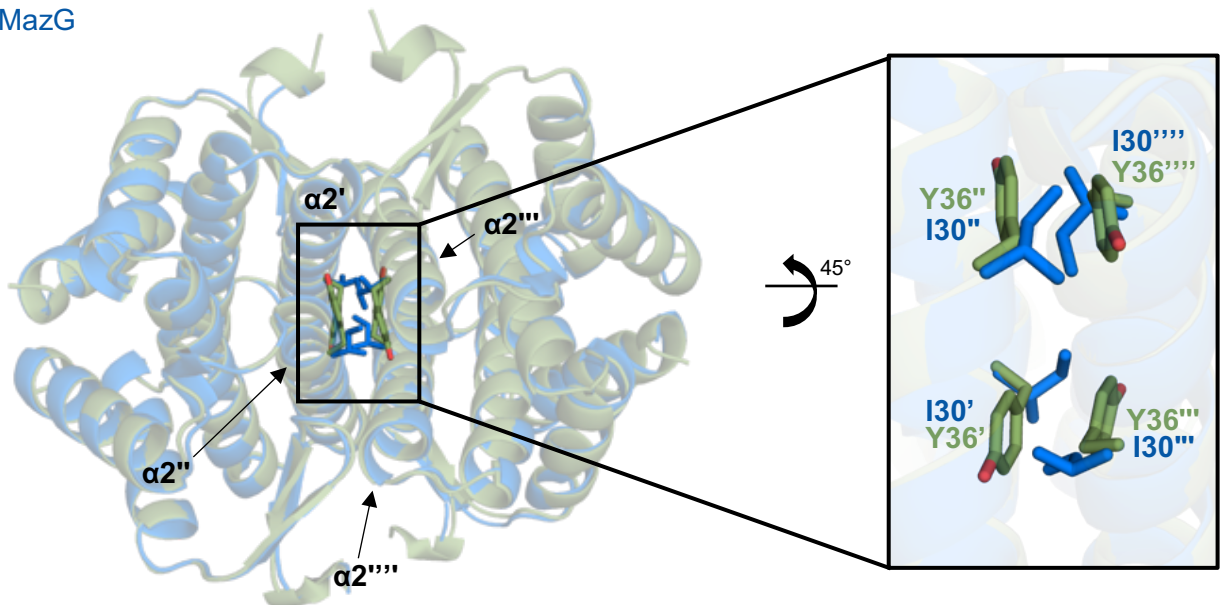


Figura R. 2.4. Superposición estructural de las enzimas EnMazG (azul) y SeMazG (verde) (Izquierda). (Derecha) Se destacan los residuos de isoleucina (azul) y tirosina (verde) de la superficie de tetramerización de EnMazG y SeMazG, respectivamente. PDBs tomados de Ciges-Tomas 2019.

Con el objetivo de comprobar esta hipótesis sobre la plasticidad se diseñaron y produjeron dos mutantes de cada MazG. Por un lado, una sustitución aminoacídica del residuo por una alanina para ver el efecto de la ausencia de la cadena lateral (EnMazG I30A, SeMazG Y36A). Por otro, el intercambio del residuo entre ambas MazG (EnMazG I30Y, SeMazG Y36I), para comprobar si los parámetros cinéticos eran alterados significativamente. Una vez los mutantes estuvieron purificados se procedió al análisis de su actividad enzimática utilizando idénticos ensayos colorimétricos basados en el reactivo Verde Malaquita (Figura R. 2.5) y se comprobó su estado oligomérico mediante SEC-MALS (Tabla R. 2.4).

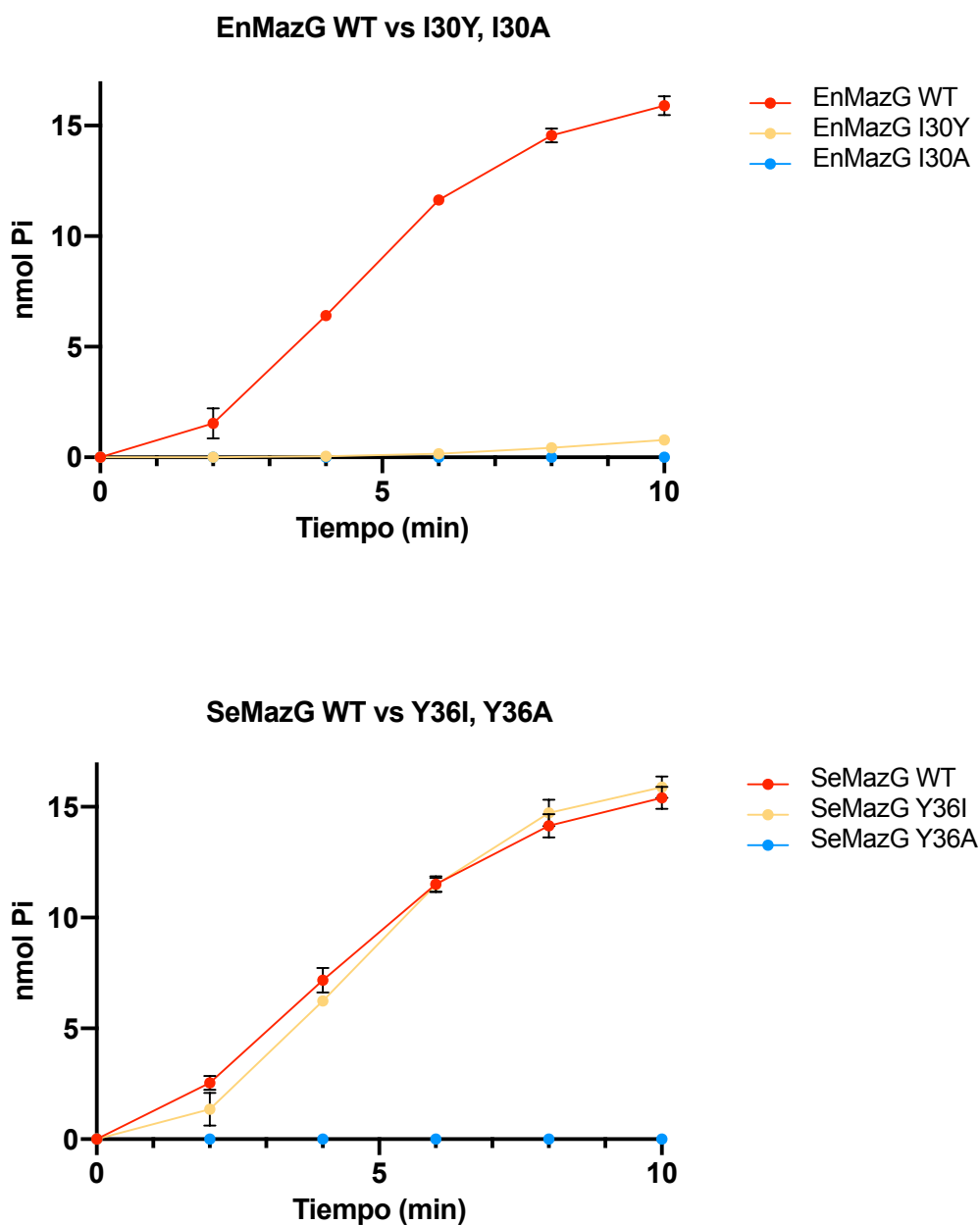


Figura R. 2. 5. Perturbación de la actividad enzimática de los mutantes puntuales EnMazG I30A/Y (*arriba*) y SeMazG Y36A/I (*abajo*). Se muestra en rojo la producción de Pi en el tiempo de ambas MazG no mutantes, en azul la de los mutantes a alanina, y en amarillo la de los mutantes del residuo de la otra MazG. Análisis realizado con *GraphPad* y dos réplicas.

Tabla R. 2.4. Pesos moleculares de EnMazG y SeMazG no mutantes y mutantes de la zona de tetramerización calculados por SEC-MALS. Se muestran los pesos moleculares teóricos calculados mediante Protparam y los experimentales con sus correspondientes errores entre paréntesis (en kDa).

Masa Molecular (kDa)	Teórica	Experimental (SEC-MALS)					
	No mutante (Tetrámero)	No mutante		Mutante a Alanina (EnMazG I30A/ SeMazG Y36A)		Mutante intercambio (EnMazG I30Y/ SeMazG Y36I)	
		Tetrámero	Dímero	Tetrámero	Dímero	Tetrámero	Dímero
EnMazG	59,2	57,32 (± 1,5 %)	-	58,5 (± 0,95 %)	-	62,7 (± 1,7 %)	-
SeMazG	62,4	67,7 (± 1,2 %)	39,6 (± 8,7 %)	-	35 (± 1,14 %)	60,42 (± 2,3 %)	29,8 (± 3,4 %)

La mutación a alanina supone la pérdida por parte de ambas enzimas de la capacidad de degradar el dUTP (Figura R. 2.5), mientras que la oligomerización tetramérica se conserva en el caso de EnMazG, pero vira a dimérica en el caso de la SeMazG. En el caso de los mutantes de intercambio de residuos, la actividad enzimática no fue alterada de igual manera. Solo en el caso de EnMazG I30Y la proteína pierde toda la actividad, mientras que en el caso de SeMazG Y36I la proteína muestra la misma capacidad catalítica que la proteína no mutante. La oligomerización de estos mutantes es idéntica a la mostrada por las proteínas no mutantes.

Ambas sustituciones de la isoleucina en EnMazG suponen la pérdida casi o total de la actividad enzimática (Figura R. 2. 5) pero no la de su conformación tetramérica. Si se comparan las estructuras no mutantes en conformación abierta y cerrada se evidencia un acercamiento llevado a cabo por los residuos de I30 tras la entrada del sustrato, las distancias entre los C α se reducen de 8,70 a 7,39 Å. Por tanto, cuando el nucleótido ocupa el centro activo y se lleva a cabo el movimiento giratorio de las hélices α 2 y α 3 junto con el estrechamiento de la zona de tetramerización para lograr la conformación cerrada y activa, en el caso del mutante a alanina (I30A) no podría mantenerse estable por la falta de los dos puntos de anclaje entre las 4 hélices α 2. En el caso del mutante a tirosina (I30Y), cuya actividad enzimática decae un 95 %, habría impedimentos estéricos que impedirían o dificultarían en gran medida el estrechamiento necesario entre dichas hélices α para que la enzima pase a la conformación activa (la distancia entre los C α de las tirosinas en el caso de SeMazG en presencia del nucleótido es 7,60 Å).

En el caso de SeMazG Y36A, la actividad enzimática se pierde completamente y la oligomerización vira a dimérica. Por tanto, la pérdida de los dos puntos de

anclaje al otro dímero formados por las tirosinas implica la pérdida de la tetramerización y con ella la transición a la conformación cerrada y activa con el correcto posicionamiento de los residuos catalíticos. La sustitución de tirosina por isoleucina (Y36I) supone la pérdida de las interacciones de apilamiento entre los anillos aromáticos, pero se mantienen las que la cadena lateral de la isoleucina establezca con su homólogo. En este caso, una mayor plasticidad de la enzima podría explicar que se lograra igualmente el estrechamiento de la zona de tetramerización y la conformación cerrada, de ahí que este mutante mantenga la misma actividad enzimática y las dos oligomerizaciones presentes en la proteína no mutante.

2.2. El uso de otros cationes divalentes en EnMazG y SeMazG.

Aunque todos los ensayos enzimáticos presentados en el apartado anterior, así como los referenciados para la caracterización de ambas enzimas fueron realizados en presencia de 5 mM de $MgCl_2$, en la bibliografía pueden encontrarse proteínas MazG capaces de degradar el sustrato en presencia de cationes distintos al Mg^{2+} . Por ello analizamos si las MazG de estos fagos pueden también utilizar otros cationes para la hidrólisis del dUTP.

2.2.1. El magnesio no puede ser sustituido por otro catión divalente para que EnMazG y SeMazG puedan hidrolizar dUTP.

Se realizaron ensayos enzimáticos idénticos a los descritos en el apartado anterior en los cuales se sustituyó el $MgCl_2$ por sales divalentes de los cationes Mn, Ni, Ca, Zn y Cu (Figura R. 2.6). La selección de cationes incluyó aquellos utilizados por otras MazG o Duts en la hidrólisis del dUTP o por otras proteínas que coordinaran metales divalentes. En la figura R. 2.6 se representa la producción de nmol de Pi en porcentaje normalizado a la cantidad total de nmol de Pi producido en presencia de $MgCl_2$ a tiempo final en la reacción (10 minutos). Tanto en el caso de EnMazG como SeMazG la actividad enzimática es totalmente inhibida en ausencia del $MgCl_2$ y en la sola presencia de cualquier catión divalente probado; con la excepción de los cationes de Ca y Ni en el caso de EnMazG, en presencia de los cuales muestra una ínfima actividad (la producción del Pi decae en más del 99 %). Por tanto, se pudo concluir que el Mg^{2+} no puede ser sustituido por otro catión divalente en las MazG para su actividad enzimática.

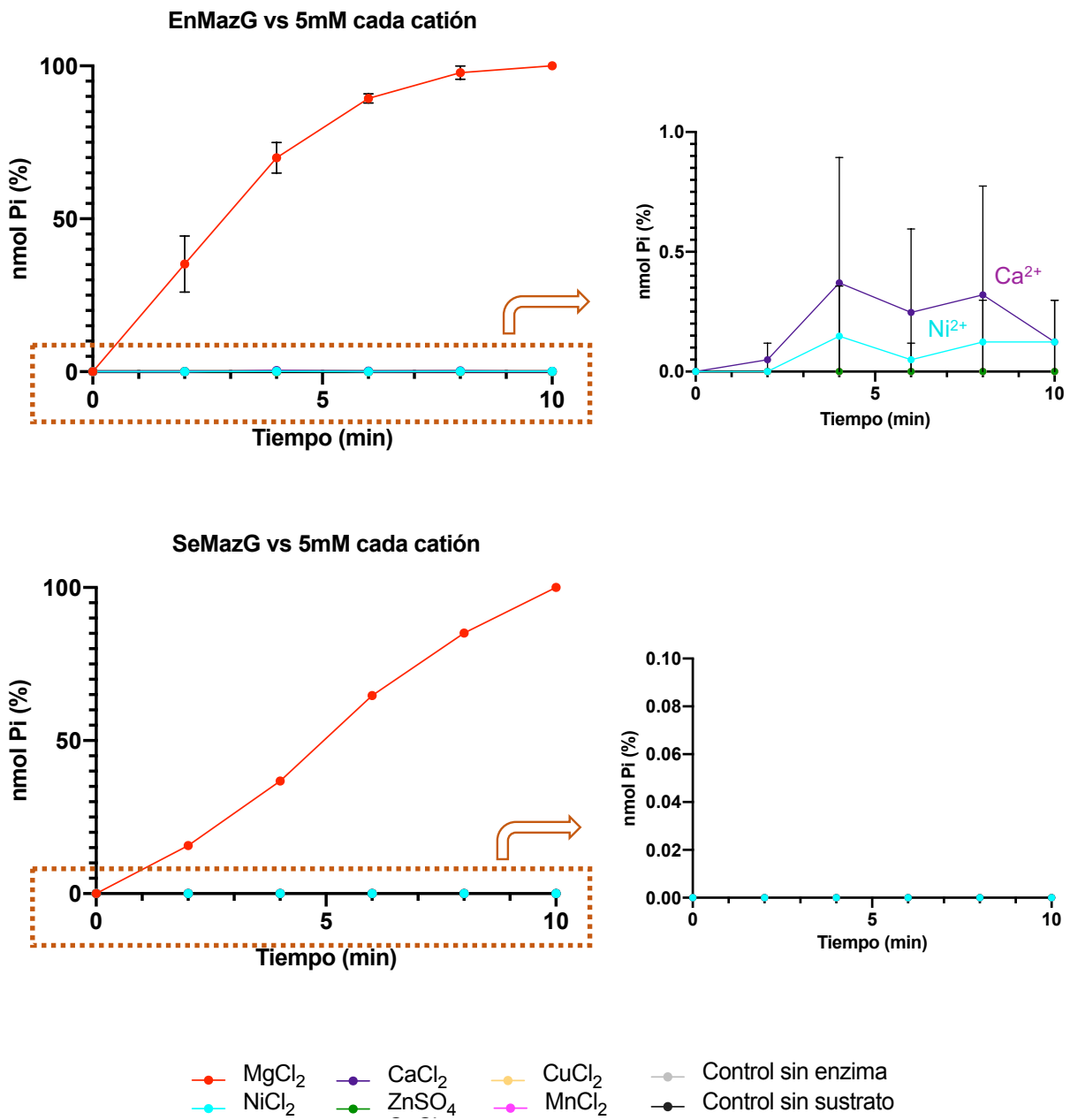


Figura R. 2. 6. Perturbación de la actividad enzimática de EnMazG (*arriba*) y SeMazG (*abajo*) en presencia de Mg²⁺ u otros cationes divalentes. Se representa la producción de nmol de Pi durante 10 minutos en porcentaje relativo respecto a la producción total en presencia de 5 mM de MgCl₂ a tiempo final. En distintos colores se representan los distintos iones probados y los controles sin enzima y sin sustrato. A la derecha se representan los mismos datos pero excluyendo las réplicas del ensayo en presencia de MgCl₂. El análisis de datos procedentes de dos réplicas se realizó en *GraphPad*. Se incluyen las barras de error.

2.2.2. La actividad enzimática de EnMazG y SeMazG es inhibida por cationes divalentes.

Con los resultados de los ensayos enzimáticos realizados se concluyó que ambas MazG no podían sustituir al Mg^{2+} para hidrolizar el dUTP, pero dado el resultado de EnMazG en presencia de Ca^{2+} y Ni^{2+} , y que en el mecanismo de hidrólisis participan 3 cationes divalentes, se planteó la posibilidad de que no todas las posiciones de los cationes fueran ocupadas por Mg^{2+} , sino por otros cationes, y que las MazG hidrolizaran el sustrato combinando el Mg^{2+} con otro catión alternativo.

Para testar esta alternativa se realizaron también ensayos enzimáticos. En estos ensayos se mantenía constante una concentración 5mM de $MgCl_2$, no limitante para la enzima, y se variaban las concentraciones del ion divalente alternativo. Las concentraciones probadas abarcaron desde la ausencia del catión alternativo (0 mM) hasta 1,2 veces más que la concentración del Mg^{2+} (7 mM).

La preparación de las muestras durante los ensayos requería de un orden concreto de adición. Las mezclas de reacción con las distintas ratios de cationes incluían el tampón de reacción y los cationes en las concentraciones a probar, y finalmente, antes de añadir la pirofosfatasa, se añadía la enzima. En caso de invertir este orden y añadir el catión alternativo, después la MazG y finalmente el Mg^{2+} , la precipitación de la proteína era frecuente.

En la figura R. 2.7 y R. 2.8 se muestran los resultados de 4 réplicas de cada ensayo, representados como el porcentaje de nmol de Pi producidos en cada reacción normalizados a la producción a tiempo final (10 minutos) en la reacción en la única presencia de Mg^{2+} . Como muestran los ensayos, todos los cationes inhiben a ambas MazG en todas las concentraciones testadas (desde una quinta parte hasta un exceso de 1,4 veces respecto a la concentración de Mg^{2+}), aunque se observaron diferencias significativas en la capacidad inhibidora. Mientras que el Mn^{2+} o el Zn^{2+} provocan la inhibición completa de ambas MazG en todas las concentraciones testadas, es necesario llegar a un 80 % de la concentración de Ni^{2+} y Cu^{2+} respecto al Mg^{2+} para observar la misma inhibición. En el caso del Ca^{2+} , con concentraciones entre el 40 y el 60 % respecto a la de Mg^{2+} es suficiente para inhibir a SeMazG y EnMazG, respectivamente. En conjunto, los ensayos muestran un idéntico efecto en ambas MazG con una graduación decreciente en la capacidad inhibidora para los cationes divalentes desde el Mn^{2+} y el Zn^{2+} , seguidos por el Ca^{2+} , y finalmente por el Ni^{2+} y el Cu^{2+} .

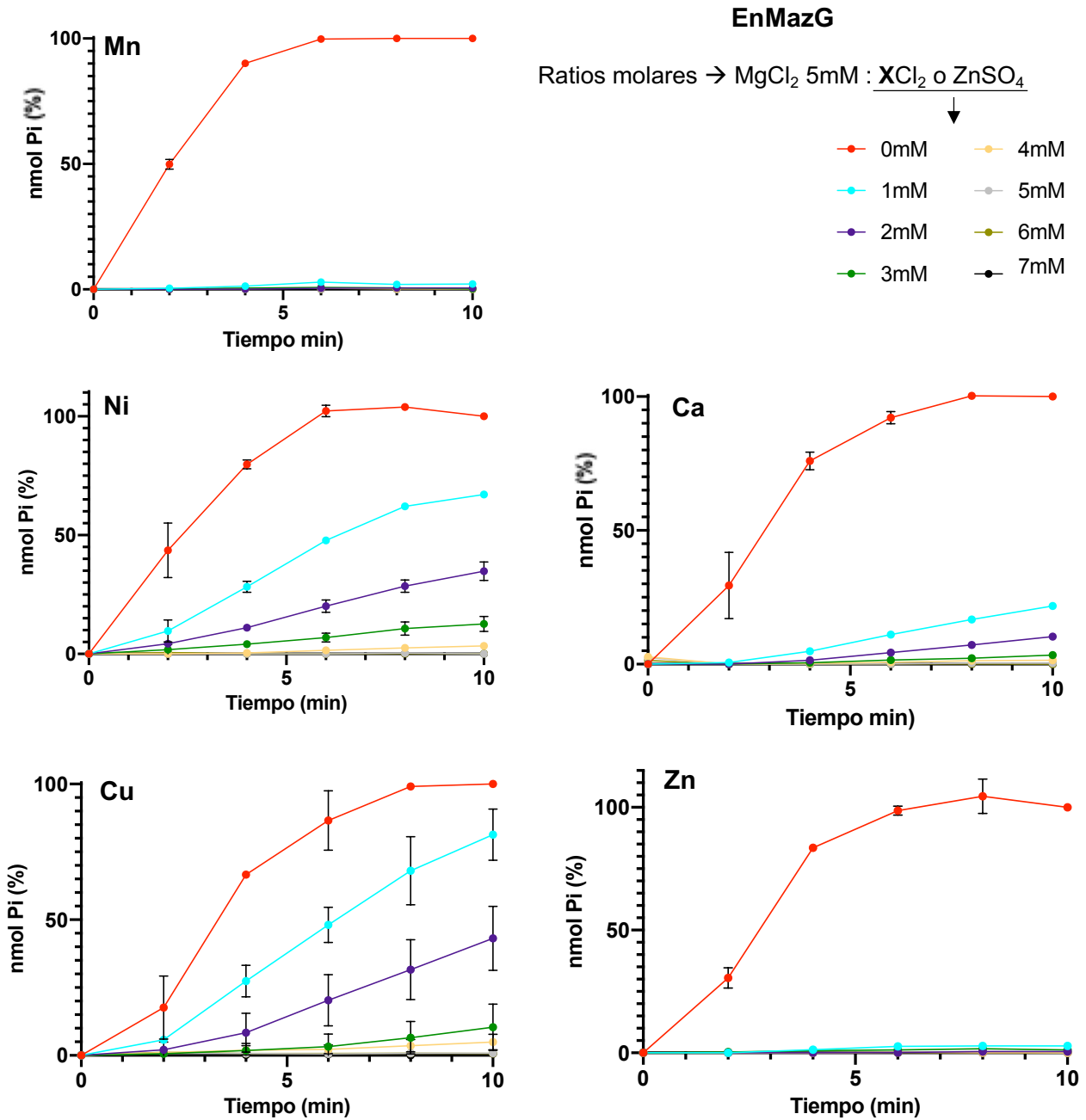


Figura R. 2.7. Perturbación de la actividad enzimática de EnMazG en presencia de Mg²⁺ y cationes divalentes alternativos en ratios molares crecientes. La concentración de MgCl₂ es constante (5 mM). Junto a cada gráfica se indica el catión divalente ensayado. Los distintos colores corresponden con las reacciones enzimáticas de las distintas ratios molares Mg²⁺: catión alternativo. El análisis se realizó con *Graphpad* y 4 réplicas. Las barras de error correspondientes se representan en negro. La X de la leyenda hace referencia al Mn, Ni, Ca o Cu según corresponda para cada ensayo.

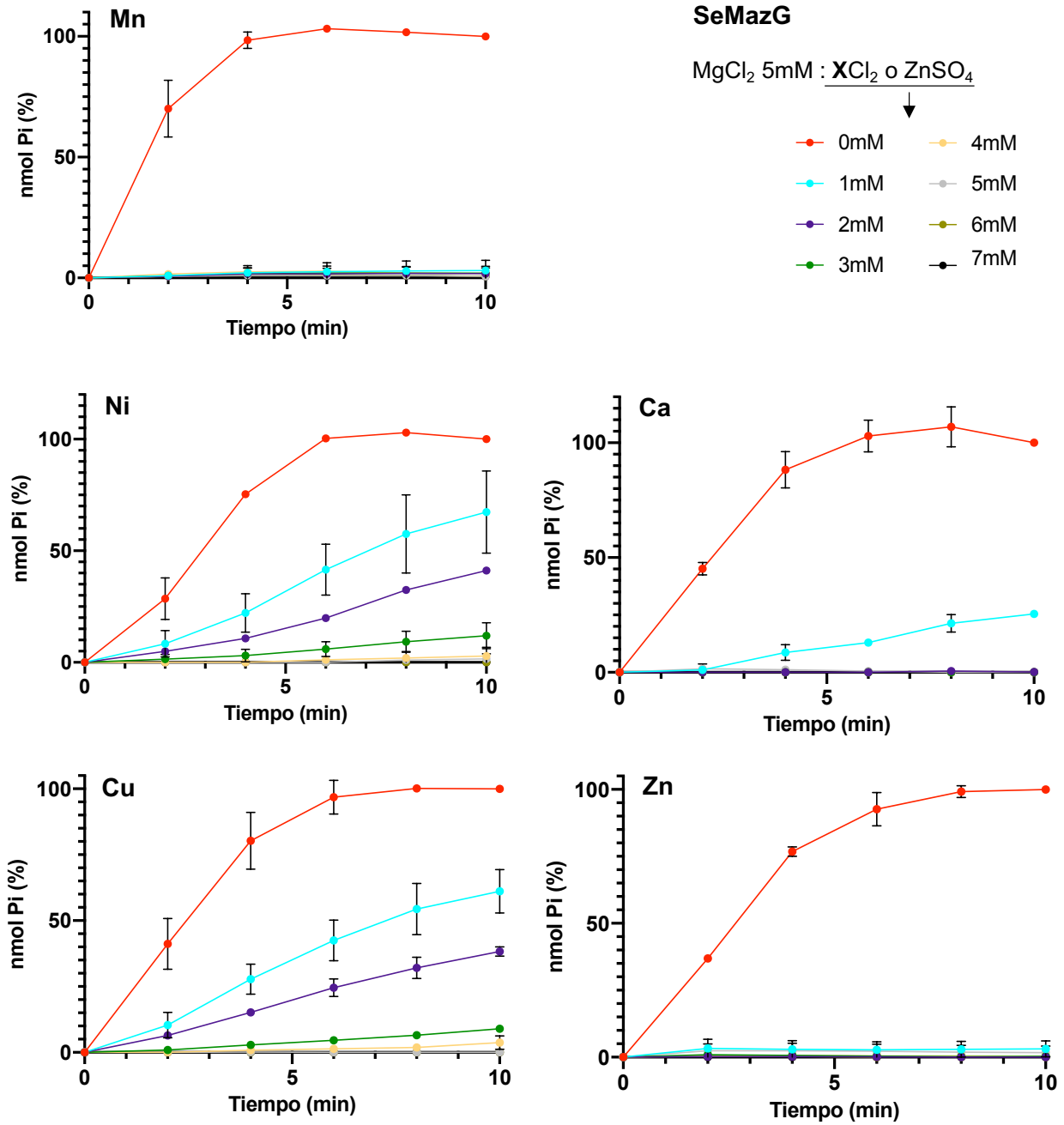


Figura R. 2.8. Perturbación de la actividad enzimática de SeMazG en presencia de Mg^{2+} y cationes divalentes alternativos en ratios molares crecientes. La concentración de $MgCl_2$ es constante (5 mM). Junto a cada gráfica se indica el catión divalente ensayado. Los distintos colores corresponden con las reacciones enzimáticas de las distintas ratios molares Mg^{2+} : catión alternativo. El análisis se realizó con *Graphpad* y 4 réplicas. Las barras de error correspondientes se representan en negro. La X de la leyenda hace referencia al Mn, Ni, Ca o Cu según corresponda para cada ensayo.

Estos resultados muestran que distintas concentraciones de distintos cationes son requeridas para obtener un mismo efecto en la enzima. Es decir, la afinidad de cada catión por el complejo MazG-dUTP es distinta.

2.2.3. Modulación de la actividad de EnMazG y SeMazG por cationes divalentes.

Con la finalidad de comprender la inhibición diferencial observada, se repitieron los ensayos previos pero incluyendo una relación de ratios molares Mg^{2+} : catión competidor mucho más amplia. Estas ratios abarcaron desde una concentración 16 veces superior del Mg^{2+} que del catión divalente alternativo, hasta 5 veces superior la concentración del ion alternativo a la del Mg^{2+} . Específicamente, las ratios molares utilizadas fueron $Mg^{2+} : X \rightarrow 1 : 0; 16 : 1; 8 : 1; 4 : 1; 2 : 1; 1 : 1; 1 : 2,5$ y $1 : 5$, donde **X** es el catión divalente alternativo correspondiente. En estos ensayos, además de los iones probados anteriormente, se incluyó el cobalto, el cual posee la misma geometría de coordinación del Mg^{2+} y se planteó como otra posible opción de sustitución o complementación al magnesio. Las concentraciones de enzima y sustrato se mantuvieron como en los experimentos anteriores, mientras que las concentraciones de los cationes fueron modificadas. El $MgCl_2$ fue utilizado a 1 mM, la concentración mínima para ambas enzimas con la que realizan actividad enzimática. Las concentraciones de los cationes alternativos fueron menores que las utilizadas en los experimentos anteriores, y en ratio molar respecto al $MgCl_2$. En las figuras R. 2. 9 y R. 2. 10 se incluyen los resultados de 4 a 6 réplicas de cada ensayo, representados como en los ensayos anteriores.

EnMazG

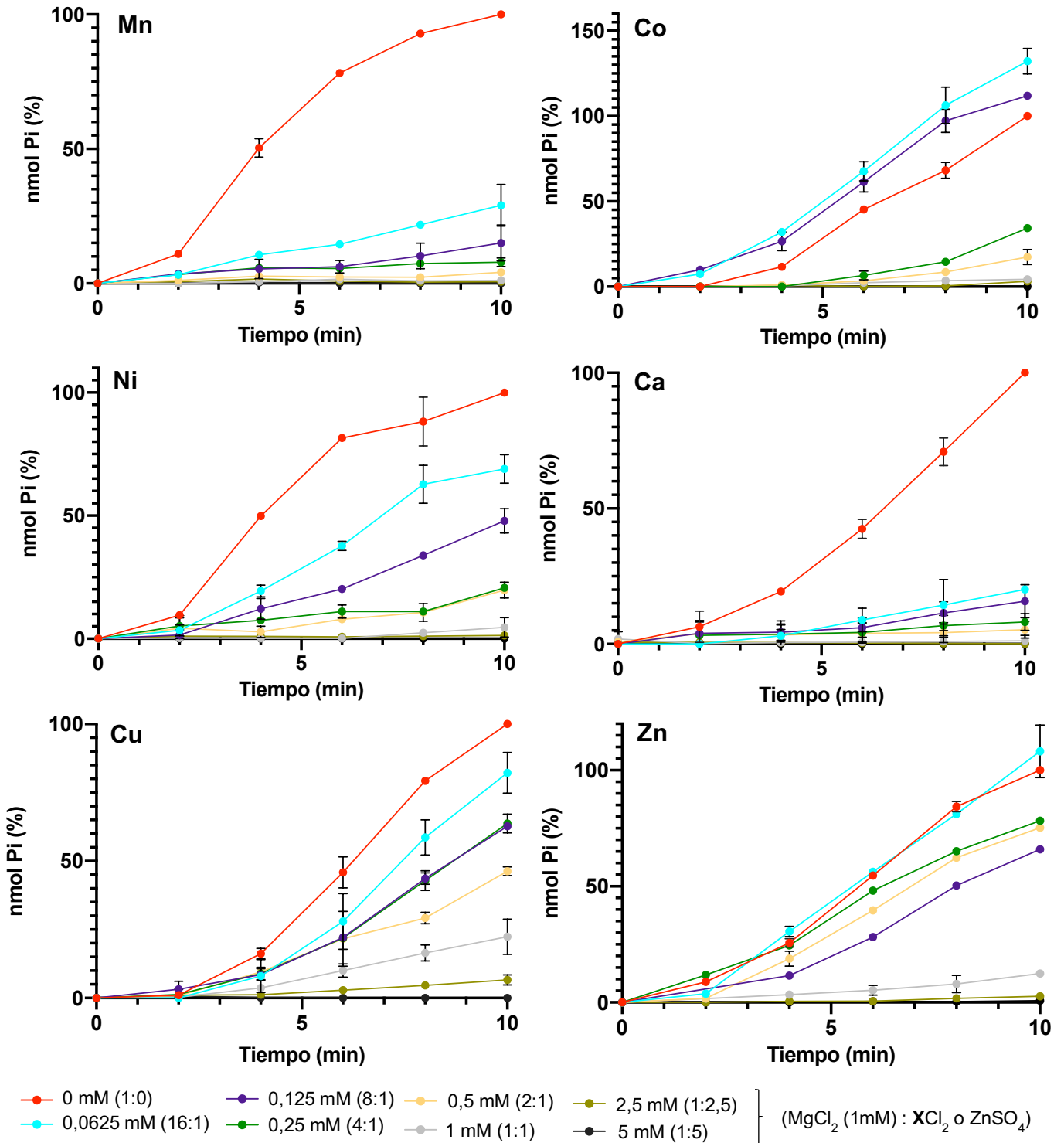


Figura R. 2.9. Actividad enzimática de EnMazG en presencia de Mg²⁺ en combinación con otros cationes divalentes en ratios molares crecientes. La concentración de MgCl₂ es constante (1 mM). Junto a cada gráfica se indica el catión divalente ensayado. Los distintos colores corresponden con los ensayos de las distintas ratios molares Mg²⁺: catión alternativo. El análisis de 4 a 6 réplicas se realizó con *Graphpad*. Las barras de error correspondientes se representan en negro. La X de la leyenda hace referencia al Mn, Ni, Ca, Co o Cu según corresponda para cada ensayo.

SeMazG

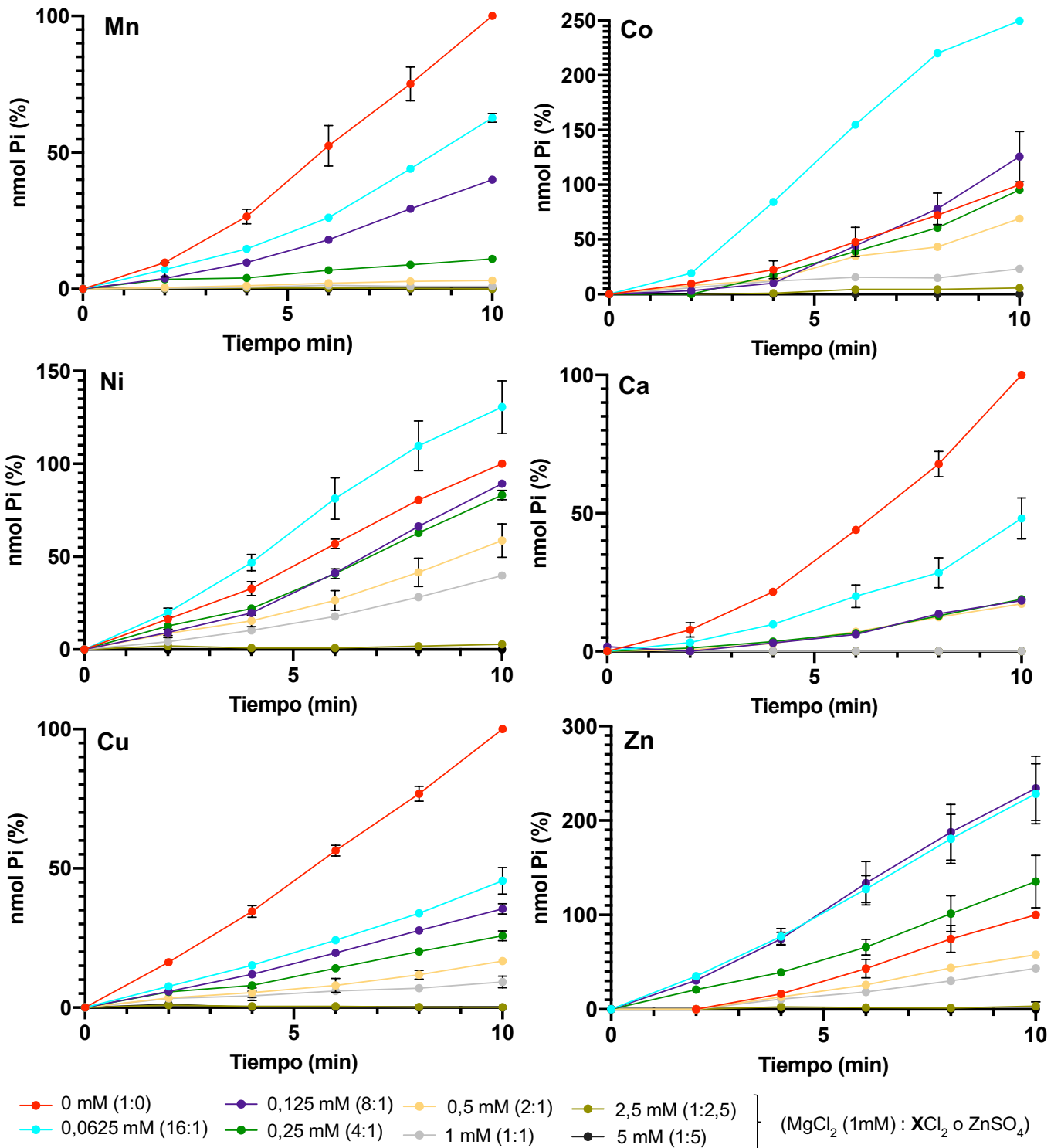


Figura R. 2.10. Actividad enzimática de SeMazG en presencia de Mg²⁺ en combinación con otros cationes divalentes en ratios molares crecientes. La concentración de MgCl₂ es constante (1 mM). Junto a cada gráfica se indica el catión divalente ensayado. Los distintos colores corresponden con los ensayos de las distintas ratios molares Mg²⁺: catión alternativo. El análisis de 4 a 6 réplicas se realizó con *Graphpad*. Las barras de error correspondientes se representan en negro. La X de la leyenda hace referencia al Mn, Ni, Ca, Co o Cu según corresponda para cada ensayo.

Estos ensayos permitieron ver grandes diferencias entre los distintos cationes alternativos. Mientras que a bajas concentraciones (16 veces menos que el Mg^{2+}) los cationes de Co^{2+} y Zn^{2+} incrementan la actividad enzimática de ambas MazG, llegando incluso a doblar la producción de nmol de Pi; otros como el Ca^{2+} y el Mn^{2+} inhiben la actividad enzimática de ambas, llegando a una disminución incluso de casi el 50 % de la actividad.

Con la finalidad de analizar cuantitativamente estos resultados se calcularon los índices IC_{50} , determinando la concentración de cada catión capaz de inhibir cada enzima al 50 % (Figuras R. 2. 11 y R. 2. 12).

En el caso de EnMazG, los cationes divalentes del Ca y Mn muestran valores IC_{50} de 0,019 y 0,024 mM respectivamente. Los cationes de Co y Zn complementan inicialmente al Mg^{2+} pero finalmente inhiben la enzima, necesitándose unas concentraciones de 0,296 y 0,498 mM respectivamente. Por su parte, el Ni^{2+} inhibe cuatro veces peor que el Mn^{2+} (IC_{50} 0,106 mM), mientras que el Cu^{2+} muestra una inhibición general similar al Co^{2+} (IC_{50} 0,339 y 0,296 mM, respectivamente) (Figura R. 2.11). Por tanto, la graduación decreciente en la capacidad inhibitoria para los cationes divalentes sería $Ca^{2+} = Mn^{2+}$, Ni^{2+} , $Co^{2+} = Cu^{2+}$, y Zn^{2+} . Por su parte, SeMazG es también inhibida al 50 % de manera similar por el Ca^{2+} y el Mn^{2+} , pero también por el Cu^{2+} , los cuales presentan unos valores IC_{50} de 0,051, 0,069 y 0,067 mM respectivamente. El resto de los iones probados, Co^{2+} , Ni^{2+} y Zn^{2+} , también complementan inicialmente al Mg^{2+} pero finalmente inhiben la actividad enzimática de la MazG, mostrando unos valores IC_{50} de 0,863, 0,706, y 1 mM respectivamente (Figura R. 2.12). La IC_{50} del catión Zn^{2+} calculada por ajuste al modelo (1,759 mM) no es fiable debido al valor extremadamente bajo del R^2 (0,2687), por lo cual se ha calculado el valor aproximado en base a los resultados reales (1 mM). Con estos resultados, la graduación decreciente en la capacidad inhibitoria para los cationes divalentes en SeMazG sería $Ca^{2+} = Mn^{2+} = Cu^{2+}$, $Co^{2+} = Ni^{2+}$, y Zn^{2+} . Por tanto, ambas MazG son inhibidas al 50 % por el calcio y el manganeso con un valor IC_{50} un orden de magnitud inferior que en caso del cobalto y níquel en SeMazG, y del cobalto, zinc y cobre en EnMazG. Es decir, para un mismo efecto inhibitorio, los cationes divalentes de cobalto, cobre, zinc y níquel requieren una concentración 10 veces superior a la del manganeso y el calcio.

Si se comparan las actividades enzimáticas de ambas MazG, a excepción del Cu^{2+} , los valores de IC_{50} para todos los cationes son de 2,5 a 6,5 veces mayores en SeMazG que en EnMazG. Debido a la gran similitud descrita anteriormente entre ambas enzimas (Ciges-Tomas 2019), se esperaría que el comportamiento enzimático en ambas fuera parecido. No obstante, esto podría ser consecuencia de una mayor plasticidad de SeMazG respecto a EnMazG. Ésta permitiría a SeMazG adaptarse a una configuración del centro catalítico que sea activa aun en presencia de estos cationes a bajas concentraciones en combinación con Mg^{2+} .

EnMazG

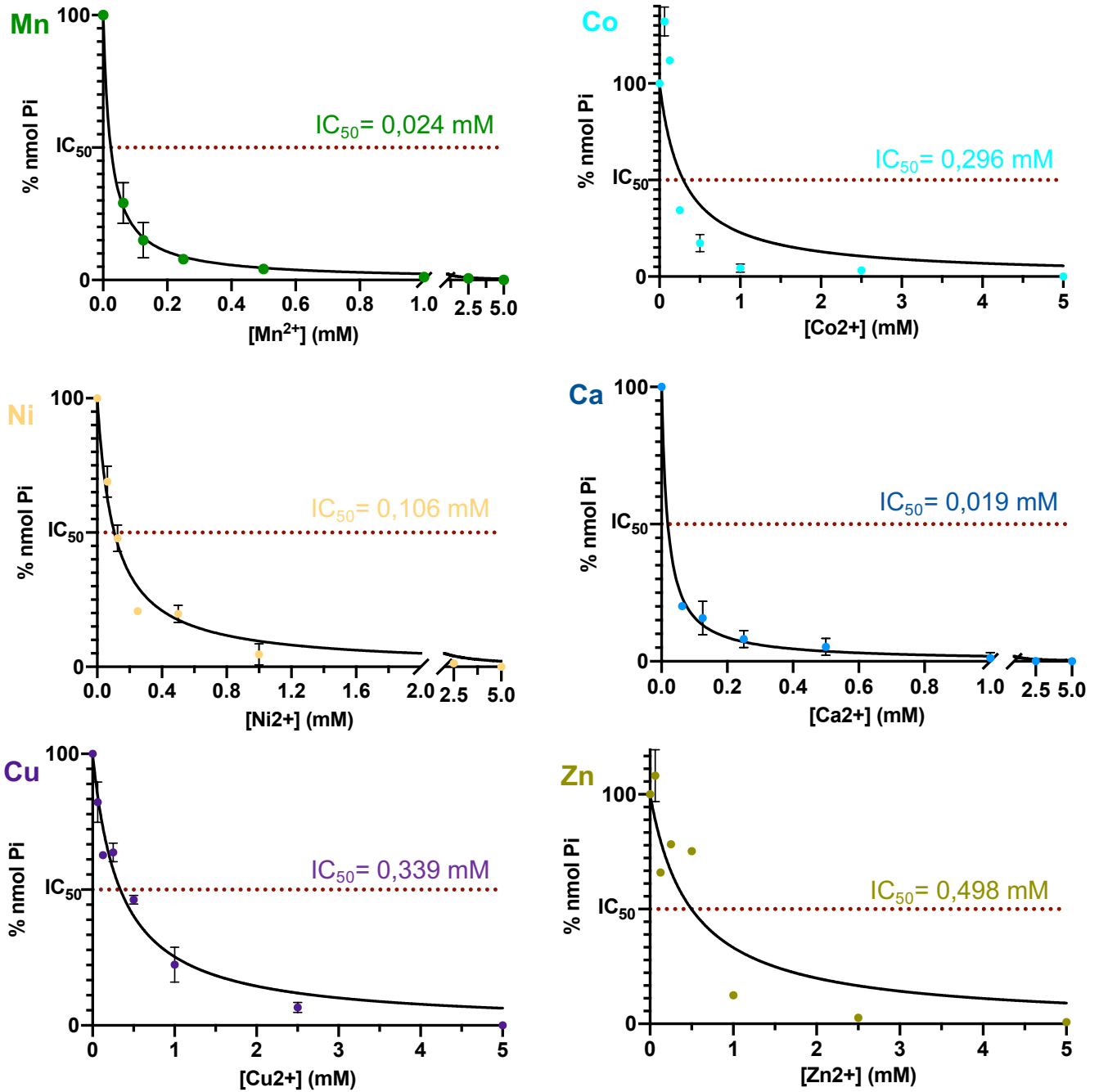


Figura R. 2. 11. Inhibición de la actividad catalítica de EnMazG en presencia de distintos cationes divalentes. Se representa el porcentaje de nmol Pi producidos a tiempo final normalizado a la correspondiente producción en la sola presencia del Mg²⁺, en función de concentraciones crecientes de distintos cationes divalentes: Mn²⁺ (verde), Co²⁺ (cian), Ni²⁺ (amarillo), Ca²⁺ (azul), Cu²⁺ (morado) y Zn²⁺ (verde militar). Junto a cada gráfica se indica el correspondiente valor IC₅₀ del catión. Estos resultados provienen de los experimentos realizados en presencia de una concentración fija de MgCl₂ (1mM).

SeMazG

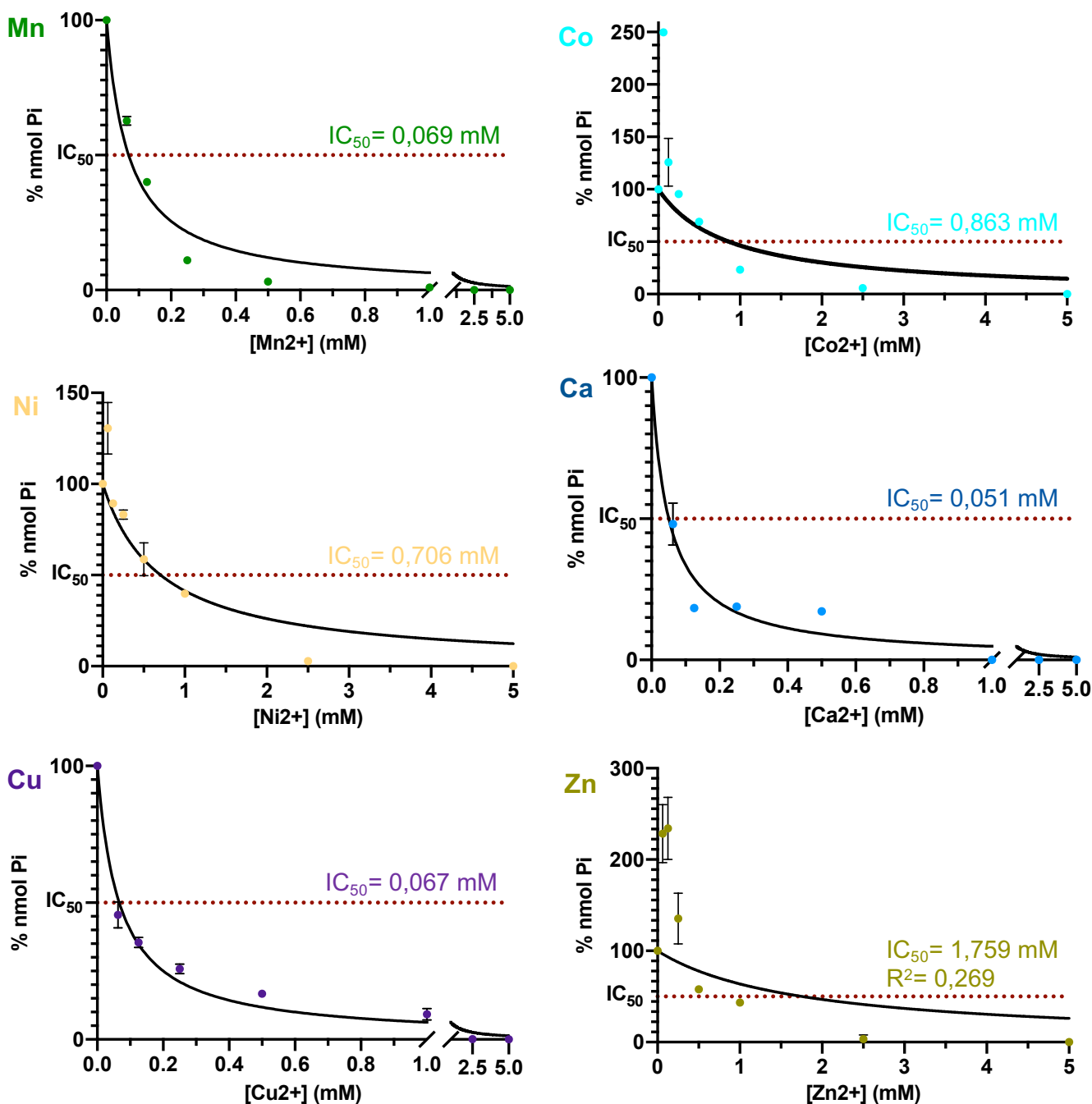


Figura R. 2. 12. Inhibición de la actividad catalítica de SeMazG en presencia de distintos cationes divalentes. Se representa el porcentaje de nmol Pi producidos a tiempo final normalizado a la correspondiente producción en la sola presencia del Mg²⁺, en función de concentraciones crecientes de distintos cationes divalentes: Mn²⁺ (verde), Co²⁺ (cian), Ni²⁺ (amarillo), Ca²⁺ (azul), Cu²⁺ (morado) y Zn²⁺ (verde militar). Junto a cada gráfica se indica el correspondiente valor IC₅₀ del catión. Estos resultados provienen de los experimentos realizados en presencia de una concentración fija de MgCl₂ (1mM).

Estas fuertes diferencias mostradas para cada catión plantearon la posibilidad que estos iones tengan un papel biológico relevante en la modulación de la actividad enzimática de estas MazG. Dado que estas enzimas tienen 3 sitios de unión no equivalente en el centro activo para estos cationes divalentes, la pregunta radicaba en si estas diferencias observadas se debían a una ocupación preferencial y/o específica de catión por cada uno de estos sitios; o era totalmente independiente, produciendo la entrada de uno solo de estos cationes alternativos el efecto sobre la actividad enzimática. Asimismo, no podía ser descartada la opción de que los efectos causados por estos iones fueran consecuencia de la unión a la enzima por sitios alostéricos alternativos al centro activo, compitiendo o no con el magnesio.

En el caso de que los iones alternativos sustituyeran al Mg^{2+} en el centro activo, se planteó la hipótesis de que su entrada en el centro activo, ya fuera total o parcialmente en las tres posiciones, podría provocar un cambio de conformación en el nucleótido que impidiera la hidrólisis. Esta hipótesis se sustentó en las observaciones realizadas en la estructura de EnMazG en presencia de dUpNHpp y Mg^{2+} obtenida previamente en el laboratorio (Ciges-Tomas 2019). Aunque se ha detallado en la introducción, recalcaremos la participación de tres cationes divalentes en la catálisis de EnMazG y la doble conformación del nucleótido (Figura R. 2. 13). La principal diferencia entre las conformaciones reside en la disposición del fosfato γ , donde la conformación "a" (γP -a) permite el correcto posicionamiento de la molécula de agua catalítica, mientras que la conformación "b" (γP -b) lo impide. Esto implicaría que la conformación "a" (γP -a) fuera catalíticamente activa, mientras que la "b" (γP -b) fuera catalíticamente inactiva.

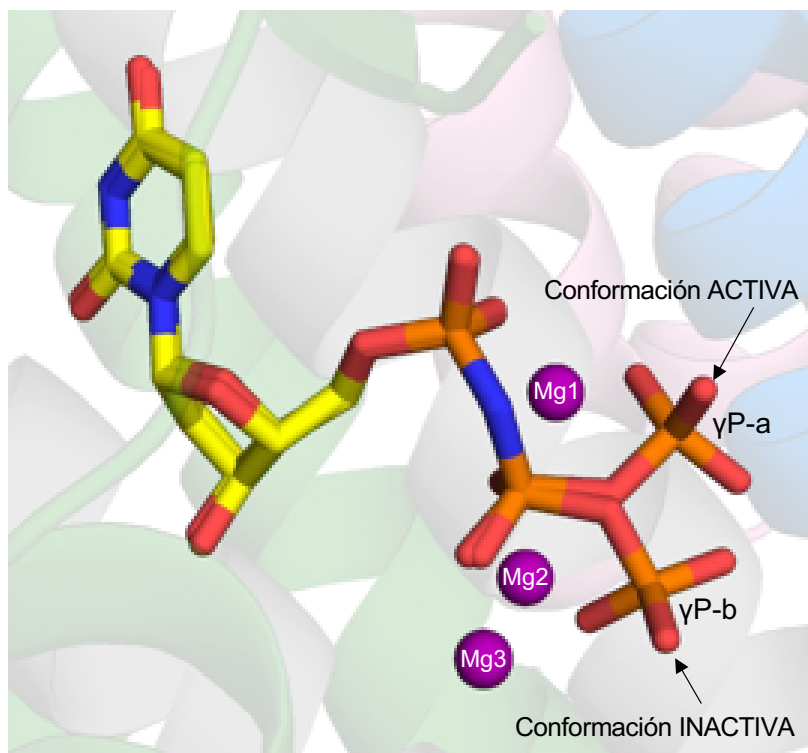


Figura R. 2. 13. Doble conformación del fosfato y observada en el nucleótido dUTP en algunos centros activos de la estructura tridimensional de EnMazG en presencia de Mg^{2+} . Se representa el nucleótido en sus dos conformaciones, identificando la catalíticamente activa (γP -a) y la inactiva (γP -b); así como los 3 cationes divalentes Mg^{2+} participantes en esferas moradas. No se representa la molécula de agua catalítica. El nucleótido está coloreado según el átomo: C en amarillo, P en naranja, O en rojo y N en azul. Estructura tridimensional tomada de *Ciges-Tomas 2019*.

2.2.4. El análisis de las estructuras tridimensionales de EnMazG en presencia de cationes divalentes permite proponer un mecanismo de modulación por el catión.

Con el objetivo de conocer el mecanismo molecular por el cual distintos cationes pueden modular la actividad catalítica de estas MazG, se llevó a cabo una aproximación estructural para visualizar a ambas enzimas en presencia de los distintos cationes.

Para ello, se utilizaron dos estrategias alternativas y complementarias: la co-cristalización y el *soaking*. En la co-cristalización, los cristales de proteína crecieron en presencia del análogo no hidrolizable dUpNHpp y de cationes alternativos (tanto de forma individual como en combinación con $MgCl_2$). En el *soaking*, los cristales crecieron en forma *apo* o en presencia de bajas concentraciones de $MgCl_2$ y de dUpNHpp, y una vez formados los cristales fueron sumergidos en distintas soluciones que contenían los cationes y el

sustrato no hidrolizable, o ratios molares conocidos de Mg^{2+} y otro catión, respectivamente, con la finalidad de que el nucleótido y los cationes difundieran dentro del cristal, y en el caso de que hubiera cationes en el centro activo, que éstos fueran desplazados por los presentes en las soluciones. Tras distintos tiempos de inmersión, los cristales fueron criocongelados.

La co-cristalización en *screenings* de condiciones fue elegida como primera opción para la obtención de las estructuras tridimensionales, ya que se presupuso que las diferencias en la actividad enzimática podrían reflejarse en cambios estructurales que, asimismo, podrían requerir de nuevas condiciones de cristalización. Para maximizar las opciones de éxito, por una parte se pusieron a cristalizar las proteínas en presencia de dUpNHpp y distintas concentraciones de los iones a estudiar, y por otra se sustituyó el análogo de sustrato por la deoxiuridina (dU) y el trifloruro de aluminio (AlF_3), o por la deoxiuridina monofosfato (dUMP) y el trifloruro de berilio (BeF_3), los cuales mimetizan el estado de transición, o por el dUMP y el pirofosfato (PPi) para mimetizar el estado post-hidrólisis.

En cuanto a los cationes, también se realizaron diferentes aproximaciones. Primeramente se fijó la concentración de uno de los cationes a 5 mM para asegurar una saturación de éste, mientras que para el segundo catión se realizaban cribados a concentraciones crecientes para poder ver el desplazamiento del catión. En el caso del $CoCl_2$, ambas MazG presentaron problemas de solubilidad a estas concentraciones, de modo que se planteó un segundo abordaje en el cual la concentración máxima total de cationes en la muestra para cristalizar se fijó a 3 mM. Ésta se alcanzaba por el conjunto de cationes participantes a diferentes ratios molares. No obstante, los problemas de solubilidad persistieron y se abordó un doble tercer planteamiento. En una de las opciones la concentración máxima total de los cationes se redujo a 1 mM. En la otra aproximación, el catión con la mayor concentración en la ratio molar a probar se fijó a 1 mM y la del otro catión era recalculada. Además de las muestras para cristalizar donde se incluían combinaciones de cationes a diferentes relaciones molares, se prepararon muestras en presencia de cada uno de los cationes en solitario (5 mM en todos los cationes a excepción del $CoCl_2$, en el cual se empleó una concentración de 1 mM).

La co-cristalización en las diferentes ratios molares de cationes fue intentada tanto en cribados de nuevas condiciones de cristalización, cribados de nuevas condiciones enriquecidos con una condición óptima seleccionada, y cribados utilizando condiciones ampliadas de las ya conocidas para obtener los cristales de ambas MazG en presencia de dUpNHpp- Mg^{2+} .

Debido a los problemas de solubilidad observados en el Co^{2+} , se decidió sustituir la co-cristalización por la técnica de *soaking* o empapado. Para ello, cristales de ambas proteínas crecidos en forma *apo* o en presencia de bajas concentraciones

de MgCl_2 y el análogo no hidrolizable dUpNHpp, fueron sumergidos en la solución en la que había crecido el cristal (solución madre) suplementada con el sustrato no hidrolizable y el Co^{2+} o con combinaciones de $\text{Mg}^{2+}:\text{Co}^{2+}$, o solo suplementada con los cationes, respectivamente. Desgraciadamente los cristales eran altamente inestables en estas condiciones, incluso a muy cortos periodos de tiempo (pocos segundos), apareciendo rápidamente fisuras en ellos y disolviéndose antes de un minuto. Los cristales supervivientes eran entonces sumergidos brevemente en una solución crioprotectora antes de ser congelados rápidamente en nitrógeno líquido para ser difractados *a posteriori*.

Tras la gran cantidad de intentos y posterior optimización, se consiguieron cristales de EnMazG en combinaciones de cationes de Mg^{2+} con Mn^{2+} , Ca^{2+} , Co^{2+} , Ni^{2+} y Zn^{2+} ; mientras que solo en presencia de MnCl_2 como catión en solitario. En cuanto a la SeMazG, solo pudieron obtenerse cristales con combinaciones de Mg^{2+} y Mn^{2+} . Cientos de ellos fueron criocongelados y testados en distintos sincrotrones, aunque un elevado porcentaje resultó en una difracción de baja resolución o sin difracción. En el caso de los cristales de Co^{2+} sometidos a *soaking*, aunque se intentó optimizar la estrategia variando concentraciones de cationes, soluciones madre y tiempos de incubación, aquellos cristales que no sufrían daño macroscópico mostraron nula o muy baja capacidad de difracción (7 Å), indicando una desestructuración interna.

2.2.4.1. El uso de la difracción anómala para completar la información obtenida con la difracción nativa.

El objetivo de obtener las estructuras tridimensionales en presencia de los distintos cationes en solitario o en combinación con el Mg^{2+} era elucidar el mecanismo molecular de modulación de la actividad enzimática de estas MazG de fago. Se plantearon tres hipótesis para explicar esta modulación:

- la entrada de cationes divalentes distintos al Mg^{2+} provocan un cambio estructural que afectan al centro activo;
- la entrada de los cationes alternativos provoca leves desplazamientos en el posicionamiento de las cadenas laterales de los residuos catalíticos, y con ello afectan a la hidrólisis del dUTP;
- los cationes alternativos se posicionan fuera del centro activo y provocan un efecto alostérico sobre la reacción enzimática.

La resolución de la estructura tridimensional podría ser suficiente para confirmar o descartar la primera hipótesis, pero en el caso de no ser consecuencia de un cambio tan drástico, se requeriría la difracción de los cristales en señal anómala para poder discernir entre cationes con geometría de coordinación y densidades electrónicas muy similares. Es decir, entre cationes que utilizan el mismo número de enlaces con otros elementos y en la misma distribución espacial para ser

estables, y que además la señal que define su posición tiene un tamaño similar, su identificación es más difícil. Por tanto, ante la posible estrategia planteada en la segunda y tercera hipótesis, se recurrió a la difracción anómala.

En la difracción anómala se excita a los electrones de un determinado átomo presente en la molécula, de modo que éstos cambian a una capa energética superior. Este cambio de capa energética implica que los pares de *Friedel* no sean iguales con símbolos opuestos, de modo que *a posteriori*, se calcula el mapa de diferencias anómalas que posiciona el origen de la señal anómala, que son los átomos de los cuales se han excitado los electrones. Es decir, permite colocar de manera precisa e inequívoca el átomo excitado en la estructura proteica. Para cada uno de los cristales testados se utilizó la energía de excitación correspondiente al catión alternativo contenido en él.

Tras la difracción de más de 300 cristales, la mejor aproximación al mecanismo de modulación fue la del catión Mn^{2+} , por ello nos centraremos primeramente en estos resultados, y dejaremos para más adelante las estructuras derivadas de cristales que contenían a otros cationes, cuyo efecto es más drástico.

2.2.4.2. El análisis de las estructuras tridimensionales de EnMazG en presencia del catión Mn^{2+} permite proponer un mecanismo de modulación.

Los diferentes ensayos de co-cristalización en presencia de concentraciones variables de Mn^{2+} fueron las más fructíferas a la hora de generar cristales con la capacidad de difracción requerida para nuestros ensayos (mejor de 3 Å de resolución). Se obtuvieron cristales de EnMazG en rangos molares $Mg^{2+}:Mn^{2+}$ desde 0,25 a 35 veces más de Mg^{2+} respecto al Mn^{2+} , y los cristales poseían principalmente tres morfologías: rombos, diamantes y varillas (Figura R. 2. 14). Las diferencias en morfología no se asociaron a las distintas ratios molares, pero sí con las condiciones de cristalización. Las condiciones que incluían polietilenglicol (PEG), indistintamente del porcentaje o número de moléculas de etilenglicol, favorecían la formación de cristales en forma de rombo y diamante; mientras que condiciones que incluían sulfato de litio o alguna sal de amonio, tanto en presencia como en ausencia de PEG, favorecían la formación de cristales en forma de varillas.

Los cristales seleccionados fueron difractados en sincrotrones, puesto que estas fuentes de luz permiten seleccionar la longitud de onda requerida para realizar ensayos de difracción anómala. En el caso del Mn^{2+} se seleccionó una longitud de onda cercana a 1.8961 Å, que corresponde a 6,55 KeV, la energía para el salto del *K-edge* de este elemento. Sin embargo, la longitud de onda exacta utilizada era determinada a través de un espectro de emisión de fluorescencia del cristal a difractar (Figura R. 2. 15), ya que el valor teórico del *K-edge* del

elemento puede variar en algunos eV por el entorno en que se encuentre. Una vez seleccionada la longitud de onda, el cristal era difractado en sus 360 ° independientemente del grupo espacial predicho, con la finalidad de obtener la máxima completitud y multiplicidad de los datos anómalos. La transmitancia del rayo X utilizada para la recogida era mayor respecto a la difracción no anómala, con la finalidad de aumentar la resolución y compensar la menor energía del rayo.

De todos los juegos de datos recogidos, finalmente se pudieron obtener estructuras a resoluciones adecuadas para EnMazG en presencia de solo manganeso y en ratios molares $Mn^{2+}:Mg^{2+} \rightarrow 1:2,5; 1:4; 1:8$. Desafortunadamente, todos los cristales testados de SeMazG, independientemente de la ratio molar de iones en que hubiesen sido obtenidos, no presentaron capacidad de difractar los rayos X o la resolución a la que lo hicieron era insuficiente para el estudio.

Los cristales de los que se obtuvieron las estructuras tridimensionales crecieron en las siguientes condiciones de cristalización:

- Ratio molar $Mn^{2+} : Mg^{2+} 1: 8 \rightarrow 170 \text{ mM Li}_2\text{SO}_4 + 100 \text{ mM Tris pH } 8,5 + 25,5 \text{ \% PEG } 4000 + 3 \text{ \% etanol}$.
- Ratio molar $Mn^{2+} : Mg^{2+} 1: 4 \rightarrow 150 \text{ mM Li}_2\text{SO}_4 + 100 \text{ mM Tris pH } 8,5 + 30 \text{ \% PEG } 300$.
- Ratio molar $Mn^{2+} : Mg^{2+} 1: 2,5 \rightarrow 15 \text{ \% etanol} + 100 \text{ mM Tris pH } 8,5 + 20 \text{ \% PEG } 300$.
- Solo $Mn^{2+} \rightarrow 30 \text{ \% propanol} + 100 \text{ mM Tris pH } 8,5 + 200 \text{ mM acetato de amonio}$.

A excepción del cristal del que se obtuvo la estructura con la ratio 1: 2,5, la morfología de estos cristales era de varillas. El cristal con la ratio 1: 2,5 tenía forma de diamante.

Los cristales se crioprotegieron antes de ser sumergidos en nitrógeno líquido para evitar roturas derivadas de la congelación con las siguientes soluciones:

- Ratio molar $Mn^{2+} : Mg^{2+} 1: 8 \rightarrow$ solución de cristalización suplementada con 10 % PEG 4000.
- Ratio molar $Mn^{2+} : Mg^{2+} 1: 4 \rightarrow$ la solución de cristalización ya era crioprotectora.
- Ratio molar $Mn^{2+} : Mg^{2+} 1: 2,5 \rightarrow$ solución de cristalización suplementada con 10 % PEG 400.
- Solo $Mn^{2+} \rightarrow$ solución de cristalización suplementada con 10 % glicerol + 10 % etilenglicol + 10 % PEG 300.

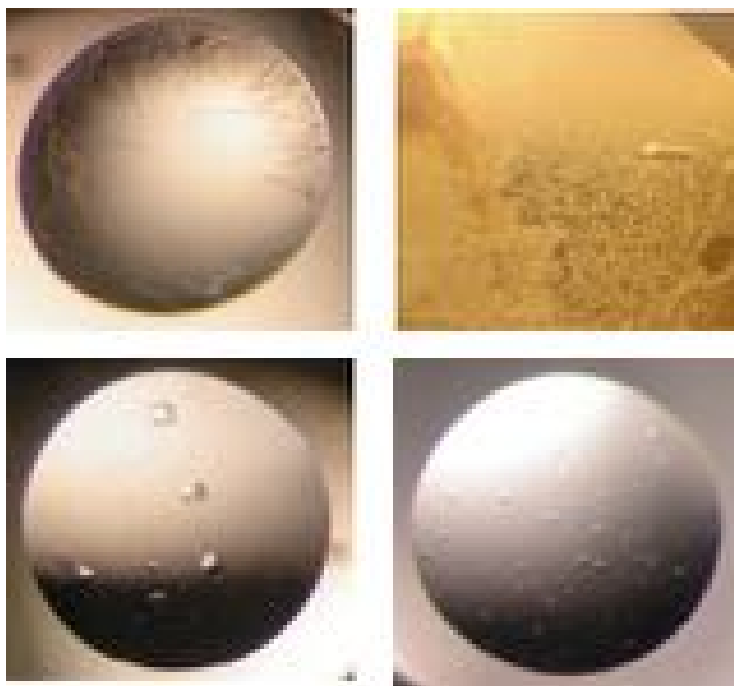


Figura R. 2. 14. Fotografías de los cristales de EnMazG en presencia de Mn^{2+} en distintas ratios molares con Mg^{2+} . Se obtuvieron en tres morfologías distintas: varillas (*arriba*), diamantes (*abajo izquierda*) y rombos (*abajo derecha*).

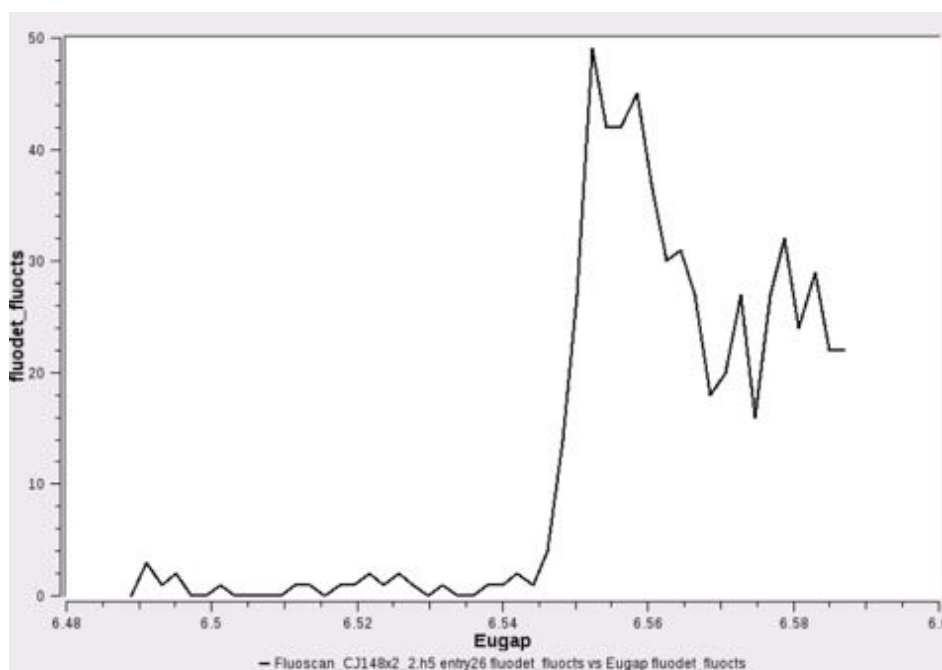


Figura R. 2. 15. Espectro de emisión de fluorescencia del Mn al excitar el cristal de proteína con un barrido de energías del haz desde 6,48 a 6,6 KeV en el sincrotrón. La energía máxima de excitación de los cationes de Mn^{2+} contenidos en el cristal se situó en 6,55 KeV, y esa energía fue la utilizada para difractar el cristal y recoger la correspondiente difracción (los 360 ° del cristal).

Tras difractar los cristales, integrar y escalar los datos recogidos, las estructuras fueron resueltas por reemplazo molecular, ya que se disponía del modelo de la estructura tridimensional de la proteína en conformación *apo* y unida al trinucleótido, ambas resueltas previamente por otro miembro del laboratorio (Ciges-Tomas 2019). Los parámetros cristalográficos de las distintas estructuras resueltas se recogen en la tabla R. 2.5.

Tabla R. 2.5. Tabla de colección de datos cristalográficos y estadísticos de refinado.

EnMazG	Mn ²⁺	Mn ²⁺ : Mg ²⁺ 1: 2,5	Mn ²⁺ : Mg ²⁺ 1: 4	Mn ²⁺ : Mg ²⁺ 1: 8
Datos de recogida				
Haz de luz	ALBA-XALOC	DLS-I24	DLS-I24	ALBA-XALOC
Energía del haz (Å)	1,892	1,896	1,896	1,892
Grupo espacial	C222 ₁	P2 ₁	P2 ₁	P2 ₁
Dimensiones de la celda (Å)	a=66,37 b=81,07 c=86,05 α=β=γ=90	a=55,14 b=86,47 c=105,04 α=γ=90 β=93,49	a=54,01 b=86,56 c=104,04 α=γ=90 β=93,76	a=52,41 b=86,35 c=103 α=γ=90 β=92,69
Resolución (Å) ^a	44,1 - 2.9 (3,004 - 2.9) ^a	50 - 2,8 (2.9 - 2,8) ^a	66,5 - 2,97 (3,076 - 2,97) ^a	66,14 - 2,57 (2,662 - 2,57) ^a
Reflexiones únicas	5334 (509)	24423 (2439)	19795 (1953)	29309 (2931)
Compleitud (%)	99,13 (98,83)	99,87 (99,71)	99,12 (98,68)	99,74 (99,97)
Multiplicidad	9,5 (10,1)	13 (12,4)	6,5 (6,2)	6,2 (6,1)
Compleitud anómalo (%)	98,9 (98,7)	99,8 (99,9)	99,5 (99,5)	97,9 (97,9)
Multiplicidad anómalo	5,1 (5,3)	6,5 (6,1)	3,3 (3,2)	3,1 (3,1)
Media I/σ (I)	8,4 (6,3)	7,8 (2,5)	6 (1,8)	5,6 (3)
R _{pim}	0,183 (0,672)	0,063 (0,240)	0,079 (0,326)	0,209 (0,580)
CC 1/2	0,848 (0,311)	0,997 (0,931)	0,995 (0,877)	0,932 (0,632)
Refinado				
R _{work}	0,1786	0,2007	0,2282	0,1983
R _{free}	0,2713	0,2716	0,3367	0,2843
Número de átomos	1623	6276	6107	6844
Proteína	1512	5891	5639	6050
Moléculas de agua	53	147	287	415
Otras moléculas	8 ^b	36 ^c	48 ^d	39 ^e
Rmsd, enlaces (Å)	0,007	0,014	0,014	0,015
Rmsd, ángulos (°)	1,38	1,83	1,91	1,90
Probabilidad de choque molecular	8,43	6,99	18,40	6,99
Diagrama de Ramachandran				
Preferidas (%)	97,96	98,59	88,68	98,34
Permitidas (%)	2,04	1,41	11,07	1,66
Outliers (%)	0	0	0,25	0

^aNúmero entre paréntesis indica los valores para la zona de alta resolución.

^bCorresponde con 2 moléculas de dUpNHpp y 6 cationes Mn²⁺.

^cCorresponde con 8 moléculas de dUpNHpp, 1 de dietilenglicol, 3 de etilenglicol y 24 cationes Mn²⁺.

^dCorresponde con 8 moléculas de dUpNHpp, 11 de dietilenglicol, 3 de etilenglicol, 7 cationes Mg²⁺ y 19 Mn²⁺.

^eCorresponde con 8 moléculas de dUpNHpp, 2 de dietilenglicol, 5 de etilenglicol, 5 cationes Mg²⁺ y 19 Mn²⁺.

Como puede verse en la tabla R. 2.5, los cristales de EnMazG obtenidos en presencia de una combinación $Mn^{2+}:Mg^{2+}$ pertenecen al grupo espacial ($P2_1$) y presentan tamaños de celda muy similar, mostrando 8 monómeros en la unidad asimétrica que se organizan en dos tetrámeros, indicando una misma estructura para todos ellos. Este grupo espacial es similar al obtenido en la estructura de EnMazG en presencia de dUpNHpp y cationes Mg^{2+} ($P2_12_12$) (Ciges-Tomas 2019), aunque con menor simetría. En ésta última, también la unidad asimétrica contiene 8 monómeros organizados en 2 tetrámeros, siendo los monómeros idénticos a excepción de la ordenación del C-terminal. No obstante, la celda unidad era distinta ($a= 109 \text{ \AA}$, $b= 87,31 \text{ \AA}$, $c= 105,35 \text{ \AA}$, $\alpha=\beta=\gamma= 90^\circ$). Por su parte, los cristales obtenidos en presencia únicamente de Mn^{2+} pertenecen al grupo espacial $C222_1$, albergando en su unidad asimétrica un dímero que formaba el tetrámero por simetría cristalográfica. Por tanto, la similitud con la estructura en presencia de dUpNHpp y cationes Mg^{2+} eran los tres ejes 2 en el grupo espacial y los ángulos de 90° de la celda unidad. En conjunto, no pudo asociarse el grupo espacial de los cristales con la morfología que poseían.

Pese a cristalizar en diferentes grupos espaciales y con diferentes celdas cristalinas, todos los tetrámeros de EnMazG eran casi idénticos como confirman los valores de RMSD inferiores a $0,6 \text{ \AA}$ para la superposición casi completa (a falta de 2 o 3 residuos en los extremos) de los tetrámeros de una misma y de distinta estructura. En el caso de la estructura de EnMazG en presencia del catión Mn^{2+} se generó el tetrámero por simetría cristalográfica y se analizó frente a los tetrámeros en el resto de las estructuras. Así mismo, el tetrámero de EnMazG en presencia del análogo no hidrolizable dUpNHpp y de Mg^{2+} (nuestra estructura de referencia) mostró también una elevada similitud en la superposición con estas estructuras, con valores de RMSD inferiores a $0,6 \text{ \AA}$ para la superposición de los tetrámeros sin el extremo C-terminal.

En la tabla R. 2. 6. se presentan todos los valores RMSD resultantes de la superposición de tetrámeros. En la primera fila se sitúan los tetrámeros de las estructuras de EnMazG en presencia del análogo de sustrato no hidrolizable y el Mg^{2+} ("Todo Mg^{2+} "), EnMazG en presencia del dUpNHpp y el Mn^{2+} ("Todo Mn^{2+} "), EnMazG en presencia del análogo de sustrato no hidrolizable y la ratio molar $Mn^{2+}: Mg^{2+} 1: 2,5$ ("Mn²⁺: Mg²⁺ 1: 2,5"), EnMazG en presencia del dUpNHpp y la ratio molar $Mn^{2+}: Mg^{2+} 1: 4$ ("Mn²⁺: Mg²⁺ 1: 4"), y EnMazG en presencia del dUpNHpp y la ratio molar $Mn^{2+}: Mg^{2+} 1: 8$ ("Mn²⁺: Mg²⁺ 1: 8"). Cada tetrámero se identifica como ABCD, o A'B'C'D' en su defecto si hay 2 tetrámeros en la unidad asimétrica. En la columna de la izquierda se incluyen las mismas estructuras que en la primera fila.

Tabla R. 2. 6. Tabla de valores de RMSD en Å para la superposición de los tetrámeros de EnMazG en complejo con el dUpNHpp y cationes Mg^{2+} , Mn^{2+} , y las ratios molares $Mn^{2+}: Mg^{2+}$ 1: 2,5, 1:4, y 1:8. Entre paréntesis se indica el número de residuos superpuestos.

EnMazG RMSD		Todo Mg^{2+}	Todo Mn^{2+}	$Mn^{2+}: Mg^{2+}$ 1: 2,5		$Mn^{2+}: Mg^{2+}$ 1: 4		$Mn^{2+}: Mg^{2+}$ 1: 8	
		ABCD	ABCD	ABCD	A'B'C'D'	ABCD	A'B'C'D'	ABCD	A'B'C'D'
Todo Mg^{2+}	ABCD	-	0,450 (378)	0,420 (379)	0,400 (378)	0,532 (378)	0,590 (379)	0,412 (379)	0,437 (379)
Todo Mn^{2+}	ABCD	-	-	0,346 (396)	0,356 (396)	0,500 (395)	0,490 (396)	0,311 (399)	0,363 (398)
$Mn^{2+}: Mg^{2+}$ 1: 2,5	ABCD	-	-	-	0,281 (396)	0,580 (398)	0,442 (396)	0,311 (397)	0,300 (395)
	A'B'C'D'	-	-	-	-	0,532 (396)	0,440 (395)	0,319 (396)	0,346 (395)
$Mn^{2+}: Mg^{2+}$ 1: 4	ABCD	-	-	-	-	-	0,513 (398)	0,517 (397)	0,515 (397)
	A'B'C'D'	-	-	-	-	-	-	0,479 (397)	0,453 (396)
$Mn^{2+}: Mg^{2+}$ 1: 8	ABCD	-	-	-	-	-	-	-	0,282 (399)
	A'B'C'D'	-	-	-	-	-	-	-	-

Una vez analizadas las superposiciones de las estructuras (Tabla R. 2. 6 y Figura R. 2. 16), que revelaron una elevada similitud entre ellas, se realizó el análisis de los centros activos, principales poseedores de la respuesta a la hipótesis planteada.

En los centros activos de las cuatro estructuras fue posible localizar los tres cationes divalentes debido a la fuerte densidad electrónica y al conocimiento previo de su posición en la estructura en presencia de Mg^{2+} . Aunque no siempre ocupaban las mismas posiciones, el análisis de la señal anómala nos permitió conocer qué posiciones eran ocupadas preferencialmente por cada tipo de ión, Mg^{2+} o Mn^{2+} . Este análisis se realizó con el software “*Finding and analysing difference map peaks*” en Phenix, especificando un valor de sigma 3 como punto de corte inferior de señal anómala (baja señal anómala es indicativa de Mg^{2+} , mientras que una señal alta es indicativa de Mn^{2+}) (Liebschner et al. 2019). En la tabla R. 2. 7 se presentan los valores de señal anómala para cada catión en las 4 estructuras de EnMazG resueltas en esta Tesis, especificándola en cada posición de las tres en cada centro activo junto con la conformación activa o inactiva adoptada por el nucleótido. Cada centro activo se identifica con la misma letra del monómero que aporta los motivos de unión de catión (residuos E31, E32, E35 y residuos D50 y D54). Los monómeros que forman un dímero se presentan enfrentados en la tabla, mientras que los dímeros que forman un tetrámero se presentan en dos columnas adyacentes. Las posiciones vacías identifican a un catión sin señal anómala, es decir, un catión Mg^{2+} .

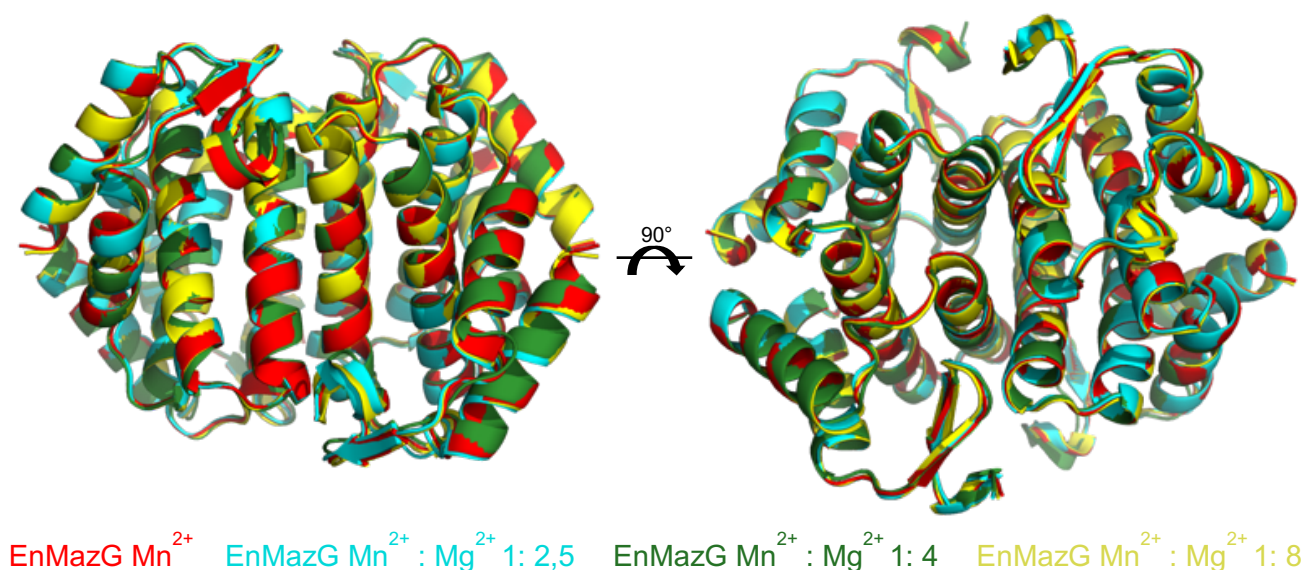


Figura R. 2. 16. Superposición de las cuatro estructuras de EnMazG en presencia de los cationes Mg²⁺ y Mn²⁺. Se muestran dos vistas ortogonales de las cuatro estructuras superpuestas: EnMazG en la sola presencia de Mn²⁺ en rojo, en presencia de la ratio molar Mn²⁺:Mg²⁺ 1 : 2,5 en cian, ratio molar 1 : 4 en verde, y ratio molar 1:8 en amarillo.

Aunque en la introducción ya se había comentado, haremos hincapié en la identificación exacta de cada una de las 3 posiciones ocupadas por cationes en el centro activo. El catión que ocupa la posición 1 es hexacoordinado por los 3 fosfatos del sustrato y los residuos del motivo de unión de metal E31, E32 y E35. El catión que ocupa la posición 2 también está hexacoordinado, participando el fosfato β , los residuos de motivos de unión de metal E32, E35 y D54, y dos moléculas de agua. Finalmente, el catión que ocupa la posición 3 es hexacoordinado mediante el fosfato β , los residuos D50 y D54, y tres moléculas de agua (Figura R. 2. 17, conformación activa). La numeración de estos cationes se sustenta principalmente en las estructuras de EnMazG en conformación *apo* (sin cationes), en presencia del producto dUMP (solo un catión y en la posición 1) y en presencia del análogo de sustrato no hidrolizable dUpNHpp (3 cationes) resueltas previamente en el laboratorio (Ciges-Tomas 2019), junto con otras estructuras de MazG disponibles en la bibliografía, en las que ahondaremos en el apartado de la discusión. Las estructuras previas de EnMazG permitieron asociar al catión 1 una función estructural (debido a su posicionamiento en presencia del sustrato como del producto), mientras que a los cationes 2 y 3 con un papel catalítico (posicionamiento solo en presencia del nucleótido).

Tabla R. 2. 7. Tablas recopilatorias de la señal anómala de cada catión en las cuatro estructuras de EnMazG en presencia de Mn^{2+} resueltas en esta Tesis. Las cuatro tablas se organizan de la siguiente manera: identificación del centro catalítico al que se refiere la señal de los cationes, la conformación adoptada por el nucleótido, y la señal anómala de cada uno de los cationes (encima o debajo de cada señal se identifica si corresponde con el catión de la posición 1 (P. 1), 2 (P. 2) o 3 (P.3). Los centros activos se identifican con la misma letra que la cadena que aporta los residuos de la péntada acídica (E31/32/35 y D50/54). En las tablas correspondientes a las ratios molares de Mn^{2+} y Mg^{2+} (*verde, azul y gris*) se representan enfrentados las cadenas que forman dímeros y en dos columnas contiguas las cuatro cadenas que forman un tetrámero. En el caso de la estructura con solo Mn^{2+} se presentan los resultados en dos columnas aunque ambos monómeros forman el dímero presente en la unidad asimétrica. Las posiciones con un guión identifican a un catión Mg^{2+} .

EnMazG Mn^{2+}					
Centro catalítico A Conf. INACTIVA			Centro catalítico B Conf. INACTIVA		
P. 1	P. 2	P. 3	P. 1	P. 2	P. 3
19,38	15,36	14,50	17,69	-	15,46

EnMazG $Mn^{2+} : Mg^{2+} 1 : 2,5$											
C. catalítico A Conf. INACTIVA			C. catalítico C Conf. INACTIVA			C. catalítico A' Conf. INACTIVA			C. catalítico C' Conf. INACTIVA		
P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3
9,97	11,97	9,13	9,66	7,65	8,91	9,18	8,01	7,01	10,31	10,00	9,54
10,95	9,71	9,44	11,04	12,67	8,82	10,21	9,30	7,46	9,42	9,97	7,25
P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3
C. catalítico B Conf. INACTIVA			C. catalítico D Conf. INACTIVA			C. catalítico B' Conf. INACTIVA			C. catalítico D' Conf. INACTIVA		

EnMazG $Mn^{2+} : Mg^{2+} 1 : 4$											
C. catalítico A Conf. INACTIVA			C. catalítico C Conf. INACTIVA			C. catalítico A' Conf. INACTIVA			C. catalítico C' Conf. INACTIVA		
P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3
7,98	8,28	5,35	7,49	7,23	4,24	9,96	7,47		7,13	7,62	5,13
6,83	6,83				6,27	6,16		6,52	8,95	5,98	
P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3
C. catalítico B Conf. INACTIVA			C. catalítico D Conf. INACTIVA			C. catalítico B' Conf. INACTIVA			C. catalítico D' Conf. INACTIVA		

EnMazG $Mn^{2+} : Mg^{2+} 1 : 8$											
C. catalítico A Conf. INACTIVA			C. catalítico C Conf. DOBLE			C. catalítico A' Conf. INACTIVA			C. catalítico C' Conf. INACTIVA		
P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3
10,52	7,41	3,84	12,88	7,70		12,25	7,41	3,51	12,31	6,84	4,58
9,33	6,51	3,47	11,80	10,81	6,90	12,12	11,01		11,62	8,97	4,05
P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3	P. 1	P. 2	P. 3
C. catalítico B Conf. INACTIVA			C. catalítico D Conf. INACTIVA			C. catalítico B' Conf. INACTIVA			C. catalítico D' Conf. INACTIVA		

La resolución de las cuatro estructuras de EnMazG en presencia de Mn^{2+} o en combinación con Mg^{2+} permitió confirmar los cambios en la disposición del centro activo, cuando el nucleótido adopta una conformación catalíticamente inactiva (Figura R. 2. 17) (Ciges-Tomas 2019). Con la superposición de un centro activo en conformación activa con uno en conformación inactiva, pudo deducirse que el paso de la conformación activa (γP -a) a inactiva (γP -b) corresponde básicamente con una rotación de unos 30° del fosfato γ sobre su enlace con el fosfato β , manteniéndose el resto del nucleótido fijo en ambas conformaciones. Debido a este movimiento, la molécula de agua catalítica no puede posicionarse en línea con el fosfato β para realizar el ataque nucleofílico, la cual originalmente se posiciona a 2,98 Å del fosfato β . Por tanto, la catálisis no puede tener lugar. Además de impedir el ataque, el movimiento del fosfato γ provoca el desplazamiento de otra molécula de agua adicional. Ésta, cuando el nucleótido adopta la conformación activa, se posiciona en el lugar que ocupa el fosfato γ y en conformación inactiva, ocurriendo lo mismo a la inversa cuando el nucleótido adopta la conformación inactiva. Quedan otras dos moléculas de agua en el centro activo, la coordinada por los residuos D50 y D54 que mantiene la misma posición, y la participante en la coordinación del agua catalítica y del catión que ocupa la posición 3, que modifica ligeramente su posición. En cuanto a los 3 iones, el posicionamiento es idéntico en las dos conformaciones, aunque los enlaces establecidos varían ligeramente. En la conformación activa el catión en posición 1 interacciona con los oxígenos de los tres fosfatos, mientras que los cationes de las posiciones 2 y 3 polarizan la molécula de agua catalítica. En la conformación inactiva el catión en posición 1 interacciona con los fosfatos α y β , mientras que el fosfato γ pasa a interaccionar con los cationes de las posiciones 2 y 3.

Aunque estos cambios habían sido vistos previamente en el laboratorio en centros activos que poseían simultáneamente ambas conformaciones (con una ocupancia del 50 %) en la estructura de EnMazG en complejo con el dUpNHpp y Mg^{2+} (Ciges-Tomas 2019), se ha confirmado con la superposición de los centros activos con las conformaciones individuales (Figura R. 2. 18).

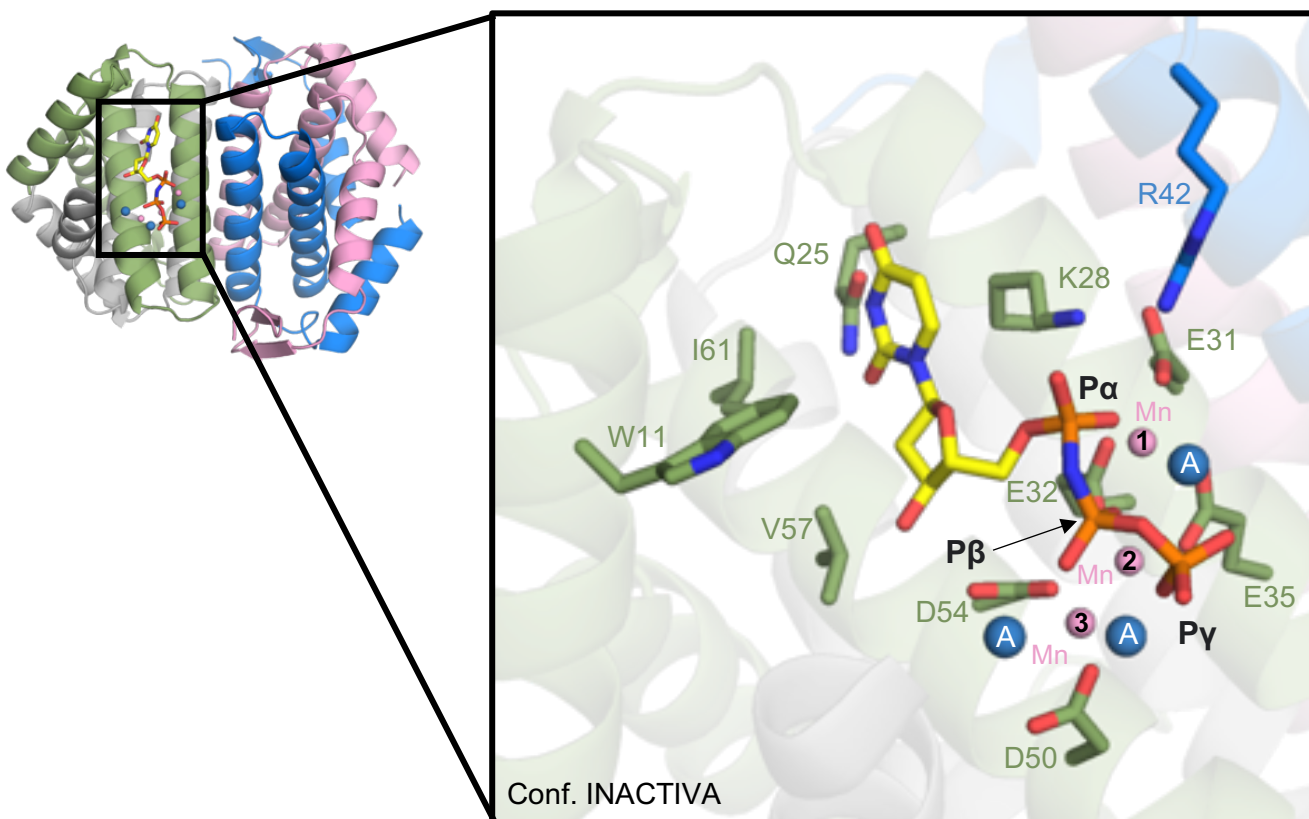
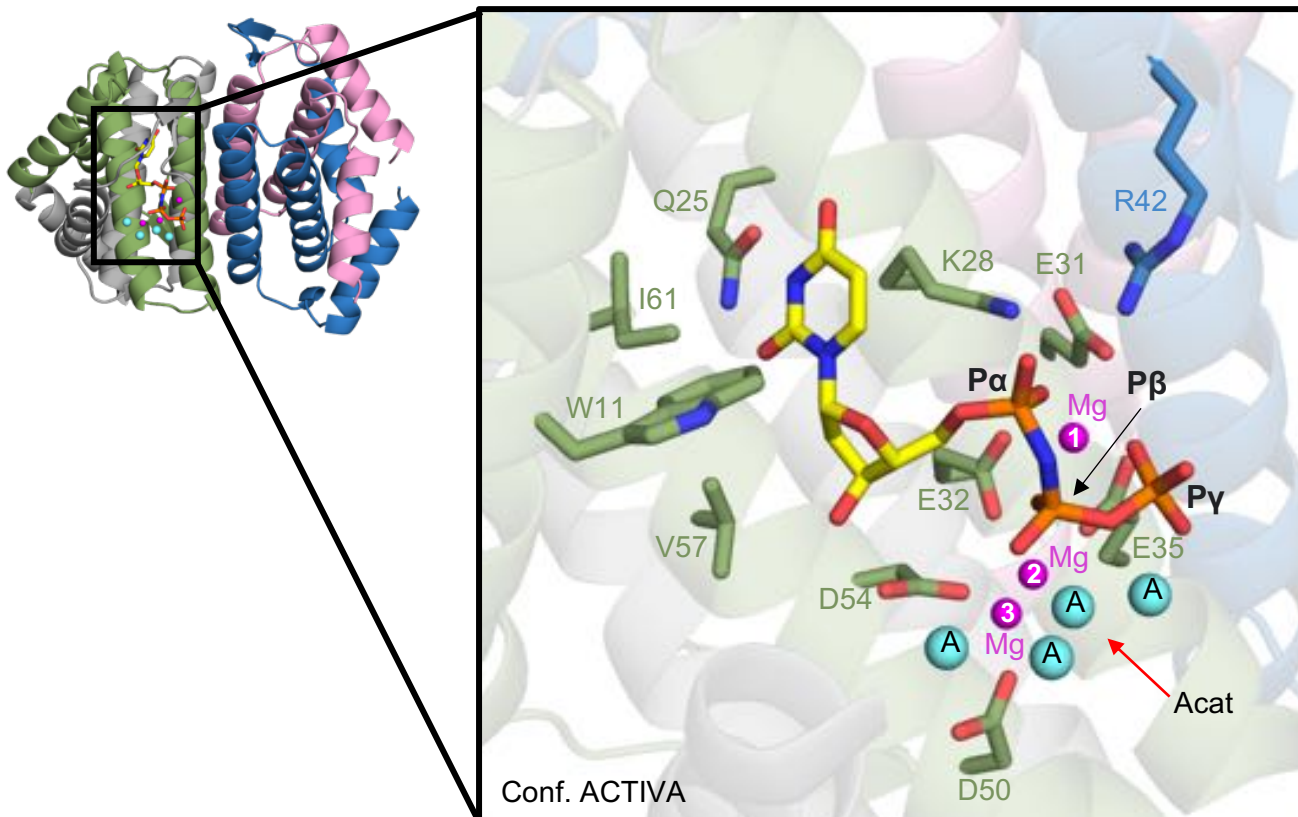


Figura R. 2. 17. Conformaciones activa e inactiva del nucleótido dUTP en el centro activo de EnMazG. Se representan en varillas las cadenas laterales de algunos de los residuos participantes en la catálisis en colores por átomos (C según el monómero al que pertenecen, O rojo, N en azul), así como el nucleótido (C en amarillo, N en azul, O en rojo, y P en naranja). En esferas se representan las moléculas de agua y los cationes (colores *cian* con una "A" y *magenta* con un número para el agua y el Mg^{2+} en conformación activa (*arriba*) y colores azul marino con una "A" y morado claro con un número para el agua y el Mn^{2+} en conformación inactiva (*abajo*), respectivamente). Los cationes son numerados según la posición que ocupan. Se identifican los tres fosfatos y el agua catalítica en la conformación activa. El PDB de la estructura para la imagen superior ha sido tomada de *Ciges-Tomas 2019*, mientras que la estructura para la imagen inferior ha sido resuelta en esta Tesis.

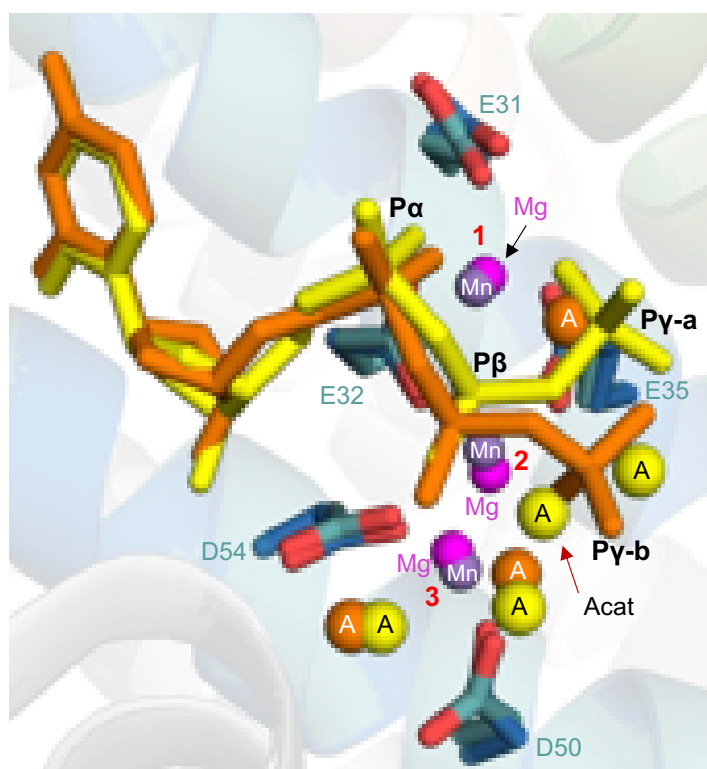


Figura R. 2. 18. Superposición de las dos conformaciones del nucleótido en EnMazG. Se superponen las estructuras de un centro catalítico de EnMazG en presencia de Mg^{2+} (conformación activa, Pγ-a) con un centro catalítico en presencia de Mn^{2+} (conformación inactiva, Pγ-b). Se representan en varillas los nucleótidos, en *amarillo* la conformación activa y en *naranja* la inactiva. También con el mismo código de color se representan en esferas las moléculas de agua (A), de modo que puede visualizarse como una molécula de agua se coloca en la posición de uno de los oxígenos del grupo fosfato y en la conformación alternativa a la que éste posee. Mediante una flecha roja se identifica el agua catalítica (Acat) en la conformación activa. Se representan en esferas moradas los cationes divalentes del Mn y en color magenta los de Mg, cuya posición es casi idéntica. También en varillas se representan las cadenas laterales de los residuos de la péntada acídica. Con números rojos se identifican las tres posiciones de los cationes.

Los resultados presentados en la tabla R. 2. 7 fueron agrupados en función de la conformación del nucleótido y las posiciones ocupadas por el catión de Mn^{2+} en el conjunto de los 26 centros activos analizados (Tabla R. 2. 8).

Tabla R. 2. 8. Resultados agrupados del análisis de señal anómala del catión Mn^{2+} en los centros activos de las cuatro estructuras de EnMazG resueltas en esta Tesis y los centros activos de la estructura de EnMazG en complejo con dUpNHpp y Mg^{2+} . En la tabla superior (azul) se agrupan los datos de la tabla R. 2. 7 y en la tabla inferior (amarillo) se agrupan los datos de la estructura de EnMazG en complejo con dUpNHpp y Mg^{2+} . El número entre paréntesis hace referencia a la cantidad de centros activos con la misma distribución de los cationes de Mn^{2+} o Mg^{2+} .

Conformación del dUTP en presencia de cationes de Mn^{2+} en las diferentes posiciones			
Posición 3 (1)	Posiciones 1 y 2 (5)	Posiciones 1 y 3 (2)	Posiciones 1, 2 y 3 (18)
100 % inactiva	80 % inactiva 20 % doble conformación	100 % inactiva	100 % inactiva

Conformación del dUTP en presencia de cationes divalentes de Mg	
Las 3 posiciones ocupadas por Mg^{2+} (8)	
75 % doble conformación	25 % activa

La presencia del catión Mg^{2+} en las tres posiciones supone que una gran mayoría de centros activos posean el nucleótido en una doble conformación, mientras que el resto está en conformación totalmente activa. Las estructuras en presencia de Mn^{2+} han permitido obtener un amplio espectro de las ocupaciones de las tres posiciones de este catión en el centro activo y las conformaciones adoptadas por el dUTP. Ninguna de las estructuras muestra la única ocupación de la posición 1 o de la posición 2 por el Mn^{2+} , sino de la posición 3 y en conformación inactiva del nucleótido. Cuando el Mn^{2+} ocupa 2 posiciones en el centro activo, podrían darse 3 situaciones, de las cuales dos han podido ser observadas en las estructuras. Cuando este catión ocupa las posiciones 1 y 2 supone que un 80 % de los centros activos asuman una conformación inactiva, mientras que el 20 % muestre una doble conformación. Cuando el Mn^{2+} ocupa las posiciones 1 y 3 todos los centros activos muestran la conformación inactiva. Finalmente, cuando las tres posiciones son ocupadas por este catión, todos los centros activos muestran una conformación inactiva. En los 34 centros activos analizados no solo se ha mantenido casi idéntico el posicionamiento de los 3

cationes independientemente de que fueran Mg^{2+} o Mn^{2+} , sino que las cadenas laterales de los residuos que forman los motivos de unión de metal también han mantenido su disposición (E31, E32, E35, D50 y D54). Los únicos cambios importantes son el no posicionamiento de la molécula de agua catalítica y el intercambio de posiciones entre uno de los oxígenos del fosfato y una molécula de agua en las conformaciones activa e inactiva, por lo cual el fosfato y pasa de ser coordinado por el catión 1 a estar coordinado por los cationes 2 y 3. Es decir, las disposiciones de los centros activos han confirmado lo que se había resaltado anteriormente en la superposición de la conformación activa e inactiva (Figura R. 2. 18).

Tras el análisis de estos resultados podría proponerse un mecanismo de inhibición de EnMazG por Mn^{2+} , en el cual la conformación del nucleótido varía en función de la ocupación de las posiciones de los iones en el centro activo por este catión. En ausencia de Mn^{2+} y en presencia de Mg^{2+} , el nucleótido se mantendría en un equilibrio entre las conformaciones activa e inactiva que se desplazaría hacia la conformación activa cuando la molécula de agua catalítica fuera posicionada correctamente, teniendo lugar el ataque nucleofílico. La entrada del catión Mn^{2+} en la posición 1 mientras las otras dos posiciones están ocupadas por Mg^{2+} , no afecta a la conformación del nucleótido, de modo que el dUTP asumiría una conformación activa (Figura R. 2. 19. A). Cuando el Mn^{2+} ocupa las posiciones 1 y 2, y el catión Mg^{2+} la posición 3, el dUTP puede asumir ambas conformaciones (Figura R. 2. 19. B), aunque mayoritariamente adopta la conformación inactiva. Cuando el Mn^{2+} ocupa las 3 posiciones disponibles, el nucleótido siempre asume la conformación inactiva (Figura R. 2. 19. C).

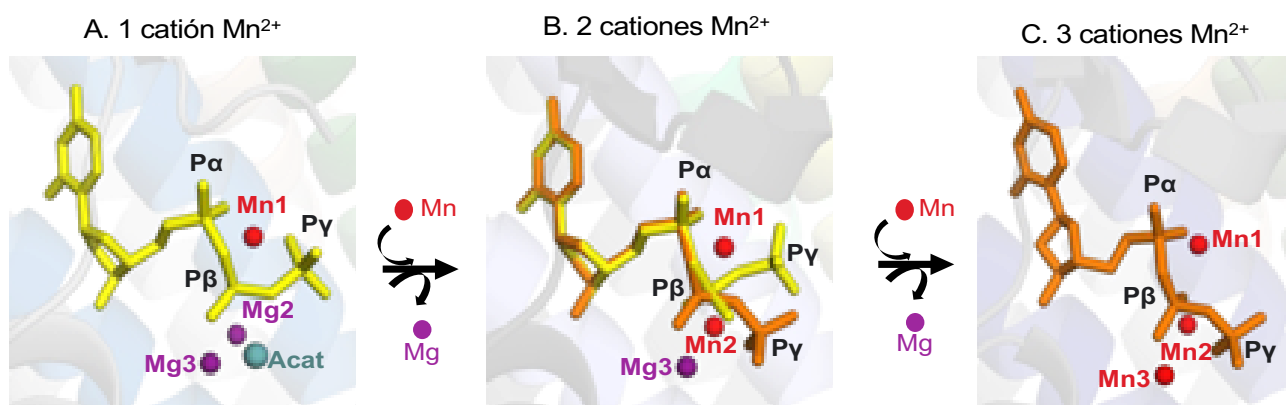


Figura R. 2.19. Mecanismo de inhibición de EnMazG por el catión divalente Mn^{2+} . Cuando el catión Mn^{2+} ocupa la posición 1 el dUTP asume una conformación activa (A); cuando el Mn^{2+} ocupa las posiciones 1 y 2, el dUTP puede asumir la conformación activa y la inactiva (B); y cuando el Mn^{2+} ocupa las 3 posiciones, el nucleótido asume la conformación inactiva (C). Se representa en varillas el dUTP, en color amarillo en conformación activa y en naranja en conformación inactiva. En esferas rojas y moradas se representan los iones Mn^{2+} y Mg^{2+} respectivamente, mientras que en cian oscuro se representa la molécula de agua catalítica.

Por tanto, en base a los resultados obtenidos y el propuesto mecanismo inhibitor del Mn^{2+} , se podría deducir que el ion que ocupa la posición 2 puede favorecer significativamente la conformación inactiva, mientras que el ion que ocupa la posición 3 la determina.

Con el objetivo de confirmar el mecanismo de interacción propuesto y arrojar luz sobre las especies de transición, se intentó obtener más cristales de EnMazG con Mg^{2+} y Mn^{2+} pero utilizando otra aproximación distinta de la co-cristalización. Para ello, se adaptó un mecanismo similar al *soaking* utilizado en otro laboratorio (Samara et al. 2017).

Se llevaron a cabo dos versiones de la misma aproximación. En una de ellas, el primer paso fue obtener cristales de EnMazG en presencia de dUTP y $CaCl_2$, de modo que la enzima estuviera inhibida completamente y no degradara el sustrato. Posteriormente, estos cristales fueron transferidos durante 30 minutos a una solución formada por los componentes de la solución en la que el cristal había crecido junto con EGTA, el cual secuestraría el exceso de Ca^{2+} . Inmediatamente después, los cristales se transfirieron a otra solución formada de nuevo por los componentes en los que crecieron los cristales junto con una combinación de iones Mg^{2+} - Mn^{2+} a altas concentraciones, y se mantuvieron en esta solución en distintos tiempos, desde 30 segundos a 10 minutos. De esta manera, los iones de Ca^{2+} que estuvieron quelados en el centro activo por la proteína-nucleótido se intercambiarían por el Mg^{2+} y Mn^{2+} . Tomar cristales a diferentes tiempos de inmersión en la tercera solución permitirían observar qué posiciones son ocupadas preferencialmente por cada catión y como esto repercute en la conformación del nucleótido. A cada tiempo de inmersión seleccionado los cristales fueron crioprotectados y congelados en nitrógeno líquido para su posterior difracción. Desafortunadamente, en este tipo de ensayos se observó que los cristales no eran estables en los diferentes pasos de *soaking*, especialmente en la inmersión en la solución con el Mg^{2+} y el Mn^{2+} . En los casos más graves los cristales se disolvían mientras que en los casos donde mantenían su integridad estructural perdían capacidad de difracción, no pudiéndose recoger datos a la resolución necesaria.

En la segunda versión, los cristales fueron obtenidos en presencia de dUTP y Mn^{2+} , y fueron transferidos a la solución en la que habían crecido en ausencia del catión divalente para eliminar el catión no unido a la proteína. Tras 30 minutos en esta solución fueron transferidos a otra solución idéntica a la anterior pero con ion Mg^{2+} , manteniéndolos allí durante distintos tiempos (desde 30 segundos a 10 minutos), con la intención de que la proteína sustituyera el Mn^{2+} por el Mg^{2+} y poder ver el proceso de entrada de este catión. Tras el tiempo deseado, los cristales se crioprotectaron y se congelaron en nitrógeno líquido para su posterior difracción. Esta segunda opción resultó, al igual que la primera, infructuosa. Los cristales pudieron soportar el paso por la solución sin cationes, pero cuando se sumergieron en la solución con el Mg^{2+} empezaron a resquebrajarse y pocos

podieron sobrevivir y ser difractados. La resolución obtenida en su difracción fue muy baja.

En ambas versiones de la aproximación, los cristales parecen desintegrarse en el mismo proceso, es decir, cuando el sitio activo empezaría a hidrolizar el dUTP por la entrada del Mg^{2+} .

2.2.4.3. El análisis de las estructuras tridimensionales de EnMazG resueltas a partir de la co-cristalización con los cationes Ca^{2+} , Zn^{2+} y Ni^{2+} permiten proponer una hipótesis sobre una modulación diferencial del mecanismo de inhibición según el catión implicado.

La co-cristalización fue utilizada también para la obtención de estructuras tridimensionales de EnMazG en presencia de otros cationes divalentes: Ca^{2+} , Zn^{2+} , Ni^{2+} y Cu^{2+} . Se utilizaron las mismas ratios molares que en el caso del Mn^{2+} (en rangos molares Mg^{2+} : otro catión desde 0,25 a 35 veces más de Mg^{2+} respecto al otro catión); y dado que el Cu^{2+} dio también problemas de solubilidad, la co-cristalización se realizó con aproximaciones que incluyeron concentraciones máximas totales de cationes de 1, 2 y 3 mM.

Se realizaron tanto nuevos cribados de condiciones de cristalización, como cribados en condiciones conocidas ampliadas. Más de 150 cristales de EnMazG en presencia de los cationes Ca^{2+} , Zn^{2+} , Ni^{2+} fueron crioprotectados y congelados para ser difractados en diferentes sincrotrones. Al igual que los cristales en presencia del Mn^{2+} , los cristales que poseían estos cationes tenían las morfologías de diamantes, rombos y varillas. Se obtuvieron unos 75 cristales que contenían Ca^{2+} con morfologías de diamante, rombos y varillas gruesas (Figura R. 2. 20); unos 35 cristales que contenían Ni^{2+} , todos con morfología de varillas finas y frecuentemente astilladas; y unos 40 cristales que contenían Zn^{2+} con una sola morfología de diamante.

Los cristales fueron difractados en las energías de excitación de los correspondientes cationes, siendo la del cesio la utilizada para el calcio (5 KeV (2,48 Å), la mínima estable en la línea XALOC, ya que la del calcio es 4,65 KeV (2,67 Å)), 8,33 KeV (1,49 Å) en el caso del níquel, y 9,66 KeV (1,28 Å) en el caso del zinc. No se obtuvieron cristales en presencia del Cu^{2+} .

La longitud de onda exacta utilizada en la difracción fue determinada de nuevo a través de un espectro de emisión de fluorescencia del cristal a difractar, ya que el valor teórico del *K-edge* del elemento puede variar en algunos eV por el efecto de apantallamiento. Una vez seleccionada la longitud de onda, el cristal era difractado en sus 360 ° independientemente del grupo espacial predicho, con la finalidad de obtener la máxima completitud y multiplicidad de los datos anómalos. La transmitancia del rayo X utilizada para la recogida era mayor respecto a la

difracción no anómala para compensar la menor energía del rayo y obtener una mayor resolución.

Tras todo este arduo trabajo experimental, tan solo 3 de los juegos de datos recogidos presentaban la calidad suficiente para poder resolver las estructuras tridimensionales. Por suerte, cada uno de ellos correspondía a un catión diferente en presencia de Mg^{2+} . Las estructuras correspondían a las ratios molares $Mg^{2+}:Ca^{2+}$ (6: 1), $Mg^{2+}:Ni^{2+}$ (1,5: 1) y $Mg^{2+}:Zn^{2+}$ (2: 1). Los tres cristales crecieron en condiciones de cristalización formadas principalmente por PEG 300, agente tamponado Tris a pH 8,5, y en el caso del Ca^{2+} y del Ni^{2+} , con un porcentaje de etanol (15 y 25 % respectivamente). Sobre ser una solución de cristalización tan similar, el cristal que contenía la ratio molar $Mg^{2+}:Ca^{2+}$ 6: 1 tenía forma de varilla gruesa, el de $Mg^{2+}:Ni^{2+}$ 1,5: 1 de varilla fina, y $Mg^{2+}:Zn^{2+}$ 2: 1 de diamante.



Figura R. 2. 20. Morfologías adoptadas por los cristales de EnMazG en presencia de Ca^{2+} y Zn^{2+} . Principalmente, los cristales adoptaban formas de varillas, rombos (*izquierda*) o diamantes (*derecha*).

Tras difractar los cristales, integrar y escalar los datos recogidos, las estructuras fueron resueltas de nuevo por reemplazo molecular, utilizando como modelo las estructuras de EnMazG en conformación *apo* y unida al trinucleótido resueltas previamente en el laboratorio (Ciges-Tomas 2019).

La estructura de la ratio molar 1:6 $Ca^{2+}:Mg^{2+}$ fue resuelta a 1,9 Å, en el grupo espacial $I4_122$. Las dimensiones de la celda unidad son $a=48,31$ Å, $b=48,31$ Å, $c=177,35$ Å, con tres ángulos de 90 °, y contenía un monómero a partir del cual se obtuvo el tetrámero por simetría cristalográfica. Los parámetros cristalográficos se presentan en la tabla R. 2. 9.

La estructura de la ratio molar 1:1,5 $Ni^{2+}:Mg^{2+}$ fue resuelta a 3 Å, en el grupo espacial $P4_32_12$. Las dimensiones de la celda unidad son $a=48,08$ Å, $b=48,08$ Å, $c=181,71$ Å, con tres ángulos de 90 °, y contenía un dímero a partir del cual se obtuvo el tetrámero por simetría cristalográfica.

La estructura de la ratio molar 1:2 $Zn^{2+}:Mg^{2+}$ fue resuelta a 2,9 Å, en el grupo espacial $P4_32_12$. Las dimensiones de la celda unidad son $a= 47,63$ Å, $b= 47,63$ Å, $c= 182,38$ Å, con tres ángulos de 90° , y contenía un dímero a partir del cual se obtuvo el tetrámero por simetría cristalográfica.

En los tres casos, tanto el grupo espacial como la celda unidad son muy similares a la EnMazG en conformación *apo*. El grupo espacial en el que cristalizó es $P4_32_12$, y la celda unidad posee las dimensiones $a= 48,62$ Å, $b= 48,62$ Å, $c= 180,26$ Å, y los ángulos $\alpha=\beta=\gamma= 90^\circ$. Además, la unidad asimétrica también contenía un dímero que por simetría cristalográfica formaba el tetrámero.

En ninguna de las estructuras con Ca^{2+} , Ni^{2+} o Zn^{2+} hay presencia del nucleótido ni de los cationes en el centro activo, ni tampoco pudo trazarse el motivo V (Figura R. 2. 21. B). La superposición del centro activo de las estructuras de EnMazG en “presencia” de Ni^{2+} , Zn^{2+} , la conformación *apo* (Figura R. 2. 21. C) y junto con la estructura con el nucleótido (Figura R. 2. 21. D), mostró que en presencia de estos cationes alternativos la proteína adoptaba una conformación con el centro activo abierto, sin sustrato ni cationes, y con el motivo V desordenado. Los valores RMSD para la superposición de la conformación *apo* con las estructuras con Ni^{2+} y Zn^{2+} muestran valores menores a 0,8 Å para la superposición de los C α de 333 residuos (no hay C-terminal), mientras que la superposición de éstas con la estructura en complejo con el sustrato muestra valores alrededor de 1,65 Å para los C α de entre 335 y 346 residuos (Tabla R. 2. 10). En el caso de la estructura en “presencia” del Ca^{2+} resulta sorprendente los valores de superposición con la estructura en conformación *apo* y en complejo con el sustrato (1,622 Å y 0,911 Å para 334 y 335 residuos, respectivamente). Es decir, esta estructura no posee en el centro activo ni sustrato ni cationes, ni tampoco tiene ordenado el extremo C-terminal, pero aun así se asemeja más a la estructura cuaternaria de la proteína en complejo con el sustrato que en ausencia de éste (Figura R. 2. 21. A, y Tabla R. 2. 10).

En estas estructuras no es posible localizar en la misma orientación a todos los residuos participantes en la catálisis (como es el caso de E35 en la estructura con Zn^{2+} o D50 en el caso de la estructura con Ca^{2+} , Figura R. 2. 21. C), incluso algunas cadenas laterales no pueden trazarse en su totalidad (por ejemplo E35 en el caso de la estructura con Ca^{2+} , Figura R. 2. 21. C), lo cual es predecible por la ausencia de nucleótido que motivaría su correcta orientación.

Por tanto, las estructuras de EnMazG en “presencia” de los cationes de Ni^{2+} y Zn^{2+} son muy similares entre sí y adoptan una conformación *apo*, con una conformación del centro activo abierta, con ausencia de cationes y de sustrato. En el caso de la EnMazG en presencia del Ca^{2+} , aunque no esté presente el sustrato ni los cationes presenta una conformación intermedia entre abierta y cerrada. “Abierta” porque no presenta ordenado el extremo C-terminal, ni todas las cadenas laterales de los residuos implicados en la catálisis están orientadas

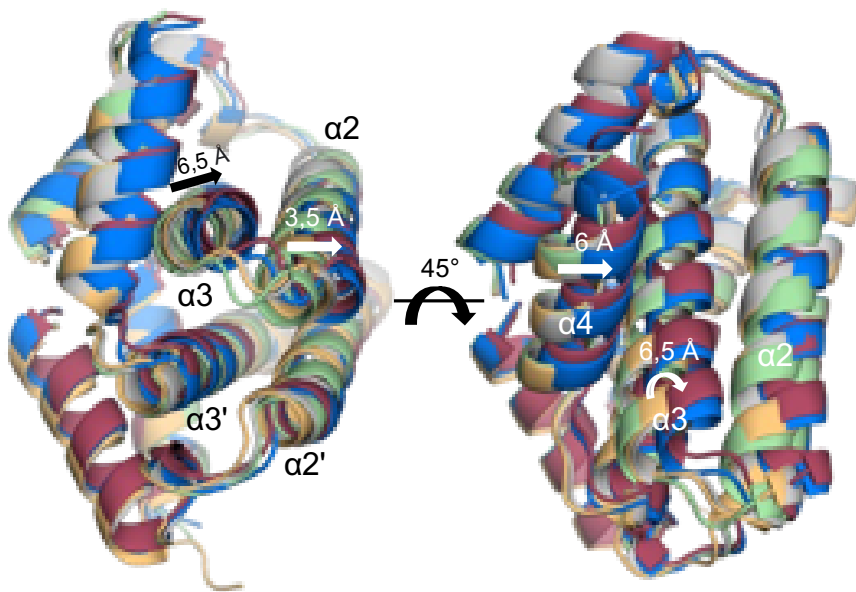
para ello, especialmente el glutámico D50 que coordina el catión 3, responsable junto al catión 2 de activar la molécula de agua que llevará a cabo el ataque nucleofílico. “Cerrada” porque sí muestra el movimiento de las hélices $\alpha 2$ y $\alpha 3$ realizado en la transición desde la conformación abierta hacia la cerrada (Figura R. 2. 21. A).

Tabla R. 2. 9. Parámetros cristalográficos de la estructura de EnMazG en “presencia” de Mg^{2+} y Ca^{2+} .

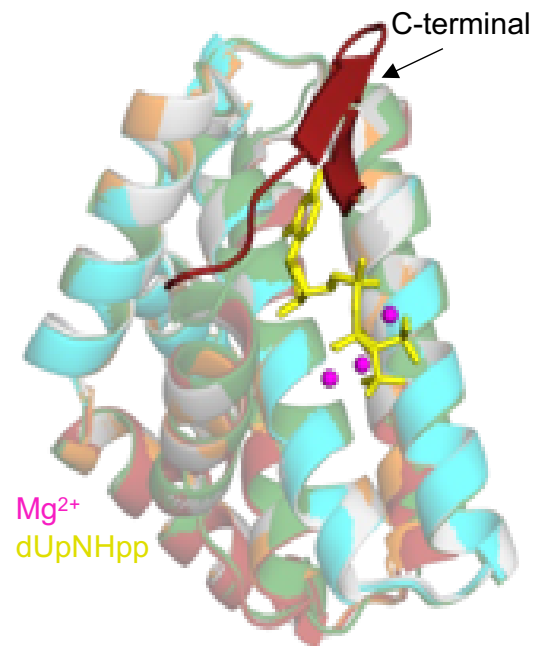
EnMazG	$Ca^{2+} : Mg^{2+} \quad 1: 6$
Datos de recogida	
Haz de luz	ALBA-XALOC
Energía del haz (Å)	2,48
Grupo espacial	I4 ₁ 22
Dimensiones de la celda (Å)	a=48,31 b=48,31 c=177,35 $\alpha=\beta=\gamma=90$
Resolución (Å) ^a	44,34– 1,9 (1,968 – 1,9)
Reflexiones únicas	8758 (843)
Compleitud (%)	99,46 (99,53)
Multiplicidad	23,8 (20,8)
Compleitud anómalo (%)	100 (100)
Multiplicidad anómalo	13,2 (11,1)
Media I/ σ (I)	24,1 (3,8)
R _{pim}	0,015 (0,190)
CC 1/2	0,999 (0,974)
Refinado	
R _{work}	0,1942 (0,2982)
R _{free}	0,2209 (0,3116)
Número de átomos	639
Proteína	609
Moléculas de agua	37
Otras moléculas	0
Rmsd, enlaces (Å)	0,018
Rmsd, ángulos (°)	2,00
Probabilidad de choque molecular	4,25
Diagrama de Ramachandran	
Preferidas (%)	98,78
Permitidas (%)	1,22
Outliers (%)	0,00

^aNúmero entre paréntesis indica los valores para la zona de alta resolución.

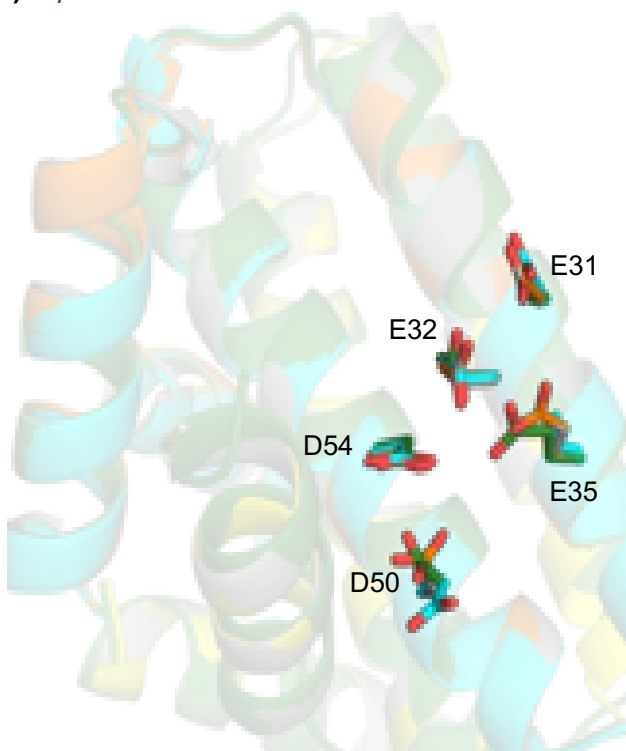
A) Apo Ca^{2+} Ni^{2+} Zn^{2+} +dUpNHpp



B) Apo Ca^{2+} Ni^{2+} Zn^{2+} +dUpNHpp



C) Apo Ca^{2+} Ni^{2+} Zn^{2+}



D) Apo Ca^{2+} Ni^{2+} Zn^{2+} +dUpNHpp

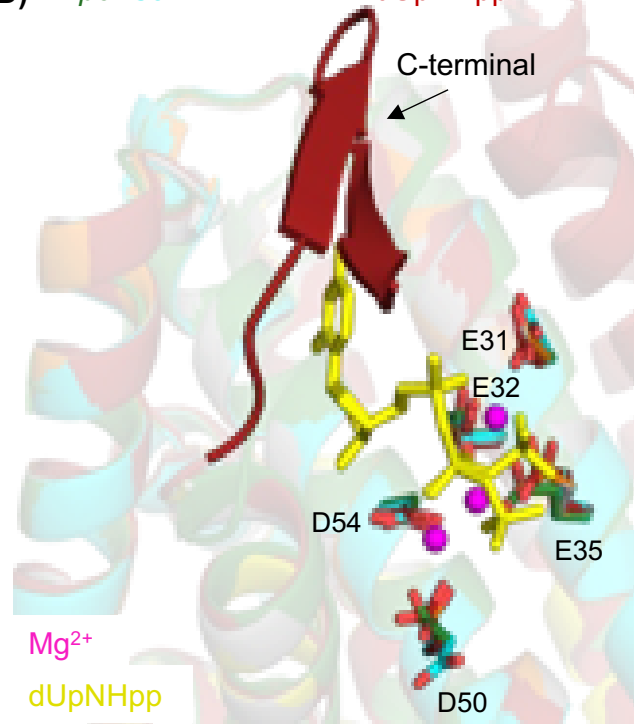


Figura R. 2. 21. Superposición de los dímeros de las estructuras de EnMazG en conformación *apo*, en complejo con el nucleótido no hidrolizable dUpNHpp, y en “presencia” de los cationes Ca^{2+} , Ni^{2+} y Zn^{2+} (colores *verde*, *rojo*, *azul/cian*, *gris* y *naranja*, respectivamente). A) Superposición de la conformación *apo*, en presencia de los cationes de Ca^{2+} , Ni^{2+} , Zn^{2+} , y en complejo con dUpNHpp. Se muestran dos vistas ortogonales y se indica el desplazamiento realizado por las hélices $\alpha 2$, $\alpha 3$ y $\alpha 4$ en presencia del Ca^{2+} y en complejo con el sustrato respecto a la conformación *apo*. B) Misma superposición que en A). El motivo C-terminal solo se ordena en presencia del sustrato. El dUpNHpp se representa en varillas amarillas y los iones Mg^{2+} en esferas color magenta. C) Vista detallada del centro activo de la superposición de la conformación *apo*, y en presencia de Ca^{2+} , Ni^{2+} y Zn^{2+} . D) Igual que (C) junto con la estructura en complejo con el sustrato. En (C) y (D) Se representan en varillas e identifican algunos residuos implicados en la catálisis del sustrato (código de colores según átomo: C mismo color de la estructura procedente, O en *rojo*).

Tabla R. 2. 10. Tabla de valores de RMSD en Å para la superposición del tetrámero de EnMazG en conformación *apo* con el tetrámero de EnMazG en presencia de los cationes divalentes de Ca, Ni y Zn, y con la estructura en complejo con el dUpNHpp y el Mg^{2+} .

EnMazG	<i>Apo</i>	Ca^{2+}	Ni^{2+}	Zn^{2+}	Mg^{2+} - dUpNHpp
<i>Apo</i>	-	1,622 (334)	0,773 (333)	0,785 (333)	1,659 (335)
Ca^{2+}	-	-	1,479 (334)	1,602 (334)	0,911 (335)
Ni^{2+}	-	-	-	0,201 (374)	1,576 (346)
Zn^{2+}	-	-	-	-	1,654 (346)
Mg^{2+} - dUpNHpp	-	-	-	-	-

El análisis de los centros activos de las 5 estructuras también incluyó a la geometría de coordinación de los cationes y la longitud de los enlaces con los que la constituye.

La geometría molecular de los cationes Mg^{2+} en el centro activo de EnMazG es octaédrica, formando 6 enlaces, como ya se había descrito anteriormente para este catión (Zheng et al. 2017). En la hexacoordinación participan los oxígenos de los fosfatos, las moléculas de agua que ocupan la posición del fosfato y en la conformación alternativa, y la péntada acídica formada por los residuos E31, E32, E35, D50 y D54 de las hélices $\alpha 2$ y $\alpha 3$ del monómero en el que se localiza el centro activo (Figura R. 2. 22).

En el caso del catión divalente del Mn, utiliza exactamente la misma hexacoordinación bipiramidal octaédrica que el Mg^{2+} , utilizando distancias

similares a éste, de modo que la sustitución de uno por el otro es muy factible, lo cual explicaría una rápida sustitución de los tres cationes divalentes Mg por Mn en el centro activo de EnMazG, y en consecuencia la rápida inhibición (valor de IC_{50} 0,024 mM) (Tabla R. 2. 11).

En el caso de los cationes Ca^{2+} y Ni^{2+} , ambos establecen una hexacoordinación como el catión Mg^{2+} , pero para ello utilizan enlaces cuya distancia con los correspondientes átomos o moléculas con los que los establecen es más pequeña (Ni^{2+} , 1,86-2,08 Å) o más grande (Ca^{2+} , 2,4-2,5 Å) que la utilizada por los cationes Mg^{2+} (1,91-2,19 Å). En el caso del Ca^{2+} , la imposibilidad de establecer las distancias necesarias para mantener la geometría de coordinación en el centro activo impediría su correcta configuración, el posicionamiento del sustrato y, en consecuencia, la actividad enzimática de la enzima. En el caso del Ni^{2+} , aunque resultaría más sorprendente por las mínimas diferencias en la distancia de los enlaces, podría estar desfavoreciendo la correcta organización del centro activo.

En el caso del catión Zn^{2+} , presenta una coordinación tetraédrica formada por 4 enlaces, con lo que su entrada en el centro activo implicaría la no interacción con los 6 elementos que establecen enlaces cada uno de los tres iones de Mg^{2+} . Esto podría no favorecer la correcta organización del centro activo.

Aunque en el caso del Co^{2+} y Cu^{2+} no se han obtenido estructuras, se ha analizado igualmente su geometría de coordinación. En el caso del Cu^{2+} , la geometría molecular es cuadrada plana, y las distancias de enlace también son un poco mayores a las establecidas por el Mg^{2+} (2,12-2,37 Å), con lo que podría ocurrir lo mismo que en el caso del Zn^{2+} . En el caso del Co^{2+} , habitualmente establece una geometría octaédrica, con 6 enlaces, al igual que el Mg^{2+} , aunque la distancia de los enlaces es un poco inferior (1,90-2,10 Å), con lo cual podría ocurrir algo igual de sorprendente que en el caso del Ni^{2+} .

Con estas pequeñas diferencias observadas en las distancias de enlace en cationes con la misma geometría octaédrica que el Mg^{2+} , se analizó el radio atómico del Ni^{2+} y del Co^{2+} para intentar comprender los resultados en EnMazG en su presencia. En la actividad enzimática, el Co^{2+} completa inicialmente al Mg^{2+} , potenciando a la enzima a bajas concentraciones, mientras que a medida que la concentración se incrementa, acaba inhibiéndola. El Ni^{2+} inhibe la actividad desde bajas concentraciones. Ambos cationes presentan un radio atómico de 1,35 Å, algo más pequeño que el del Mg^{2+} (1,50 Å) (Slater, J.C. 1964), con lo cual, el radio atómico no podría explicar el comportamiento diferencial en la actividad enzimática.

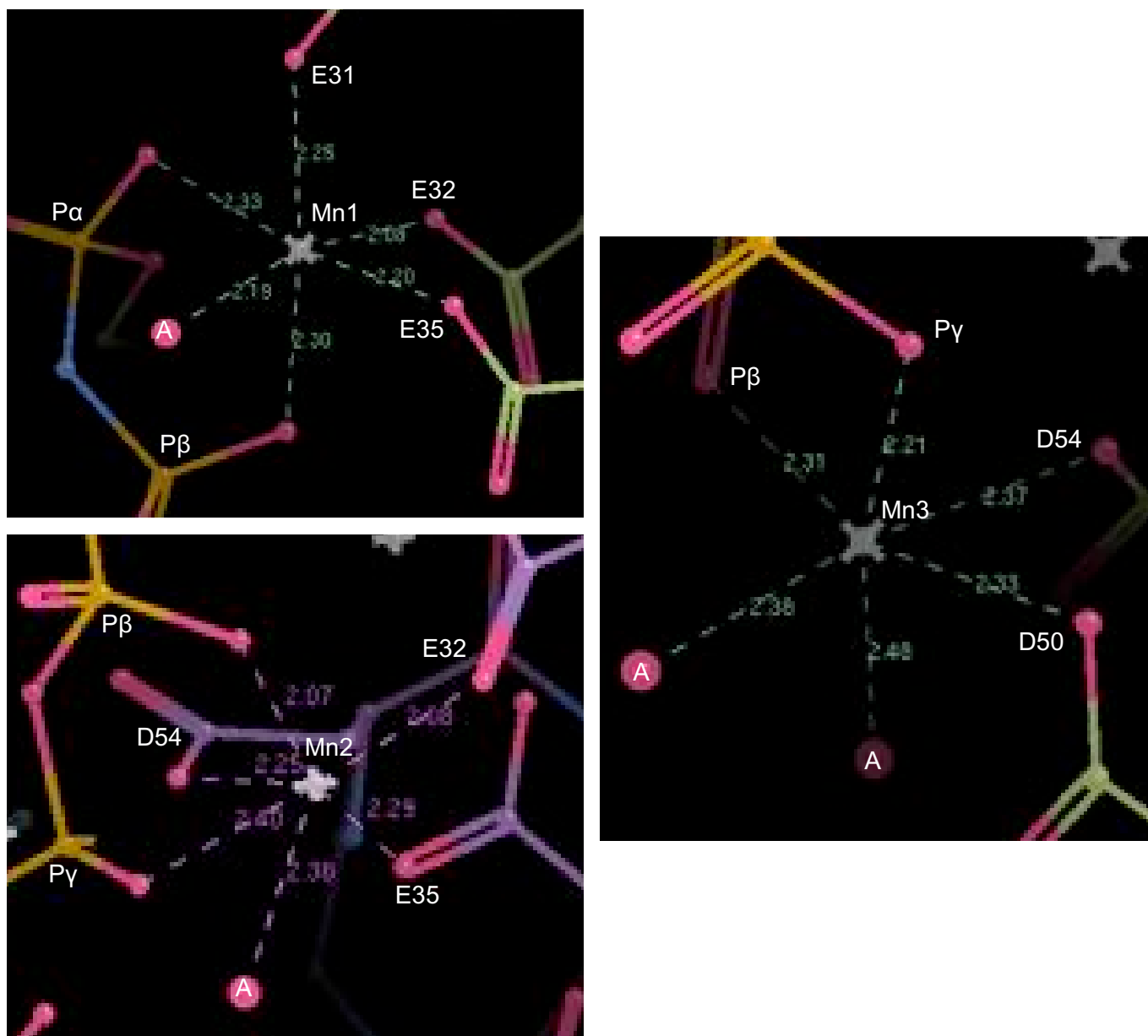


Figura R. 2. 22. Coordinación de los iones Mn^{2+} en el centro activo de EnMazG. Los tres cationes presentan una hexacoordinación con geometría octaédrica. En la coordinación del Mn1 (*arriba izquierda*) participan oxígenos de los fosfatos α y β , los residuos de glutámico 31, 32 y 35, y una molécula de agua. En la coordinación del Mn2 (*abajo izquierda*) participan oxígenos de los fosfatos β y γ , los residuos de glutámico 32 y 35, el residuo aspártico 54, y una molécula de agua. En la coordinación del Mn3 (*derecha*) participan oxígenos de los fosfatos β y γ , los residuos de aspártico 50 y 54, y dos moléculas de agua. Los iones se representan en estrellas grises, las moléculas de agua en esferas rosas, los átomos de oxígeno también en rosa, los de fósforo en naranja y los carbonos en morado o verde. Se identifican los residuos participantes y se indican las distancias entre los átomos deseados. Imagen creada con COOT.

Tabla R. 2 .11. Geometrías moleculares de los distintos cationes divalentes utilizados en esta Tesis, junto con las distancias de los enlaces establecidos entre éstos y átomos de oxígeno provenientes de un carbono (M-OC), y entre éstos y átomos de oxígeno provenientes de moléculas de agua (M-H₂O). Figura adaptada de *Zheng et al. 2017*.

<u>Catión divalente</u>	<u>Geometría molecular</u>	<u>Distancias M-OC/M-H₂O (Å)</u>
Mg ²⁺	Octaédrica, 6 enlaces	1,91-2,19 / 2,19
Mn ²⁺	Octaédrica, 6 enlaces	1,91-2,19 / 2,19
Co ²⁺	Normalmente octaédrica, 6 enlaces	1,90-2,10 / 2,10
Ni ²⁺	Octaédrica, 6 enlaces Plano cuadrado, 4 enlaces	1,86-2,07 / 2,08
Ca ²⁺	Octaédrica, 6 enlaces	2,4-2,5 / 2,4-2,5
Zn ²⁺	Tetraédrica, 4 enlaces	2,15 / 2,09
Cu ²⁺	Plano cuadrado, 4 enlaces	2,12 / 1,97-2,37

Por tanto, los resultados obtenidos parecen indicar que EnMazG muestra una gran preferencia por el catión Mg²⁺ para llevar a cabo la hidrólisis del sustrato. Además, en combinación con los cationes Zn²⁺ y Co²⁺ a bajas concentraciones su capacidad enzimática se potencia, mientras que en presencia de otros cationes con mismas geometrías de coordinación o distancias de enlace muy similares, como el Mn²⁺, su actividad es inhibida incluso a bajas concentraciones.

En el caso de SeMazG, de la cual no se dispone de estructuras en presencia de estos cationes divalentes, se esperarían resultados similares debido a la gran similitud descrita anteriormente entre ambas enzimas (Ciges-Tomas 2019).

CAPÍTULO III

3.1. Duts de *Mycobacterium* tienen actividad moonlighting.

A nivel de secuencia, las Duts de las bacterias del género *Mycobacterium* presentan como característica diferencial una corta inserción específica de 5 residuos de longitud, ausente de forma general en el resto de Duts, desde *E. coli* a la humana (Figura R. 3.1 y R. 3.2). Las estructuras tridimensionales de la Dut de *M. tuberculosis* muestran que esta inserción, ubicada en la secuencia entre los motivos conservados IV y V, se pliega como un corto *loop* cerca del extremo C-terminal, y su función fisiológica y funcional es desconocida (Pecsi et al. 2012).

Utilizando a *Mycobacterium smegmatis* como un organismo modelo de crecimiento rápido de *M. tuberculosis*, se demostró que esta inserción de pequeño tamaño resultaba clave en la supervivencia de la bacteria. Aunque la enzima era totalmente funcional para la hidrólisis del dUTP en ausencia de dicho *loop*, cepas bacterianas que portaban Duts deletacionales en esa inserción no eran viables en condiciones normales de laboratorio (Pecsi et al. 2012).

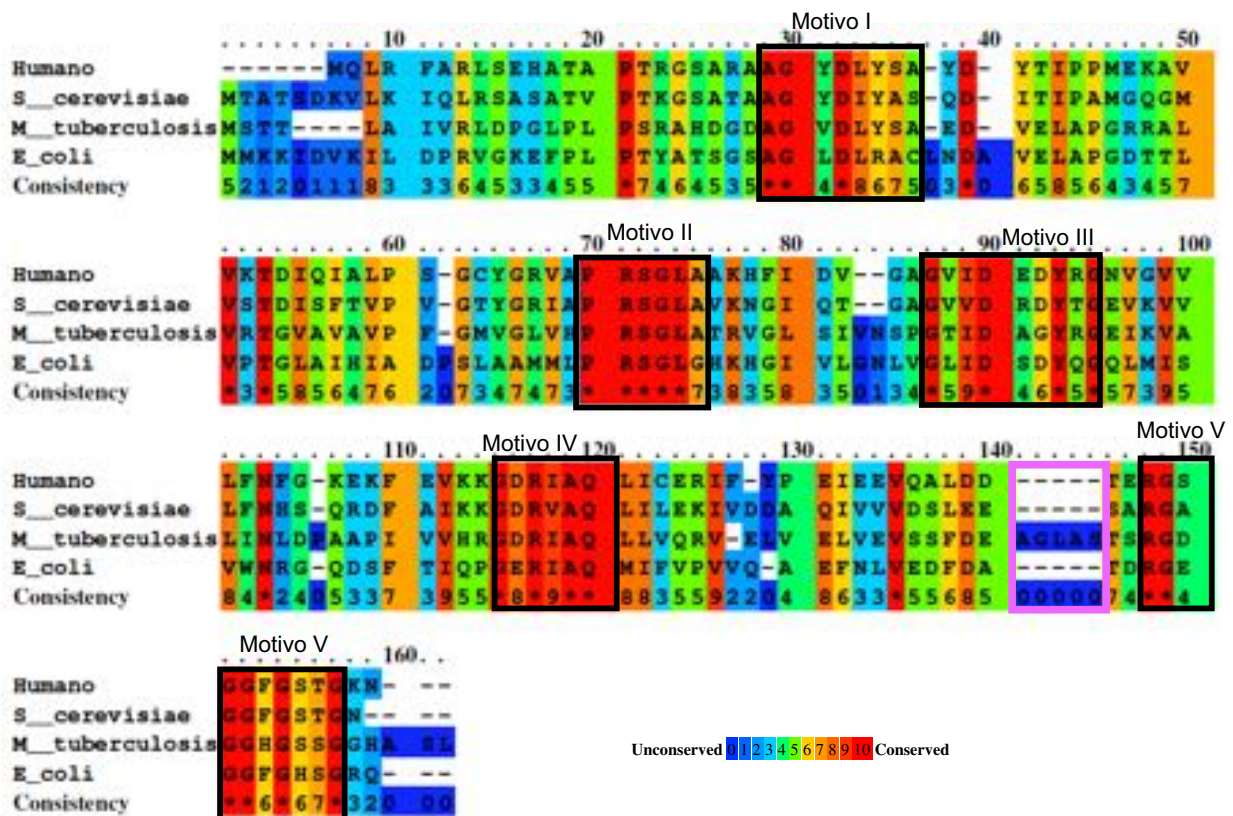


Figura R. 3. 1. La inserción del género *Mycobacterium* no está presente en otras Duts. Se alinean las secuencias aminoacídicas de las Duts de humano, *S. cerevisiae*, *M. tuberculosis* y *E. coli*. Se identifican en cuadrados negros los 5 motivos conservados de las Duts triméricas. En un recuadro rosa se identifica el *loop* específico del género *Mycobacterium* en la Dut de *M. tuberculosis* (secuencia de residuos: AGLAS). Alineamiento realizado con *Praliné*, con los residuos coloreados en función del grado de conservación.

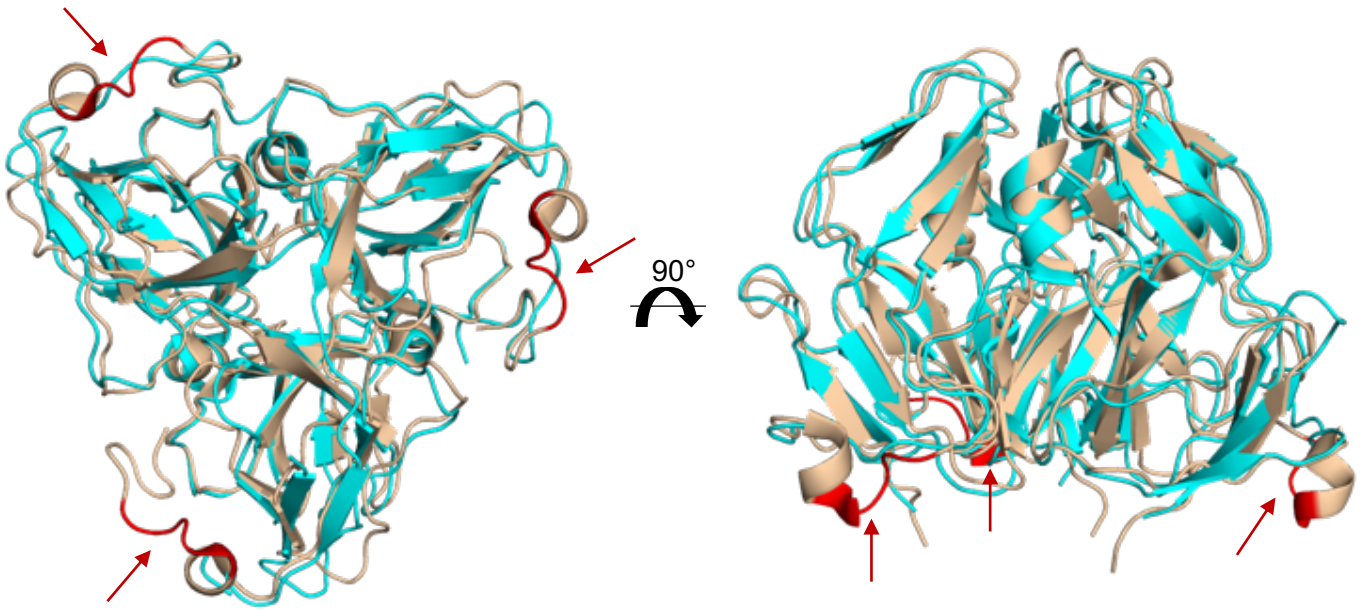


Figura R. 3. 2. Localización estructural del *loop* específico del género *Mycobacterium*. Superposición de las estructuras de Dut de *M. tuberculosis* (color trigo) y humana (color cian). El *loop* de 5 residuos presente en mDut se destaca en color rojo en los tres monómeros y se señala con flechas rojas. Códigos PDB: 1MQ7 para *M. tuberculosis*, y 1Q5U para humano.

Este sorprendente hallazgo, sumado a la propuesta del laboratorio en la cual las funciones reguladoras *moonlighting* de Duts tendrían un mecanismo análogo a las GTPasas (Tormo-Más et al. 2013; Maiques et al. 2016), hizo plantear la hipótesis de que la Dut de *M. tuberculosis* también estuviera realizando otras funciones a parte de la puramente hidrolítica del dUTP.

Si las Duts de *Mycobacterium* tenían actividad *moonlighting*, se hipotetizó que la llevarían a cabo por interacción proteína-proteína, al igual que las Duts de fagos de *S. aureus*, y posiblemente la inserción específica estaría implicada en dichas interacciones. Con el fin de confirmar esta hipótesis y conocer quiénes serían los compañeros de la mDut en sus actividades *moonlighting*, se encargó a una empresa externa (*Hybrigenics services*) un ensayo de doble híbrido en levadura, del cual se obtuvieron las proteínas candidatas de interacción. El análisis clasificó a las candidatas de A a D en función de la probabilidad de interacción, siendo A las de mayor probabilidad y D las de menor.

El grupo de las dianas de muy alta probabilidad de interacción está formado por dos proteínas, Rv3202c y Rv3524, y el grupo de alta probabilidad está formado por otras dos proteínas, Rv1429 y Rv3354. Dado que el objetivo final era conocer si el *loop* específico de especie estaba implicado en las posibles actividades *moonlighting* de mDut, el primer paso fue obtener las proteínas candidatas en forma soluble para poder realizar ensayos de interacción para identificar las

candidatas que sí interaccionaran, y con ellas proceder a los ensayos estructurales.

Con esta finalidad, se analizó para cada una de las 4 proteínas candidatas la información contenida en la base de datos UniProt y Pfam, se realizaron en algunas de ellas modelados *in-silico* con el servidor *Phyre2* y se analizaron con detalle las estructuras secundarias predichas, tanto en el modelado como en el servidor *Xtal-Pred-RF*.

3.2. Una ADN helicasa dependiente de ATP (Rv3202c).

La búsqueda en UniProt (ID 053348) identifica a Rv3202c como una ADN helicasa con función ATPasa localizada en el citosol, aunque también se la relaciona con actividad exonucleasa. El análisis de dominios se realizó con Pfam, y se identificaron 3: dos dominios helicasa (UvrD y UvrD_C), y un dominio endonucleasa (PDDEXK-1) (Figura R. 3. 3).



Figura R. 3.3. Distribución de dominios de la proteína Rv3202c según Pfam.

El análisis de doble híbrido identificó la región de interacción con mDut en los residuos 1 a 79, que corresponden al dominio UvrD helicasa N-terminal. Con el objetivo de obtener la proteína en forma soluble se llevó a cabo una primera aproximación en la cual se generaron dos plásmidos de expresión (pETNKI 1.1 y 1.5) que incluían el dominio UvrD entero (residuos M1 a E295), precedido de una etiqueta de poli-histidinas o una *Strep-tag* en el extremo N-terminal que permitiría su purificación por cromatografía de afinidad. Las construcciones se prepararon a partir del ADN genómico de *M. tuberculosis* H37Rv (los cebadores utilizados pueden consultarse en la tabla M.1).

Las pruebas iniciales de expresión incluyeron dos concentraciones de inductor (100 μ M y 1 mM de IPTG), dos temperaturas y tiempos de inducción (37 ° C durante 3 horas o 20 ° C durante 16 horas), así como varios medios de cultivo (LB, *Hyper*, *Superior*, *Turbo* y *Terrific*). Todas ellas mostraron un alto rendimiento, y se seleccionaron como condiciones óptimas de expresión las células BL21(DE3) inducidas con 100 μ M de IPTG a 20 ° C durante 16 horas en el medio LB. Con estas condiciones establecidas, se continuó con las pruebas de solubilidad. Teniendo en cuenta el punto isoeléctrico de la proteína (8,76) se probó el tampón PBS (pH es 7,4) en presencia y ausencia de 5 mM MgCl₂. Dado que la proteína estaba en los cuerpos de inclusión, se intentó solubilizar utilizando el mismo tampón PBS con 5 mM MgCl₂ pero incluyendo el agente

caotrópico urea a 4 M. Aunque se consiguió solubilizar, tras el intento de replegado de la proteína con sucesivos pasos de diálisis, la proteína precipitó a una concentración inferior a 3 M urea.

Tras estos infructuosos resultados, se hipotetizó la posibilidad de que la insolubilidad del constructo fuera consecuencia de la exposición de ciertos fragmentos de la proteína al solvente o que el plegamiento no era el correcto por la ausencia de otros dominios de la proteína. Por tanto, se decidió utilizar una segunda aproximación. Tras una nueva consulta en la base de datos Pfam se diseñó una nueva construcción que incluía los dos dominios UvrD (residuos M1 a D781), los cuales se encuentran ampliamente distribuidos de manera conjunta en la naturaleza como una única proteína. Esto podría ser indicativo de la necesidad de ambos dominios para que el constructo fuese estable. De modo que en este constructo se prescindió del dominio PDDEXK-1 C-terminal, el cual parece un anexo que solo aparece en algunos miembros de esta familia y que además se encuentra unido a la porción UvrD por un largo conector sin estructura secundaria predicha. Tras varios intentos fallidos en la clonación debido al elevado contenido en GCs en el gen (Anexo I), se decidió obtener la construcción de manera sintética. Con ella, se volvieron a realizar las pruebas de expresión en los distintos medios de cultivo utilizados anteriormente, y se seleccionó el medio *Hyper*[®]. Las concentraciones del inductor y los tiempos de expresión fueron las mismas utilizadas en la construcción previa (células BL21(DE3) inducidas con 100 μ M de IPTG a 20 ° C durante 16 horas). La posterior prueba de solubilidad con distintos tampones, variando tanto las concentraciones salinas como los componentes del tampón (Tabla R. 3. 1) resultó en la proteína en la fracción insoluble, con lo que se recurrió a la resolubilización desde 4 M urea. No obstante, aún en estas condiciones la proteína seguía siendo no soluble. Por ello se abordó la producción de la proteína completa como tercera aproximación.

Tras diversos intentos de clonado de la proteína por fragmentos (los cebadores utilizados se especifican en el Anexo I), se decidió obtener el gen sintético y optimizado para su producción en *E. coli*. Con ello se redujo el elevado porcentaje del contenido en GC del gen, el cual probablemente estaba dando lugar a la formación de estructuras secundarias que impedían la clonación e incluso la secuenciación posterior de los clones. La construcción sintética fue utilizada para la transformación de células BL21 (DE3). Tras la selección de las condiciones de expresión óptimas (inducción con 100 μ M de IPTG a 20 ° C durante 16 horas en medio LB), se intentó solubilizar la proteína en varios tampones (Tabla R. 3. 1) pero una ínfima cantidad de proteína parecía obtenerse en la fracción soluble (las muestras de sobrenadante se sometieron a Western-Blot pero la señal del anticuerpo anti-His obtenida no era concluyente).

Tabla R. 3. 1. Tabla recopilatoria de los tampones probados sin éxito para solubilizar la proteína ADN helicasa dependiente de ATP en las tres aproximaciones. Se indican las tres construcciones, junto a al tamaño en kDa, el punto isoeléctrico predicho y los tampones probados.

Proteína	Tamaño (kDa)	Punto isoeléctrico (pI)	Tampones probados en prueba de solubilidad sin éxito
Rv3202c (1 dominio)	30,82	8,76	PBS 1X pH 7,4
			PBS 1X pH 7,4 + 5 mM MgCl ₂
			PBS 1X pH 7,4 + 5 mM MgCl ₂ + Urea (1 - 4M)
Rv3202c (2 dominios)	81,14	9,40	200 mM Tris-HCl pH 8 + 200 mM NaCl
			10 mM Tris-HCl pH 8 + 1 mM EDTA + 200 mM NaCl + 50% glicerol + 2mM DTT + 5mM MgCl ₂
			10 mM Tris-HCl pH 8 + 200 mM NaCl + 20% glicerol + 1 mM β-mercaptoetanol + 5 mM MgCl ₂
			50 mM Tris-HCl pH 8 + 1 M KCl + 1 mM EDTA + 2 mM DTT
			200 mM Tris-HCl pH 8 + 200 mM NaCl + 5 mM MgCl ₂
			200 mM Hepes pH 7 + 400 mM NaCl + 5 mM MgCl ₂
Rv3202c (entero)	111,50	8,72	200 mM Hepes pH 7 + 400 mM NaCl + 5 mM MgCl ₂ + 1 mM β-mercaptoetanol + 20 % glicerol
			200 mM Hepes pH 7 + 400 mM NaCl + 5 mM MgCl ₂ + 20 % glicerol

Por tanto, se decidió purificar la proteína a partir de una expresión de 3L de medio LB. Se realizaron dos intentos, uno con cada uno de los dos tampones seleccionados ((A) 200 mM Hepes pH 7 + 200 mM NaCl + 5 mM MgCl₂; y (B) 200 mM Hepes pH 7 + 200 mM NaCl + 5 mM MgCl₂ + 1 mM EDTA + 50 % glicerol + 2 mM DTT). No obstante, la proteína de interés quedó en la fracción insoluble.

3.3. Una proteína de membrana que incluye un β -propeller (Rv3524).

La búsqueda en UniProt (ID I6X7J6) identifica a Rv3524 como una proteína de membrana probablemente conservada. El análisis de secuencia en Pfam identificó 3 repeticiones NHL, un motivo estructural conservado presente en una gran familia de reguladores de crecimiento. Mediante BLASTp, se encontró esta misma proteína en *Mycobacterium canetti*.

Junto a la búsqueda en las bases de datos, en el modelado *in-silico* con *Phyre2* se identificó un dominio β -propeller de 6 palas formadas por 3 o 4 hebras β ordenadas antiparalelamente, además de tres hebras β antiparalelas que parecen ser el punto de anclaje a la membrana (Figura R. 3. 4). Cada una de las palas del dominio β -propeller sería una repetición NHL, de modo que en lugar de 3 realmente habría 6 repeticiones.

El análisis de doble híbrido identificó la región de interacción con mDut en los residuos 234 a 343, que corresponde con las 3 últimas repeticiones NHL. No obstante, la primera aproximación fue el diseño y preparación de una construcción génica que incluyó solo el dominio β -propeller completo (residuos G79 al extremo C-terminal de la proteína), es decir, básicamente toda la proteína a excepción del fragmento teóricamente de inserción en la membrana bacteriana. De nuevo, la construcción está precedida por una etiqueta de histidinas, y se clonó en el vector pETNKL 1.1. Los cebadores utilizados en la extracción a partir del ADN genómico de *M. tuberculosis* y la clonación pueden consultarse en la tabla M. 1. Una vez comprobada la corrección de la construcción genética, se transformaron células *E. coli* BL21 (DE3) con el plásmido resultante. Se realizó la prueba de expresión y solubilidad y se seleccionaron las mejores condiciones y tampones para la posterior purificación. Las condiciones de expresión seleccionadas fueron el crecimiento en medio LB y la inducción posterior del cultivo con 100 μ M del inductor IPTG a 20 ° C durante 16 horas. Con el tampón de lisis celular seleccionado, 200 mM Tris-HCl pH 8 y 200 mM NaCl, la proteína se obtuvo mayoritariamente en la fracción soluble (el punto isoeléctrico de la proteína es 4,44). El sobrenadante fue sometido a cromatografía de afinidad, cuya elución se llevó a cabo con el mismo tampón de lisis pero suplementado con 500 mM de imidazol. La proteína eluida de mayor

pureza fue sometida a cromatografía de exclusión molecular con el tampón final 100 mM Tris-HCl pH 8 y 150 mM NaCl.

Una vez purificada y concentrada la proteína, se recurrió a los geles de poliacrilamida en condiciones nativas como primera aproximación para detectar interacción entre el β -*propeller* y la mDut. Se llevó a cabo un ensayo de titración fijando la concentración de mDut (11 μ M) e incrementando en ratios molares desde 0,5 a 4 la concentración de Rv3524, en presencia de 1 mM del sustrato no hidrolizable de la Dut, o en presencia de 5 mM del producto dUMP, o en la sola participación de ambas proteínas. Aunque se realizó en ambos sentidos, es decir, fijando la concentración de mDut o la del candidato, no se detectó interacción mediante esta aproximación.

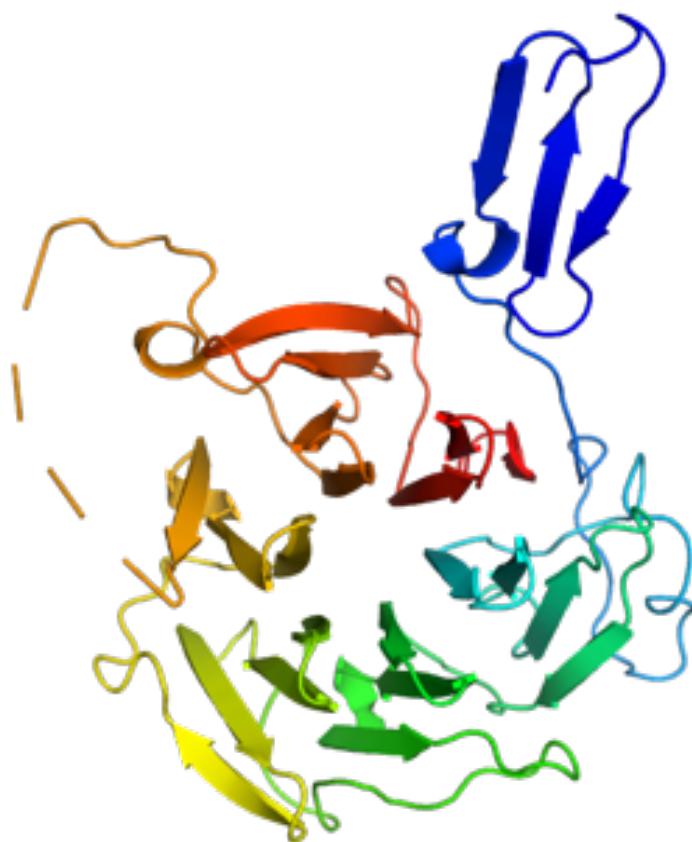


Figura R. 3.4. Modelado *in-silico* de la proteína Rv3524 obtenido en el servidor *Phyre2*.

Dado que estos resultados podrían deberse a una interacción débil, o a una interacción cuyos valores de K_{on} y K_{off} son demasiado altos para poder ver la formación del complejo por esta técnica, se perseveró en el intento de detectar interacción entre ambas proteínas. La siguiente aproximación fue comprobar si el β -*propeller* se unía a la mDut y ello afectaba la capacidad catalítica de la enzima. El resultado fue igualmente negativo, no detectando ningún cambio en

la cinética de mDut que pudiera ser atribuido a la interacción entre ambas proteínas. No obstante, este resultado no descarta totalmente la interacción, ya que ésta podría no involucrar al centro activo de la Dut y por ende no ver ningún efecto sobre su actividad enzimática.

Ante estos resultados, se decidió intentar una tercera aproximación, en este caso mediante una técnica cuantitativa que nos permitiera detectar constantes de afinidad bajas o cinéticas de unión rápida. Para ello se recurrió a la técnica BLI. La prueba inicial fijando en el sensor mDut a una concentración de 300 nM no detectó interacción con el β -*propeller* que actuaba como analito a una concentración de 3000 nM. Para descartar que la inmovilización de mDut en el sensor impidiera la unión, se realizó de nuevo el ensayo pero se inmovilizó el β -*propeller* en el sensor a una concentración de 600 nM y se utilizó mDut como analito a una concentración de 6000 nM. Desafortunadamente, se obtuvo el mismo resultado de no interacción. Las condiciones específicas de tiempo en cada ensayo pueden consultarse en la tabla M. 9.

Dada la elevada confianza de interacción detectada en el ensayo de doble híbrido entre este candidato y mDut, y los resultados negativos de las tres aproximaciones, se planteó la posibilidad de que el plegamiento de la proteína candidata no fuera el correcto y que ello provocara su incapacidad de interacción con la Dut. Para comprobar esta hipótesis, se recurrió a la técnica de termofluor para tener más datos sobre el plegamiento de la proteína (Figura R. 3. 5). En este ensayo, el aumento progresivo de la temperatura provoca la desnaturalización gradual de la proteína, lo cual a su vez incrementa la señal de fluorescencia emitida por las moléculas de un fluoróforo (SYPRO-orange) unidas a los residuos hidrofóbicos a medida que éstos quedan expuestos al solvente. Mientras que en proteínas correctamente plegadas el perfil de termofluor describe una fluorescencia normalizada creciente a medida que se incrementa la temperatura, en el caso de la proteína β -*propeller* mostró un perfil exactamente contrario.

No obstante, se decidió ratificar la hipótesis del mal plegamiento por otra técnica: dicroísmo circular. El análisis de los datos, asistido por el responsable del aparato, mostró que un 50 % de la estructura podría tener la estructura secundaria en hojas β esperada, pero el otro 50 % de la proteína serían otras estructuras secundarias (Figura R. 3. 6). Por tanto, ambas técnicas confirmaron la hipótesis: la proteína β -*propeller* no estaba bien plegada.

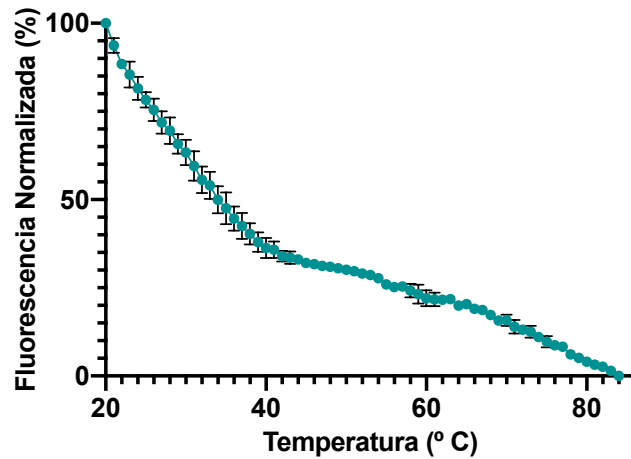


Figura R. 3. 5. Perfil de termofluor de la proteïna Rv3524 (β -propeller). Se muestra el valor de fluorescencia normalizada en función del incremento gradual de la temperatura. Se representan los resultados de dos réplicas del ensayo, con las correspondientes barras de error.

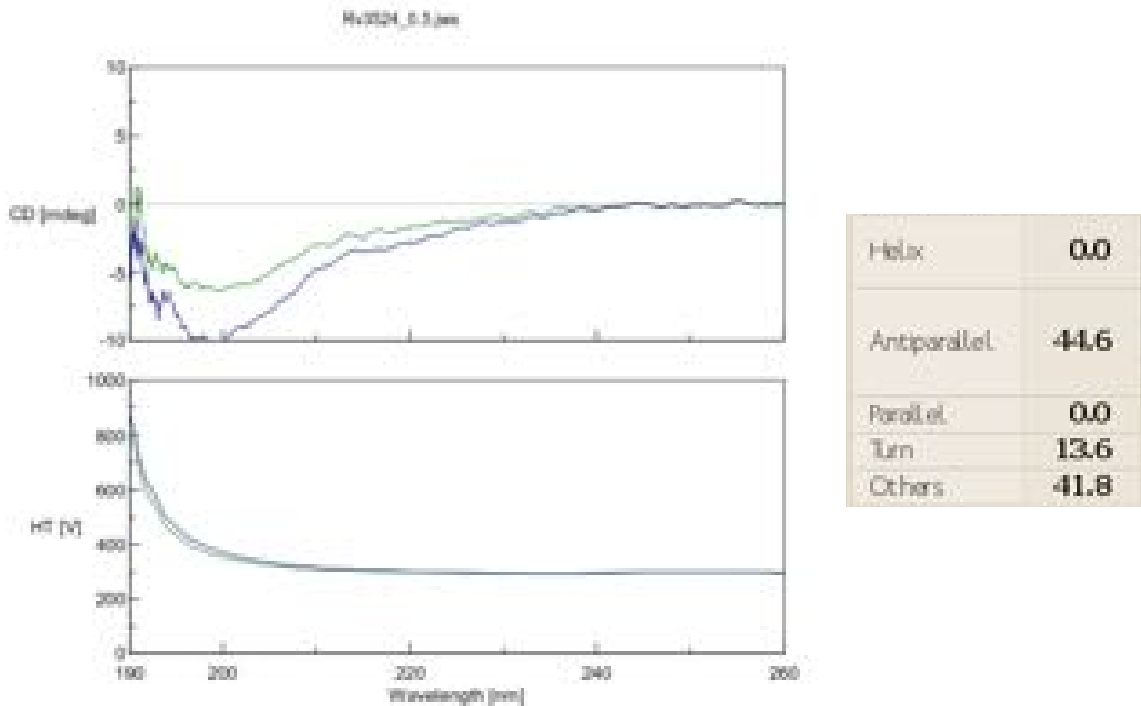


Figura R. 3. 6. Espectro de diàstro circular de la proteïna Rv3524 (β -propeller) (izquierda) y resumen de las estructuras secundarias encontradas (derecha). Según los picos invertidos obtenidos (gráfica de arriba) se hace una estimación de si la proteïna tiene estructura secundaria. El resumen de las estructuras secundarias extraído a partir de los datos introducidos en el software *BeStSel* (derecha) mostró que el 50 % de la proteïna poseía las hojas β esperadas, pero el otro 50 % poseía otras estructuras.

Finalmente, para descartar que los problemas del plegamiento fueran debidos a las condiciones de purificación, se analizó la bibliografía para buscar condiciones óptimas para proteínas con plegamientos similares. Se utilizaron las condiciones descritas por el Dr. Marcin y compañeros (Marcin, Neira, and Bravo 2015) para purificar el dominio β -*propeller* de la proteína Erb1, un factor de ensamblaje participante en la correcta maduración de la subunidad ribosómica 60S en levadura. La composición del tampón utilizado fue 50 mM Hepes pH 7,5 + 500 mM NaCl + 5 mM β -mercaptoetanol + 0,1 % Triton X-100 + 10 % glicerol. Casi el 100 % de la proteína quedó en la fracción insoluble.

3.4. Un regulador transcripcional de la familia PucR (Rv1429).

La búsqueda en UniProt (ID O06829) identifica a Rv1429 como una proteína conservada. Tanto en esta base de datos como en Pfam, se identifican 2 dominios en esta proteína: un dominio de regulador transcripcional de la familia PucR con una hélice-vuelta-hélice (HTH), y un dominio tipo CdaR – GGDEF, el cual puede encontrarse en un regulador de diácidos de carbohidratos (CdaR), en proteínas PutR (un activador transcripcional) y en proteínas relacionadas. Con la búsqueda en BLASTp, se encontró una identidad del 99 % con reguladores transcripcionales de la familia PucR presentes en *Mycobacterium canetti* y *Mycobacterium mungi*.

El análisis de doble híbrido identificó la región de interacción con mDut en los residuos 178 a 255. No obstante, debido al tamaño de la proteína (425 residuos, 47,7 kDa), en la primera aproximación se produjo la proteína completa fusionada en su extremo N-terminal a una etiqueta *StrepII-tag* que permitiera la posterior purificación por cromatografía de afinidad. El gen se obtuvo del ADN genómico de *M. tuberculosis*, y los cebadores utilizados pueden consultarse en la tabla M.1.

Una vez preparada la construcción en el vector pETNKI 1.5, células *E. coli* BL21 (DE3) fueron transformadas para las posteriores pruebas de expresión y solubilidad. Las pruebas de expresión se realizaron en distintos medios (LB, *Hyper*, *Superior*, *Turbo* y *Terrific*), con dos concentraciones de inductor (IPTG, 100 μ M y 1mM) y dos tiempos y temperaturas de inducción (37 ° C durante 3 h o 20 ° C durante 16 horas). Las pruebas de solubilidad se realizaron con 6 tampones distintos, que incluían diferentes agentes tamponantes, pHs, concentraciones de NaCl y la presencia de 5 mM MgCl₂ (tabla R. 3. 2). Tras las pruebas, el crecimiento del cultivo se llevaría a cabo en medio *Hyper*® y la

inducción de la expresión se realizaría con 1 mM de IPTG a 20 ° C durante 16 horas. No obstante, en todos los tampones probados la cantidad de proteína en la fracción soluble era muy similar a la de otros contaminantes presentes.

Ante estos resultados, la siguiente aproximación fue la solubilización de la proteína en presencia del agente caotrópico urea, observándose que a concentraciones no muy elevadas como 2 M se conseguía una elevada solubilización, y la proteína se mantenía soluble tras la posterior eliminación casi total de este agente por diálisis (0,5 M) (Figura R. 3. 7). Por tanto, el tampón para la purificación seleccionado es 200 mM Tris pH 8 + 500 mM NaCl + 5 mM MgCl₂ + 2 M urea.

Tabla R. 3. 2. Tabla recopilatoria de los tampones en la prueba de solubilidad sin éxito para la proteína Rv1429 (regulador transcripcional de la familia PucR).

Proteína	Tamaño (KDa)	Punto isoeléctrico (pI)	Tampones probados en prueba de solubilidad sin éxito
Rv1429 (entero)	47,70	6,08	150 mM Tris pH 8,5 + 300 mM NaCl
			75 mM Hepes pH 7,5 + 400 mM NaCl + 5 mM MgCl ₂
			200 mM Tris pH 8 + 200 mM NaCl + 5 mM MgCl ₂
			200 mM Tris pH 8 + 500 mM NaCl + 5 mM MgCl ₂
			200 mM Hepes pH 7 + 200 mM NaCl + 5 mM MgCl ₂
			200 mM Hepes pH 7 + 500 mM NaCl + 5 mM MgCl ₂

Tras la prueba de solubilidad con urea y posterior replegado de la proteína, se sometieron las muestras de las concentraciones 2, 1, 0,5 y 0 M del agente caotrópico a un ensayo de termofluor para determinar a qué concentración de urea mínima la proteína ya tenía un buen plegamiento (Figura R. 3. 8). De este ensayo pudo deducirse que la proteína replegada con 0,5 M de urea ya poseía un buen plegamiento.

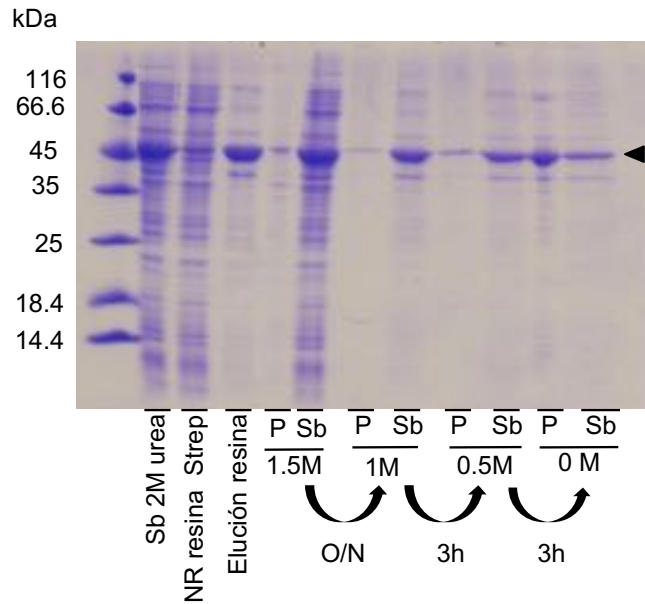


Figura R. 3. 7. Resultado de la electroforesis en gel de acrilamida en condiciones desnaturizantes para las muestras de la prueba de solubilidad desde 2 M urea de la proteína Rv1429. El orden de las muestras de izquierda a derecha es: marcador de tamaños (kDa), sobrenadante inicial tras la sonicación del pellet celular con tampón con 2 M urea, muestra del no retenido de la columna de afinidad, elución de la resina de afinidad, las ocho muestras correspondientes con las 4 concentraciones de urea (muestras de las fracciones de *pellet* y sobrenadante de cada concentración, respectivamente). Con las flechas curvas de indican los tiempos de diálisis entre los pasos sucesivos (toda la noche (O/N) o 3 h).

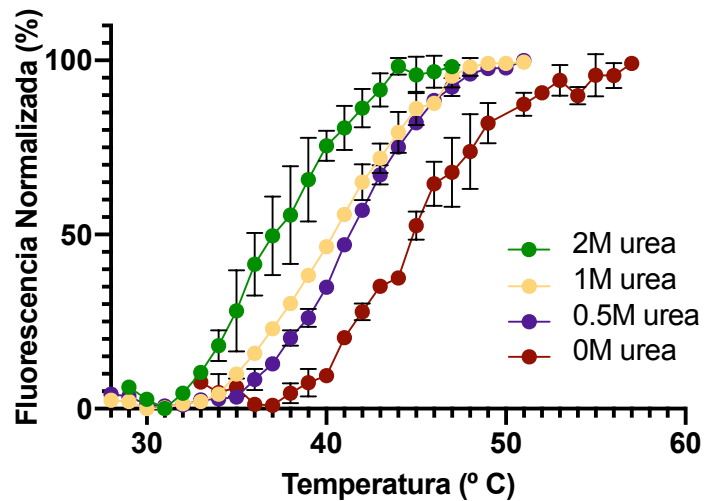


Figura R. 3. 8. Perfil de termofluor de las muestras del regulador transcripcional de la familia PucR a distintas concentraciones de urea (0, 0,5, 1 y 2 M). El aumento progresivo de la temperatura provoca la desnaturalización gradual de la proteína e incremento de la señal de fluorescencia emitida por el fluoróforo unido a los residuos hidrofóbicos. Se incluyen los datos de dos réplicas en el análisis y las correspondientes barras de error. Cada concentración se representa en un color distinto.

De las muestras de 0,5 y 0 M de urea obtenidas de la prueba de solubilidad se llevó a cabo una electroforesis en gel de acrilamida en condiciones nativas para intentar probar la interacción entre esta proteína candidata y la Dut de *M. tuberculosis*. Aunque no pudo observarse interacción, la muestra de la proteína candidata avanzó por el gel como una banda arrastrada o *smear* que no proporcionaba conclusiones verosímiles. Por tanto, se repetirá la purificación con el tampón 200 mM Tris-HCl pH 8 + 500 mM NaCl + 2 M urea para la lisis celular y unión a la resina de la columna de la cromatografía de afinidad, y el tampón 200 mM Tris-HCl pH 8 + 500 mM NaCl + 2 M urea + 2,5 mM destiobiotina para la elución de la misma columna. Posteriormente se realizará la eliminación de la urea de manera progresiva hasta una concentración final de 0,5 M mediante pasos sucesivos de diálisis. Con la proteína pura se intentará detectar interacción por otras técnicas (BLI, termoforesis...).

3.5. Una proteína de función desconocida (Rv3354).

La búsqueda en UniProt (ID O50383) identifica a R3354 como una proteína con un dominio de función desconocida. No obstante, la búsqueda mediante BLASTp resultó en una lipoproteína de membrana putativa de *M. tuberculosis* con una identidad del 88 %, identificada como LprJ. En la bibliografía se le asigna una función accesoria o como proteína de unión a ligando capaz de interactuar con la proteína KdpD, participante en la vía de transducción de señal Kdp, asociada con la respuesta primaria al estrés osmótico (Steyn, Joseph, and Bloom 2003).

La base de datos Pfam identificó 2 dominios en esta proteína, siendo uno de ellos un dominio con función desconocida (DUF 732, residuos 36 a 107), y el otro un péptido señal en el extremo N-terminal (residuos 1 a 32). Está anotada como una proteína con localización extracelular. Dado que el servidor *Phyre2* no pudo proporcionar un modelado de la proteína completa, para el diseño de las construcciones génicas se recurrió principalmente a la predicción de las estructuras secundarias. El ensayo de doble híbrido había identificado como zona de interacción los residuos incluidos entre la V48 a la P127 (línea azul en la figura R.3.9). Por tanto, se decidió preparar una construcción que incluía desde la G47 a la S129 (extremo C-terminal) (línea roja en la figura R. 3.9), de modo que se eliminaba el péptido señal y no se interrumpía *a priori* ningún elemento de estructura secundaria. La construcción, introducida en el vector pETNKI 1.1, y con el gen extraído del ADN genómico de *M. tuberculosis*, incluía una etiqueta de histidinas en el extremo N-terminal para la posterior purificación por cromatografía de afinidad. Los cebadores utilizados pueden consultarse en la tabla M.1.

Las pruebas de expresión se realizaron en distintos medios (LB, *Hyper*, *Superior*, *Turbo* y *Terrific*), con dos concentraciones de inductor (IPTG, 100 µM y 1mM), y en dos tiempos y temperaturas de inducción (37 ° C durante 3 h o 20 ° C durante 16 horas). Las condiciones óptimas de expresión seleccionadas fueron 100 µM IPTG a 37 ° C durante 3 horas. No obstante, a diferencia de las otras proteínas candidatas, la construcción fue expresada en células *Rosseta 2*, ya que los codones del aminoácido prolina están codificados por el triplete CCC. Éste constituye un codón raro para *E. coli*, y las células BL21 (DE3) no están optimizadas para su expresión, al contrario que las *Rosseta2* que suplementan con tRNAs específicos para suplir a los codones AUA, AGG, AGA, CUA, CCC, y GGA. Aunque respecto al total de codones, el triplete CCC representa el 3 %, ya se había descrito con anterioridad la alteración de la expresión proteica en *E. coli* con una presencia del 2,3 % (Kane 1995).

En la prueba de solubilidad se incluyeron 5 tampones con distintas composiciones pero no pudo obtenerse la proteína en la fracción soluble con ninguno de ellos (Tabla R. 3.3). Asimismo, se intentó la solubilización a partir de los cuerpos de inclusión con urea, utilizando hasta 4 M de dicho agente caotrópico, pero la proteína se mantuvo en la fracción insoluble.

Tabla R. 3. 3. Tabla recopilatoria de los tampones en la prueba de solubilidad sin éxito para la proteína Rv3354.

Proteína	Tamaño (KDa)	Punto isoeléctrico (pI)	Tampones probados en prueba de solubilidad sin éxito
Rv3354 (G47 a final)	8,27	3,90	200 mM Tris-HCl pH 8 + 200 mM NaCl
			200 mM Hepes pH 7 + 400 mM NaCl
			200 mM Tris-HCl pH 8 + 400 mM NaCl
			200 mM Tris-HCl pH 8,5 + 400 mM NaCl
			200 mM Tris-HCl pH 8 + 200 mM NaCl + 4 M urea

Tras estos resultados, se decidió diseñar y preparar una nueva construcción génica. En esta nueva aproximación se incluyeron los residuos N33 a la S129, es decir, toda la proteína a excepción del péptido señal (línea morada en la figura R.3.9).

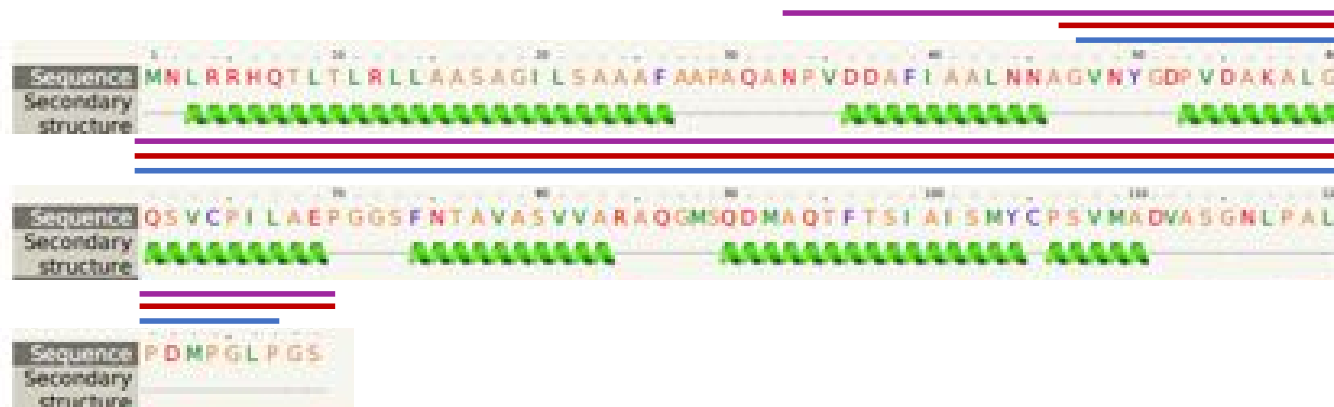


Figura R. 3.9. Predicción de la estructura secundaria de la proteína Rv3354 por el servidor *Phyre2*. La línea azul identifica la región de interacción identificada en el ensayo de doble híbrido, la línea roja identifica la región incluida en la primera construcción génica (residuo G47 a S129), y la línea morada identifica la región incluida en la segunda construcción génica (N33 a S129).

Esta segunda construcción fue utilizada para la transformación de células *Rosetta2*, y las mismas condiciones de expresión que en la construcción anterior fueron utilizadas (100 μ M IPTG a 37 ° C durante 3 horas). La prueba de solubilidad incluyó los mismos tampones probados en la construcción anterior, pero no se recurrió al agente caotrópico urea. Se consiguió obtener proteína en la fracción soluble y se confirmó por Western-Blot (Figura R. 3. 10), seleccionando como tampón de purificación 200 mM Tris-HCl pH 8 + 200 mM NaCl. El siguiente paso será la purificación en estas condiciones, y con la proteína pura se intentará detectar interacción con mDut por diversas técnicas (Nativo-PAGE, BLI, termoforesis...).

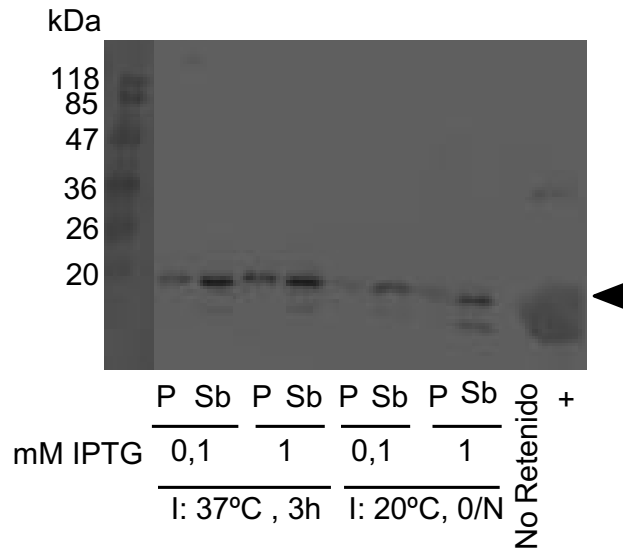


Figura R. 3. 10. Membrana revelada tras someter las muestras de la prueba de solubilidad de la proteína Rv3353 (N33 a S129) a Western-Blot. Las muestras incluyen las fracciones insoluble (P) y soluble (Sb) de las 4 condiciones de expresión probadas y con el tampón de solubilización seleccionado. Las muestras, de izquierda a derecha, son: marcador molecular (kDa), muestra de la fracción insoluble y soluble de la expresión a 37 °C durante 3 h con 0,1 mM de IPTG, muestra de la fracción insoluble y soluble de la expresión a 37 °C durante 3h con 1 mM de IPTG, muestra de la fracción insoluble y soluble de la expresión a 20 °C durante toda la noche con 0,1 mM de IPTG, muestra de la fracción insoluble y soluble de la expresión a 20 °C durante toda la noche con 1 mM de IPTG, muestra de no retenido por la columna de afinidad, y control positivo.

DISCUSIÓN

1. El represor StI de SaPIbov1 como un inhibidor de Duts universal.

1.1 El mimetismo como garantía de no escape.

La interacción de StI con Duts triméricas y diméricas ha sido descrita a nivel genético y bioquímico por diversos laboratorios (Szabó et al. 2014; Benedek et al. 2019; 2018; Nyíri et al. 2018; Tormo-Más et al. 2013; Alite et al. 2017; Hirmondó et al. 2015; Donderis et al. 2017; Tormo-Más et al. 2010; Maiques et al. 2016; Hill and Dokland 2016). Su base molecular, tanto con Duts diméricas como triméricas de fagos de *S. aureus*, fue caracterizada por el laboratorio donde se ha realizado esta Tesis (Rafael Ciges-Tomas et al. 2019).

Para interactuar con Duts de fagos de *S. aureus*, StI de la SaPIbov1 utiliza una estrategia de mimetización del sustrato, interactuando con residuos del centro activo de la Dut. En este trabajo de Tesis se ha analizado si este mecanismo es también utilizado para reconocer Duts triméricas de organismos procariontes y eucariotes con los cuales los fagos de *S. aureus* no mantienen un contacto evolutivo, pero para las que sí se ha descrito interacción. Con este fin, se han resuelto las estructuras tridimensionales de los complejos formados por StI y Duts representativas de estos grupos, la Dut de *M. tuberculosis* y la humana, ya caracterizadas estructural y funcionalmente (Mol, Harris, McIntosh, et al. 1996; Varga et al. 2007; Chan et al. 2004; Hirmondó et al. 2015), y cuya interacción con StI había sido demostrada (Nyíri et al. 2018; Hirmondó et al. 2015; Pecsí et al. 2012). El análisis comparativo de estas estructuras con el complejo biológico de StI con la Dut del fago $\phi 11$ de *S. aureus*, y con la estructura descrita durante la realización de esta Tesis de StI con la de Dut del crustáceo *Litopenaeus vannamei* (Wang et al. 2021), ha permitido confirmar una estrategia de interacción común entre el represor StI y las Duts triméricas de fago, bacteria y eucariota.

En esta estrategia de mimetización del sustrato llevada a cabo por el represor, éste interactúa con la mayor parte de los residuos implicados en el reconocimiento del sustrato y/o en la catálisis. Esto implica que la unión de StI a la enzima produzca la inhibición de su actividad enzimática, ya que ocupa el lugar del sustrato y ejerce una inhibición competitiva con el dUTP. De hecho, se han descrito inhibiciones del 40 % en *Drosophila melanogaster*, 80 % en *M. tuberculosis*, y hasta el 100 % en la Dut del fago $\phi 11$ de *S. aureus* (Benedek et al. 2018; Hirmondó et al. 2015; Nyíri et al. 2018).

La gran cantidad de residuos de Dut implicados en la interacción y su necesidad de conservación para llevar a cabo la actividad enzimática, impide o dificulta que la enzima pueda presentar mutaciones para evitar la unión a StI, ya que comprometería su actividad. La capacidad de interactuar con un amplio número de residuos catalíticos que son conservados en estas enzimas y

representan su firma de secuencia, ha permitido al represor ser capaz de interactuar con otras Duts extremadamente alejadas evolutivamente de su diana biológica.

De hecho, Stl ha sido propuesto como un buen candidato como inhibidor proteico de Duts (Nyíri et al. 2018), capaz de actuar como un péptido inhibidor, interactuando con múltiples residuos simultáneamente que aseguran la unión del inhibidor a la proteína diana. Además, siendo conocidas las estructuras de los dominios de Stl implicados en la interacción de Duts diméricas y triméricas, podría ser utilizado como modelo para diseñar otros péptidos inhibidores de estas Duts.

1.2. Stl interacciona con las 3 familias de Duts.

Esta estrategia de mimetización le ha permitido a Stl, mediante el reclutamiento de dos dominios distintos, la interacción con Duts de dos de las tres familias descritas: triméricas y diméricas. La tercera familia de Duts, las monoméricas, son exclusivas de herpesvirus, pero con una gran similitud en el centro activo con Duts triméricas. Dada la estrategia de mimetización de la diana-nucleótido, se planteó la posibilidad de que el represor también interactuara con esta familia de Duts mediante el mismo dominio utilizado en triméricas y que igualmente las inhibiera, aunque las diferencias entre ambas familias podrían impedir la unión. En este trabajo de Tesis se ha confirmado la interacción entre ambas proteínas. Para ello se ha utilizado como modelo la Dut monomérica del virus de Epstein-Barr como representativa de la familia ya que se dispone de datos estructurales y caracterización previa. Estos datos estructurales nos han permitido generar *in silico* el modelo del complejo con Stl^{N-ter}. Este modelo era congruente con un mecanismo de interacción muy similar al empleado en Duts triméricas: las cuatro tirosinas del represor (Y105, Y112, Y113 y Y116) mimetizarían el sustrato dUTP interactuando con los residuos conservados en las Duts monoméricas que lo reconocen. De este modo, la Y113 de Stl ocuparía el lugar del uracilo, la Y112 el espacio de la desoxirribosa, la Y105 se situaría cerca del sitio ocupado por el fosfato γ , y la Y116 también se situaría en el centro activo, pero algo alejada de la posición que ocuparía el fosfato β (Figura D. 1). Este modelo *in silico* ha sido apoyado experimentalmente por ensayos bioquímicos que han mostrado que la actividad enzimática de la Dut del virus de Epstein-Barr es inhibida al 60 % por Stl, comparable con la inhibición del 70 y el 80 % inducida por el represor en hDut y mDut, respectivamente. Por tanto, en esta Tesis ha sido mostrada la inhibición de una Dut monomérica, y por tanto, se confirma que Stl puede inhibir a todas las familias de Duts.

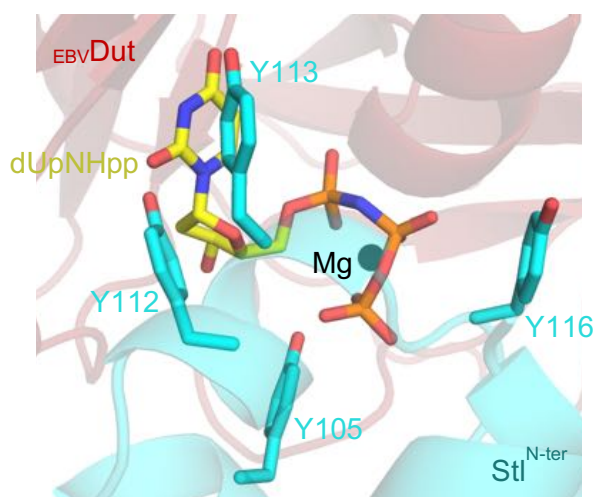


Figura D. 1. El Stl^{N-ter} mimetizaría el nucleótido en el centro activo de la EBVDut. Superposición de la estructura de EBVDut en complejo con el sustrato no hidrolizable dUpNHpp (PDB 2BT1) con el modelo *in-silico* del complejo EBVDut-Stl^{N-ter}. Se representan en varillas de color cian los residuos de Stl que mimetizarían el sustrato, mientras que el sustrato se representa en varillas y en color según átomo (C en amarillo, O en rojo, N en azul, P en naranja), y el ión Mg²⁺ en una esfera negra.

Además, el alineamiento de más de 10.300 secuencias de Duts triméricas y monoméricas (Figura R.1.19) permitió concluir que en la mayoría de los casos los residuos de los motivos de las Duts que participan en la interacción con Stl están altamente conservados (a excepción del motivo I), ratificando que Stl posiblemente podrá reconocer a un gran número de estas Duts y por lo tanto el carácter universal de Stl como inhibidor de Duts.

Debido a la gran amplitud de la familia de Duts, en este alineamiento fueron eliminadas las secuencias de Duts de gran tamaño, como la Dut de *Caenorhabditis elegans* y *Drosophila virilis*. En ambas especies, la Dut es monomérica, pero su secuencia muestra la presencia de 3 copias muy similares de un pseudo-gen de Dut trimérica. En el caso de *Drosophila*, el modelado *in-silico* muestra un plegamiento mimetizando una Dut trimérica, donde cada copia del gen correspondería con un monómero de una Dut trimérica, con segmentos peptídicos para la conexión covalente entre las pseudo-subunidades (Figura D. 2) (Benedek et al. 2016). En el caso de Dut de *C. elegans*, el modelador Phyre2 solo fue capaz de modelar un 30 % de la secuencia, constituyendo un monómero de Dut trimérica, con lo cual se esperaría la misma estructura que en el caso de *D. virilis* (Figura D. 2). Cada una de las tres repeticiones fusionadas mantienen los 5 motivos catalíticos de las Duts triméricas (Figura D. 3) y pequeñas variaciones en secuencia, pero mantienen todos los residuos de las Duts triméricas asociados a la interacción con el represor. Por tanto, cabría esperar que Stl interaccionara con estas Duts, un caso de pseudo-trímero, cuyo origen

podría asociarse con dos duplicaciones independientes a partir de una de las subunidades.

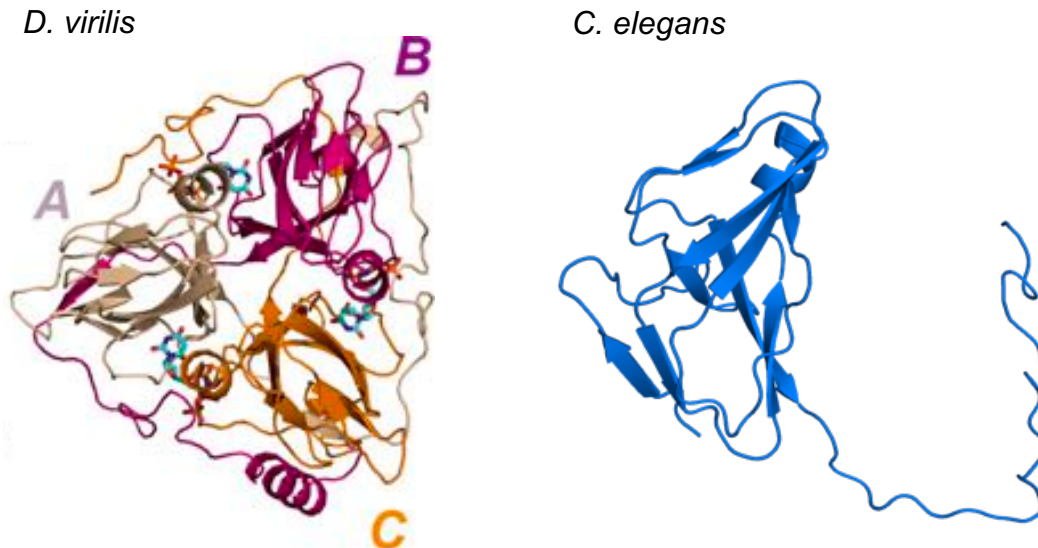


Figura D.2. Modelados *in-silico* de las Duts de *D. virilis* (izquierda) y *C. elegans* (derecha). En el caso de Dut de *D. virilis* se identifican en tres colores y letras distintos las 3 copias del gen (marrón claro, magenta y naranja; A, B y C). En el caso de Dut de *C. elegans* solo se modelizó un monómero con el 30 % de la secuencia.

Si aún en estos casos que podrían considerarse excepciones de Dut podría esperarse que StI pudiera interactuar con ellas, este represor ha mostrado como la evolución ha diseñado una proteína profundamente especializada. Estas observaciones podrían abrir la puerta a que StI inhibiera otros miembros de la familia NTP-hidrolasas que guardan homología mecánica con las Duts.

De hecho, en esta Tesis se ha trabajado con las proteínas MazG de un fago de *S. epidermidis* y uno de *E. faecalis*, cuyo centro activo es muy similar al de una Dut dimérica. Por tanto, no sería ilógico pensar que la interacción entre estas MazG y StI fuera posible. Aunque los ensayos *in vitro* realizados en una Tesis anterior mostraron que no había interacción, no podría descartarse que otros miembros de la familia MazG sí interaccionaran con este represor.

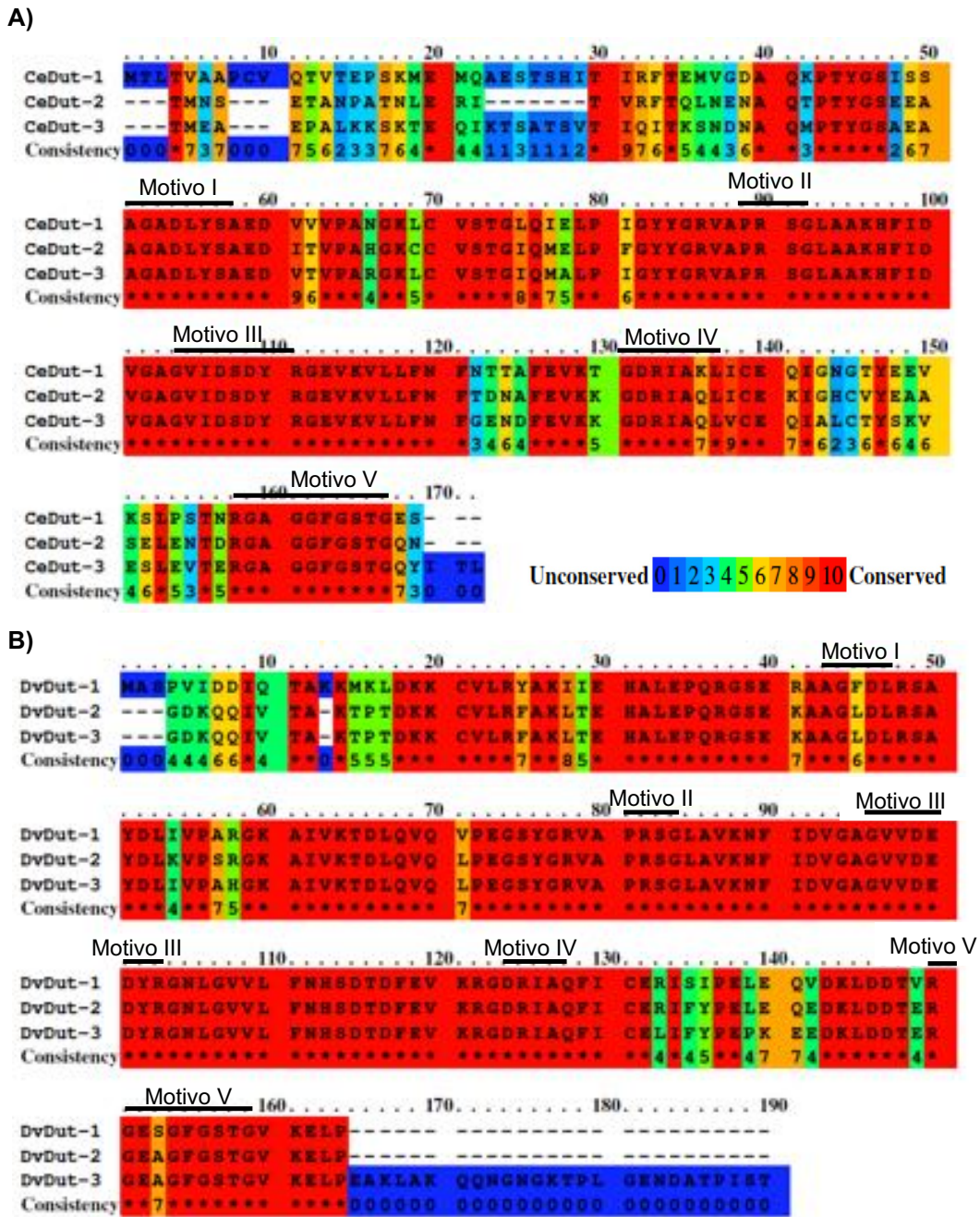


Figura D. 3. Alineamientos de las 3 copias de los genes de Dut de *C. elegans* (A) y *D. virilis* (B). Cada una de las copias de ambos genes conservan los 5 motivos catalíticos, y las pequeñas variaciones en secuencia no incluyen a los residuos que interaccionan con Stl.

1. 3. El motivo V de las Duts triméricas juega un papel activo en la interacción con el represor StI de SaPIbov1.

Estudios previos con las Duts triméricas de los fagos de *S. aureus* $\phi 80\alpha$ y $\phi 11$ habían arrojado resultados contradictorios sobre la implicación del motivo V en el reconocimiento y unión a StI. Ambas Duts son antirepresoras de islas con represores StI de la familia de SaPIbov1. El estudio de varios mutantes puntuales de Dut $\phi 80\alpha$ dio lugar a la conclusión inicial de que la ordenación del motivo V era necesaria para la interacción con StI y consecuente inducción de la SaPI, y dado que el dUTP tenía un papel esencial al facilitar la ordenación del motivo V, su participación era requerida (Tormo-Más et al. 2013). No obstante, este planteamiento sobre el papel del dUTP fue refutado por otro laboratorio trabajando con la Dut del fago $\phi 11$ (Szabó et al. 2014), proponiendo que la entrada del sustrato en el centro activo de la Dut y la interacción de la Dut con StI eran procesos excluyentes.

Tras esta publicación y la controversia surgida, se volvió a plantear el papel del sustrato y del motivo V en la interacción de Dut con StI. Se diseñaron nuevos y específicos mutantes, tanto de la $\phi 80\alpha$ Dut como de la $\phi 11$ Dut, y se realizó un ensayo comparativo de la capacidad de estos mutantes en el reconocimiento de StI, en presencia y ausencia del sustrato, así como en la inducción de la movilidad de la isla (Maiques et al. 2016). Los ensayos con los mutantes mostraron que la $\phi 80\alpha$ Dut se une más débilmente a StI y la contribución del motivo V es clave para estabilizar la unión, de modo que en su ausencia la baja afinidad impide que *in-vivo* la Dut desreprima la isla. Al mismo tiempo, los ensayos con los mutantes de $\phi 11$ Dut, con una afinidad 5 veces mayor por el represor, mostraron una menor contribución del motivo V para alcanzar el grado de unión a StI necesario para la inducción de la SaPI. Contrariamente a la participación diferencial del motivo V según la Dut en la interacción con StI, este motivo siempre sería necesario para la inhibición por sustrato, ya que su plegamiento sobre el centro activo ocupado por el dUTP impediría la interacción de StI con la Dut. De estos resultados se propuso un modelo convergente de interacción entre StI y Duts triméricas, en el cual la forma Dut *apo* es la competente para interaccionar con el represor, los motivos IV y VI se encargan del reconocimiento de StI, y el motivo V estabiliza el complejo. El dUTP actúa a modo de interruptor, permitiendo o impidiendo la interacción (Figura I. 27) (Maiques et al. 2016).

Más adelante, con la resolución de la estructura del complejo de Dut del fago $\phi 11$ con StI^{N-ter}, quedó confirmada la participación de los motivos conservados I, II, III, IV y VI de Dut, esclareciéndose las interacciones específicas entre los residuos de estos motivos con este fragmento del represor, y quedando pendiente las interacciones con el motivo V, el cual estaba desordenado en la estructura y no pudo visualizarse (Rafael Ciges-Tomas et al. 2019).

La resolución de la estructura de Stl^{N-ter} en complejo con hDut en esta Tesis permitió visualizar como el motivo V interviene en la unión de Stl, pudiendo observar el extremo C-terminal en su proyección desde la Dut hasta la inserción en Stl, y con ello confirmando la participación de todos los motivos catalíticos de la Dut en la interacción con el represor. Los residuos R130, S132 y F135 del extremo C-terminal de hDut juegan un papel crucial en la interacción con el represor. Especialmente relevante es el residuo F135, que se introduce en un bolsillo hidrofóbico localizado entre el dominio de unión al ADN y el dominio intermedio del Stl. Aunque otros residuos del motivo V a parte de estos tres participan de manera directa, la F135 es el principal punto de anclaje al Stl, además de estabilizar el *P-loop* (Figura R. 1.15).

La participación del motivo V de hDut en la formación del complejo con Stl se suma a otras dos observaciones de este mecanismo de interacción descritas durante el desarrollo de esta Tesis y que corresponden a los complejos de la misma porción N-terminal de Stl (residuos 1-156) con las Duts de fago de *S. aureus* $\phi 80\alpha^{E112K}$, resuelta en el laboratorio, y la Dut de *L. vannamei* (Alite Hernández 2021; Wang et al. 2021). Ambos complejos muestran que residuos similares del motivo V a los observados en hDut-Stl^{N-ter} son utilizados para interaccionar con Stl. La Dut del fago $\phi 80\alpha^{E112K}$ presenta una mutación puntual en el motivo VI que incrementa un orden de magnitud su afinidad por Stl respecto a la Dut WT, y fue utilizada en otra Tesis con el fin de conocer la contribución del motivo V (Alite Hernández 2021).

No obstante, en estas dos estructuras las densidades electrónicas correspondientes a los motivos V eran más pobres, apoyando la naturaleza altamente flexible del extremo C-terminal. En el caso del complejo con la Dut de *L. vannamei* no fue posible observar la densidad electrónica en todas las subunidades de Stl, mientras que en el complejo con Dut de fago $\phi 80\alpha^{E112K}$ no fue posible visualizar la totalidad del motivo, no obteniéndose la densidad electrónica para algunos residuos localizados entre el cuerpo de la Dut y el bolsillo hidrofóbico de Stl. Estas observaciones confirmaron que el motivo V actuaría a modo de cierre de seguridad, asegurando la interacción con Stl una vez se ha producido el reconocimiento entre ambas proteínas (Figura D. 4).

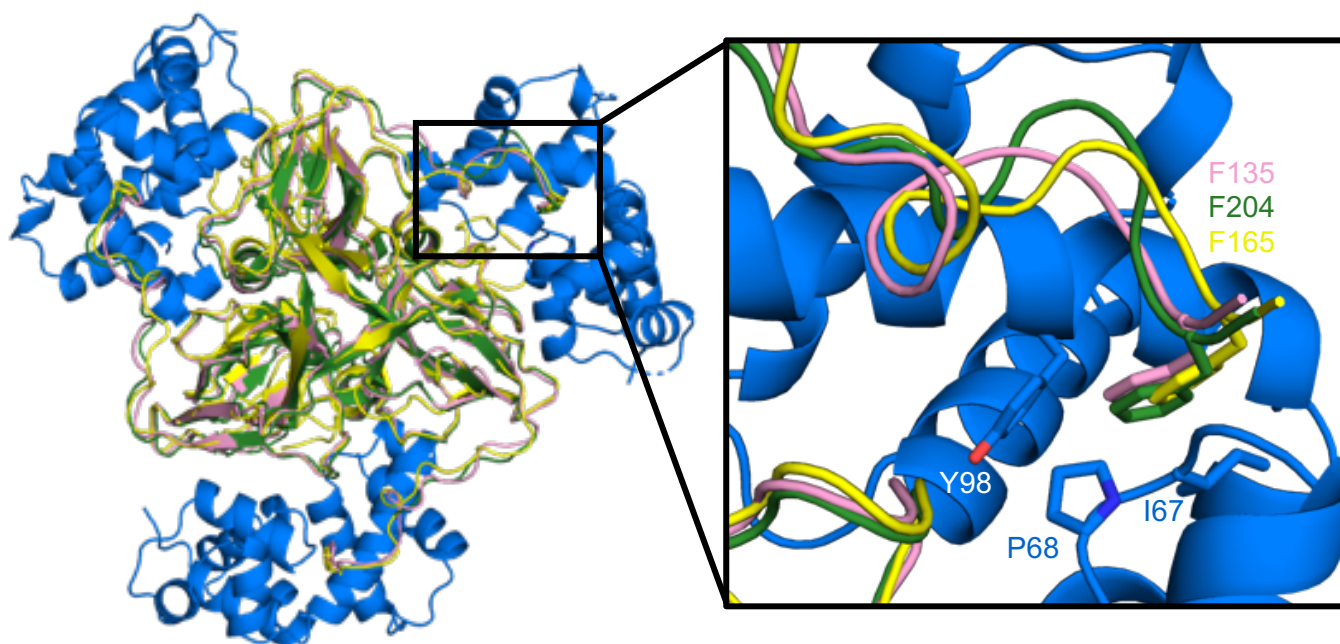


Figura D. 4. Localización del extremo C-terminal en la estructura de tres complejos de Stl^{N-ter} con tres Duts triméricas. Se muestran las estructuras de los complejos de Stl^{N-ter} con la hDut (rosa), $\phi 80\alpha\text{Dut}^{\text{E112K}}$ (amarillo) y lvDut (verde). Se destaca en varillas la fenilalanina conservada de las Duts que se inserta en el bolsillo hidrofóbico del represor, del cual se destacan en varillas tres residuos participantes (I67, P68, Y98). Código PDB de la estructura lvDut con Stl^{N-ter} 7DLV. Estructura de $\phi 80\alpha\text{Dut}^{\text{E112K}}$ con Stl^{N-ter} tomada de *Alite Hernández 2021*. Estructura hDut con Stl^{N-ter} resuelta en esta Tesis.

La esencialidad del residuo de fenilalanina del motivo V en la interacción con Stl se había intentado probar previamente mediante medidas de afinidad con mutantes puntuales en $\phi 11\text{Dut}$. En concreto se utilizó un mutante que sustituía esta fenilalanina por un triptófano (F164W). Esta mutación puntual no tuvo efectos en la actividad catalítica ni en la afinidad de la Dut por el sustrato (Leveles et al. 2013), y el análisis de su unión a Stl, calculada por microbalanza de cristal de cuarzo, tampoco mostró diferencias, apoyando la no implicación en la unión al represor (Szabó et al. 2014). Sin embargo, los resultados del laboratorio mostraron posteriormente que la participación del motivo V de la $\phi 11\text{Dut}$ en la interacción con Stl es mínima. En el caso de Duts con menor afinidad por Stl, como $\phi 80\alpha\text{Dut}$ o hDut, parece que esta contribución es mucho más relevante. Nuestros ensayos con la Dut de *M. tuberculosis*, en la cual este residuo está sustituido por una histidina, apoya esta propuesta. La sustitución de esta histidina por un residuo de fenilalanina (mDut^{H145F}) incrementa la afinidad de la Dut por el represor 5 veces.

Dado que la sustitución de la fenilalanina por una histidina en la Dut de *M. tuberculosis* no es única en la naturaleza, y no tiene un efecto importante sobre

la actividad enzimática, surge la duda de porqué los fagos de *S. aureus* no han llevado a cabo la misma sustitución aminoacídica para escapar de la interacción con Stl y consecuente expresión de la SaPI.

1.4. ¿Es Stl un inhibidor universal absoluto de Duts?

En línea con este posible pequeño cambio que podrían generar Duts con capacidad de escape a la unión a Stl, se han descrito Duts cuya estructura anticipa que serían refractarias a esta unión, y por lo tanto a su inhibición por Stl. Este es el caso de Duts de ciertos Lactobacillales como *Streptococcus* y *Enterococcus*, o la Dut del virus de mono de Mason-Pfizer, como se podría deducir a partir de sus estructuras tridimensionales (*E. faecalis* (Ciges-Tomas 2019), *S. mutans* (código PDB 3H6X), y Mason-Pfizer (código PDB 2D4N).

Esta asunción se basa en que estas Duts disponen al motivo V de una manera no canónica, ya que no se proyecta desde la tercera subunidad del trímero sobre el centro activo que conforman las otras dos, sino que se repliega acomodándose sobre el centro activo de la misma subunidad. Aunque es posible que en estas Duts la participación del motivo V sea mínima, como en el caso de la Dut del fago $\phi 11$, y por tanto éste no supondría un impedimento para la interacción con Stl; también podría ser posible que fuera requerido como en el caso de la Dut $\phi 80\alpha$, de modo que la interacción con Stl sería muy débil o no podría ser posible.

En las Duts de ciertos Lactobacillales, los residuos correspondientes al motivo V están conservados, así como su posicionamiento sobre el centro activo (Ciges-Tomas 2019). No obstante, junto a la disposición no canónica del motivo V que impediría la estabilización del complejo Dut-Stl, poseen una inserción de 6 residuos (*lb-loop*, aminoácidos NNENNE) junto al motivo III, muy próxima al centro activo (Figura D. 5). La estructura tridimensional de la Dut de *E. faecalis* muestra que la disposición de este *lb-loop* impone impedimentos estéricos a la interacción con el Stl, ya que las hélices $\alpha 5$, $\alpha 7$, $\alpha 8$ y $\alpha 9$ del represor (las principales implicadas en la interacción con las Duts triméricas) ocuparían las posiciones de esta inserción, de acuerdo con los complejos descritos (Figura D. 6). En el caso de la estructura tridimensional de Dut de *S. mutans*, no puede visualizarse el extremo C-terminal, pero la presencia del *lb-loop*, que también supondría un impedimento estérico para la interacción con el represor Stl, así como la longitud y secuencia de los extremos C-terminales indican una idéntica conformación a la Dut de *E. faecalis*.

En la Dut del virus del mono Mason-Pfizer (MPV) ocurre algo similar (Németh-Pongrácz et al. 2007). Se ha propuesto que el motivo V, el cual conserva la mayoría de los residuos presentes en otras Duts triméricas, se ordena sobre el centro activo de la misma subunidad y no en el de la subunidad adyacente. El cambio de orientación podría deberse al menor número de residuos presentes

entre el motivo IV y V (21 residuos en hDut, 17 residuos en la Dut del virus del mono *Mason-Pfizer*), que imposibilitaría que el motivo V pudiera establecer interacciones con el ligando del centro activo de la subunidad adyacente. Por tanto, en cada centro activo participan 2 y no las 3 subunidades de la Dut. Estos resultados no pueden visualizarse en la estructura tridimensional, ya que solo es visible el inicio del C-terminal asumiendo una conformación alternativa a la canónica (Figura D. 7). No obstante, incluso en una conformación completamente extendida no podría cubrir los 36 Å para poder rodear y llegar al centro activo contiguo. Por tanto, al igual que en el caso de la Dut de *E. faecalis*, cabría esperar que en el caso de que el represor StI interaccionara con esta Dut, no podría estabilizarse el complejo con el extremo C-terminal.

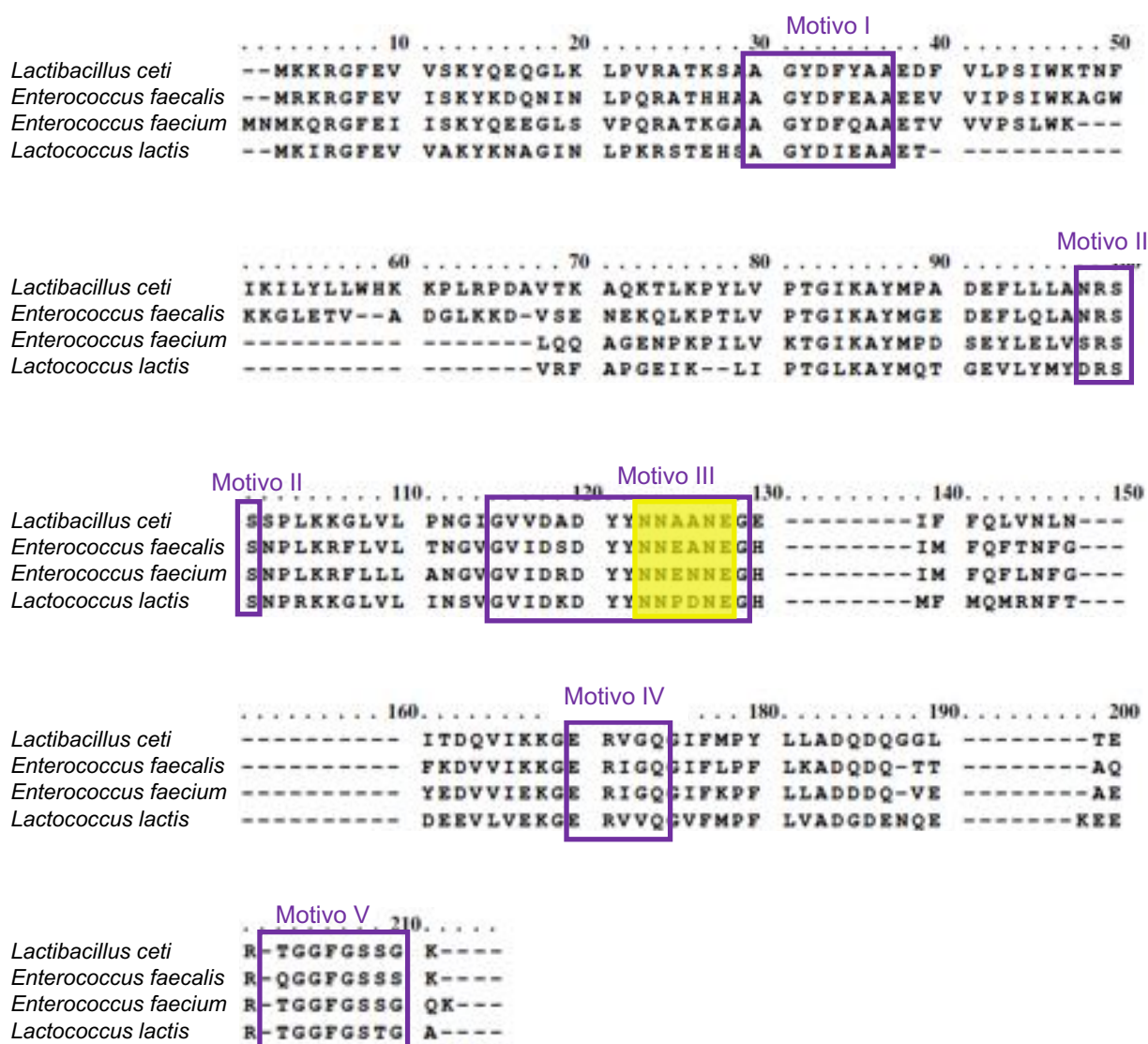


Figura D. 5. Localización del *lb-loop*. Se muestra el alineamiento de Duts de bacterias de distintos géneros del orden Lactobacillales. En cuadrados morados se identifican los 5 motivos conservados de Duts triméricas y en un recuadro amarillo de resalta la inserción de 6 residuos característica de este orden (*lb-loop*).

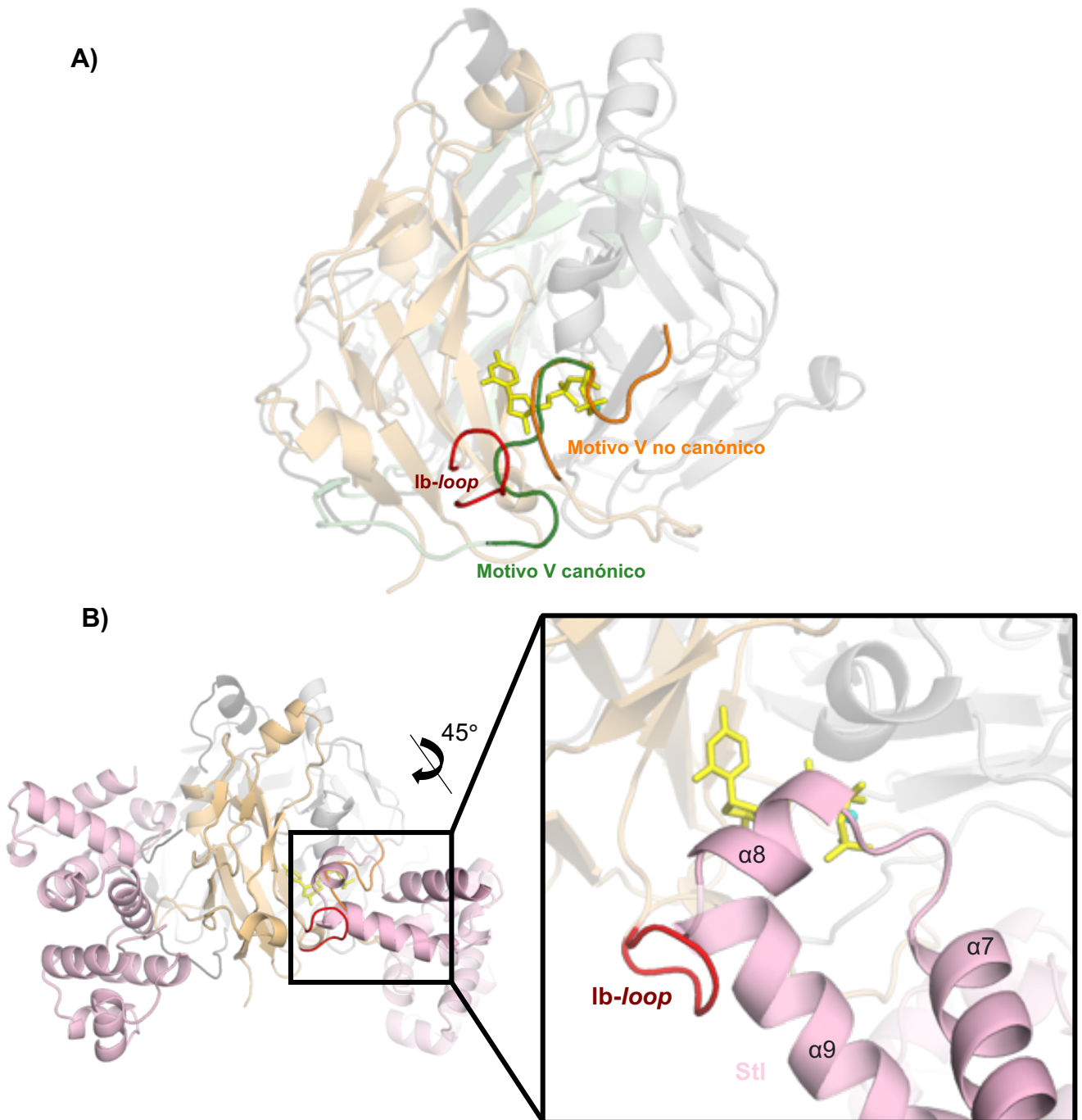


Figura D. 6. La Dut de *E. faecalis* ordena el motivo V sobre el centro activo de la misma subunidad, y el *Ib-loop* podría suponer un impedimento estérico para la interacción con StI. A) Superposición de la Dut de *E. faecalis* sobre hDut. Solo se representa un monómero de hDut y se identifica en él el motivo V que se ordena de forma canónica en Duts triméricas sobre el centro activo de la subunidad contigua (monómero y motivo V en verde); en rojo se resalta el *Ib-loop*, y en naranja el motivo V de Dut de *E. faecalis* ordenado sobre el mismo centro activo (monómero en naranja). Estructura tomada de *Ciges-Tomas 2019*. B) Superposición de la Dut de *E. faecalis* sobre el complejo de hDut-StI^{N-ter}. Se representan los monómeros de StI (rosa) y de Dut de *E. faecalis* (naranja y grises). En rojo se resalta el *Ib-loop*. Se identifican las hélices α de StI implicadas en la interacción con Duts triméricas. Los nucleótidos se representan en varillas color amarillo y los Mg^{2+} como esferas color cian.

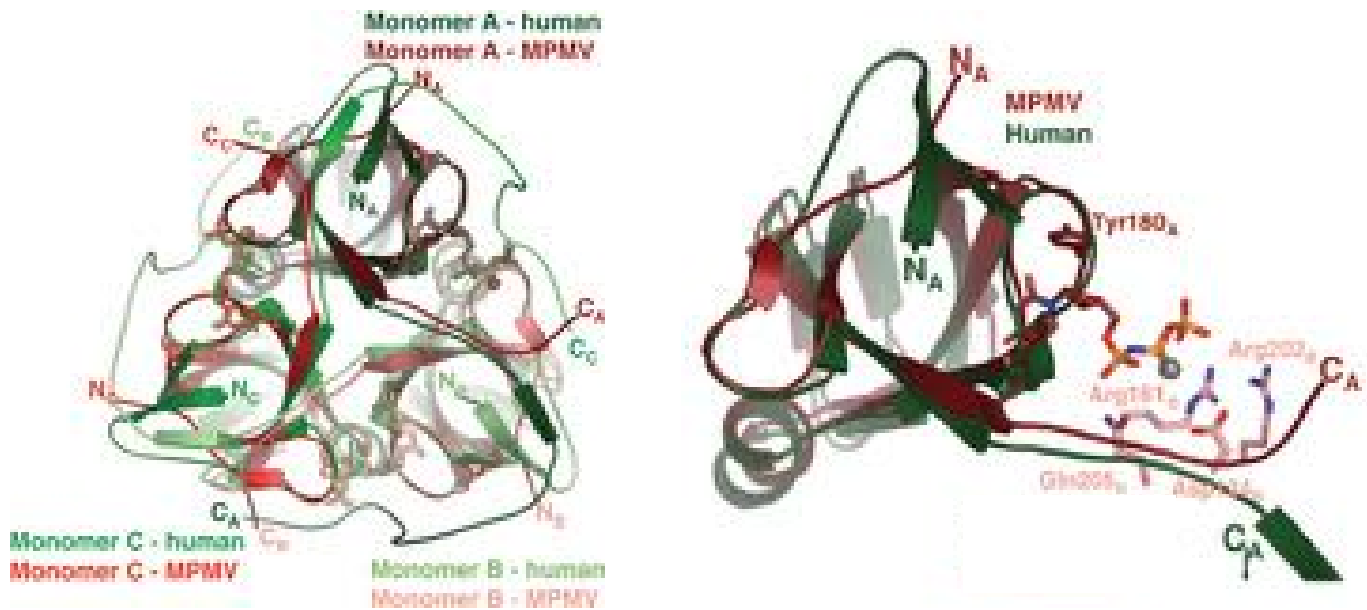


Figura D. 7. Superposición de las estructuras tridimensionales de la hDut y la Dut del MPV en complejo con el sustrato no hidrolizable α,β -imido-dUTP. (Izquierda) Superposición de los trímeros de ambas Duts (códigos PDB 1Q5U y 2D4N, respectivamente). El trímero de hDut se representa en tonalidades verdes y el trímero de la Dut de MPV en tonalidades rojas. Los sustratos y los iones de magnesio se representan en varillas y esferas del mismo color de la subunidad a la que pertenecen. (Derecha) Superposición del monómero A de cada Dut, resaltando la distinta ordenación tomada por los extremos N y C-terminal (N_A y C_A en rojo para la Dut de MPV y en verde para la hDut, respectivamente), aunque solo puede visualizarse el inicio en el caso del C-terminal. Las cadenas laterales de los residuos conservados del monómero B de la Dut de MPV, implicados en la acomodación del sustrato en el centro activo que forma con el monómero A, se representan en varillas y en colores según el átomo, al igual que el sustrato (C en rojo claro-salmón, N en azul, O rojo, P en naranja). El ion magnesio se representa como una esfera en gris. Figura tomada de *Németh-Pongrácz et al. 2007*.

Como ocurre en el caso del cambio de la fenilalanina del motivo V en las Duts representadas por mDut, estos cambios en la disposición del motivo V plantean la posibilidad de que existan variaciones en las Duts que, manteniendo su actividad catalítica, les permitan escapar de la interacción con StI y consecuente inducción de la SaPI. Al mismo tiempo, estos datos abren otra posibilidad. Si los fagos evolucionan tan rápido y seleccionan aquello que les beneficia, ¿Por qué no han seleccionado una Dut como la $\phi 80\alpha$ en todos los fagos y conservan la $\phi 11$ Dut? ¿O por qué no han seleccionado una Dut como la de *Enterococcus faecalis*?

La inducción de la SaPI por la interacción de la Dut del fago con el represor StI parece algo totalmente pernicioso para el fago, que pierde gran parte de las

cápsides en las que empaquetaría su ADN por la encapsidación del ADN de la SaPI. No obstante, la Dut del fago de $\phi 80\alpha$ posee una afinidad por el StI 5 veces menor a la que posee la $\phi 11$ Dut, y también es mucho peor inductora de la isla que ésta última. Por tanto, como fruto de la adaptación y evolución propia, los fagos deberían haber seleccionado esta Dut, cuya menor capacidad de inducción de la isla se traduciría en una mayor producción de la progenie propia. Y en la misma línea iría la adopción de una Dut de *Enterococcus*, que le permitiría seguir con su actividad enzimática, pero también escapar de la interacción con StI. Además, hay fagos que no se ven afectados ante la delección de Dut, con lo que podría ser una opción no disponer de ella. No obstante, dado que las Duts han sido propuestas como proteínas con funciones señalizadoras (Penadés et al. 2013), podría ser una opción que la Dut del fago ayudante interaccionara con otra proteína de un modo similar a como lo hace con StI, y esta interacción fuera beneficiosa para el fago y la Dut no pudiera ser sustituida.

Por otro lado, es posible que la inducción de la SaPI no sea un proceso totalmente pernicioso para el fago, y por ello ha decidido mantener la Dut en su genoma. Recientemente ha sido descrito que islas cromosómicas inducibles por fago (PICI) de *S. aureus* tendrían una relación de mutualismo con los fagos ayudantes, y no una relación de parasitismo como siempre se ha considerado (Fillol-Salom et al. 2022). El fago ayudante perdería parte de su progenie al encapsidar ADN de la PICI en lugar de su propio ADN, pero una vez integrados en la bacteria huésped, los sistemas de protección codificados en la PICI proporcionarían protección tanto al fago ayudante como a la bacteria, bloqueando la entrada de otros fagos o elementos genéticos móviles en el huésped. Por tanto, el fago ayudante sacrificaría un porcentaje de la progenie pero se aseguraría el porcentaje restante. Dado que las SaPI son consideradas la familia prototípica de las PICI, la Dut convendría ser conservada por el fago ayudante para beneficiarse de la SaPI.

2. El análisis del mecanismo molecular de MazG, con similitudes al de Duts diméricas, podría revelar variabilidad en el mecanismo catalítico.

Las Duts triméricas, con capacidad para hidrolizar el dUTP pero no el dUDP, utilizan un catión metálico para la unión y posicionamiento del sustrato (dUTP) para la hidrólisis, y para la posterior estabilización del estado de transición. En las Duts diméricas (capaces de degradar el dUTP y el dUDP), participan dos cationes metálicos en la unión del sustrato dUTP, la activación del nucleófilo, y

la estabilización del estado de transición; es decir, un mecanismo catalítico diferente. Sin embargo, para la hidrólisis del dUDP se ha propuesto el uso de tres cationes metálicos en lugar de dos (Hemsworth, González-Pacanowska, and Wilson 2013; Hemsworth et al. 2011). La superposición de los sustratos trifosfato y difosfato en el centro activo de Duts diméricas muestra que las disposiciones de los cationes presentes con dUTP (cationes 1 y 2) conservan la misma posición cuando el dUDP está unido, mientras que el tercer metal (catión 3) se posiciona cerca del lugar ocupado por uno de los oxígenos del fosfato y del dUTP (Figura D. 8). Usando como modelo la Dut dimérica de *Campylobacter jejuni*, se ha propuesto que este tercer catión, en sinergia con el catión de la posición 2, ayudaría en la coordinación de la molécula de agua catalítica para el ataque nucleofílico sobre el enlace α - β -fosfato (Hemsworth et al. 2011).

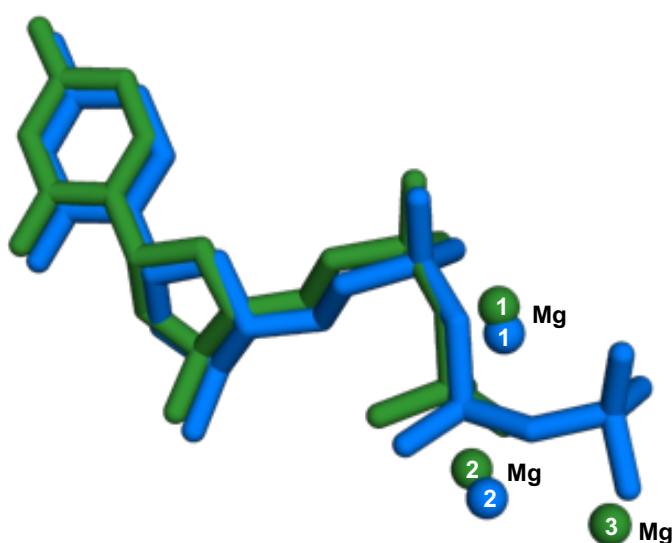


Figura D. 8. Localización del catión 3 en Duts diméricas. Se superponen los sustratos de las estructuras de Dut de *C. jejuni* en complejo con dUpNHpp y 2 cationes Mg^{2+} (azul) y de la misma Dut en complejo con dUpNHp y 3 cationes Mg^{2+} (verde). Se representan en varillas y en esferas los sustratos y los cationes, respectivamente (Códigos PDB: 2C1C y 1W2Y, respectivamente).

Las MazG de fago de *E. faecalis* y *S. epidermidis*, cuya actividad enzimática en presencia de distintos cationes ha sido estudiada en esta Tesis, utilizan 3 cationes Mg^{2+} para llevar a cabo la hidrólisis del dUTP. La similitud del centro activo entre estas proteínas y las Duts diméricas (conservación de dos motivos de unión a metal en cada centro activo, y una elevada homología estructural aunque no en secuencia (Figura D. 9)) (Gonçalves, de Sanctis, and McSweeney 2011; Moroz et al. 2005), permitió realizar una comparación estructural del posicionamiento de los cationes por la superposición de los respectivos

nucleótidos en los diferentes enzimas (dUDP en el caso de Duts diméricas y dUTP en EnMazG). La superposición reveló que los cationes 1 y 2 conservan las posiciones en el centro activo, mientras que el catión 3 difiere en su localización quedando más cerca del fosfato β en el caso de la MazG respecto a la posición ocupada en la Dut dimérica (cerca del fosfato γ). Además, tomando como referencia el dUDP de la Dut, el uracilo en la MazG se desplaza unos 30° hacia el fosfato α , mientras que el azúcar de la desoxirribosa se desplaza igualmente pero alejándose del mismo fosfato (Figura D.10).



Figura D. 9. Alineamiento de secuencias de Duts diméricas (de fago de *Staphylococcus aureus* ϕ DI y ϕ O11) y MazG (de fago de *E. faecalis* y *S. epidermidis*, EnMazG y SeMazG, respectivamente). Sobre el alineamiento se representan las estructuras secundarias de Dut ϕ DI en morado y en la parte inferior las de SeMazG en verde. Los residuos participantes en la interacción del sustrato se identifican con un asterisco, los residuos con una elevada homología en rosa, los residuos conservados en rojo, y en cuadrados rojos los dos motivos de unión de metal. El lazo característico de Duts diméricas se identifica en azul y el extremo C-terminal de MazG en verde oscuro. Figura tomada de Ciges-Tomas 2019.

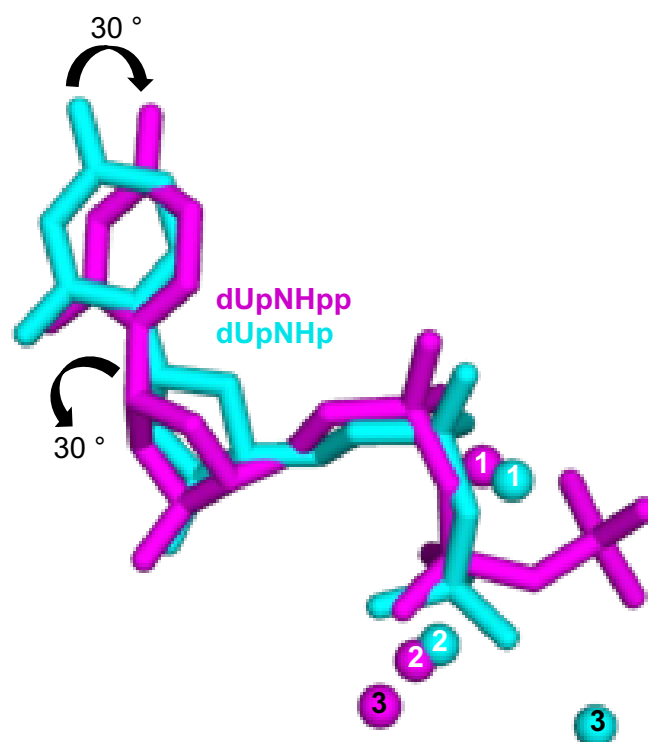


Figura D.10. Localización del catión 3 en Duts diméricas y en EnMazG. Se superponen los sustratos de las estructuras de Dut de *L. major* en complejo con dUpNHp y 3 Mg²⁺ (cian) y de EnMazG en complejo con dUpNHpp y 3 Mg²⁺ (magenta). Se representan en varillas y en esferas los sustratos y los cationes, respectivamente. El uracilo y el azúcar de ribosa se desplazan en el caso de EnMazG respecto a la Dut dimérica. Se identifican las posiciones de los tres cationes en ambos centros activos. Código PDB Dut de *L. major*: 2CJE.

Ampliando este análisis del posicionamiento de los cationes de los centros activos de Duts diméricas y MazG, en los que se conservan las posiciones de los cationes 1 y 2, se incluyó otra MazG con la finalidad de conocer si también entre MazG se mantenían las posiciones de los cationes. Para ello, se superpusieron la EnMazG (3 cationes) y la MazG de *Deinococcus radiodurans* (DR2231) (2 cationes) en complejo con análogos no hidrolizables de dUTP, con la Dut dimérica de *Leishmania major* en complejo con un análogo no hidrolizable de dUDP (3 cationes) (Figura D. 11). En ninguna de las estructuras descritas de MazG de *D. radiodurans* (DR2231) se ha identificado un tercer catión en el centro activo en presencia del análogo no hidrolizable del sustrato dUTP, pero tampoco se ha encontrado ninguna otra estructura de MazG con estas características. La superposición confirmó la conservación de las posiciones 1 y 2 en el centro activo, mientras que la posición ocupada por el catión 3 en EnMazG no está ocupada en el caso de DR2231, y cerca de la posición ocupada por el catión 3

en Dut dimérica se sitúa en DR2231 una molécula de agua (la cual no es el agua catalítica por la distancia excesiva al sustrato).

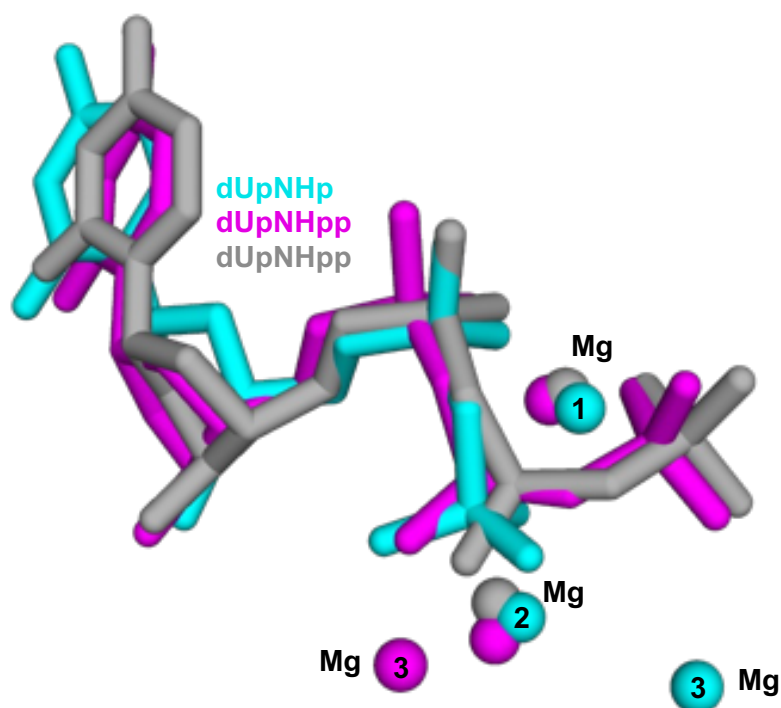


Figura D.11. Localización de los cationes en Duts diméricas y diferentes MazG. Se superponen los sustratos de las estructuras de Dut de *L. major* en complejo con dUpNHp y 3 Mg²⁺ (cian), EnMazG en complejo con dUpNHpp y 3 Mg²⁺ (magenta), y MazG de *D. radiodurans* en complejo con dUpNHpp y 2 Mg²⁺ (gris). Se representan en varillas y en esferas los sustratos y los cationes, respectivamente, éstos últimos con el mismo color que el sustrato con el que se coordinan. Se identifican las posiciones de los tres cationes en ambos centros activos. Código PDB Dut de *L. major*: 2CJE, MazG de *D. radiodurans*: 5HVA.

Este análisis confirma que tanto para las Duts diméricas como para las MazG las posiciones 1 y 2 se mantienen constantes, lo cual apoya las funciones asociadas a ellas. El catión que ocupa la posición 1, que coordina los tres fosfatos del sustrato, juega un papel estructural clave en el posicionamiento del sustrato, la estabilización del estado de transición y la posterior salida del producto dUMP. El catión 2 activaría la molécula de agua que realizará el ataque nucleofílico en línea al fosfato β (Hemsworth, González-Pacanowska, and Wilson 2013; Hemsworth et al. 2011; Moroz et al. 2004). En el caso de algunas MazG, como las estudiadas en esta Tesis (de fago de *S. epidermidis* y de fago de *E. faecalis*), el mecanismo catalítico incluye un tercer catión que, en sinergia con el catión 2, se ha propuesto que polarizaría a la molécula de agua para que lleve a cabo el ataque nucleofílico (Ciges-Tomas 2019), como había sido propuesto para el

tercer catión en las Duts diméricas en la hidrólisis del dUDP (Hemsworth et al. 2011).

En ninguna de las estructuras analizadas de MazG de *D. radiodurans* (DrMazG) se ha identificado la participación de un tercer catión en la catálisis, lo cual podría indicar una mayor similitud del mecanismo de hidrólisis de esta MazG con Duts diméricas respecto a EnMazG y SeMazG, utilizando dos cationes en la hidrólisis del dUTP en lugar de 3, ya que los residuos formadores de los motivos de unión a metal se conservan completamente entre Duts diméricas y DrMazG, mientras que difieren con EnMazG y SeMazG. El residuo en discordia sería el D50 en EnMazG, el cual es un E67 y E79 en la Dut dimérica ϕ DI y en DrMazG, respectivamente. El distinto posicionamiento y orientación de la cadena lateral del aspártico podría obligar a una conformación del centro activo ligeramente distinta, donde el posicionamiento y activación de la molécula de agua catalítica requeriría de un tercer catión cerca del residuo aspártico, para actuar en sinergia con el catión en posición 2. En el caso de la Dut dimérica y DrMazG, la cadena lateral del glutámico (E67 y E79, respectivamente) permitiría el posicionamiento y activación de la molécula de agua catalítica mediante la coordinación con el catión en posición 2 (Figura D. 12). Por tanto, la MazG de *D. radiodurans* tendría una mayor similitud con Duts diméricas en su mecanismo catalítico que con las MazGs estudiadas en esta Tesis para la hidrólisis del dUTP a dUMP y PPi. Podría, entonces, constituir un ejemplo de MazG a caballo entre Duts diméricas y SeMazG/EnMazG.

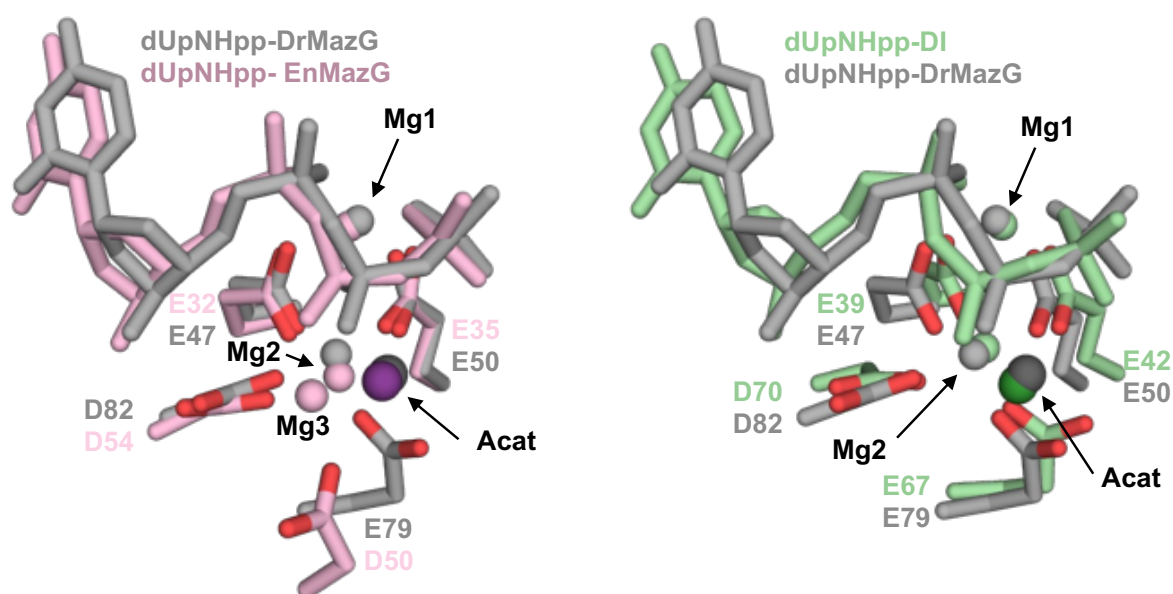


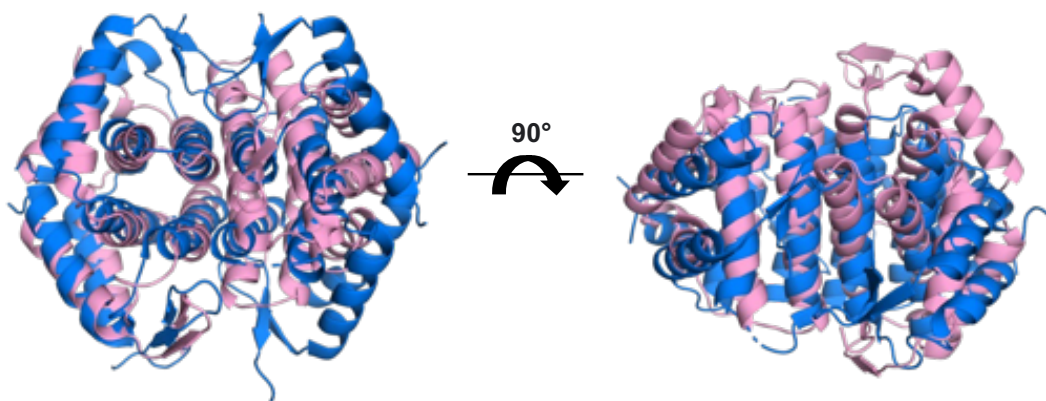
Figura D. 12. Superposición de los análogos no hidrolizables del sustrato dUTP en la estructura en complejo con DrMazG (gris) con la estructura en complejo con EnMazG (rosa, izquierda), y con la estructura en complejo con Dut dimérica del fago ϕ DI de *S. aureus* (verde, derecha). (Izquierda) Tres residuos del dominio de unión al metal en DrMazG (E47, E50 y D82) se orientan de la misma manera que en EnMazG (E32, E35 y D54), al igual que los cationes 1 y 2 (Mg1 y Mg2), y las moléculas de agua catalítica (representadas en esferas gris oscuro y morado si pertenecen a DrMazG o EnMazG respectivamente, e indicadas como Acat). (Derecha) Cuatro residuos del dominio de unión al metal que coordinan los cationes 1 y 2 en DrMazG (E47, E50, E79 y D82) se orientan de la misma manera que en la Dut ϕ DI (E39, E42, E67 y D70), al igual que los dos cationes divalentes. Las moléculas de agua catalítica se posicionan de manera similar en las tres estructuras (esferas gris oscuro, morado o verde oscuro si pertenecen a DrMazG, EnMazG o Dut ϕ DI respectivamente, e indicadas como Acat). Códigos PDB: DrMazG 5HVA y Dut ϕ DI 5MIL.

Junto a DrMazG, se ha analizado otra proteína con un dominio MazG, la proteína MazZ del cianofago S-2L (Czernecki et al. 2021), con capacidad para degradar (d)GTP hasta (d)GMP. Estructuralmente tiene una mayor similitud con EnMazG que con una Dut dimérica (Figura D. 13), pero su mecanismo catalítico parece ser algo más parecido al de Dut. Esta proteína fue cristalizada en presencia del sustrato dGTP, pero los autores solo obtuvieron la estructura de MazZ en presencia de dGDP y dos cationes Mn^{2+} . Esta estructura, sumada al análisis con el HPLC de los resultados de los experimentos de actividad sobre dGTP y GTP, en el cual se obtuvieron dGMP y GMP respectivamente, llevó a proponer que el mecanismo catalítico implicaba dos desfosforilaciones sucesivas, y no a la hidrólisis directa del sustrato en el nucleótido monofosfato y PPI. El centro activo es casi idéntico al de Duts diméricas en presencia de dUDP, con una orientación

muy similar de las cadenas laterales de los residuos del motivo de unión a metal, y con el mismo posicionamiento de la molécula de agua catalítica (Figura D. 14). Dado que la disposición del centro activo en Duts diméricas es similar con el dUTP y con el dUDP como sustrato (con la participación del 3^{er} catión como principal diferencia en la hidrólisis del dUDP respecto al dUTP, y un pequeño desplazamiento de la molécula de agua catalítica (Figura D. 15)), la proteína MazZ podría llevar a cabo la hidrólisis del dGTP en dos desfosforilaciones sucesivas utilizando en la primera de ellas un mismo número de cationes (2) con una disposición similar a una Dut dimérica en presencia del dUTP, mientras que podría utilizar en la segunda desfosforilación un mismo número de cationes (3) y una disposición similar a la Dut dimérica en presencia del dUDP. No obstante, se requeriría de la estructura de MazZ en presencia del dGTP para poder confirmar esta hipótesis, ya que podría ser otra opción que las 2 desfosforilaciones las realizara en presencia de 3 cationes Mn²⁺. Además, la estructura de MazZ en presencia de dGDP fue obtenida también en ausencia de cationes, con lo cual es posible que la presencia de este producto y no del dGMP sea consecuencia de una catálisis muy ralentizada por la ausencia de los cationes, los cuáles no posicionarían a la molécula de agua catalítica correctamente para la hidrólisis del enlace fosfoanhidro entre los fosfatos α y β , y ésta quedaría en una posición donde le es más factible atacar el enlace entre los fosfatos β y γ .

Por tanto, la variabilidad descrita en la bibliografía en los sustratos hidrolizados por distintas MazG podría, además de incluir las diferencias en los residuos determinantes de la especificidad de nucleótido y la cantidad de cationes utilizados, incluir una variedad en el mecanismo de catálisis, pudiendo llevar a cabo la producción del nucleótido monofosfato en una o dos reacciones consecutivas, pero conservando las posiciones 1 y 2 de los cationes.

A. MazZ
EnMazG



B. MazZ
Dut ϕ DI

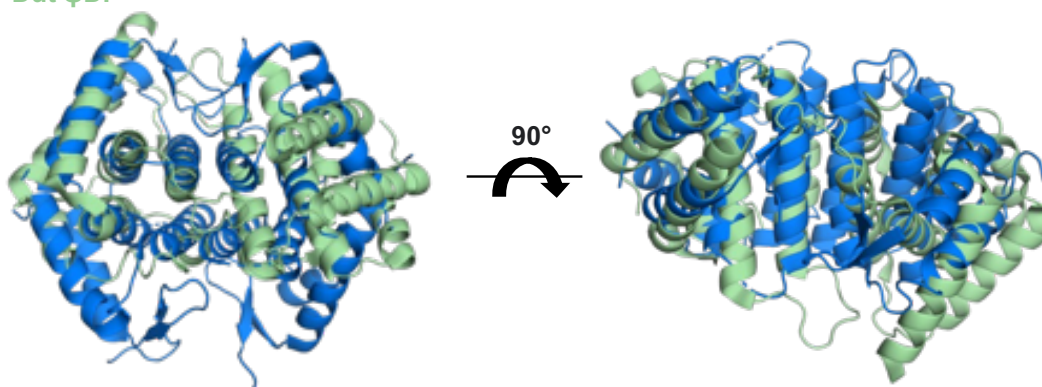


Figura D. 13. MazZ tiene una mayor similitud con EnMazG que con Dut ϕ DI. A) Superposición de la estructura de MazZ (azul) con la estructura de EnMazG (rosa). B) Superposición de la estructura de MazZ (azul) con la estructura de Dut ϕ DI (verde). En ambos casos se muestran dos vistas ortogonales. Códigos PDB: MazZ 7ODY y Dut ϕ DI 5MIL.

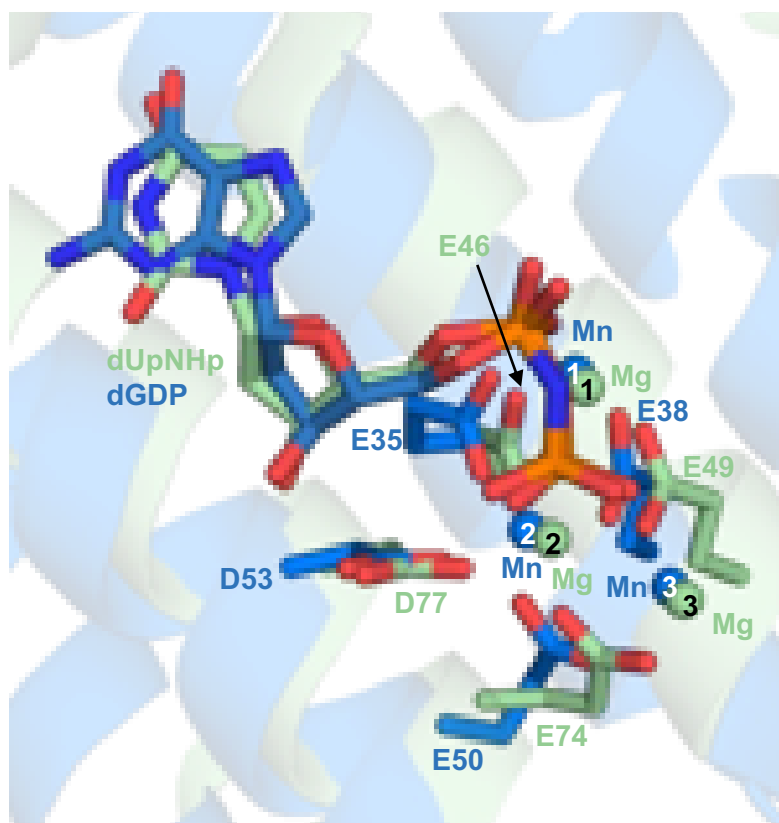


Figura D. 14. Superposición del análogo de sustrato no hidrolizable dUpNHp, cationes y residuos coordinantes de éstos últimos en las estructuras de Dut de *C. jejuni* y MazG del cianofago S-2L (MazZ). El sustrato y las cadenas laterales de los residuos implicados en la coordinación de los cationes se representa en varillas, mientras que los cationes se representan en esferas. Todos los elementos pertenecientes a la estructura de Dut de *C. jejuni* están en verde, mientras que los de MazZ están en azul marino. Las cadenas laterales y el nucleótido están coloreados según el átomo (C azul marino o verde, O en rojo, N en azul, P en naranja). Códigos PDB 1W2Y (Dut *C. jejuni*) y 7ODY (MazZ).

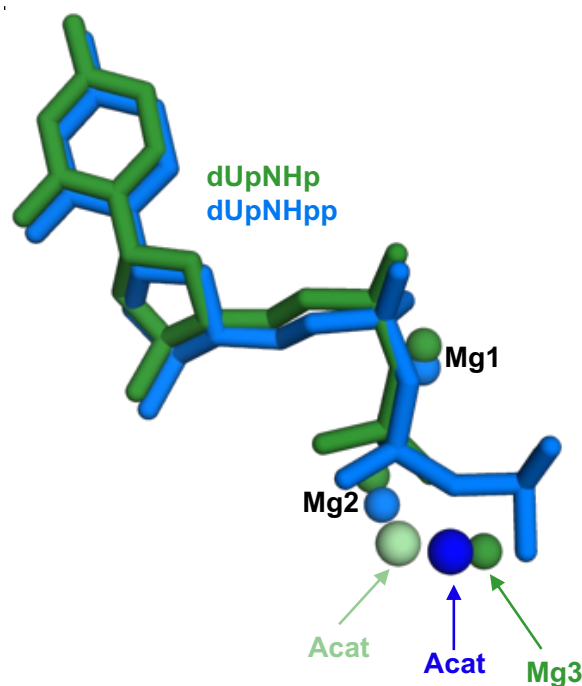


Figura D. 15. La disposición del centro activo en Duts diméricas en ambos sustratos difiere en la participación del tercer catión en el caso de dUDP y en un pequeño desplazamiento de la molécula de agua catalítica (Acat). Superposición de los sustratos (en varillas) de las estructuras de Dut de *C. jejuni* en complejo con dUpNHpp y 2 cationes Mg^{2+} (azul) y de la misma Dut en complejo con dUpNHp y 3 cationes Mg^{2+} (verde). Se representa la molécula de agua catalítica para cada sustrato (en esferas y en color verde en el caso de dUDP, y en color azul en el caso de dUTP). Códigos PDB: sustrato dUpNHp 1W2Y, sustrato dUpNHpp 2CIC.

2.1. El catión que ocupa la posición 3 podría ser determinante en la catálisis del sustrato.

En la hidrólisis del ARN llevada a cabo por la ARNasa H1 de *Bacillus halodurans* es necesaria la participación de tres cationes divalentes Mg (Samara and Yang 2018). Aunque en el sitio activo entren dos cationes Mg^{2+} , no se produce la reacción, sino que es necesaria la entrada del tercer catión. Dos cationes Mg^{2+} actúan en primer lugar coordinando el fosfato a escindir, y una vez posicionado correctamente, se produce la entrada del tercer catión, que en sinergia con otro de los cationes previamente posicionados, activa la molécula de agua catalítica para el ataque nucleofílico. Por tanto, el catión que actúa en tercer lugar resulta determinante para la actividad ARNasa de la enzima. En el caso de las MazG estudiadas en esta Tesis, se ha propuesto que el catión que ocupa la posición 3 actúa en sinergia con el catión de la posición 2 y activan la molécula de agua que llevará a cabo el ataque nucleofílico. De este modo, podría resultar determinante de la misma manera que en la ARNasa H1 que la entrada del catión

de la posición 3 *a posteriori* de la entrada de los otros dos, que posicionarían el nucleótido en el centro activo, determinara la conformación final (activa o inactiva, con el cambio de orientación del fosfato y que permite o no el posicionamiento de la molécula de agua catalítica, respectivamente). La visualización de los tres cationes participantes en las estructuras de EnMazG resueltas a lo largo de esta Tesis refleja un proceso estático, consecuencia de la cristalización en presencia del nucleótido no hidrolizable dUpNHpp, que “atrapa” al centro activo de la proteína en la disposición requerida para iniciar la catálisis pero no permite que se realice. No obstante, para poder confirmar esta hipótesis habría que obtener las estructuras del centro activo a distintos tiempos de la catálisis del sustrato original, desde antes de la entrada del sustrato hasta la salida del producto, y con ello confirmar la participación detallada de los iones. Para ello se podría recurrir al protocolo descrito por Samara y colaboradores (Samara and Yang 2018), utilizado en esta Tesis para un objetivo similar con EnMazG y concentraciones variables de iones Mg^{2+} y Mn^{2+} . Los resultados con este protocolo no nos permitieron “atrapar” estadios intermedios, requiriendo de una optimización que no pudo llevarse a cabo por limitaciones temporales. Sin embargo, las estructuras de EnMazG en presencia de distintos ratios $Mg^{2+}:Mn^{2+}$ sí han permitido proponer un mecanismo de inhibición de esta MazG por Mn^{2+} . En éste, la ocupación de la posición 2 por el catión Mn^{2+} predispone a una conformación inactiva del nucleótido, pero es la ocupación por este catión de la posición 3 la que la determina. Por tanto, estos resultados podrían apoyar que la posición 3 fuera la última en ocuparse y ser, además, la más determinante en la catálisis. La optimización de los experimentos realizados permitirá esclarecer el papel del catión que ocupa la posición 3.

3. La modulación de EnMazG y SeMazG por cationes divalentes diferentes al Mg^{2+} podría señalar al fago en la bacteria huésped.

La homeostasis de los iones metálicos en las bacterias patogénicas es necesaria para resistir a los esfuerzos del huésped por manipular su disponibilidad y toxicidad cuando es infectado. La respuesta adaptativa de la bacteria implica el tránsito de metales y el uso de proteínas que perciben la presencia de éstos (Braymer and Giedroc 2014). Por tanto, una regulación fina de metales es crucial para sobrevivir en el huésped y colonizarlo. En el caso de *S. aureus*, éste ha desarrollado un mecanismo para superar las limitaciones de Mn^{2+} impuestas por el huésped mediante la expresión de transportadores de elevada afinidad de este catión (Cassat and Skaar 2012).

Del mismo modo, las bacterias utilizan múltiples mecanismos inmunes para defenderse de la infección por fago, cuyo ciclo vital depende de la homeostasis

de la bacteria hospedadora, a la cual deben adaptarse. Los sistemas de restricción – modificación y CRISPR – Cas han sido descritos como las principales líneas de defensa, ya que son las mejores caracterizadas. No obstante, recientes estudios han puesto de manifiesto la existencia de un amplísimo y variado arsenal de sistemas de defensa (Millman et al., 2022). Se han descrito proteínas dCTP desaminasas de bacteria que convierten el dCTP en dUTP en respuesta a la infección por fago, así como una dGTPasa con reminiscente homología a SAMHD1, que degrada el dGTP en desoxiguanosina (Tal et al. 2022). Estas proteínas de bacteria tendrían actividad defensiva frente a fagos al eliminar desoxinucleótidos específicos como el dCTP y el dGTP del *pool* de nucleótidos durante la infección vírica, limitando al fago de los bloques de construcción del ADN esenciales, y con ello deteniendo su replicación. En base a estos datos, podría ser posible que un fago que infecta una bacteria que posee Dut lleve consigo la MazG para asegurarse la eliminación de todo el dUTP del *pool* de la bacteria, de modo que el retraso en la replicación de su genoma provocado por la privación por parte del huésped de dGTP, dCTP u otro nucleótido, no se vea aún más retrasado por la incorporación errónea del uracilo y la consecuente respuesta de los sistemas de reparación. En caso que la bacteria infectada no disponga de sistemas defensivos, el fago se asegura la replicación con una mínima cantidad de uracilo incorporado erróneamente en el ADN. Además, el dUMP producto de la actividad de MazG es el precursor para la síntesis del dTTP, con lo que se aseguraría este bloque esencial en la replicación de su ADN.

Los estudios enzimáticos y estructurales realizados en esta Tesis muestran una modulación de la actividad catalítica de hidrólisis del dUTP de EnMazG y SeMazG en presencia de varios cationes divalentes diferentes al Mg^{2+} . Esto podría tratarse de una cuestión de supervivencia. Es decir, cuando el fago poseedor de MazG entra en la bacteria, puede que las concentraciones de cationes como el Zn^{2+} , Mn^{2+} y Cu^{2+} estén en concentraciones elevadas respecto a las habituales como consecuencia de que la bacteria, a su vez, se encuentre en plena lucha con el sistema de defensa de un huésped al que ha infectado, el cual puede estar sometiéndola a una explosión oxidativa, o simplemente que la bacteria no está en las condiciones óptimas requeridas por el fago para su replicación genómica y producción de progenie. Ante estas elevadas concentraciones la MazG se inhibiría como consecuencia de la entrada de estos iones en el centro activo, y a modo de sensor no higienizaría el *pool* de nucleótidos, ya que de nada va a servir el esfuerzo si no va a poder llevar a cabo el ciclo lítico, o en caso de producir progenie podría no encontrar otras bacterias que infectar. Además, al igual que algunas Duts, la inhibición de MazG podría señalar de alguna manera al fago para que quedase a la espera (en fase lisogénica) de que la bacteria recuperara, si es posible, las condiciones óptimas para la replicación del fago, donde las concentraciones de estos iones metálicos fuesen bajas, y entonces iniciaría el saneamiento del *pool* de nucleótidos que prepararía a la bacteria para una máxima producción de su progenie. Sin

embargo, esto podría suponer un arma de doble filo, porque la bacteria podría defenderse de la infección utilizando esta inhibición e impedir la replicación del fago, al igual que lo hacen como defensa antibacteriana algunos organismos huéspedes.

Dado que en esta Tesis también se ha descrito la modulación de ambas MazG por parte de otros cationes divalentes como el Ca^{2+} , el Co^{2+} y el Ni^{2+} , podría ser posible que elevadas concentraciones de estos cationes estuvieran también relacionadas con condiciones de la bacteria huésped no ideales para la replicación del fago; de modo que la modulación de la actividad de ambas MazG por estos cationes podría tener una finalidad similar a la propuesta.

Estas observaciones podrían ser otro ejemplo de la exquisita relación entre fago y huésped. Por ello habría que confirmarse *in vivo* el papel que ejercen los iones metálicos en el ciclo vital de estos fagos, y si las MazG están implicadas en él.

4. Toda investigación es útil, aunque no se consigan los resultados previstos.

Como en todo proyecto experimental, la realización de los ensayos no siempre rinde los resultados esperados. Uno de los objetivos de Tesis consistía en confirmar la capacidad reguladora de la Dut de *M. tuberculosis* utilizando cuatro posibles candidatos. La consecución de este objetivo supuso una fuerte inversión de trabajo y podría verse como un fracaso porque no produjo resultados experimentales que confirmaran nuestras hipótesis. Sin embargo, y desde el punto de vista de lo que supone un Tesis, podemos considerar que sí se ha obtenido cierto éxito, ya que ha sido un proceso de aprendizaje continuo para el análisis crítico de los resultados y el incansable planteamiento de nuevas opciones. Todas las técnicas en biología molecular para el clonaje de los genes, el diseño de construcciones génicas, las alternativas para solubilizar las proteínas, su posterior purificación, su detección y estudio del plegamiento, han servido de base para aplicarlas a otras líneas de investigación de esta Tesis que han resultado exitosas, compensando todo el esfuerzo que se había empleado en este proyecto.

Aunque es cierto que podríamos haber utilizado aproximaciones experimentales alternativas a las realizadas, como ensayos de complementación o ensayos *in vivo* de doble híbrido en bacteria, las limitaciones temporales de una Tesis hicieron imposible acometerlas. Los experimentos realizados en esta línea de investigación quedan registrados en esta memoria de Tesis para un posible uso posterior en el laboratorio, con la finalidad de que sirva como base a partir de la cual diseñar nuevos experimentos que arrojen luz sobre posibles funciones *moonlighting* de esta Dut.

CONCLUSIONES

CONCLUSIONES

1. Las estructuras de las Dut humana y de *M. tuberculosis* en complejo con la porción N-terminal del represor StI de la SaPIBov1 confirman un mismo mecanismo de reconocimiento e interacción con Duts triméricas de organismos eucariotas y procariotas similar al de su diana biológica vírica.
2. El motivo V de las Duts triméricas participa en la unión del represor mediante la estabilización del complejo una vez StI ha reconocido el centro activo de la Dut. En este proceso juegan un papel relevante residuos conservados del motivo: el residuo de fenilalanina, que se posiciona sobre el anillo de sustrato y en el complejo se introduce en un bolsillo hidrofóbico de StI, y varios residuos de glicina que proveen de la flexibilidad requerida al motivo V para rodear al represor y alcanzar este bolsillo.
3. El modelado *in-silico* del complejo StI-Dut monomérica y estudios computacionales proponen que StI será capaz de reconocer a esta familia de Duts e interactuar de una forma similar a la utilizada con Duts triméricas. Los ensayos enzimáticos apoyan esta interacción.
4. El represor StI ha optimizado una estrategia de mimetismo molecular que le capacita para interactuar con las tres familias de Duts: monoméricas, diméricas y triméricas, ratificando a StI como un inhibidor proteico universal de Duts.
5. Las proteínas MazG del fago ϕ FL2A de *Enterococcus faecalis* (EnMazG) y del fago ϕ IPLA7 de *Staphylococcus epidermidis* (SeMazG) degradan específicamente el dUTP, por tanto, pueden ser consideradas dUTPasas (Duts). En cada uno de los centros activos participan 3 de las 4 subunidades, y el único residuo procedente del tercer monómero, una arginina que ayuda a la estabilización del sustrato, resulta clave para una óptima actividad enzimática.
6. EnMazG y SeMazG muestran preferencia por el catión Mg^{2+} para su actividad enzimática. No obstante, en combinación a bajas concentraciones con los cationes Zn^{2+} y Co^{2+} su capacidad enzimática se potencia, mientras que en cationes como el Mn^{2+} , con geometrías de coordinación y distancias de enlace muy similares, su actividad es inhibida. SeMazG también potencia su actividad enzimática en combinación a bajas concentraciones con el catión Ni^{2+} .

CONCLUSIONES

7. El dUTP puede adoptar 2 conformaciones en el centro activo de EnMazG, donde la principal diferencia es la disposición del fosfato γ . La conformación activa (γ P-a) permite el posicionamiento de la molécula de agua catalítica, y la inactiva (γ P-b) impide el ataque nucleofílico. La disposición adoptada por el fosfato γ podría estar modulada por los 3 cationes metálicos que participan en el mecanismo catalítico.
8. Se propone un mecanismo de inhibición del catión Mn^{2+} en la proteína EnMazG en el cual la entrada del catión en la posición 1, con un papel estructural, no determina la conformación del sustrato; mientras que la entrada en las posiciones 2 y 3, ambas con un papel catalítico, influye hacia la disposición catalíticamente no competente del fosfato γ , aunque la posición 3 posee el papel determinante.

ANEXOS

Anexo I. Tabla detallada de los intentos de clonación de la ADN helicasa dependiente de ATP de *M. tuberculosis* realizados en esta Tesis. Se destacan en negrita los **overhangs** necesarios en el inserto para la clonación (Fragmento Sec. es fragmento de secuenciación, explicado en el texto principal).

Construcción Rv3202c	Vector	Sistema clonación	Cebadores clonación 5' → 3'	Cebadores de secuenciación 5' → 3'
2 dominios (2343 pb)	pETNKI 1.1 / 1.5	LIC	CAGGGACCCGGT ATGTCACATATCTGGGGTGTCTG (FW) CGAGGAGAAGCCCGGTTAG TCGGCACCTGGCACACCG (RV)	Cebadores T7 forward y reverse
	pLIC-SGC 1		TACTTCCAATCC ATGTCACATATCTGGGGTGTCTG (FW) TATCCACCTTTACTGTCAG TCGGCACCTGGCACACCG (RV)	
	pNIC28			
	pETNKI 1.1	<i>Kpn I</i> y <i>Blp I</i>	GGTACCTT ATGTCACATATCTGGGGTGTCTG (FW) GCTCAGC GTGGCACCTGGCACACC (RV)	
		Gibson	AGAACCTGTA CTTCCAATCC ATGTCACATATCTGGGGTGTCTGAGG (FW) GCCCAAGGGGTTATGCTAG TCATGGCGCCGACCCTCG (RV) CTAGCATAACCCCTTGGG (FW linealizar plásmido) GGATTGGAAGTACAGGTTT (RV linealizar plásmido)	
Entero (3168 pb) Fragmentos	pLIC-SGC 1	Gibson	Fragmento A CCAATCCATAAGC ATGTCACATATCTGGGGTGTCTGAGGC (FW) CCCCATCGATCAGG TGCGCGCGTCTCGCAGCGC (RV) Fragmento B CGACGCGCGCAC CTGATCGATGGGGTGCCTTG (FW) GTTCCGCCATCGCC GTGGCCCACTGGGCAATC (RV) Fragmento C CAGTGGGCCGAC GGCGATGGCGAACCCGTC (FW) GGGGTTATGCTAG TCATGGCGCCGACCCTCG (RV) Linealizar plásmido TCGGCGCCATGACTAG CATAACCCCTTGGG (FW) GATATGTGACATGCTTATGGATTGGAAGTACAG (RV)	<u>Secuenciación Fragmento Sec. A</u> Cebador T7 (FW) CGAAGGCCTCCAAAGCT (RV) <u>Secuenciación Fragmento Sec. AB</u> CTGGACCCGCTGGAG (FW) CTGGCCACCCAGACTT (RV) <u>Secuenciación Fragmento Sec. BC</u> CGATTGGTGGTGTGGACC (FW) GCACAACGGATCACTCGTG (RV) <u>Secuenciación Fragmento Sec. C</u> CGACGGCGATGGCGAAC (FW) CAAAAACCCCTCAAGACCCGT (RV)

Anexo II. Tampones generales utilizados en esta Tesis.

1. **Tampón PBS** (solución salina tamponada con fosfato) → 138 mM NaCl + 3 mM KCl + 8,1 mM Na₂HPO₄ + 1,5 mM KH₂PO₄ (pH 7,4).
2. **Solución de teñido** → solución azul Coomassie brillante 93,7 μM + ácido clorhídrico al 0,11 %.
3. **Tampón para el lavado de la membrana** (Western-Blot) → TBS-T (tampón Tris salino pH 7,5 + 0,1 % Tween-20).
4. **Tampón para el bloqueo de la membrana** (Western-Blot) → TBS-T + 2 % de BSA al 25 %.
5. **Tampón cátodo y ánodo del SDS-PAGE** → 25 mM Tris-HCl pH 8,3 + 192 mM de glicina y 0,1 % de SDS.
6. **Tampón cátodo y ánodo del PAGE** → 25 mM Tris-HCl pH 8,3 + 192 mM de glicina.
7. **Tampón parte resolutive del SDS-PAGE y PAGE** → 1,5 M Tris pH 8,8.
8. **Tampón parte apilamiento del SDS-PAGE y PAGE** → 0,5 M Tris pH 6,8.
9. **Tampón TBE** → 90 mM Tris + 90 mM ácido bórico + 2 mM EDTA
10. **Tampón de carga 6X para ADN** → 60 % glicerol + 60 % v/v 1X TAE + azul de bromofenol.
11. **Tampón TAE** → 40 mM Tris + 20 mM ácido acético + 1 mM EDTA

Anexo III. Tablas recopilatorias de los valores RMSD (Å) tras la superposición de los monómeros de Dut y Stl^{N-ter} de la estructura tridimensional del complejo mDut-Stl^{N-ter} (tabla superior e inferior, respectivamente) (PDB 7PWX). Entre paréntesis se indica el número residuos cuyos C α han sido superpuestos.

mDut	mDut Trímero 1			mDut Trímero 2		
	Monómero A	Monómero B	Monómero C	Monómero A'	Monómero B'	Monómero C'
Monómero A	-	0,440 (133)	0,349 (133)	0,337 (133)	0,311 (133)	0,388 (133)
Monómero B	-	-	3,469 (131)	0,387 (131)	0,350 (131)	0,720 (131)
Monómero C	-	-	-	0,337 (131)	0,345 (131)	3,508 (131)
Monómero A'	-	-	-	-	0,409 (135)	0,348 (135)
Monómero B'	-	-	-	-	-	0,283 (136)
Monómero C'	-	-	-	-	-	-

Los valores en azul son debido a pequeños desplazamientos en algunas regiones de los monómeros C y C', pero éstas son de no interacción con Stl^{N-ter} o las interacciones que se establecen con él no se ven alteradas.

RMSD (Å)		Monómero D	Monómero E	Monómero F	Monómero D'	Monómero E'	Monómero F'	
Stl ^{N-ter}	Trímero 1	Monómero D	-	1,416 (125)	1,464 (125)	1,676 (126)	1,612 (125)	1,571 (110)
		Monómero E	-	-	0,627 (146)	0,383 (146)	0,572 (145)	0,956 (113)
		Monómero F	-	-	-	0,671 (146)	0,611 (145)	0,923 (113)
	Trímero 2	Monómero D'	-	-	-	-	0,473 (145)	0,873 (112)
		Monómero E'	-	-	-	-	-	0,943 (114)
		Monómero F'	-	-	-	-	-	-

Los valores en azul se corresponden con la superposición con el monómero D con el resto de los monómeros de Stl. Este monómero no ha podido ser trazado completamente, sino en tres fragmentos, y las grandes diferencias son los residuos cercanos a las regiones sin densidad electrónica. Ninguno de ellos está participando en la interacción con la Dut.

Anexo IV. Tablas recopilatorias de los valores RMSD (Å) tras la superposición de los monómeros de Dut y Stl^{N-ter} de la estructura tridimensional del complejo hDut-Stl^{N-ter} (tabla superior e inferior, respectivamente) (PDB 7PWJ). Entre paréntesis se indica el número residuos cuyos C α han sido superpuestos.

RMSD (Å)		Monómero A	Monómero B	Monómero C
hDut (hDut-Stl ^{N-ter})	Monómero A	-	0,278 (134)	0,236 (134)
	Monómero B	-	-	0,350 (140)
	Monómero C	-	-	-

RMSD (Å)		D	E	F
Stl^{N-ter} (hDut-Stl ^{N-ter})	D	-	0,407 (144)	0,595 (143)
	E	-	-	0,833 (146)
	F	-	-	-

Anexo V. Tabla recopilatoria de los valores RMSD (Å) de superposición de los monómeros de Stl^{N-ter} de las estructuras en complejo con hDut y mDut, y con la estructura de referencia (ϕ 11Dut-Stl^{N-ter}). Cada una de las letras hace referencia a un monómero de Stl^{N-ter}, y los números entre paréntesis son la cantidad de residuos cuyos C α han sido superpuestos.

Stl ^{N-ter}		mDut-Stl ^{N-ter}						hDut-Stl ^{N-ter}			ϕ 11Dut-Stl ^{N-ter}			
		D	E	F	D'	E'	F'	D	E	F	D	E	F	G
mDut-Stl ^{N-ter}	D	-	1.416 (125)	1.464 (125)	1.676 (126)	1.612 (125)	1.571 (110)	1.306 (123)	1.169 (124)	1.254 (123)	1.34 (125)	1.343 (125)	1.322 (125)	1.306 (125)
	E	-	-	0.627 (146)	0.383 (146)	0.572 (145)	0.956 (113)	0.801 (142)	0.915 (142)	0.765 (141)	0.726 (145)	0.898 (146)	0.704 (146)	0.68 (144)
	F	-	-	-	0.671 (146)	0.611 (145)	0.923 (113)	0.915 (142)	1.063 (142)	0.872 (141)	0.593 (145)	0.719 (146)	0.561 (146)	0.553 (144)
	D'	-	-	-	-	0.473 (145)	0.873 (112)	0.786 (141)	1.039 (142)	0.831 (141)	0.78 (145)	0.937 (146)	0.759 (146)	0.734 (144)
	E'	-	-	-	-	-	0.943 (114)	0.814 (142)	1.017 (142)	0.761 (141)	0.672 (145)	0.795 (145)	0.674 (145)	0.634 (144)
	F'	-	-	-	-	-	-	1.012 (110)	1.146 (110)	1.09 (111)	1.018 (114)	0.978 (113)	0.881 (113)	0.931 (112)
hDut-Stl ^{N-ter}	D	-	-	-	-	-	-	-	0.407 (144)	0.595 (143)	0.732 (142)	0.868 (142)	0.781 (142)	0.701 (142)
	E	-	-	-	-	-	-	-	-	0.833 (146)	0.827 (142)	0.948 (143)	0.903 (143)	0.796 (142)
	F	-	-	-	-	-	-	-	-	-	0.693 (141)	0.832 (142)	0.731 (142)	0.657 (141)
ϕ 11Dut-Stl ^{N-ter}	D	-	-	-	-	-	-	-	-	-	-	0.296 (145)	0.203 (145)	0.193 (144)
	E	-	-	-	-	-	-	-	-	-	-	-	0.318 (147)	0.333 (144)
	F	-	-	-	-	-	-	-	-	-	-	-	-	0.204 (144)
	G	-	-	-	-	-	-	-	-	-	-	-	-	-

Anexo VI. Interacciones intermoleculares en el complejo SaPIbovl-Stl^{N-ter} – mDut.

Bovl-Stl ^{N-ter}			mDut			Distancia (Å)
Elemento estructural	Residuo	Átomo	Elemento estructural	Residuo	Átomo	
α4	55 Val	CG2	β1	9 Arg	NH2	3,8
		CG1	Lβ8-final	126 Glu	OE2	3,5
	56 Asn	ND2			O	2,8
	59 Glu	OE1	Lβ1β2	18 Ser	OG	3,0
α7	102 Asn	OD1		21 His	NE2	3,6
	106 Tyr	OH	24 Asp		CB	2,9
		107 Asn		O	Lβ1β2	87 Arg
	88 Gly		3,8			
α8	109 Gly	N	β6	91 Lys	NZ	3,5
		O				3,7
	110 Asp	OD1				3,1
	112 Tyr	CG	Lβ1β2	86 Tyr	CE1	3,3
		OH	β5	81 Thr	O	3,8
	113 Tyr		O	Lβ5β6	83 Asp	N
		α1		66 Gly	3,5	
		CB	Lβ4α1	65 Ser	O	3,7
			α1	69 Thr	OG1	2,6
		OH	β3	44 Leu	CD2	3,8
β6			91 Lys	N	2,8	
CE2	Lβ5β6	89 Glu	O	3,5		
86 Tyr		86 Tyr	OH	3,5		
114 Ser	O	α1	70 Arg	NE	3,7	
α9	116 Tyr		OH			3,1
			CZ	67 Leu	CD1	3,3
	O	Lβ4α1	64 Arg	NH2	2,8	
117 Asp	OD2	β7	110 Arg	NE	3,4	
α10	152 Leu	O	α1	70 Arg	NH1	3,4

Anexo VII. Interacciones intermoleculares en el complejo SaPIbovl-Stl^{N-ter} - hDut.

BovI-Stl ^{N-ter}			hDut			Distancia (Å)
Elemento estructural	Residuo	Átomo	Elemento estructural	Residuo	Átomo	
α4	56 Asn	OD1	β8	121 Glu	N	2,9
	59 Glu	CD		119 Ile	CG2	3,7
	66 Gly	O	Lβ11-final	136 Gly	N	3,2
Lα4α5	67 Ile	CG1		135 Phe	CB	3,9
	68 Pro	CD		CE1	3,4	
α5	70 Tyr	CB	Lβ1β2	16 Thr	CB	3,9
		CE1		18 Gly	CA	3,9
	74 Arg	NH1	β2	28 Tyr	OH	2,8
α7	98 Tyr	CD2	Lβ11-final	135 Phe	CZ	3,3
	99 Ser	CB		133 Gly	C	3,5
	102 Asn	ND2		134 Gly	CA	3,8
	106 Tyr	CD1	Lβ5β6	130 Arg	O	2,9
		O		81 Asp	CB	3,7
				83 Arg	N	2,8
	107 Asn	OD1	Lβ11-final	84 Gly	N	3,7
ND2		127 Asp		OD1	3,2	
α8	109 Gly	N	Lβ5-β6	82 Tyr	OH	3,5
	112 Tyr	CG			CE1	3,2
		OH	β5	77 Val	O	3,5
				79 Asp	N	3,0
	113 Tyr	CE1	Lβ5-β6	82 Tyr	OH	3,7
		CZ		85 Asn	O	3,9
		OH	β6	87 Gly	CA	3,5
					O	2,6
	C	α1	64 Gly	CA	3,9	
	114 Ser	OG		68 Lys	CG	3,5
α9	116 Tyr	O	Lβ4α1	62 Arg	NH1	2,6
		OH	Lβ6-β7	104 Asp	OD1	3,6
	117 Asp	OD2		105 Arg	N	2,8
				NH1	3,0	

Anexo VIII. Residuos de reconocimiento e interacción de Sfl^{III} en los complejos con p11Out, mOut, nOut y aa-Out.

Sfl ^{III}	p11Out	mOut	nOut	aa-Out
55 Val				118 Glu
56 Arg	149 Gln	126 Glu	121 Glu	117 Glu
59 Glu		18 Ser		115 Gln
66 Glu			119 Ile	
67 Ile			128 Gly	
68 Pro			135 Phe	
70 Tyr	18 Glu		18 Thr	
	20 Asn		18 Gly	
74 Arg	18 Glu			
			28 Tyr	
77 Asp	15 Arg			
84 Tyr	20 Asn			
			135 Phe	
89 Ser			133 Gly	
			134 Gly	
102 Asn	20 Asn			
		21 His		128 Tyr
			130 Arg	
106 Tyr	21 His	21 His		
	24 Asp	24 Asp		
			81 Asp	78 Gly
			83 Arg	85 Thr
107 Asn			88 Gly	
			84 Gly	
			127 Asp	
			131 Gly	
109 Gly	84 Tyr	86 Tyr	82 Tyr	
110 Asp		91 Lys		84 Arg
112 Tyr	84 Tyr	86 Tyr	82 Tyr	79 Tyr
	72 Lys	81 Thr	77 Val	74 Arg
	81 Asp	83 Asp	78 Asp	
	86 Gly	86 Gly		
	85 Ser	85 Ser		
	89 Ser	89 Thr		
113 Tyr		84 Leu		
	89 Gly	91 Lys	87 Gly	
		89 Glu	85 Asn	
		86 Tyr	82 Tyr	
			64 Gly	
114 Ser			68 Lys	
	70 Lys	70 Arg		
				173 Gly
		67 Leu		
116 Tyr	64 Arg	64 Arg	62 Arg	121 Arg
	115 Ile		Absence motif VI	
	113 Lys			
			104 Asp	
117 Asp	133 Lys		105 Arg	210 Arg
132 Leu	70 Lys	70 Arg		

ANEXO XIX. Participación en proyectos relacionados con el COVID-19.

Justo a la mitad de la realización de esta Tesis, la pandemia causada por el COVID-19 llevó al desarrollo de numerosos proyectos de investigación relacionados con distintos aspectos del virus, con muchas y diversas aproximaciones experimentales en todo el mundo.

En el centro de investigación donde se ha desarrollado esta Tesis se llevaron a cabo diversos proyectos, y tuve la suerte de poder participar activamente en dos de ellos. Pude aportar mi formación previa en el diseño de construcciones génicas, en la purificación de proteínas, y en las técnicas de medida de afinidad de unión entre proteínas, entre otras.

Como resultado de la participación en estos proyectos, he podido formar parte de una publicación científica en la revista PLoS Pathogens, con título “*The structural role of SARS-CoV-2 genetic background in the emergence and success of spike mutations: The case of the spike A222V mutation*” (Ginex et al. 2022).

Además, queda pendiente la publicación de un segundo artículo sobre la creación de un mutante de la proteína humana Ace2, la responsable del reconocimiento y unión del virus, con una afinidad de unión por el dominio de unión al receptor de la proteína *spike* (RBD) del virus superior a la de la proteína Ace2 *wild-type*. De este modo Ace2 mutante actuaría como señuelo en la entrada del virus en las células humanas, evitando la infección.

BIBLIOGRAFÍA

- Abrantes, M. C., Lopes, M.F., and Kok, J. 2011. "Impact of Manganese, Copper and Zinc Ions on the Transcriptome of the Nosocomial Pathogen *Enterococcus Faecalis* V583." *PLoS ONE* 6 (10).
- Alite, C., Humphrey, S., Donderis, J., Maiques, E., Ciges-Tomas, J.R, Penadés, J.R., and Marina, A. 2017. "Dissecting the Link between the Enzymatic Activity and the SaPI Inducing Capacity of the Phage 80 α dUTPase." *Scientific Reports* 7 (1): 1–12.
- Alite, C. 2021. "Bases Moleculares de La Señalización Mediada Por dUTPasas Triméricas de Bacteriófagos." Valencia: Instituto de Biomedicina de Valencia - CSIC.
- Ariza, M.E., and Marshal, W. 2011. "A Human Endogenous Retrovirus K dUTPase Triggers a TH 1, T H 17 Cytokine Response: Does It Have a Role in Psoriasis." *Journal of Investigative Dermatology* 131 (12): 2419–27.
- Ariza, M.E., Ronald Glaser, P., Kaumaya, T.P., Jones, C. and Marshall, W. 2009. "The EBV-Encoded dUTPase Activates NF- κ B through the TLR2 and MyD88-Dependent Signaling Pathway 1." *The Journal of Immunology*. Vol. 182.
- Baldo, A. M, and McClure, M. A. 1999. "Evolution and Horizontal Transfer of dUTPase-Encoding Genes in Viruses and Their Hosts." *Journal of Virology*.
- Barabás, O., Pongrácz, V., Kovári, J., Wilmanns, M., and Vértessy, B. G. 2004. "Structural Insights into the Catalytic Mechanism of Phosphate Ester Hydrolysis by dUTPase." *Journal of Biological Chemistry* 279 (41): 42907–15.
- Benedek, A., Horváth, A., Hirmondó, R., Ozohanics, O., Békési, A., Módos, K., Révész, A., Vékey, K., Nagy, G. N., and Vértessy, B. G. 2016. "Potential Steps in the Evolution of a Fused Trimeric All- β dUTPase Involve a Catalytically Competent Fused Dimeric Intermediate." *FEBS Journal*, September, 3268–86.
- Benedek, A., Pölöskei, I., Ozohanics, O., Vékey, K., and Vértessy, B. G. 2018. "The StI Repressor from *Staphylococcus aureus* Is an Efficient Inhibitor of the Eukaryotic Fruitfly dUTPase." *FEBS Open Bio* 8 (2): 158–67.
- Benedek, A., Temesváry-Kis, F., Khatanbaatar, T., Leveles, I., Surányi, E. V., Szabó, J. E., Wunderlich, L., and Vértessy, B. G. 2019. "The Role of a Key Amino Acid Position in Species- Specific Proteinaceous dUTPase Inhibition." *Biomolecules* 9 (6).

- Bowring, J., Neamah, M., Donderis, J., Mir-Sanchis, I., Alite, C., Ciges-Tomas, J.R., Maiques, E., Medmedov, I., Marina, A., and Penadés, J.R. 2017. "Pirating Conserved Phage Mechanisms Promotes Promiscuous Staphylococcal Pathogenicity Island Transfer." *ELife* 6.
- Braymer, J.J., and Giedroc, D. P. 2014. "Recent Developments in Copper and Zinc Homeostasis in Bacterial Pathogens." *Current Opinion in Chemical Biology*.
- Cassat, J. E., and Skaar, E. P. 2012. "Metal Ion Acquisition in *Staphylococcus aureus*: Overcoming Nutritional Immunity." *Seminars in Immunopathology* 34 (2): 215–35.
- Chan, S., Segelke, B., Lakin, T., Krupka, H., Cho, U.H., Kim, M., So, M. et al. 2004. "Crystal Structure of the *Mycobacterium tuberculosis* dUTPase: Insights into the Catalytic Mechanism." *Journal of Molecular Biology* 341 (2): 503–17.
- Chu, R., Lin, Y., Rao, M.S., and Reddy, J.K. 1996. "Cloning and Identification of Rat Deoxyuridine Triphosphatase as an Inhibitor of Peroxisome Proliferator-Activated Receptor α ." *Journal of Biological Chemistry* 271 (44): 27670–76.
- Ciges-Tomas, J.R. 2019. "dUTPasas, Una Nueva Familia de Proteínas Señalizadoras." Valencia: Instituto de Biomedicina de Valencia (CSIC).
- Ciges-Tomas, J.R., Alite, C., Humphrey, S., Donderis, J., Bowring, J., Salvatella, X., Penadés, J.R., and Marina, A. 2019. "The Structure of a Polygamous Repressor Reveals How Phage-Inducible Chromosomal Islands Spread in Nature." *Nature Communications* 10 (1).
- Czernecki, D., Frédéric B., Kaminski, P. A., and Delarue, M. 2021. "Characterization of a Triad of Genes in Cyanophage S-2L Sufficient to Replace Adenine by 2-Amino adenine in Bacterial DNA." *Nature Communications* 12 (1).
- Dearborn, A. D., and Dokland, T. 2012. "Mobilization of Pathogenicity Islands by *Staphylococcus aureus* Strain Newman Bacteriophages." *Bacteriophage* 2 (2): 70–78.
- Donderis, J., Bowring, J., Maiques, E., Ciges-Tomas, J.R., Alite, C., Mehmedov, I., Tormo-Mas, M.A., Penadés, J.R., and Marina, A. 2017. "Convergent Evolution Involving Dimeric and Trimeric dUTPases in Pathogenicity Island Mobilization." *PLoS Pathogens* 13 (9).

- Dubrovay, Z., Gáspári, Z., Hunyadi-Gulyás, E., Medzihradszky, K. F., Perczel, A., and Vértessy, B.E. 2004. "Multidimensional NMR Identifies the Conformational Shift Essential for Catalytic Competence in the 60-KDa *Drosophila melanogaster* dUTPase Trimer." *Journal of Biological Chemistry* 279 (17): 17945–50.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. 2010. "Features and Development of Coot." *Acta Crystallographica Section D: Biological Crystallography* 66 (4): 486–501.
- Evans, P. R. 2011. "An Introduction to Data Reduction: Space-Group Determination, Scaling and Intensity Statistics." *Acta Crystallographica Section D: Biological Crystallography* 67 (4): 282–92.
- Evans, P. R., and Murshudov, G.N. 2013. "How Good Are My Data and What Is the Resolution?" *Acta Crystallographica Section D: Biological Crystallography* 69 (7): 1204–14.
- Folta-Stogniew, E, and Weeks, E.R. n.d. "Introduction to Light Scattering Theory."
- Freeman, L., Buisson, M., Tarbouriech, N., van der Heyden, A., Labbé, P., and Burmeister, W.P. 2009. "The Flexible Motif V of Epstein-Barr Virus Deoxyuridine 5'-Triphosphate Pyrophosphatase Is Essential for Catalysis." *Journal of Biological Chemistry* 284 (37): 25280–89.
- Fillol-Salom, A., Rostøl, J.T., Ojiogu, A. D., Chen, J., Douce, G., Humphrey, S., and Penadés, J.R. 2022. "Bacteriophages Benefit from Mobilizing Pathogenicity Islands Encoding Immune Systems against Competitors." *Cell* 185 (17): 3248-3262.e20.
- Gibson, D., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C.A., and Smith, H.O. 2009. "Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases." *Nature Methods* 6 (5): 343–45.
- Ginex, T., Marco-Marín, C., Wieczór, M., Mata, C. P., Krieger, J., Ruiz-Rodriguez, P., López-Redondo, M.L., et al. 2022. "The Structural Role of SARS-CoV-2 Genetic Background in the Emergence and Success of Spike Mutations: The Case of the Spike A222V Mutation." *PLoS Pathogens* 18 (7 July).
- Glaser, R., Litsky, M.L., Padgett, D.A., Baiocchi, R.A., Yang, E., Chen, M., Yeh, P.E., Green-Church, K. B., Caligiuri, M.A., and Williams, M. 2006. "EBV-Encoded dUTPase Induces Immune Dysregulation: Implications for the Pathophysiology of EBV-Associated Disease." *Virology* 346 (1): 205–18.
- Gonçalves, A.M., de Sanctis, D., and McSweeney, S. M. 2011. "Structural and Functional Insights into DR2231 Protein, the MazG-like Nucleoside Triphosphate." *Journal of Biological Chemistry* 286 (35): 30691–705.

- Prasad, G.S. 2001. "Glycine Rich P-Loop Motif in Deoxyuridine Pyrophosphatase." *Current Protein & Peptide Science* 2 (4): 301–11.
- Harkiolaki, M., Dodson, E. J., Bernier-Villamor, V., Turkenburg, J.P., González-Pacanowska, D. and Wilson, K.S. 2004. "The Crystal Structure of *Trypanosoma cruzi* dUTPase Reveals a Novel dUTP/dUDP Binding Fold." *Structure* 12 (1): 41–53.
- Hellman, L.M., and Fried, M.G., 2007. "Electrophoretic Mobility Shift Assay (EMSA) for Detecting Protein-Nucleic Acid Interactions." *Nature Protocols* 2 (8): 1849–61.
- Hemsworth, G. R., González-Pacanowska, D. and Wilson, K. S., 2013. "On the Catalytic Mechanism of Dimeric dUTPases." *Biochemical Journal* 456 (1): 81–88.
- Hemsworth, G. R., Moroz, O. V., Fogg, M.J., Scott, B., Bosch-Navarrete, C., González-Pacanowska, D., and Wilson, K.S. 2011. "The Crystal Structure of the *Leishmania major* Deoxyuridine Triphosphate Nucleotidohydrolase in Complex with Nucleotide Analogues, dUMP, and Deoxyuridine." *Journal of Biological Chemistry* 286 (18): 16470–81.
- Hermoso Domínguez, J. A., and Martínez Ripoll, M. 2003. "Estructura de Proteínas Por Cristalografía de Rayos X." In *Estructura de Proteínas*, edited by Editorial Ariel, 1st edition, 1:253–85. Barcelona.
- Hidalgo-Zarco, F., Camacho, A. G., Bernier-Villamor, V., Nord, J., Ruiz-Pérez, L.M., and González-pacanowska, D. 2001. "Kinetic Properties and Inhibition of the Dimeric dUTPase-dUDPase from *Leishmania major*." *Journal of Biological Chemistry* 276 (18): 16470–81.
- Hill, R. L.L., and Dokland, T. 2016. "The Type 2 dUTPase of Bacteriophage PdbINM1 Initiates Mobilization of *Staphylococcus aureus* Bovine Pathogenicity Island 1." *Journal of Molecular Biology* 428 (1): 142–52.
- Hirmondó, R., Szabó, J. E., Nyíri, K., Tarjányi, S., Dobrotka, P., Tóth, J. and Vértessy, B. G. 2015. "Cross-Species Inhibition of dUTPase via the Staphylococcal StI Protein Perturbs dNTP Pool and Colony Formation in *Mycobacterium*." *DNA Repair* 30: 21–27.
- Humphrey, S., San Millán, A., Toll-Riera, M., Connolly, J., Flor-Duro, A., Chen, J., Ubeda, C., MacLean, R. C., and Penadés, J.R. 2021. "Staphylococcal Phages and Pathogenicity Islands Drive Plasmid Evolution." *Nature Communications* 12 (1).
- Kane, J. F. 1995. "Effects of Rare Codon Clusters on High-Level Expression of Heterologous Proteins in *Escherichia coli*." *Current Opinion in Biotechnology*. 6: 494-500.

- Kerepesi, C., Szabó, J.E., Papp-Kádár, V., Dobay, O., Szabó, D., Grolmusz, V., and Vértessy, B.G. 2016. "Life without dUTPase." *Frontiers in Microbiology* 7: 1768.
- Kim, M., and Hong, M. 2016. "Crystal Structure of the Bacillus-Conserved MazG Protein, a Nucleotide Pyrophosphohydrolase." *Biochemical and Biophysical Research Communications* 472 (1): 237–42.
- Krissinel, E. 2012. "Enhanced Fold Recognition Using Efficient Short Fragment Clustering." *Journal of Molecular Biochemistry* 1: 76 – 85.
- Lee, S., Kim, M. H., Kang, B. S., Kim, J., Kim, G., Kim, Y., and Kim, K. J. 2008. "Crystal Structure of *Escherichia coli* MazG , the Regulator of Nutritional Stress Response. *The Journal of Biological Chemistry* 283 (22): 15232–40.
- Leveles, I., Németh, V., Szabó, J.E., Harmat, V., Nyíri, K., Bendes, A. A., Papp-Kádár, V. et al. 2013. "Structure and Enzymatic Mechanism of a Moonlighting dUTPase." *Acta Crystallographica Section D: Biological Crystallography* 69 (12): 2298–2308.
- Liebschner, D., Afonine, P., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze, B., et al. 2019. "Macromolecular Structure Determination Using X-Rays, Neutrons and Electrons: Recent Developments in Phenix." *Acta Crystallographica Section D: Structural Biology* 75 (October): 861–77.
- Lu, L. D. , Sun, Q., Yong Fan, X., Zhong, Y., Yao, Y. F., and Zhao, G. P. 2010. "Mycobacterial MazG Is a Novel NTP Pyrophosphohydrolase Involved in Oxidative Stress Response." *Journal of Biological Chemistry* 285 (36): 28076–85.
- Lyu, L. D, Tang, B., Fan, X., Ma, H., and Zhao, G. 2013. "Mycobacterial MazG Safeguards Genetic Stability via Housecleaning of 5-OH-dCTP" 9 (12).
- Madrid, A.S., and Ganem, D. 2012. "Kaposi's Sarcoma-Associated Herpesvirus ORF54/dUTPase Downregulates a Ligand for the NK Activating Receptor NKp44." *Journal of Virology* 86 (16): 8693–8704.
- Maiques, E., Quiles-Puchalt, N., Donderis, J., Ciges-Tomas, J.R., Alite, C., Bowering, J., Humphrey, S., Penadés, J.R., and Marina, A. 2016. "Another Look at the Mechanism Involving Trimeric dUTPases in *Staphylococcus aureus* Pathogenicity Island Induction Involves Novel Players in the Party." *Nucleic Acids Research* 44 (11).
- McCoy, A.J., Grosse-Kunstleve, R. W., Adams, P.D., Winn, M. D., Storoni, L.C., and Read, R.J. 2007. "Phaser Crystallographic Software." *Journal of Applied Crystallography* 40 (4): 658–74.

- Millman, A., Melamed, S., Leavitt, A., Doron, S., Bernheim, A., Hör, J., Garb, J., et al. 2022. "An Expanded Arsenal of Immune Systems That Protect Bacteria from Phages." *Cell Host and Microbe* 30 (11): 1556-1569.e5.
- Mol, C.D., Harris, J.M., McIntosh, E.M., and Tainer, J.A. 1996. "Human dUTP Pyrophosphatase: Uracil Recognition by a Hairpin and Active Sites Formed by Three Separate Subunits," *Structure* 4 (9): 1077–92.
- Moroz, O. V., Harkiolaki, M., Galperin, M. Y., Vagin, A. A., González-Pacanowska, D., and Wilson, K.S. 2004. "The Crystal Structure of a Complex of *Campylobacter jejuni* dUTPase with Substrate Analogue Sheds Light on the Mechanism and Suggests the 'Basic Module' for Dimeric d(C/U)TPases." *Journal of Molecular Biology* 342 (5): 1583–97.
- Moroz, O. V., Murzin, A.G., Makarova, K.S., Koonin, E. V., Wilson, K.S., and Galperin, M.Y. 2005. "Dimeric dUTPases, HisE, and MazG Belong to a New Superfamily of All- α NTP Pyrophosphohydrolases with Potential 'House-Cleaning' Functions," 243–55.
- Mota, C.S., Gonçalves, A.M., and de Sanctis, D. 2016. "*Deinococcus radiodurans* DR2231 Is a Two-Metal-Ion Mechanism Hydrolase with Exclusive Activity on dUTP." *FEBS Journal* 283 (23): 4274–90.
- Murshudov, G. N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R. A. Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. 2011. "REFMAC5 for the Refinement of Macromolecular Crystal Structures." *Acta Crystallographica Section D: Biological Crystallography* 67 (4): 355–67.
- Mustafi, D., Bekesi, A., Vertessy, B.G., and Makinen, M. W. 2003. "Catalytic and Structural Role of the Metal Ion in dUTP Pyrophosphatase." *PNAS* 100 (10): 5670-5675.
- Németh-Pongrácz, V., Barabás, O., Fuxreiter, M., Simon, I., Pichová, I., Rumlová, M., Zábranská, H., et al. 2007. "Flexible Segments Modulate Co-Folding of dUTPase and Nucleocapsid Proteins." *Nucleic Acids Research* 35 (2): 495–505.
- Nyíri, K., Mertens, H., Tihanyi, B., Nagy, G.N., Kohegyi, B., Matejka, J., Harris, M. J., et al. 2018. "Structural Model of Human dUTPase in Complex with a Novel Proteinaceous Inhibitor." *Scientific Reports* 8 (1): 1–15.
- Nyman, P.O. 2001. "Introduction to dUTPases." *Current Protein and Peptide Science* 2: 277–85.

- Papp-Kádár, V., Szabó, J.E., Nyíri, K., and Vertessy, B. G. 2016. "In Vitro Analysis of Predicted DNA-Binding Sites for the StI Repressor of the *Staphylococcus aureus* SaPIBov1 Pathogenicity Island." *PLoS ONE* 11 (7).
- Pecsi, I., Hirmondo, R., Brown, A. C., Lopata, A., Parish, T., Vertessy, B.G., and Tóth, J. 2012. "The dUTPase Enzyme Is Essential in *Mycobacterium smegmatis*." *PLoS ONE* 7 (5): 1–11.
- Pécsi, I., Szabó, J.E., Adams, S. D., Simon, I., Sellers, J.R., Vértessy, B.G., and Tóth, J. 2011. "Nucleotide Pyrophosphatase Employs a P-Loop-like Motif to Enhance Catalytic Power and NDP/NTP Discrimination." *PNAS* 108 (35): 14437–42.
- Pegan, S.D., Tian, Y., Sershon, V., and Mesecar, A.D. 2010. "A Universal, Fully Automated High Throughput Screening Assay for Pyrophosphate and Phosphate Release from Enzymatic Reactions." *Combinatorial Chemistry & High Throughput Screening*. Vol. 13.
- Penadés, J.R., and Christie, G. E. 2015. "The Phage-Inducible Chromosomal Islands: A Family of Highly Evolved Molecular Parasites." *Annual Review of Virology*. Annual Reviews Inc.
- Penadés, J. R., Donderis, J., García-Caballer, M., Tormo-Más, M.A., and Marina, A. 2013. "dUTPases, the Unexplored Family of Signalling Molecules." *Current Opinion in Microbiology* 16 (2): 163–70.
- Powell, H. R. 2017. "X-Ray Data Processing." *Bioscience Reports*. Portland Press Ltd.
- Requena, C. E., Pérez-Moreno, G., Ruiz-Pérez, L.M., Vidal, A.E., and González-Pacanowska, D. 2014. "The NTP Pyrophosphatase DCTPP1 Contributes to the Homeostasis and Cleansing of the dNTP Pool in Human Cells." *Biochemical Journal* 459 (1): 171–80.
- Rihtman, B., Bowman-Grahl, S., Millard, A., Corrigan, R.M., Clokie, M.R.J., and Scanlan, D.J. 2019. "Cyanophage MazG Is a Pyrophosphohydrolase but Unable to Hydrolyse Magic Spot Nucleotides." *Environmental Microbiology Reports* 11 (3): 448–55.
- Robinson, A., Guilfoyle, A.P., Harrop, S.J., Boucher, Y., Stokes, H.W., Curmi, P.M.G., and Mabbutt, B.C. 2007. "A Putative House-Cleaning Enzyme Encoded within an Integron Array : 1 . 8 Å Crystal Structure Defines a New MazG Subtype" *Molecular Biology*, 66 (3): 610–21.

- Samara, N.L., Gao, Y., Wu, J., and Yang, W. 2017. "Detection of Reaction Intermediates in Mg²⁺-Dependent DNA Synthesis and RNA Degradation by Time-Resolved X-Ray Crystallography." In *Methods in Enzymology*, 592:283–327. Academic Press Inc.
- Samara, N.L., and Yang, W. 2018. "Cation Trafficking Propels RNA Hydrolysis." *Nature Structural and Molecular Biology* 25 (8): 715–21.
- Sanz-Frasquet, C., Ciges-Tomas, J.R., Alite, C., Penadés, J.R., and Marina, A. 2023. "The Bacteriophage–Phage-Inducible Chromosomal Island Arms Race Designs an Interkingdom Inhibitor of dUTPases." *Microbiology Spectrum*, 11 (1).
- Steyn, A.J.C., Joseph, J., and Bloom, B.R. 2003. "Interaction of the Sensor Module of *Mycobacterium tuberculosis* H37Rv KdpD with Members of the Lpr Family." *Molecular Microbiology* 47 (4): 1075–89.
- Sultana, A., and Lee, J.E. 2015. "Measuring Protein-Protein and Protein-Nucleic Acid Interactions by Biolayer Interferometry." *Current Protocols in Protein Science* 2015: 19.25.1-19.25.26.
- Szabó, J. E., Németh, V., Papp-Kádár, V., Nyíri, K., Leveles, I., Bendes, A., Zagyva, I., et al. 2014. "Highly Potent dUTPase Inhibition by a Bacterial Repressor Protein Reveals a Novel Mechanism for Gene Expression Control." *Nucleic Acids Research* 42 (19): 11912–20.
- Takács, E., Nagy, E., Leveles, I., Harmat, V., Lopata, A., Tóth, J., and Vértessy, B.G. 2010. "Direct Contacts between Conserved Motifs of Different Subunits Provide Major Contribution to Active Site Organization in Human and Mycobacterial dUTPases." *FEBS Letters* 584 (14): 3047–54.
- Tal, N., Millman, A., Stokar-Avihail, A., Fedorenko, T., Leavitt, A., Melamed, S., Yirmiya, E., et al. 2022. "Bacteria Deplete Deoxynucleotides to Defend against Bacteriophage Infection." *Nature Microbiology* 7 (8): 1200–1209.
- Tarbouriech, N., Buisson, M., Seigneurin J.M., Cusack, S. and Burmeister, W.P. 2005. "The Monomeric dUTPase from Epstein-Barr Virus Mimics Trimeric dUTPases." *Structure* 13 (9): 1299–1310.
- Tormo-Más, M.A., Donderis, J., García-Caballer, M., Alt, A., Mir-Sanchis, I., Marina, A., and Penadés, J.R. 2013. "Phage dUTPases Control Transfer of Virulence Genes by a Proto-Oncogenic G Protein-like Mechanism." *Molecular Cell* 49 (5): 947–58.

- Tormo-Más, M.A, Mir, I., Shrestha, A., Tallent, S.M., Campoy, S., Lasa, I., Barbé, J., Novick, R.P., Christie, G.E., and Penadés, J.R. 2010. "Moonlighting Bacteriophage Proteins Derepress Staphylococcal Pathogenicity Islands." *Nature* 465 (7299): 779–82.
- Tóth, J., Varga, B., Kovács, M., Málnási-Csizmadia, A., and Vértessy, B.G. 2007. "Kinetic Mechanism of Human dUTPase, an Essential Nucleotide Pyrophosphatase Enzyme." *Journal of Biological Chemistry* 282 (46): 33572–82.
- Úbeda, C., Maiques, E., Barry, P., Matthews, A., Tormo-Mas, M.A., Lasa, I., Novick, R.P., and Penadés, J.R. 2008. "SaPI Mutations Affecting Replication and Transfer and Enabling Autonomous Replication in the Absence of Helper Phage." *Molecular Microbiology* 67 (3): 493–503.
- Vardakou, M., Salmon, M., Faraldos, J.A., and O'Maille, P.E. 2014. "Comparative Analysis and Validation of the Malachite Green Assay for the High Throughput Biochemical Characterization of Terpene Synthases." *MethodsX* 1: e187–96.
- Varga, B., Barabás, O., Kovári, J., Tóth, J., Hunyadi-Gulyás, E., Klement, E., Medzihradszky, K.F., Tölgyesi, F., Fidy, J., and Vértessy, B.G. 2007. "Active Site Closure Facilitates Juxtaposition of Reactant Atoms for Initiation of Catalysis by Human dUTPase." *FEBS Letters* 581 (24): 4783–88.
- Vértessy, B.G., and Tóth, J. 2009. "Keeping Uracil out of DNA: Physiological Role, Structure and Catalytic Mechanism of dUTPases." *Accounts of Chemical Research* 42 (1): 97–106.
- Wang, F., Liu, C., Wang, C., Wang, Y., Zang, K., Wang, X., Liu, X., Li, S., Li, F., and Ma Q. 2021. "Structural Basis of Staphylococcal StI Inhibition on a Eukaryotic dUTPase." *International Journal of Biological Macromolecules* 184 (2021): 821–30.
- Wegrecki, M., Neira, J.L., and Bravo, J. 2015. "The Carboxy-Terminal Domain of Erb1 Is α Seven-Bladed β -Propeller That Binds RNA." *PLoS ONE* 10 (4).
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., et al. 2011. "Overview of the CCP4 Suite and Current Developments." *Acta Crystallographica Section D: Biological Crystallography* 67 (4): 235–42.
- Zhang, J., and Inouye, M. 2002. "MazG, a Nucleoside Triphosphate Pyrophosphohydrolase, Interacts with Era, an Essential GTPase in *Escherichia coli*" 184 (19): 5323–29.

- Zhang, J., Zhang, Y., and Inouye, M. 2003. “*Thermotoga maritima* MazG Protein Has Both Nucleoside Triphosphate Pyrophosphohydrolase and Pyrophosphatase Activities” 278 (24): 21408–14.
- Zheng, H., Cooper, D.R., Porebski, P.J., Shabalin, I.G., Handing, K.B., and Minor, W. 2017. “CheckMyMetal: A Macromolecular Metal-Binding Validation Tool.” *Acta Crystallographica Section D: Structural Biology* 73: 223–33.

