

Alejandro del Río García

INGENIERÍA DE PROTEÍNAS PARA AUMENTAR LA TERMOESTABILIDAD DE LA POLIMERASA DEL BACTERIÓFAGO ϕ 29

Trabajo final de grado

Tutor: Dr. Gerard Pujadas Anguiano

Biotecnología



UNIVERSITAT ROVIRA I VIRGILI

Tarragona

Curso: 2017-2018

INDICE

1. RESUMEN.....	3
2.INTRODUCCIÓN.....	4
3. HIPOTESIS Y OBJETIVOS	9
4.MATERIALES Y MÉTODOS	10
4.1) Búsqueda de la estructura de las proteínas e identificación de la unidad funcional de la polimerasa ϕ 29	10
4.2) Análisis de los factores que contribuyen a la termoestabilidad de una enzima	11
4.3) Aumento de la termoestabilidad mediante “alanine scanning”	12
4.4) Aumento de la termoestabilidad mediante adición de un puente disulfuro	13
4.5) Análisis de la termoestabilidad de la polimerasa ϕ 29 de (Povilaitis et al., 2016)	14
5.RESULTADOS Y DISCUSIÓN	15
5.1) Análisis de los factores que contribuyen a la termoestabilidad de una enzima	16
5.2) Aumento de la termoestabilidad de la polimerasa ϕ 29 mediante “alanine scanning”	17
5.3) Aumento de la termoestabilidad de la polimerasa ϕ 29 por adición de un puente disulfuro.....	24
5.4) Combinación de ambas estrategias: Pol ϕ 29 mut2+puente disulfuro	27
5.5) Análisis de la termoestabilidad de la polimerasa ϕ 29 de (Povilaitis et al., 2016)	28
6.CONCLUSIONES.....	31
7.AUTOEVALUACION	32
8.BIBLIOGRAFIA.....	33
9.ANEXOS.....	37

1. RESUMEN

La polimerasa $\phi 29$ (Pol $\phi 29$) es la replicasa del bacteriófago $\phi 29$ que infecta a *Bacillus subtilis*. Es una polimerasa muy usada en PCR isotérmica y en técnicas de secuenciación de genoma completo (WGA). Además es una enzima mesófila y a temperaturas elevadas su actividad catalítica disminuye. En este estudio, se aumenta la termoestabilidad de la polimerasa $\phi 29$ incrementando las interacciones entre los residuos de la proteína. Para llevar a cabo este objetivo, se analizan dos estrategias diferentes: “alanine scanning” y introducción de un puente disulfuro. Los valores de energía muestran como el doble mutante Pol $\phi 29$ K371C-A382C es el que aumenta más su termoestabilidad en comparación con la Pol $\phi 29$ *wild type* (Pol $\phi 29$ -WT) (-366,57 kcal/mol y -124,93 kcal/mol respectivamente). La combinación de ambas estrategias en cambio desestabilizan la enzima. La polimerasa mutante que se presenta en este estudio se podría utilizar en futuros experimentos in vitro para analizar su estabilidad a temperaturas elevadas para aumentar la eficiencia de la reacción en PCR isotérmica.

PALABRAS CLAVE: Ingeniería de proteínas, polimerasa $\phi 29$, termoestabilidad, puente disulfuro, interacciones.

2.INTRODUCCIÓN

Las enzimas son biomoléculas presentes en los organismos vivos. Son biocatalizadores (normalmente proteicos) que disminuyen la energía de activación de una reacción química aumentando su velocidad.

Existen muchas enzimas importantes en el metabolismo de los seres vivos. Una de las más estudiadas es la ADN polimerasa replicativa. Esta enzima tiene un papel clave en la división celular, más concretamente en la replicación del genoma. Es la enzima encargada de la síntesis de una cadena de ADN complementaria a partir de una cadena molde. De esta forma el genoma de la célula va replicándose cuando esta entra en división. Las ADN polimerasas replicativas también se llaman nucleotidil transferasas por su actividad catalítica (Fig1). Se clasifican en diferentes familias según la homología de su estructura primaria. Así pues, las ADN polimerasas replicativas de eucariotas se encuentran en la familia B, las bacterias en la familia A y C y las arqueas en la B y D (Johansson and Dixon, 2013). La mayoría de polimerasas tienen una estructura terciaria que sigue el modelo conocido como “mano derecha”, con tres dominios conocidos como “dedos” (*fingers*), “palma” (*palm*) y “pulgar” (*thumb*). En el dominio de la “palma” se localiza la actividad replicasa. En él, el extremo 3' hidroxilo libre de la cadena que se está sintetizando actúa como nucleófilo que ataca el grupo fosfato del deoxinucleótido (dNTPs) que es complementario en la cadena molde. El dominio “dedo” posiciona los nucleótidos entrantes en relación con la hebra patrón y induce cambios conformacionales cuando el nucleótido se posiciona cerca de dos iones metálicos divalentes (Mg^{+2}) que actúan como cofactores. Por último, el dominio “pulgar” participa en la procesividad y translocación del ADN, manteniendo en su lugar el dúplex de ADN en elongación (Johansson and Dixon, 2013).

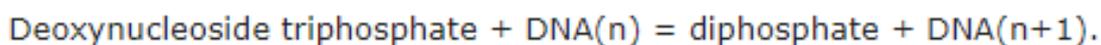


Fig. 1 Reacción química simplificada de las ADN polimerasas. A partir de la cadena simple de ADN y del sustrato (los deoxinucleótidos trifosfatos o dNTPs) se genera difosfato y una cadena de ADN con un nucleótido incorporado (n+1) que es complementario a la cadena molde (*polA* - DNA polymerase I, *thermostable* - *Thermus aquaticus* - *polA* gene & protein, no date)

Las ADN polimerasas se usan en múltiples aplicaciones biotecnológicas en investigación básica y en diagnóstico. Son muy importantes en las técnicas de secuenciación de genomas. En el año 1977 Sanger y Coulson usaron la ADN polimerasa I de *Escherichia coli* para secuenciar el ADN del bacteriófago ϕ X174 descubriendo lo que se conoce como secuenciación de Sanger (Aschenbrenner and Marx, 2017) . A partir de aquí el campo de la secuenciación ha crecido a pasos agigantados con la aparición de las técnicas de secuenciación masiva de última generación (NGS). Estas técnicas permiten obtener la secuencia nucleotídica de ácidos nucleicos en un menor tiempo y con un bajo coste. La más destacada es la plataforma *Illumina* que utiliza una variante de la polimerasa 9^oN que posee mutaciones en los dominios de la palma y los dedos. Por otro lado, la plataforma *IonTorrent* utiliza la *Bst* ADN polimerasa para la secuenciación del genoma (Aschenbrenner and Marx, 2017).

La técnica más importante del uso de las polimerasas es la *Polymerase Chain Reaction* más conocida como PCR. Esta aplicación se basa en el uso de estas enzimas para obtener múltiples copias de fragmentos específicos de ADN a partir de una muestra pequeña. Con el paso del tiempo surgieron múltiples adaptaciones y mejoras de la técnica inicial como la PCR cuantitativa (qPCR), PCR isotérmica, la PCR transcriptasa (RT-PCR) y la amplificación de genoma entre otras. Cada una de estas variantes se usa para diferentes propósitos lo que conlleva al uso de polimerasas con unas características específicas (Aschenbrenner and Marx, 2017). Por ejemplo, para la detección de patógenos se necesitan polimerasas con una gran sensibilidad para poder detectar cantidades pequeñas de ácido nucleico del patógeno en la muestra. En cambio para procesos de clonaje se necesitan polimerasas con una tasa de error muy baja (Aschenbrenner and Marx, 2017).

En la PCR transcriptasa, se parte de una molécula de ARN para obtener la cadena complementaria de ADN (cDNA) mediante la transcriptasa inversa. Posteriormente se amplifica el ADN por PCR (Farrell and Farrell, 2010). Por ejemplo, esta técnica se utiliza para estudiar los genes que se expresan en las células en determinadas condiciones de estrés (Farrell and Farrell, 2010). Por otro lado, la PCR cuantitativa se basa en el uso de fluorocromos para monitorizar los ciclos de PCR. De esta forma a la vez que se amplifica se detecta la cantidad de ADN en la muestra. Esta técnica es muy importante en la detección de patógenos como virus que causan infecciones

congénitas como el varicela zoster virus, virus del herpes 1 y 2 y virus de la inmunodeficiencia humana 1 y 2 (Nagy, 2015). También es muy útil para el diagnóstico de enfermedades como la detección de trisomías usando marcadores de SNP y la detección de deleciones como la $\Delta F508$ en fibrosis quística. En una hora se puede amplificar y detectar la presencia de esta deleción en 32 muestras diferentes (Nagy, 2015). En la industria alimentaria esta técnica también es muy importante en la detección de patógenos en los alimentos como *Listeria monocytogenes*, *Salmonella enterica*, *Campylobacter jejuni* entre otros (Cocolin *et al.*, 2011). Además, la PCR también es muy importante en el análisis forense de muestras biológicas, en cadáveres y en test de paternidad.

De todas las variantes de PCR, la PCR clásica es la técnica más extendida en los laboratorios básicos de biología molecular. El proceso de PCR clásica consta de una serie de ciclos de temperatura. Cada ciclo consta de 3 pasos a temperaturas diferentes: 1) desnaturalización de la doble hélice de ADN a temperaturas superiores a 90°C. 2) Se reduce la temperatura para facilitar la hibridación de los cebadores, que delimitarán la región a amplificar. 3) Durante la fase de elongación de la cadena por parte de la polimerasa, la temperatura dependerá de las condiciones óptimas de esta. Este ciclo de temperaturas se produce en los llamados termocicladores. Son aparatos que permiten calentar y enfriar los tubos de PCR y automatizar así los ciclos de temperatura. Tanto la temperatura de los ciclos como el tiempo de desnaturalización son dos factores limitantes que pueden desactivar la enzima (Terpe, 2013).

Las temperaturas tan elevadas a las que se somete la reacción, hace necesario el uso de ADN polimerasas termoestables. Estas enzimas son capaces de mantener su actividad catalítica a temperaturas muy elevadas. Además, tienen un tiempo de vida medio y una eficacia de amplificación mayor. Una de las más utilizadas en PCR es la Taq polimerasa aislada de la bacteria termófila *Thermus aquaticus* (Terpe, 2013). Esta polimerasa posee actividad exonucleasa. Su temperatura óptima de elongación es entre 75-80°C. Además, la vida media de esta enzima a 95°C es de 40 min (Terpe, 2013). La eficiencia de esta enzima es de aproximadamente un 80% en fragmentos más cortos que 1kb y con un contenido en CG entre 45-65%. A medida que el tamaño del amplicón es mayor de 1kb la eficiencia va disminuyendo

(Terpe, 2013). Por último, la tasa de error (frecuencia de mutación x bp x duplicación) de la Taq polimerasa es de $1.8 \times 10^{-4} - 8.0 \times 10^{-6}$ (Terpe, 2013).

En los últimos años, la aparición de la variante de PCR isotérmica ha ganado terreno. La PCR isotérmica posibilita que la reacción de amplificación se produzca sin la necesidad de los ciclos de temperatura del termociclador. En definitiva, la reacción se realiza a temperatura constante sin la necesidad de ciclos. Muchas polimerasas son adecuadas para esta variante de PCR pero la ADN polimerasa $\phi 29$ posee unas propiedades muy valiosas y únicas en comparación con el resto de polimerasas comerciales. Esta polimerasa se aisló y se identificó en el bacteriófago $\phi 29$ que infecta la bacteria mesófila *Bacillus subtilis* (Lovmar and Syvänen, 2006). Es una polimerasa de la familia B y es una de las más pequeñas que se conocen. Tiene una tasa de síntesis de ADN muy alta (50-200 bases/s) (Takahashi *et al.*, 2014). A parte de la actividad de replicasa, también posee actividad $3' \rightarrow 5'$ exonucleasa que permite la alta eficiencia en la síntesis de ADN. Otra característica importante es que la propia polimerasa tiene actividad helicasa (desplazamiento de la cadena de ADN). Esta función la hace idónea para técnicas de amplificación isotérmica como *rolling circle amplification (RCA)* o *multiple displacement amplification (MDA)* y puede usarse en genomas circulares y en moléculas lineales de ADN (Takahashi *et al.*, 2014). Además, esta característica del desplazamiento de cadena permite amplificar fragmentos muy grandes 10-12kb con un contenido en GC de hasta un 80%. Así pues, esta enzima se suele utilizar en técnicas de amplificación de genomas completos (WGA). Por último, la tasa de error de esta polimerasa es una de las más bajas como consecuencia de las características comentadas anteriormente ($< 3.0 \times 10^{-6}$) (Lovmar and Syvänen, 2006).

Si se comparan las dos polimerasas, la $\phi 29$ tiene una tasa de error menor, tiene actividad helicasa intrínseca por lo que no se necesita la intervención de proteínas accesorias para separar las hebras de ADN. Además, tiene una eficiencia de síntesis mayor y es capaz de amplificar fragmentos mucho más largos. Sin embargo, la Taq polimerasa es termoestable y su temperatura óptima de actividad como ya se ha dicho es de 75-80°C. En cambio la temperatura óptima de la polimerasa del fago $\phi 29$ es de 30°C y se inactiva a 65°C (Gadkar and Rillig, 2005).

La velocidad de una reacción enzimática es un factor muy importante y depende de la temperatura (a mayor temperatura, mayor velocidad de reacción). Muchos trabajos han demostrado que aumentando la termoestabilidad de una enzima se puede trabajar a temperaturas más elevadas y aumentar así la vida media de la enzima. (Le *et al.*, 2012) demostraron que introduciendo un puente disulfuro en los residuos A162-K308 de la lipasa B de *Candida antarctica* se aumentaba la termoestabilidad. Además, la vida media a 50°C de la lipasa mutante era 4,5 veces más que la *wild-type* (220 min a 49 min respectivamente). Por otro lado, (Kim *et al.*, 2011) mutaron los aminoácidos F180Y y A752K de la ADN polimerasa de *Thermococcus celericrescens*. La vida media del doble mutante a 94°C fue de 42 min, mucho más mayor que la del *wild-type* (*wt*) (6 min). (Pezeshgi Modarres *et al.*, 2015) aumentaron la termoestabilidad de la ADN ligasa de *Thermococcus sp* con las mutaciones A287K, G304D, S364I, y A387K. El tiempo de inactivación a 94°C del mutante fue mucho mayor que el *wt* (41 min y 8 respectivamente). Muchas de estas mutaciones coinciden con residuos que forman enlaces adicionales con los aminoácidos cercanos.

Como ya se ha comentado, la polimerasa del fago ϕ 29 no es termoestable (temperatura óptima entre 30-35°C). Por lo tanto su estabilidad a temperaturas elevadas es un aspecto a mejorar. En el 2016, (Povilaitis *et al.*, 2016) desarrollaron una polimerasa ϕ 29 termoestable (7 mutaciones) que catalizaba la reacción a 40-42°C de manera óptima. Además, el tiempo de reacción disminuyó de 16 a 4h respecto la polimerasa *wt*.

3. HIPOTESIS Y OBJETIVOS

Un aumento de las interacciones entre los residuos de la enzima podría incrementar la termoestabilidad de la polimerasa $\phi 29$. A partir de esta hipótesis, los objetivos en este trabajo son los siguientes:

- ✓ Evaluar con programas bioinformáticos qué mutaciones son idóneas para aumentar la termoestabilidad de la polimerasa del bacteriófago $\phi 29$ y proponer un mutante final que sea termoestable
- ✓ Comparar las diferentes estrategias utilizadas y analizar los motivos que hacen la polimerasa del trabajo de (Povilaitis *et al.*, 2016) más termoestable que la *wt*.

La finalidad de este estudio es desarrollar una polimerasa mutante termoestable para poder usarla en PCR isotérmica a temperaturas elevadas para aumentar así la eficiencia de la reacción.

4.MATERIALES Y MÉTODOS

4.1) Búsqueda de la estructura de las proteínas e identificación de la unidad funcional de la polimerasa ϕ 29

Para obtener los códigos (PDB ID) y las estructuras de las polimerasas Taq y ϕ 29 se consultó la base de datos UniProt [<http://www.uniprot.org/>; (Bateman *et al.*, 2017)]. En el buscador se introducen las palabras clave “*Taq polymerase*” y “*Bacillus phage 29 polymerase*” respectivamente. Se selecciona la entrada que interesa: P19821 para la Taq polimerasa y P03680 para la polimerasa ϕ 29. En el apartado “*3D structure databases*” se muestran numerosos códigos para las polimerasas. Se seleccionará aquella que corresponda con la estructura de la polimerasa y no esté unida a su sustrato o un inhibidor. Las polimerasas con códigos 1TAQ (polimerasa Taq) y 1XHX (polimerasa ϕ 29) cumplen con los requisitos establecidos. Una vez obtenidos los códigos se descargarán sus estructuras de la base de datos *Protein Data Bank* (PDB) [<https://www.rcsb.org/>; (Berman *et al.*, 2000)].

Para detectar la unidad funcional a partir de la unidad asimétrica se utiliza el programa de visualización molecular *RasMol* <http://www.RasMol.org>. La estructura de la proteína se visualiza en formato *spacefill* y se pinta según la cadena polipeptídica. A continuación se restringe la visualización de la proteína a aquellos residuos que están a una distancia de 3 Å y 5 Å entre las dos cadenas para analizar si existe contacto entre las dos. La unidad funcional se puede obtener en el PDB [<https://www.rcsb.org/>; (Berman *et al.*, 2000)] en la sección *biological assembly*.

4.2) Análisis de los factores que contribuyen a la termoestabilidad de una enzima

Para el análisis de los diferentes parámetros que figuran en el apartado 5.1 se han utilizado diferentes programas/servidores bioinformáticos. Para el cálculo de los puentes salinos y los puentes disulfuro se usó el servidor *WHAT IF Web Interface* [<https://swift.cmbi.umcn.nl/servers/html/index.html>]; (Hekkelman *et al.*, 2010)]. Los resultados de los puentes salinos muestran puentes entre los mismos residuos varias veces. Esto se debe a que hay interacciones entre diferentes átomos de los residuos, pero todas estas interacciones componen el mismo puente salino. Por eso, se utilizó el programa *Cygwin* (<https://www.cygwin.com/>) para eliminar las repeticiones. El servidor *CAPTURE* [<http://capture.caltech.edu/>]; (Gallivan and Dougherty, 1999)] se utilizó para analizar las interacciones catión-pi y para cuantificar el número de glicinas se utilizó el programa *Swiss-PdbViewer* [<http://www.expasy.org/spdbv/>]; (Guex and Peitsch, 1997)]

Por último, para el análisis de los residuos hidrofóbicos con un empaquetaje subóptimo se utilizó la base de datos *PDB_REDO* [http://www.cmbi.ru.nl/pdb_redo/]; (Joosten *et al.*, 2014)]. Para cada una de las polimerasas en el apartado *WHAT CHECK* de la base de datos se indica los residuos con un empaquetaje subóptimo. Se seleccionan aquellos que son hidrofóbicos. A continuación para asegurarse que estos resultados no son debidos a un mal encaje con el mapa de densidad electrónica se utilizará el programa *Jmol* (*Jmol: un visor Java de código abierto para estructuras químicas en tres dimensiones*, no date) para verificarlo. La normalización se hace para asegurarse que los resultados no estén influenciados por la diferencia de longitud entre las dos polimerasas y se calcula por la siguiente ecuación:

$$\frac{\text{cantidad (número) del parámetro a estudiar}}{\text{residuos de la polimerasa}}$$

4.3) Aumento de la termoestabilidad mediante “alanine scanning”

Para el análisis del “alanine scanning” se trabajará con la estructura depositada en la base de datos del *PDB_REDO* [http://www.cmbi.ru.nl/pdb_redo/; (Joosten *et al.*, 2014)] de la polimerasa del fago. Esta base de datos mejora los modelos estructurales depositados en el PDB [<https://www.rcsb.org/>; (Berman *et al.*, 2000)] utilizando programas más modernos. Por eso, es mejor trabajar con las coordenadas de esta base de datos puesto que son más fiables.

Antes de realizar el “alanine scanning” se reparará el archivo del *PDB_REDO* [http://www.cmbi.ru.nl/pdb_redo/; (Joosten *et al.*, 2014)] de la polimerasa. Esto permite optimizar aquellos residuos de la proteína que tienen una energía demasiado elevada y generar un modelo mejor.

El programa *FoldX 3.0 Beta 5.1* (Schymkowitz *et al.*, 2005) permite: realizar la reparación del archivo del *PDB_REDO* [http://www.cmbi.ru.nl/pdb_redo/; (Joosten *et al.*, 2014)] 1XHX, analizar el “alanine scanning”, mutar los residuos candidatos por el resto de 19 aminoácidos, generar el mutante que recoge las mutaciones escogidas y calcular la diferencia de energía entre la proteína nativa y el mutante generado.

Los resultados de los residuos candidatos mutados por los 19 aminoácidos restantes después del “alanine scanning” aparecen como valores de energía absoluta. *FoldX 3.0 Beta 5.1* (Schymkowitz *et al.*, 2005) funciona muy bien en el cálculo de energías relativas (*Energía absoluta del mutante – energía absoluta de la proteína nativa*) pero en cambio no es muy fiable en el cálculo de energías absolutas. Por eso, antes de analizar los resultados se calculará las energías relativas.

Para visualizar la localización de los residuos en la estructura tridimensional de la proteína 1XHX reparada se utiliza el programa *RasMol* <http://www.RasMol.org>. Se visualiza la proteína en forma de esqueleto, se seleccionan los residuos que se quiere analizar y se visualiza en formato *spacefill* y de color rojo.

El análisis de las interacciones entre residuos para las 8 mutaciones se realiza con el servidor *Dynamut* (Rodrigues, Pires and Ascher, 2018).

SWISS-MODEL [<https://swissmodel.expasy.org/>; (Waterhouse *et al.*, 2018)] permite realizar el modelado 3D por homología para conocer la estructura del mutante. Se utiliza como “*template*” la estructura conocida de la proteína nativa 1XHX. Para la superposición de las dos estructuras se descarga el modelo generado por SWISS-MODEL [<https://swissmodel.expasy.org/>; (Waterhouse *et al.*, 2018)] y se visualiza con el programa *Swiss-PdbViewer* [<http://www.expasy.org/spdbv/>; (Guex and Peitsch, 1997)]. Las dos estructuras se visualizan en formato *ribbon* y se colorea según la capa.

4.4) Aumento de la termoestabilidad mediante adición de un puente disulfuro

Los programas usados para la predicción de los puentes disulfuro son: *MODIP* [<http://caps.ncbs.res.in/iws/modip.html>; (Sowdhamini *et al.*, 1989)] y *Disulfide by Design 2.0 v 2.12* [<http://cptweb.cpt.wayne.edu/DbD2/>; (Dombkowski, Craig and Dombkowski, 2013)]. El valor del ángulo de χ^3 escogido es el que aparece por defecto ($97^\circ \pm 30$). En el caso de la polimerasa $\phi 29$ *MODIP* predijo 134 puentes disulfuro diferentes posibles. Este programa clasifica los posibles puentes en categorías: A, B, C, D. Aquellos que están en la categoría A son los puentes disulfuro más probables de producirse. Es por eso que se cogen los que pertenecen a esta categoría (24 posibles). Por último, de esos 24 posibles pares de residuos solo se cogerán aquellos que sean predichos simultáneamente por *Disulfide by Design 2.0 v 2.12* [<http://cptweb.cpt.wayne.edu/DbD2/>; (Dombkowski, Craig and Dombkowski, 2013)] quedando los resultados de la tabla 6.

Los factores de temperatura (*B-factor*) para cada residuo de la polimerasa $\phi 29$ se obtienen del servidor *WHAT IF Web Interface* [<https://swift.cmbi.umcn.nl/servers/html/listavb.html>; (Hekkelman *et al.*, 2010)]. Posteriormente, para obtener el *B-factor* de cada pareja de residuos candidatos simplemente se hace el sumatorio de los *B-factor* individuales de cada residuo que forman la pareja candidata a la mutación. Para generar el mutante y calcular las energías de los dobles mutantes y la proteína nativa se utiliza *FoldX 3.0 Beta 5.1* (Schymkowitz *et al.*, 2005).

La situación de los residuos que forman parte del puente disulfuro se analiza con *Swiss-PdbViewer* [<http://www.expasy.org/spdbv/>; (Guex and Peitsch, 1997)]. Las estructuras secundarias se muestran en formato *ribbon* y los residuos que participan del puente disulfuro en formato esqueleto.

4.5) Análisis de la termoestabilidad de la polimerasa ϕ 29 de (Povilaitis *et al.*, 2016)

Para visualizar la localización de los siete residuos escogidos por Povilaitis se utiliza el programa *RasMol* <http://www.RasMol.org>. Se visualiza la proteína en formato esqueleto, se seleccionan los siete residuos y se visualizan en formato *spacefill* y de color rojo. Con *FoldX 3.0 Beta 5.1* (Schymkowitz *et al.*, 2005) se genera el mutante con las siete mutaciones y después se realiza el “*alanine scanning*” para analizar la contribución de estos residuos en la estabilidad de la enzima.

Para el cálculo de las interacciones por puentes de hidrógeno de los siete residuos se utiliza el programa *Swiss-PdbViewer* [<http://www.expasy.org/spdbv/>; (Guex and Peitsch, 1997)]. Se restringe la visualización a aquellos aminoácidos que se encuentran a una distancia máxima de 5 Å del residuo mutado y se analizan los enlaces.

5.RESULTADOS Y DISCUSIÓN

Para este estudio se necesita obtener la estructura de la polimerasa $\phi 29$. Además, para el primer apartado de los resultados se necesitará también la estructura de la polimerasa Taq termoestable.

Las estructuras de las polimerasas Taq y $\phi 29$ se obtienen de la base de datos *Protein Data Bank* (PDB) [<https://www.rcsb.org/>; (Berman *et al.*, 2000)] con códigos 1TAQ (Kim *et al.*, 1995) y 1XHX (Kamtekar *et al.*, 2004) respectivamente. Las estructuras depositadas en formato PDB corresponden a la unidad asimétrica que no tiene por qué coincidir con la unidad funcional. La unidad asimétrica hace referencia a un concepto cristalográfico. En cambio, como biotecnólogo interesa trabajar con las coordenadas de la proteína biológicamente activa (unidad funcional). Por ejemplo, una proteína puede estar cristalizada en forma de tetramero (unidad asimétrica) pero la unidad biológicamente activa o (unidad funcional) ser un dímero. Esto significaría que la unidad asimétrica, en este caso, estaría formada por dos proteínas iguales o unidades funcionales. Es decir la proteína activa estaría formada por dos cadenas y en ningún caso por cuatro.

La unidad asimétrica y la funcional coinciden en el caso de la Taq polimerasa (PDB ID: 1TAQ) pero en la polimerasa $\phi 29$ (PDB ID: 1XHX) la unidad asimétrica está formada por más de una cadena. Para verificar si la unidad asimétrica es la misma que la funcional se analizará si existe contacto y superposición entre las diferentes cadenas polipeptídicas de la unidad asimétrica gracias al programa de visualización molecular *RasMol* <http://www.RasMol.org>. Los resultados se muestran en la (Fig2).

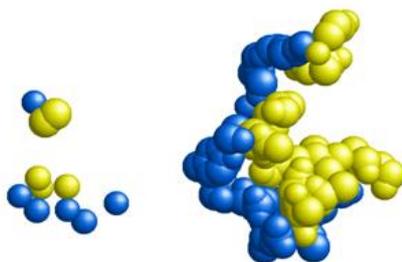


Fig. 2 Se visualizan los residuos con una distancia de 3,0 Å entre las dos cadenas (izquierda). Residuos con una distancia de 5,0 Å entre las dos cadenas (derecha). Cada color (azul y amarillo) corresponde con residuos de una cadena diferente. Programa: RasMol <http://www.RasMol.org>

Como puede verse en la (Fig2) no existe contacto directo entre las dos cadenas y los átomos no se superponen entre ellos. Así pues, se puede afirmar que la unidad funcional de la polimerasa $\phi 29$ está formada por una cadena polipeptídica ya que si dos cadenas no están totalmente en contacto no pueden formar parte de la misma proteína.

5.1) Análisis de los factores que contribuyen a la termoestabilidad de una enzima

Una de las características intrínsecas de las proteínas es su flexibilidad. La flexibilidad de una proteína juega un papel muy importante en su estabilidad a temperaturas elevadas. Si una proteína es muy flexible, tendrá poca estabilidad y será propensa a desnaturalizarse y a inactivarse a temperaturas moderadas. Por el contrario, una proteína rígida y por lo tanto poco flexible soporta temperaturas más elevadas. Cualquier interacción que exista entre residuos y átomos de la misma proteína aumenta la rigidez. Por lo tanto, si somos capaces de incrementar las interacciones entre residuos de la misma proteína podremos aumentar su termoestabilidad.

Para empezar con la investigación, se analizará el número de estas interacciones en nuestra polimerasa *target* mesófila $\phi 29$ y la polimerasa Taq termófila. También se analizará el número de glicinas ya que este aminoácido aumenta la movilidad localmente. Por último, también se analizará los residuos hidrofóbicos con un empaquetaje subóptimo puesto que un número elevado provoca una disminución de la estabilidad de la proteína. Dada la diferencia considerable en la longitud entre la polimerasa $\phi 29$ (575 aminoácidos) (Kamtekar *et al.*, 2004) y la Taq polimerasa (832 aminoácidos) (Kim *et al.*, 1995) se normalizarán los valores para que sean fiables. Los resultados se presentan en la Tabla 1.

Tabla 1 Análisis del número de las diferentes interacciones presentes en la polimerasa ϕ 29 y la polimerasa Taq.

	Polimerasa ϕ 29 (1XHX)		Polimerasa Taq (1TAQ)	
	Resultado	Normalización	Resultado	Normalización
Número de puentes salinos	79	0,14	190	0,23
Número de interacciones catión-pi	13	0,023	13	0,016
Número de puentes disulfuro	0	-	0	-
Número de glicinas	41	0,071	55	0,066
Número de residuos hidrofóbicos de la proteína con un empaquetaje subóptimo	0	-	2	0,002

5.2) Aumento de la termoestabilidad de la polimerasa ϕ 29 mediante “alanine scanning”

Viendo los resultados anteriores es difícil predecir qué estrategia seguir para continuar con el trabajo. No hay diferencias significativas entre los resultados de las dos polimerasas. Para seguir con la investigación se realizará un “alanine scanning”. Esta estrategia muta todos los aminoácidos de la proteína a estudiar por alaninas y analiza como varía la energía del mutante respecto a la proteína nativa. Este análisis permite averiguar como contribuye cada aminoácido en la estabilidad de la proteína. Una proteína es más estable conforme su energía total es menor (más negativo). El resultado del “alanine scanning” se presenta como una diferencia de energías relativas (*Energía del mutante – energía de la proteína nativa*). Si esta diferencia es negativa significa que la energía del mutante es más negativa y entonces el mutante es más estable que la proteína nativa. Por el contrario, si esta diferencia es positiva significa que la energía del mutante es mayor y por lo tanto el aminoácido original contribuye en la estabilidad de la proteína. Si queremos aumentar la termoestabilidad de la polimerasa del bacteriófago, los residuos candidatos a mutar serán aquellos que muestren una diferencia de energía más negativa. Los resultados simplificados para la polimerasa ϕ 29 se presentan en la Tabla 2.

Tabla 2 Resultado del "alanine scanning" para Pol ϕ 29-WT. Se muestran solo los 7 aminoácidos que tienen una diferencia de energía más negativa. PROGRAMA: FoldX 3.0 Beta 5.1 (Schymkowitz *et al.*, 2005)

FoldX 3.0 Beta 5.1 (2011)							
by the FoldX Consortium							
Jesper Borg, Frederic Rousseau, Joost Schymkowitz,							
Luis Serrano and Francois Stricher							

PDB file analysed: RepairPDB_1xhx_final.pdb							
Output type: alascan							
ASP	249	to	ALA	energy	change	is	-2,39019 kcal/mol
ASP	458	to	ALA	energy	change	is	-1,29263 kcal/mol
THR	368	to	ALA	energy	change	is	-1,03685 kcal/mol
SER	194	to	ALA	energy	change	is	-0,969414 kcal/mol
ARG	23	to	ALA	energy	change	is	-0,866879 kcal/mol
ARG	236	to	ALA	energy	change	is	-0,774941 kcal/mol
THR	16	to	ALA	energy	change	is	-0,768795 kcal/mol

La Tabla 2 muestra los 7 aminoácidos con una diferencia de energía menor y por lo tanto candidatos a ser mutados para aumentar la estabilidad de la enzima. La mutación en los aminoácidos ASP249 y ASP458 es muy probable que produzca una gran pérdida en la funcionalidad de la polimerasa. Estudios han demostrado que estos residuos están muy conservados porque tienen un papel crítico en la actividad replicasa de la polimerasa (Blasco *et al.*, 1993; Kamtekar *et al.*, 2004). Por esta razón, estos residuos se descartan y no serán estudiados.

Es importante observar si los residuos candidatos están alejados en la estructura tridimensional de la proteína. Si lo están, los resultados de las mutaciones individuales pueden ser acumulativos y es de esperar que el incremento

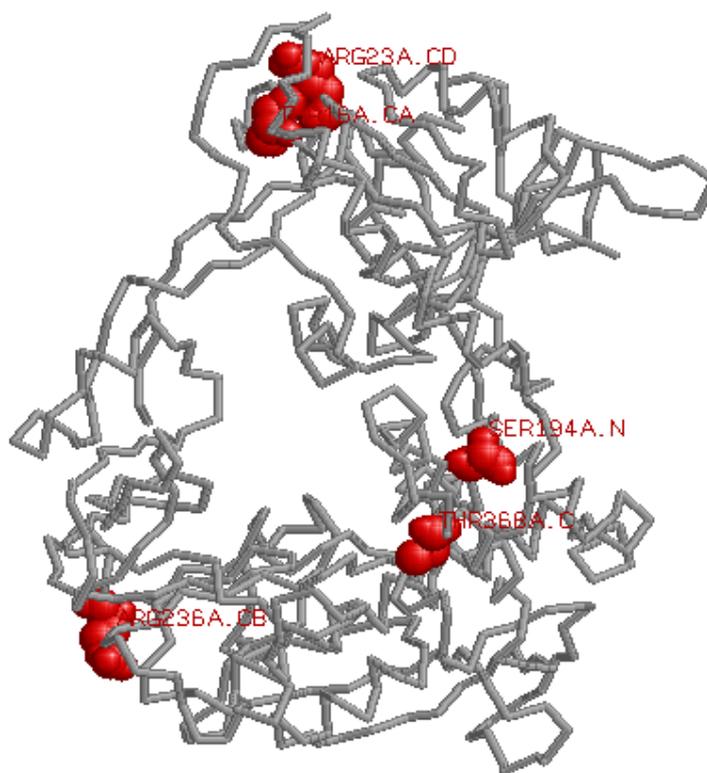


Fig. 3 Representación en esqueleto de la polimerasa ϕ 29. Se indica en rojo la situación de los 5 aminoácidos candidatos a ser mutados: arriba (ARG23, THR16). Abajo y de izquierda a derecha (ARG236, THR368, SER194). PROGRAMA: RasMol <http://www.RasMol.org>

de la termoestabilidad sea mayor. Por el contrario, si están muy próximos, las mutaciones simultáneas podrían no incrementar la estabilidad de la proteína e incluso afectar su funcionalidad. En la (Fig3) se muestra la localización de los cinco residuos candidatos a ser mutados.

Como se puede observar, los residuos se concentran en 4 áreas rojas. ARG23 y THR16 se encuentran muy próximos. Para evitar los problemas mencionados solo se elegirá ARG23 ya que estabiliza más la proteína que THR16 (-0,86 y -0,76 kcal/mol respectivamente). Los tres residuos restantes (SER194, ARG236 y THR368) están suficientemente separados para que todos ellos aumenten simultáneamente la estabilidad de la proteína.

Los cuatro residuos anteriores se escogen para ser mutados. Es importante estudiar la mutación de estos residuos por los aminoácidos restantes (aparte de alanina). Puede ser que alguno de ellos genere una mayor estabilidad en la proteína mutante. Los resultados muestran que las mejores mutaciones para aumentar la termoestabilidad de la polimerasa del fago $\phi 29$ son: ARG23-PHE23 (-1,95 kcal/mol), SER194-ALA194 (-0,84 kcal/mol), ARG236-ALA236 (-0,71 kcal/mol) y THR368-MET368 (-2,99 kcal/mol). Se genera el mutante con las cuatro mutaciones (Pol $\phi 29$ mut) y se calcula la diferencia de energía entre este mutante y el *wt* (Pol $\phi 29$ -WT). Los resultados se muestran en la Tabla 3.

Tabla 3. Energías totales o absolutas de la polimerasa $\phi 29$ con las mutaciones PHE23, ALA194, ALA236 y MET368 (Pol $\phi 29$ mut) y la polimerasa *wt* (Pol $\phi 29$ -WT). También se muestra la diferencia de energía entre el mutante y la *wt*

	ENERGÍA TOTAL (kcal/mol)	DIFERENCIA ENERGIA (kcal/mol)
Pol $\phi 29$ mut	-132,12 kcal/mol	-5,52 kcal/mol
Pol$\phi 29$-WT	-126,60 kcal/mol	

Como se puede observar en la Tabla 3 la energía total de Pol $\phi 29$ mut es menor que la del *wt* (-132,12 kcal/mol y -126,60 kcal/mol respectivamente). Es decir, las mutaciones han disminuido la energía del mutante aumentando así su estabilidad.

Más concretamente la estabilidad que provocan las cuatro mutaciones a la vez es de -5,52 kcal/mol.

Así pues, queda demostrado que las cuatro mutaciones anteriores estabilizan la polimerasa de forma notable. A continuación se analizará si aún se puede aumentar más esta característica. Se realizará el mismo procedimiento anterior del “alanine scanning” pero con el mutante Polφ29 mut que ya ha demostrado ser más estable que la *wt*. De esta forma, se pretende buscar nuevas mutaciones que aún puedan aumentar más la termoestabilidad de la polimerasa. La Tabla 4 muestra los 10 residuos candidatos a ser mutados por tener los niveles de energía menores.

Tabla 4. Resultados del “alanine scanning” para Polφ29 mut . Solo se muestran los 10 residuos con valores de energía más negativos. PROGRAMA: FoldX 3.0 Beta5.1 (Schymkowitz *et al.*, 2005)

FoldX 3.0 Beta 5.1 (2011)							
by the FoldX Consortium							
Jesper Borg, Frederic Rousseau, Joost Schymkowitz,							
Luis Serrano and Francois Stricher							

PDB file analysed: MUT_RepairPDB_lxhx_final.pdb							
Output type: alascan							
GLY	191	to	ALA	energy	change	is	-1,05233 kcal/mol
GLY	197	to	ALA	energy	change	is	-0,674305 kcal/mol
SER	388	to	ALA	energy	change	is	-0,648973 kcal/mol
ASP	200	to	ALA	energy	change	is	-0,624489 kcal/mol
ASP	569	to	ALA	energy	change	is	-0,57471 kcal/mol
SER	10	to	ALA	energy	change	is	-0,560519 kcal/mol
GLY	217	to	ALA	energy	change	is	-0,558631 kcal/mol
ASP	169	to	ALA	energy	change	is	-0,527626 kcal/mol
GLY	391	to	ALA	energy	change	is	-0,477778 kcal/mol
LYS	305	to	ALA	energy	change	is	-0,47291 kcal/mol

Si se observan los resultados del “alanine scanning” de los residuos que previamente se ha visto que aumentaban la estabilidad en Polφ29 mut (PHE23, ALA194, ALA236 y MET368) (*data not shown*) muestran como su mutación a Ala no es favorable para la estabilidad de la proteína. Por ejemplo, en el caso del residuo 23, el cambio de PHE23-ALA23 da un resultado de 1,13 kcal/mol (*data not shown*). Esto confirma que la mutación ARG23-PHE23 predicha antes en Polφ29 mut estabiliza la polimerasa puesto que ahora la diferencia de energía es positiva.

La localización de los 10 residuos candidatos a ser mutados (Fig 4) muestra como existen 6 áreas distintas. Por un lado, los residuos SER388, GLY191, y GLY391 se encuentran muy próximos en el espacio. Por lo tanto, solo se mutará el residuo GLY191 por ser aquel que contribuye menos a la estabilidad de la proteína (-1,05 kcal/mol). En el caso del clúster formado por ASP200 y GLY197 se mutará solo este último por tener una menor energía (-0,67 kcal/mol). Por último, del clúster formado por SER10 y ASP200 solo se mutará el residuo 10 (-0,64 kcal/mol). El resto de residuos están separados para poder aumentar todos ellos la estabilidad de la proteína.

Por lo tanto, los residuos a mutar de Pol ϕ 29 mut son: SER10, GLY191, GLY197, GLY217, LYS305 y ASP569. El análisis posterior de estos residuos muestra que las mejores mutaciones para aumentar la termoestabilidad de Pol ϕ 29 mut son: SER10-VAL10 (-0,71 kcal/mol), GLY191-GLU191 (-2,25 kcal/mol), GLY197-LEU197 (-2,62 kcal/mol), GLY217-PRO217 (-2,06 kcal/mol), LYS305-PRO305 (-1,35 kcal/mol) y ASP569-TYR569 (-2,24 kcal/mol). Se puede observar como la Gly es el aminoácido más común para ser mutado. Este resultado es de esperar puesto que la glicina es un aminoácido muy flexible que desestabiliza la proteína (Pace and Scholtz, 1998).

Las seis mutaciones anteriores son las óptimas para aumentar la termoestabilidad del mutante Pol ϕ 29 mut. Se genera este segundo mutante (Pol ϕ 29 mut2) y se calculan las energías. La Tabla 5 muestra las energías totales para la polimerasa *wt* inicial (Pol ϕ 29-WT) y Pol ϕ 29 mut2.

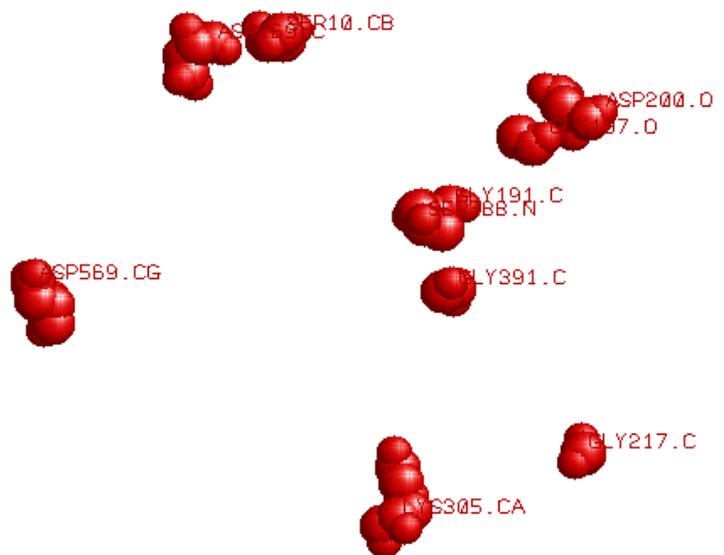


Fig. 4 Localización en rojo de los 10 residuos candidatos a ser mutados. PROGRAMA: RasMol <http://www.RasMol.org>

Tabla 5 Energías totales o absolutas de Pol ϕ 29 mut2 y Pol ϕ 29-*WT*. También se muestra la diferencia de energía entre este mutante y la *wt*

	ENERGÍA TOTAL (kcal/mol)	DIFERENCIA ENERGIA (kcal/mol)
Pol ϕ 29 mut2	-140,67 kcal/mol	-14,07 kcal/mol
Pol ϕ 29- <i>WT</i>	-126,60 kcal/mol	

La diferencia de energía entre Pol ϕ 29 mut2 y Pol ϕ 29-*WT* es de -14,07 kcal/mol. Esto demuestra que las seis mutaciones adicionales aumentan significativamente la estabilidad de la proteína *wt*. Además, el efecto de las mutaciones es acumulativo puesto que existe una diferencia de energía entre el mutante Pol ϕ 29 mut2 (-140,67 kcal/mol) y Pol ϕ 29 mut (-132,12 kcal/mol) de - 8,55 kcal/mol.

Las diez mutaciones de Pol ϕ 29 mut2 que aumentan la estabilidad coinciden con aminoácidos que forman enlaces no covalentes adicionales por su naturaleza química. En el **Anexo 1** se puede ver como la mutación SER10-VAL10 produce un incremento en el mutante del número de interacciones con los residuos situados alrededor. Esto confirma la hipótesis de que al aumentar el número de enlaces entre los aminoácidos de la proteína se incrementa la estabilidad de la polimerasa.

Para reforzar los resultados obtenidos se analiza el efecto de las diez mutaciones con un nuevo servidor: *DynaMut* (Rodrigues, Pires and Ascher, 2018). Como se puede observar en el **Anexo 2** todas las mutaciones excepto dos provocan un aumento en la estabilidad de la proteína ($\Delta\Delta G > 0$). Los residuos que estabilizan la proteína predichos por *FoldX 3.0 Beta 5.1* (Schymkowitz *et al.*, 2005) y *DynaMut* (Rodrigues, Pires and Ascher, 2018) se sugiere que son más probables de producirse. Por eso, los residuos 23 y 236 que desestabilizan la proteína según *DynaMut* (**Anexo2**) se descartarían.

Por último, se evalúa si las 8 mutaciones de Pol ϕ 29 mut2 afectan en la estructura tridimensional de la polimerasa. Se realiza un modelado de estructura 3D por homología utilizando como “*template*” la estructura de la proteína nativa 1XHX. Los resultados muestran un 98,77% de similitud en la secuencia. El valor QMEAN (*Qualitative Model Energy ANalysis*) nos da información sobre el grado de similitud estructural entre el modelo (Pol ϕ 29 mut2 en este caso) y la estructura “*template*”

(1XHX). Valores próximos a cero indican una gran similitud estructural. Los resultados para Pol ϕ 29 mut2 da un valor de 0,21 demostrando la semejanza entre la proteína mutante y la *wt*.

En la (Fig5) se muestra la superposición de la estructura predicha para Pol ϕ 29 mut2 y la estructura *wt* (1XHX). Como se observa, las dos estructuras están perfectamente superpuestas. El análisis de los aminoácidos mutados muestra como estos no desestabilizan las estructuras secundarias. En definitiva, las 8 mutaciones de Pol ϕ 29 mut2 no provocan cambios significativos en la estructura tridimensional de la polimerasa ϕ 29. Así pues, las 8 mutaciones encontradas es probable que no provoquen una pérdida en la funcionalidad de la enzima.

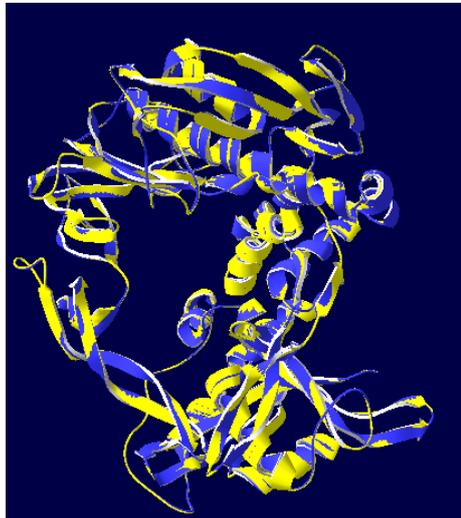


Fig. 5 Superposición de las dos estructuras: en amarillo Pol ϕ 29 mut2 y en azul *wt*. PROGRAMA: *Swiss-PdbViewer* [<http://www.expasy.org/spdbv/>; (Guex and Peitsch, 1997)].

En definitiva, las mutaciones finales en el mutante Pol ϕ 29 mut2 escogidas son: SER10-VAL10, GLY191-GLU191, SER194-ALA194, GLY197-LEU197, GLY217-PRO217, LYS305-PRO305, THR368-MET368 y ASP569-TYR569.

5.3) Aumento de la termoestabilidad de la polimerasa $\phi 29$ por adición de un puente disulfuro

Un puente disulfuro es un enlace covalente que se produce entre dos Cys cuando sus cadenas laterales están próximas. La introducción de un puente disulfuro extra en una proteína provoca un aumento de la estabilidad de forma significativa. (Tan *et al.*, 2014) aumentaron la termoestabilidad de la lipasa de *Penicillium cyclopium* reemplazando los residuos Val248 y Thr251 por cisteínas formándose un puente disulfuro. La vida media a 35°C de la lipasa mutante fue 12,8 mayor que la *wt*. En este apartado se estudiará si la introducción de un puente disulfuro en la polimerasa de estudio aumenta la termoestabilidad.

Para la predicción de los puentes disulfuro se utilizan dos programas que tienen el mismo objetivo pero que parten de diferente base teórica. De esta forma, aquellos puentes disulfuro que son predichos por ambos programas tienen una probabilidad mucho más alta de producirse (un total de 12 puentes disulfuro predichos). Para reducir las parejas de residuos candidatas se usó otro criterio El aumento de la termoestabilidad implica que la proteína mutante es más rígida (menos flexible) que la *wt*. Se sugirió que aquellos pares de residuos candidatos que fueran más flexibles aportarían una mayor estabilización de la proteína cuando se creara un puente disulfuro. Los factores de temperatura (*B-factor*) nos permiten analizar la flexibilidad de los pares de residuos. Los resultados se muestran en la Tabla 6.

Tabla 6. Se muestran las 12 dobles mutaciones predichas simultáneamente por los dos programas: MODIP [<http://caps.ncbs.res.in/iws/modip.html>]; (Sowdhamini *et al.*, 1989)] y Disulfide by Design [<http://cptweb.cpt.wayne.edu/DbD2/>]; (Dombkowski, Craig and Dombkowski, 2013)]. Además, se muestra entre paréntesis el valor del factor de temperatura (*B-factor*) para cada pareja de mutaciones posibles.

SER10 ALA174 (43,84)	LYS18 ASP21 (92,48)	MET30 TYR38 (38,08)	LEU63 ASP121 (55,44)	LEU253 ASP458 (50,02)	TYR281 LYS352 (48,72)
HIS284 LYS352 (44,84)	PRO300 TYR315 (56,65)	LEU328 LEU333 (49,97)	LYS371 ALA382 (65,38)	LYS479 TYR482 (78,63)	CYS530 MET533 (75,12)

Cuanto mayor sea el valor del *B-factor*, mayor es la flexibilidad de los residuos y es de esperar que el incremento en la estabilidad sea también mayor. Por eso, para continuar con el experimento se elegirán los cuatro pares de residuos con los *B-factor* más altos y uno con el *B-factor* bajo (control): Lys18-Asp21, Lys479-Tyr482,

Cys530-Met533, Lys371-Ala382 y Met30-Tyr38 (control). Se generan los dobles mutantes y se calculan las correspondientes energías (Tabla7).

Tabla 7 Energías de los cinco dobles mutantes y de la polimerasa *wt*

	ENERGÍA TOTAL (kcal/mol)
Polφ29 K18C-D21C	-366,56 kcal/mol
Polφ29 K479C-Y482C	-366,52 kcal/mol
Polφ29 C530-M533C	-366,53 kcal/mol
Polφ29 K371C-A382C	-366,57 kcal/mol
Polφ29 M30C-Y38C (control)	-366,50 kcal/mol
Polφ29-WT	-124,93 kcal/mol

Como se puede observar en la tabla 7 todos los dobles mutantes tienen un valor de energía mucho más negativo (3 veces más aprox) que la Polφ29-WT. Es decir, cualquier doble mutante con el puente disulfuro es más termoestable que la polimerasa *wt* confirmando así la hipótesis inicial. La localización de los nuevos puentes disulfuro muestra como los dos residuos del puente 18-21 forman parte de tiras-β antiparalelas que se unen mediante las nuevas cisteínas (Fig6). De igual forma ocurre con el puente 30-38 (*data not shown*). Las cisteínas del puente 371-382 unen dos hélices α (**Anexo3**). En el caso del puente 479-482 el primer residuo es una cisteína libre y el segundo forma parte de una tira-β (**Anexo3**). Se sugiere que la introducción extra de los puentes disulfuro estabilizan la unión entre las tiras-β y las hélices α lo que se traduce en un aumento en la termoestabilidad de la polimerasa. El puente 530-533 no forma parte de ninguna estructura secundaria.

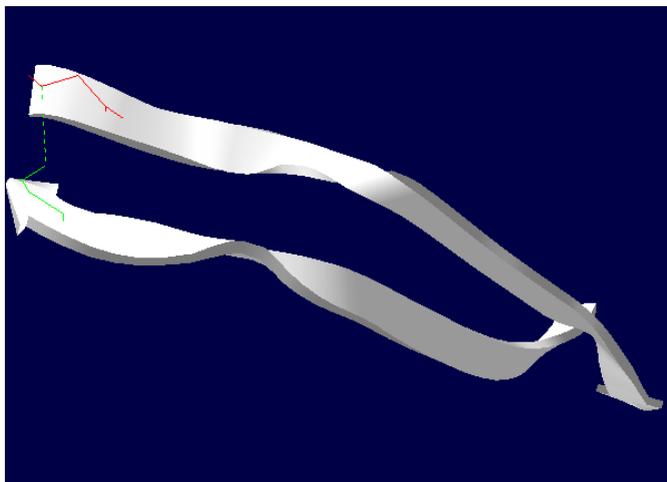


Fig. 6 Situación de las Cys del puente disulfuro (28-21). En rojo Cys21. En verde Cys18. El enlace que se forma se muestra con la línea verde discontinua. PROGRAMA: *Swiss-PdbViewer* [<http://www.expasy.org/spdbv/>; (Guex and Peitsch, 1997)].

Cabe esperar que cuanto más alejados en la estructura primaria estén los pares de residuos el incremento en estabilidad sea mayor. En este caso, todos los residuos que generan el puente disulfuro están muy próximos en la estructura primaria de la proteína (18-21, 479-482, 530-533, 371-382 y 30-38). En cambio, el incremento es significativo. Existen estudios en los que los residuos involucrados en un puente disulfuro están próximos en la estructura primaria de la proteína. (van Beek *et al.*, 2014) diseñaron un puente disulfuro L323C-A325C para aumentar la termoestabilidad de la ciclohexanona monooxigenasa. Así pues, los resultados son fiables.

Por otro lado, no hay diferencias entre los dobles mutantes y el control. Esto sugiere que en este caso, la flexibilidad de los residuos candidatos no influye mucho en el resultado final. El puente disulfuro por sí solo genera una gran estabilidad en la proteína aun cuando los residuos candidatos no son muy flexibles (control). Además, hay que tener en cuenta las mutaciones en los que uno de los dos residuos a mutar es ácido o base. El motivo es que podría formar parte de un puente salino que al producirse la mutación se rompería. Como ya se ha mencionado el puente salino es una interacción que estabiliza una proteína. En este caso, la Lys479 forma un

punto salino en la proteína nativa (*data not shown*) por lo que se descartaría el doble mutante Lys479-Tyr482 en futuros experimentos in vitro.

5.4) Combinación de ambas estrategias: Pol ϕ 29 mut2+puente disulfuro

En este trabajo, se han visto dos estrategias fiables para aumentar la termoestabilidad de la polimerasa ϕ 29. En este apartado se analizará si la combinación de ambas incrementa la termoestabilidad. Como puede verse en el **Anexo4** las diferencias de energía entre los mutantes y Pol ϕ 29 mut2 son positivas. Es decir, las energías de los mutantes son mayores y en consecuencia son proteínas más flexibles que no incrementan su termoestabilidad respecto Pol ϕ 29 mut2.

Estos resultados confirman que las dos estrategias conjuntas desestabilizan la proteína. Se sugiere que la introducción de los diferentes puentes disulfuro en la estructura de Pol ϕ 29 mut2 generaría impedimentos estéricos con los átomos cercanos que desestabilizaría la enzima. Por otro lado, si se comparan las dos estrategias, la energía total de los dobles mutantes es menor que la del mutante Pol ϕ 29 mut2 (-366 kcal/mol y -140,67 kcal/mol respectivamente). Es decir, los dobles mutantes son más estables y más rígidos y en consecuencia los que más han incrementado la termoestabilidad de la polimerasa. Un enlace covalente es difícil de romper y produce una gran rigidez en la estructura de una macromolécula. Es por esta razón que el carácter covalente del puente disulfuro adicional produce un mayor incremento en la termoestabilidad de la proteína que aquellas interacciones no covalentes (enlace más débil) que se producen en el “*alanine scanning*”.

En definitiva, el mutante Pol ϕ 29 K371C-A382C es el que se escoge en este trabajo. De todos los mutantes analizados es el que más aumenta la termoestabilidad (-366,57 kcal/mol).

de energía será positiva. Esta suposición coincide con los resultados (Tabla 8) excepto con los residuos THR97 y ASP197 (-0,55 kcal/mol y -2,24 kcal/mol respectivamente). Para estos residuos, el programa no predice un incremento en la estabilidad:

- El residuo THR97 se encuentra en el dominio exonucleasa de la proteína muy próximo al sustrato de ADN. Esta mutación provoca una mayor interacción entre el complejo proteína-ADN pero no incrementa la termoestabilidad respecto a la proteína wt (Povilaitis *et al.*, 2016). Por eso, los resultados del “alanine scanning” para este residuo no predicen un incremento en la termoestabilidad de la polimerasa.

- En el caso del residuo 197 la mutación G197D estabiliza la hélice α (Povilaitis *et al.*, 2016). Como se ha comentado anteriormente, la glicina es un aminoácido muy flexible que disminuye la estabilidad de la proteína (Pace and Scholtz, 1998). Por lo tanto, es previsible que la mutación de este residuo aumente la termoestabilidad. Sin embargo, los resultados del “alanine scanning” no detectan este incremento en la estabilidad (-2,2 kcal/mol). Hay que tener en cuenta que los resultados de los programas bioinformáticos son predicciones. Además, no son perfectos y son más fiables los resultados de un experimento *in vitro* que *in silico*.

Tabla 8 Resultados del “alanine scanning” de los siete residuos mutados que aumentan la termoestabilidad de la polimerasa en el artículo de (Povilaitis *et al.*, 2016). PROGRAMA: FoldX 3.0 Beta 5.1 (Schymkowitz *et al.*, 2005)

```
FoldX 3.0 Beta 5.1 (2011)
by the FoldX Consortium
Jesper Borg, Frederic Rousseau, Joost Schymkowitz,
Luis Serrano and Francois Stricher
```

```
-----
Output type: alascan
```

```
ARG 8 to ALA energy change is 1.62768 kcal/mol
ALA 51 to ALA energy change is 0 kcal/mol
THR 97 to ALA energy change is -0,555208 kcal/mol
ASP 197 to ALA energy change is -2,24392 kcal/mol
LYS 221 to ALA energy change is 0,616968 kcal/mol
PRO 497 to ALA energy change is 2,22078 kcal/mol
LEU 526 to ALA energy change is 1,29371 kcal/mol
```

Por último, se analizó como afectaba al entorno de la proteína las siete mutaciones descritas en el artículo. En el **Anexo 5** se muestran las interacciones de cada uno de los siete residuos mutados. De todos los rotámeros disponibles para cada residuo se ha escogido aquel que encaja mejor en la estructura de la proteína. Como se puede ver, todas las mutaciones del artículo son favorables y forman interacciones por puentes de hidrógeno. Estos enlaces estabilizarían la proteína lo que se traduce en un aumento en la termoestabilidad.

(Povilaitis *et al.*, 2016) sugieren que el aumento de la termoestabilidad de la polimerasa $\phi 29$ de su estudio se debe a diversos mecanismos moleculares: La sustitución de M8 por un residuo con carga positiva podría estabilizar el complejo proteína-ADN. La mutación V51A incrementaría la estabilidad de la hélice α . El cambio E221K introduciría un enlace iónico adicional entre este residuo y GLU272. Por otro lado, la sustitución Q497P incrementaría la rigidez del esqueleto de la proteína. Por último, el residuo F526 es un residuo hidrofóbico que se encuentra en la superficie de la proteína desestabilizando así la proteína.

Viendo estos resultados se demuestra que una forma para aumentar la termoestabilidad de la polimerasa consiste en incrementar el número de interacciones entre los residuos de la proteína, tal y como se había propuesto en el presente estudio.

6.CONCLUSIONES

Las conclusiones finales del presente estudio son las siguientes:

- ✓ Un aumento de las interacciones entre los residuos de la proteína provoca un incremento en la termoestabilidad de la polimerasa $\phi 29$.
- ✓ Las dos estrategias estudiadas han aumentado la termoestabilidad de la polimerasa: mediante “alanine scanning” se han desarrollado 8 mutaciones para Pol $\phi 29$ mut2 (SER10-VAL10, GLY191-GLU191, SER194-ALA194, GLY197-LEU197, GLY217-PRO217, LYS305-PRO305, THR368-MET368 y ASP569-TYR569). Por otro lado, se han obtenido cinco dobles mutantes con un puente disulfuro extra y que han aumentado de forma significativa la termoestabilidad de la polimerasa $\phi 29$.
- ✓ De todos los mutantes estudiados Pol $\phi 29$ K371C-A382C es el que más ha incrementado su termoestabilidad (-366,57 kcal/mol) y por lo tanto es el mutante final escogido.
- ✓ El análisis de ambas estrategias ha demostrado que la introducción de un puente disulfuro ha incrementado más la termoestabilidad de la polimerasa por su carácter de enlace covalente. Además, la combinación de las dos estrategias conjuntas desestabiliza la polimerasa. El análisis de la polimerasa termoestable de (Povilaitis *et al.*, 2016) muestra como las mutaciones están distribuidas por todos los dominios de la proteína y forman interacciones con los residuos de alrededor.

Este estudio cuenta con algunas limitaciones: Al ser un trabajo *in silico* puede que los resultados obtenidos no sean satisfactorios una vez se reproducen en el laboratorio. Hay que tener en cuenta que los programas bioinformáticos nos realizan predicciones fiables pero puede que no coincidan con los resultados *in vitro*. Por otro lado, habría que hacer un estudio más exhaustivo del efecto de las mutaciones en la estructura y la funcionalidad de la enzima.

7.AUTOEVALUACION

La realización del presente estudio ha sido muy gratificante. Por un lado, he entendido lo importante que es realizar una búsqueda bibliográfica previa a la realización de un proyecto de investigación. Además, a nivel teórico he podido profundizar en el campo de la ingeniería de proteínas que había estudiado durante el grado en Estructura y Función de Biomoléculas, Ingeniería de Proteínas y Bioinformática.

Durante el proceso experimental he utilizado una gran cantidad de recursos bioinformáticos que ya conocía: *Protein Data Bank* (PDB), *UniProt*, *SWISS-MODEL* etc. Por otro lado, también he aprendido nuevos programas como *DynaMut*. He sabido recopilar toda la información que los servidores daban para obtener los resultados de mi experimento. Sin embargo, he tenido ciertas dificultades en el uso y en la interpretación de los resultados de algunos programas bioinformáticos que he podido solucionar con éxito.

Este estudio pone de manifiesto la gran importancia de la bioinformática. Ahora, soy más consciente de la repercusión de este campo en un proyecto de investigación. Las herramientas computacionales sirven de guía para poder llegar al éxito en el laboratorio. En mi caso, este estudio *in silico* mostraría las mutaciones más idóneas para aumentar la termoestabilidad de la polimerasa $\phi 29$ en futuros estudios *in vitro*.

En definitiva, este estudio ha sido muy enriquecedor de cara a mi vida profesional como biotecnólogo ya que he elaborado y realizado mi propio proyecto de investigación. Además, he desarrollado aptitudes que creo que son necesarias para mi futuro como la capacidad crítica y el razonamiento científico.

8. BIBLIOGRAFIA

- Aschenbrenner, J. and Marx, A. (2017) 'DNA polymerases and biotechnological applications', *Current Opinion in Biotechnology*, 48, pp. 187–195. doi: 10.1016/j.copbio.2017.04.005.
- Bateman, A. *et al.* (2017) 'UniProt: the universal protein knowledgebase', *Nucleic Acids Research*. Oxford University Press, 45(D1), pp. D158–D169. doi: 10.1093/nar/gkw1099.
- van Beek, H. L. *et al.* (2014) 'Stabilization of cyclohexanone monooxygenase by a computationally designed disulfide bond spanning only one residue', *FEBS Open Bio*, 4(1), pp. 168–174. doi: 10.1016/j.fob.2014.01.009.
- Berman, H. M. *et al.* (2000) 'The Protein Data Bank', *Nucleic Acids Research*. Oxford University Press, 28(1), pp. 235–242. doi: 10.1093/nar/28.1.235.
- Blasco, M. A. *et al.* (1993) 'Phi 29 DNA polymerase active site. Residue ASP249 of conserved amino acid motif "Dx2SLYP" is critical for synthetic activities.', *The Journal of biological chemistry*, 268(32), pp. 24106–13. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8226957> (Accessed: 26 May 2018).
- Cocolin, L. *et al.* (2011) 'The challenge of merging food safety diagnostic needs with quantitative PCR platforms', *Trends in Food Science & Technology*. Elsevier, 22, pp. S30–S38. doi: 10.1016/J.TIFS.2011.02.009.
- Dombkowski, C., Craig, D. B. and Dombkowski, A. A. (2013) 'Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins', 14. Available at: <http://www.biomedcentral.com/1471-2105/14/346> (Accessed: 10 May 2018).
- Farrell, R. E. and Farrell, R. E. (2010) 'RT-PCR', in *RNA Methodologies*. 4th Edition. Elsevier, pp. 391–395. doi: 10.1016/B978-0-12-374727-3.00018-8.
- Gadkar, V. and Rillig, M. C. (2005) 'Application of Phi29 DNA polymerase mediated whole genome amplification on single spores of arbuscular mycorrhizal (AM) fungi', *FEMS Microbiology Letters*, 242(1), pp. 65–71. doi: 10.1016/j.femsle.2004.10.041.
- Gallivan, J. P. and Dougherty, D. A. (1999) 'Cation-pi interactions in structural biology.', *Proceedings of the National Academy of Sciences of the United States of*

America. National Academy of Sciences, 96(17), pp. 9459–64. doi: 10.1073/PNAS.96.17.9459.

Guex, N. and Peitsch, M. C. (1997) 'SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling', *Electrophoresis*, 18(15), pp. 2714–2723. doi: 10.1002/elps.1150181505.

Hekkelman, M. L. *et al.* (2010) 'WIWS: a protein structure bioinformatics Web service collection.', *Nucleic acids research*, 38(Web Server issue), pp. W719-23. doi: 10.1093/nar/gkq453.

Jmol: un visor Java de código abierto para estructuras químicas en tres dimensiones (no date). Available at: <http://jmol.sourceforge.net/> (Accessed: 17 April 2018).

Johansson, E. and Dixon, N. (2013) 'Replicative DNA Polymerases', *Cold Spring Harbor Perspectives in Biology*, 5(6), pp. a012799–a012799. doi: 10.1101/cshperspect.a012799.

Joosten, R. P. *et al.* (2014) 'The PDB_REDO server for macromolecular structure model optimization.', *IUCrJ*, 1(Pt 4), pp. 213–20. doi: 10.1107/S2052252514009324.

Kamtekar, S. *et al.* (2004) 'Insights into Strand Displacement and Processivity from the Crystal Structure of the Protein-Primed DNA Polymerase of Bacteriophage phi29', *Mol.Cell*, 16, pp. 609–618. doi: 10.2210/PDB1XHX/PDB.

Kim, K. P. *et al.* (2011) 'Improved thermostability and PCR efficiency of *Thermococcus celericrescens* DNA polymerase via site-directed mutagenesis', *Journal of Biotechnology*, 155(2), pp. 156–163. doi: 10.1016/j.jbiotec.2011.06.022.

Kim, Y. *et al.* (1995) 'Crystal structure of *Thermus aquaticus* DNA polymerase.', *Nature*, 376, pp. 612–616. doi: 10.2210/PDB1TAQ/PDB.

Le, Q. A. T. *et al.* (2012) 'Development of thermostable *Candida antarctica* lipase B through novel in silico design of disulfide bridge', *Biotechnology and Bioengineering*, 109(4), pp. 867–876. doi: 10.1002/bit.24371.

Lovmar, L. and Syvänen, A.-C. (2006) 'Multiple displacement amplification to create a long-lasting source of DNA for genetic studies', *Human Mutation*, 27(7), pp. 603–614. doi: 10.1002/humu.20341.

- Nagy, B. (2015) 'Application of real-time polymerase chain reaction in the clinical genetic practice', *Journal of Pediatric Genetics*, 02(01), pp. 001–008. doi: 10.3233/PGE-13042.
- Pace, C. N. and Scholtz, J. M. (1998) 'A helix propensity scale based on experimental studies of peptides and proteins.', *Biophysical journal*. The Biophysical Society, 75(1), pp. 422–7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9649402> (Accessed: 21 May 2018).
- Pezeshgi Modarres, H. *et al.* (2015) 'Understanding and Engineering Thermostability in DNA Ligase from *Thermococcus* sp. 1519', *Biochemistry*, 54(19), pp. 3076–3085. doi: 10.1021/bi501227b.
- polA - DNA polymerase I, thermostable - Thermus aquaticus - polA gene & protein* (no date). Available at: <http://www.uniprot.org/uniprot/P19821> (Accessed: 17 March 2018).
- Povilaitis, T. *et al.* (2016) 'In vitro evolution of phi29 DNA polymerase using isothermal compartmentalized self replication technique', *Protein Engineering Design and Selection*, 29(12), pp. 617–628. doi: 10.1093/protein/gzw052.
- Rodrigues, C. H., Pires, D. E. and Ascher, D. B. (2018) 'DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability', *Nucleic Acids Research*. doi: 10.1093/nar/gky300.
- Schymkowitz, J. *et al.* (2005) 'The FoldX web server: an online force field', *Nucleic Acids Research*, 33(Web Server), pp. W382–W388. doi: 10.1093/nar/gki387.
- Sowdhamini, R. *et al.* (1989) 'Stereochemical modeling of disulfide bridges. Criteria for introduction into proteins by site-directed mutagenesis.', *Protein engineering*, 3(2), pp. 95–103. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/2594728> (Accessed: 10 May 2018).
- Takahashi, H. *et al.* (2014) 'Preparation of Phi29 DNA polymerase free of amplifiable DNA using ethidium monoazide, an ultraviolet-free light-emitting diode lamp and trehalose.', *PloS one*. Public Library of Science, 9(2), p. e82624. doi: 10.1371/journal.pone.0082624.

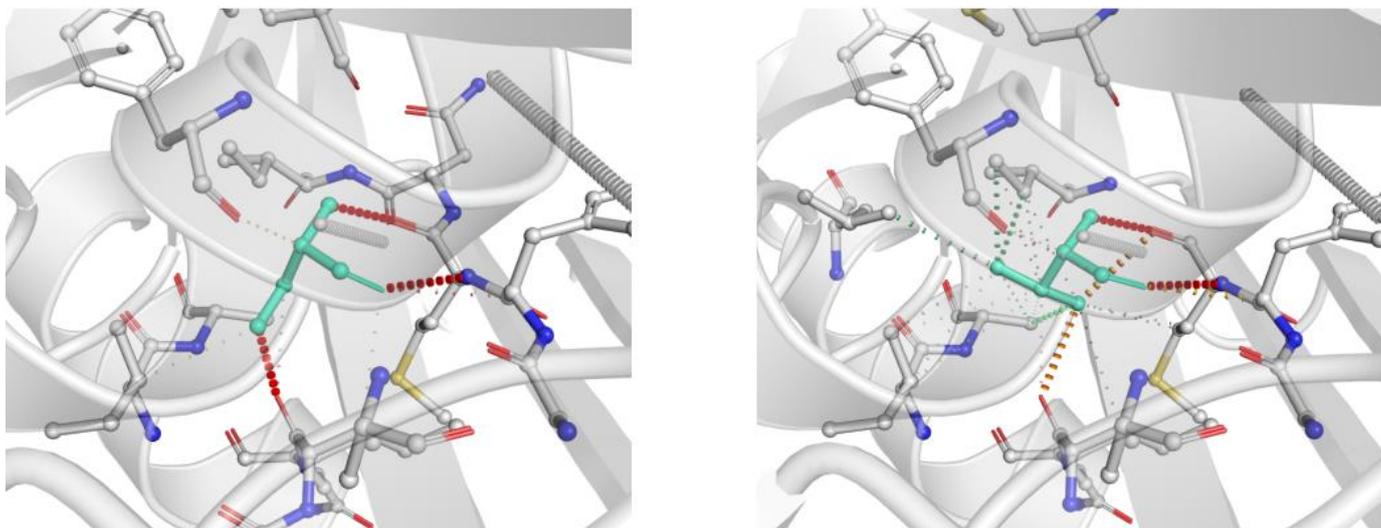
Tan, Z. *et al.* (2014) 'Enhancing the Thermostability of a Cold-Active Lipase from *Penicillium cyclopium* by In Silico Design of a Disulfide Bridge', *Applied Biochemistry and Biotechnology*, 173(7), pp. 1752–1764. doi: 10.1007/s12010-014-0962-7.

Terpe, K. (2013) 'Overview of thermostable DNA polymerases for classical PCR applications: from molecular and biochemical fundamentals to commercial systems', *Applied Microbiology and Biotechnology*, 97(24), pp. 10243–10254. doi: 10.1007/s00253-013-5290-2.

Waterhouse, A. *et al.* (2018) 'SWISS-MODEL: homology modelling of protein structures and complexes', *Nucleic Acids Research*. doi: 10.1093/nar/gky427.

9.ANEXOS

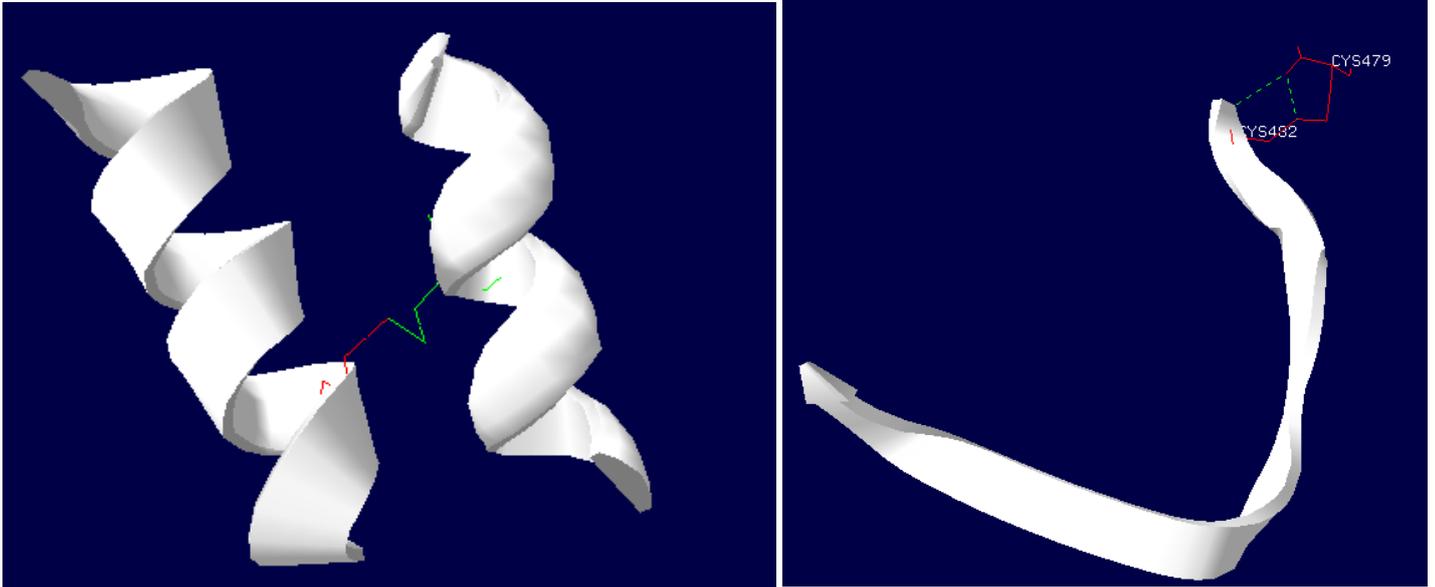
Anexo 1. Interacciones en el residuo 10 (pintado de azul claro). Izquierda: *wild type* (SER10). Derecha : mutante (VAL10). Las líneas discontinuas de colores indican diferentes interacciones entre el residuo y los aminoácidos de su alrededor.



Anexo 2. Análisis del efecto de las 8 mutaciones en Polφ29 mut2. PROGRAMA: DynaMut (Rodrigues, Pires and Ascher, 2018).

SER10-VAL10	$\Delta\Delta G$: 1.724 kcal/mol (Stabilizing)
ARG23-PHE23	$\Delta\Delta G$: -0.344 kcal/mol (Destabilizing)
GLY191-GLU191	$\Delta\Delta G$: 1.135 kcal/mol (Stabilizing)
SER194-ALA194	$\Delta\Delta G$: 1.267 kcal/mol (Stabilizing)
GLY197-LEU197	$\Delta\Delta G$: 2.392 kcal/mol (Stabilizing)
GLY217-PRO217	$\Delta\Delta G$: 0.357 kcal/mol (Stabilizing)
ARG236-ALA236	$\Delta\Delta G$: -0.667 kcal/mol (Destabilizing)
LYS305-PRO305	$\Delta\Delta G$: 0.382 kcal/mol (Stabilizing)
THR368-MET368	$\Delta\Delta G$: 2.036 kcal/mol (Stabilizing)
ASP569-TYR569	$\Delta\Delta G$: 0.835 kcal/mol (Stabilizing)

Anexo 3 Situación de los puentes disulfuro. El enlace se muestra en líneas discontinuas. Izquierda: CYS371 (rojo)-CYS382(verde). Derecha: CYS479-CYS482. PROGRAMA: *Swiss-PdbViewer* [<http://www.expasy.org/spdbv/>; (Guex and Peitsch, 1997)]

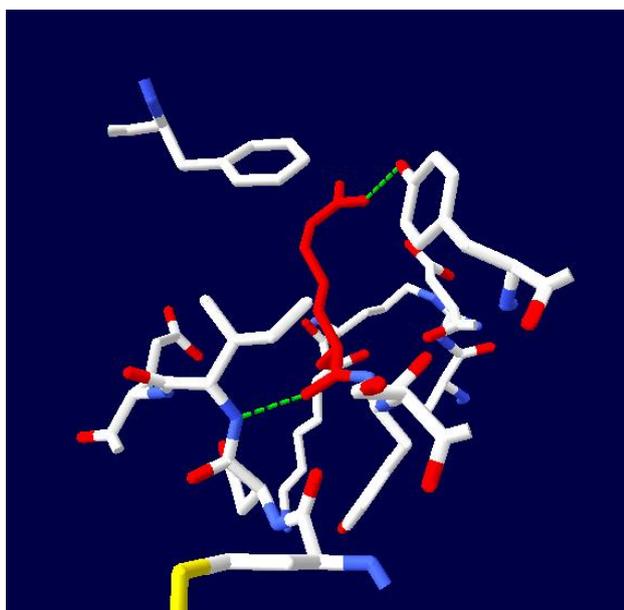


Anexo 4 Análisis del efecto de las dos estrategias conjuntas. Se muestra la diferencia de energía de los mutantes respecto a Pol ϕ 29mut2. PROGRAMA: *FoldX 3.0 Beta 5.1* (Schymkowitz *et al.*, 2005)

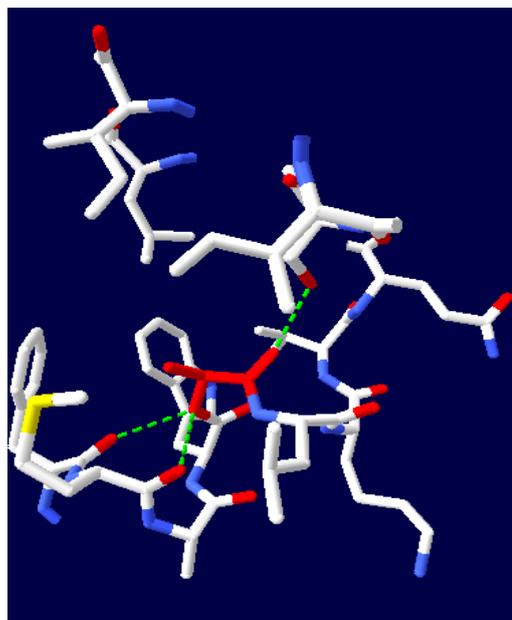
	DIFERENCIA ENERGIA Pol ϕ 29 mut2
Pol ϕ29 mut2 K18C-D21C	0,06 kcal/mol
Pol ϕ29 mut2 C530-M533C	3,41 kcal/mol
Pol ϕ29 mut2 K371C-A382C	1,45 kcal/mol
Pol ϕ29 mut2 M30C-Y38C	4,67 kcal/mol

Anexo 5. Interacciones por puentes de hidrógeno (líneas verdes puenteadas) que se producen para cada una de las siete mutaciones (residuo en rojo) del artículo de (Povilaitis *et al.*, 2016) PROGRAMA: *Swiss-PdbViewer* [<http://www.expasy.org/spdbv/>; (Guex and Peitsch, 1997)]

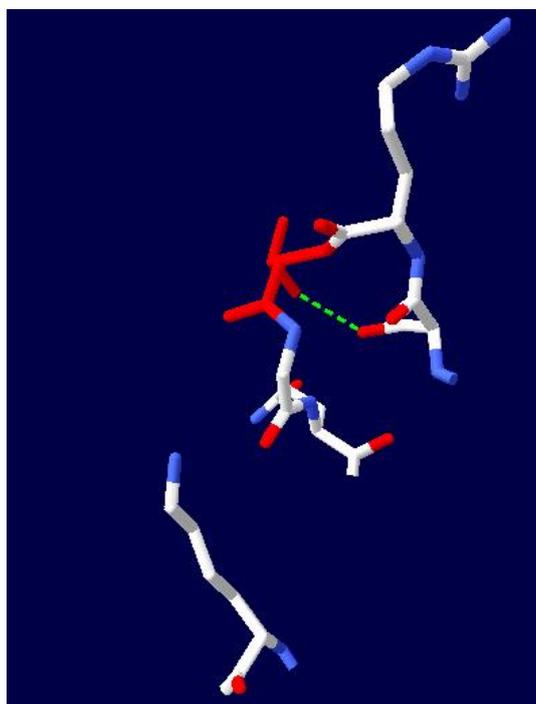
M8R



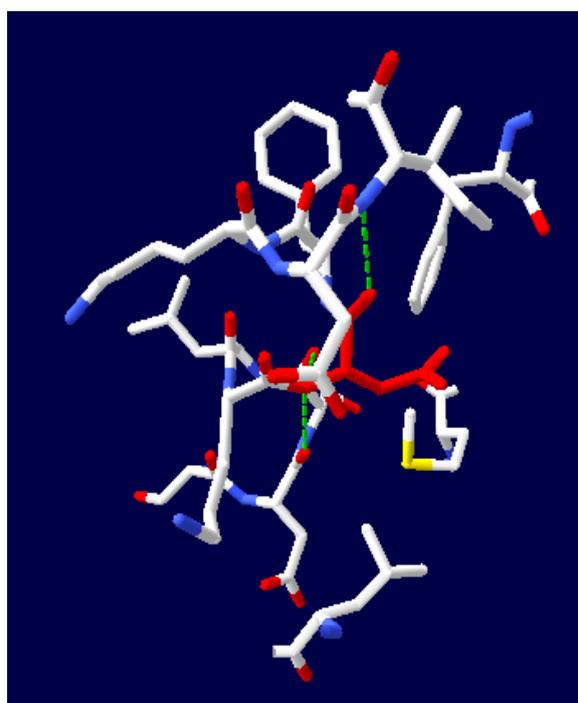
V51A



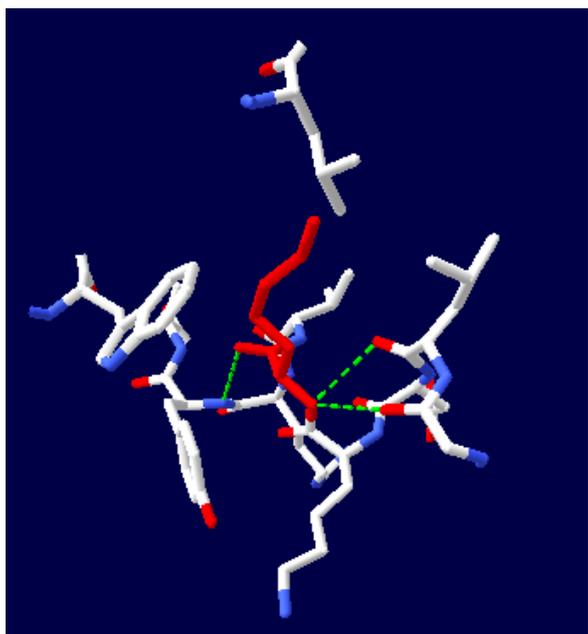
M97T



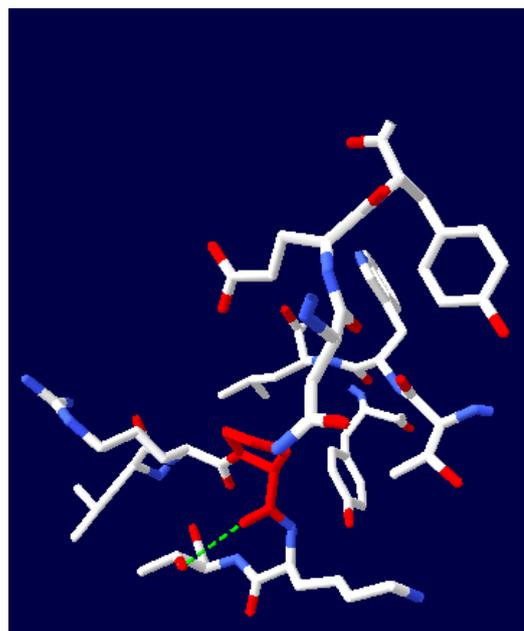
G197D



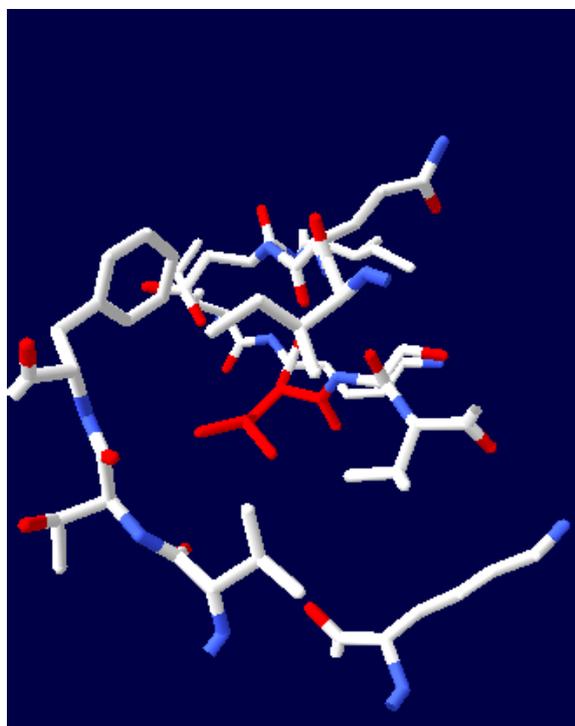
E221K



Q497P



F526L



ANNEX 2

FITXA DE SEGUIMENT DEL TUTOR/A del TFG

Nom i Cognoms de l'Alumne/a: Alejandro del Río García

Nom i Cognoms del Tutor/a: Gerard Fijadas Amiguano

Data de la entrevista amb l'alumne: 10/04/2018

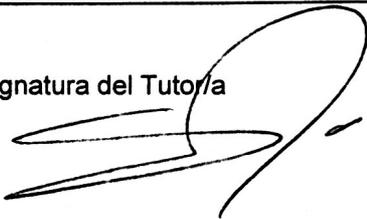
Recomanacions durant el seguiment: Ensenyar el treball com un article científic. En la part, posar un apartat d'abstract i les hipòtesis i objectius posar-los en un apartat pròpiament al final de la introducció Posar Materials i mètodes al final de l'article.

Observacions: _____

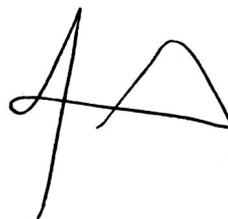
Observacions Darrera revisió:

És la primera revisió

Signatura del Tutor/a



Signatura del Alumne/a



Tarragona a 10 de abril 2018

ANNEX 2

FITXA DE SEGUIMENT DEL TUTOR/A del TFG

Nom i Cognoms de l'Alumne/a: Aljondro del Río García

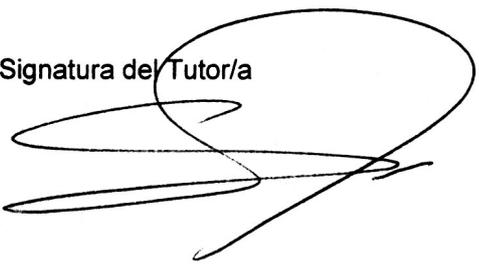
Nom i Cognoms del Tutor/a: GERARDO PUJADAS ANGUIANO

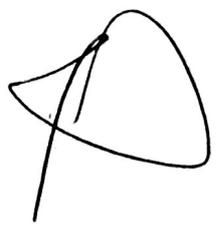
Data de la entrevista amb l'alumne: 30- MAIG- 2018

Recomanacions durant el seguiment: modificar la forma de referenciar els programes usats. No la seta commemorar en el text materials y métodos A Amis, posar resultats importants que no lo posen a la memoria escrita per que són molt extensos

Observacions: Resultats inicials positius en l'experiment

Observacions Darrera revisió:

Signatura del Tutor/a 

Signatura del Alumne/a 

Tarragona a 10 de maig 2018

ANNEX 2

FITXA DE SEGUIMENT DEL TUTOR/A del TFG

Nom i Cognoms de l'Alumne/a: Alexandro del Río García

Nom i Cognoms del Tutor/a: Gerard Pujadas Amarguero

Data de la entrevista amb l'alumne: 28/05/2018

Recomanacions durant el seguiment: Anàlisis per disjuntiu dels resultats dels ponts disulfur. Puxar bé per donar a les conclusions perquè estratègia escollida per augmentar l'estabilitat i les mutacions escollides.

Observacions: Especificar l'angle de torçió de X_3 en el programa DBD.

Observacions Darrera revisió:

Bé referenciat ~~amb~~ programes usats

Signatura del Tutor/a



Signatura del Alumne/a



Tarragona a 28 de maig 2018