

Universidad de Almería

Escuela politécnica superior y Facultad de Ciencias Experimentales

Área de Química Física

Departamento de Química y Física



Cribado de las condiciones de cristalización de la Ubiquitina in silico y en el laboratorio

Marina Plaza Garrido
Julio 2016



MÁSTER INTERUNIVERSITARIO EN QUÍMICA

MEMORIA DEL TRABAJO FIN DE MÁSTER

CURSO ACADÉMICO:

CONVOCATORIA: 1

2

UNIVERSIDAD:

TÍTULO DEL TRABAJO:

Cribado de las condiciones de cristalización de la Ubiquitina in silico y en el laboratorio

ESTUDIANTE:

Marina Plaza Garrido

TUTOR(es):

Dra. Ana María Cámara Artigas

Fecha: Almería, 1 de Julio de 2016

VºBº Tutor(es)

Fdo:

El Estudiante

Fdo:

Resumen/Abstract

En este Trabajo fin de Máster se ha trabajado con la proteína ubiquitina. La ubiquitina es una proteína globular pequeña que tiene un número considerable de aminoácidos de lisina en su superficie, lo que provoca que exista una alta entropía superficial que impide la formación de las interacciones necesarias para formar un cristal. Hasta la fecha, sólo se han resuelto unas pocas estructuras de la forma nativa de la ubiquitina y la mayoría de ellas fueron determinadas a partir de cristales obtenidos en presencia de cationes metálicos divalentes. En este trabajo se ha determinado una nueva estructura de la ubiquitina en presencia del catión magnesio. Los cristales pertenecen al grupo espacial triclinico ($a=29,96 \text{ \AA}$, $b=30,18 \text{ \AA}$, $c=41,41 \text{ \AA}$, $\alpha=88,52^\circ$, $\beta=79,12^\circ$, $\gamma=67,37^\circ$). También se han analizado las bases de datos existentes que tratan de predecir si una proteína es o no cristalizable, aplicándolas tanto a la ubiquitina nativa como a sus diferentes mutaciones de los aminoácidos de lisinas por serinas. De esta manera, se ha realizado un cribado de las condiciones de cristalización de la ubiquitina tanto de forma computacional, como de forma experimental.

In this Master's Thesis we have worked with the ubiquitin protein. Ubiquitin is a small globular protein that has a considerable number of lysine residues on its surface. The result of this is a high surface entropy that precludes the formation of crystal-packing interactions. Nowadays, only a few structures of the native form of ubiquitin has been solved, and the most of the crystals that led to these structures were obtained in the presence of different divalent metal cations. In this work, a new crystallographic structure of human ubiquitin has been solved from crystals that grown in the presence of the magnesium cation. The crystals belonged to a triclinic space group ($a=29,96 \text{ \AA}$, $b=30,18 \text{ \AA}$, $c=41,41 \text{ \AA}$, $\alpha=88,52^\circ$, $\beta=79,12^\circ$, $\gamma=67,37^\circ$). The exiting databases that try to predict if a protein is crystallizable or non crystallizable have been analyzed. The databases have been applied to the native ubiquitin and to the lysine-serine mutants. Thus, it has been performed a screening of the crystallization conditions of the ubiquitin both computationally, as experimentally.

Contenido

Resumen/Abstract.....	5
1. Introducción.....	9
1.1 Ubiquitina	9
1.1.1 Estructura y funciones de la ubiquitina.....	9
1.1.2 Estructuras cristalinas.....	12
1.2 Cristalización de proteínas.....	14
1.2.1 Fundamento.....	15
1.2.2 Técnicas de cristalización.....	17
1.3 Cristalografía computacional.....	19
2. Objetivos.....	23
4. Materiales y métodos.....	25
4.1 Materiales.....	25
4.2 Métodos	27
4.2.1 Clonación, expresión y purificación de la ubiquitina.....	27
4.2.2 Electroforesis en SDS-PAGE.....	31
4.2.3 Cristalización de proteínas	33
4.2.4 Capturado de las imágenes de los cristales	34
4.2.5 Difracción de los cristales, procesado, resolución de la estructura y refinamiento	35
5. Resultados y discusión	37
5.1 Purificación de la ubiquitina	37
5.2 Cristalización de la ubiquitina	38
5.3 Estructura de la ubiquitina	43
5.4 Análisis de las bases de datos	47
6. Conclusiones/Conclusions.....	57
7. Bibliografía.....	59

1. Introducción

1.1 Ubiquitina

1.1.1 Estructura y funciones de la ubiquitina

La ubiquitina (figura 1) es una pequeña proteína globular de 76 aminoácidos y un peso molecular de 8565 Da. Su secuencia polipeptídica está altamente conservada y además se encuentra en la mayoría de los tejidos de los organismos eucariotas. De su ubicuidad procede su nombre y es una proteína que tiene un papel clave en algunos procesos que tienen lugar en el interior de la célula. Por ello, es una proteína altamente estudiada.¹ Su función principal es dirigir el reciclaje de proteínas,² ya que puede asociarse a las mismas y marcarlas para su posterior destrucción en el proteosoma mediante el proceso de ubiquitinación.¹

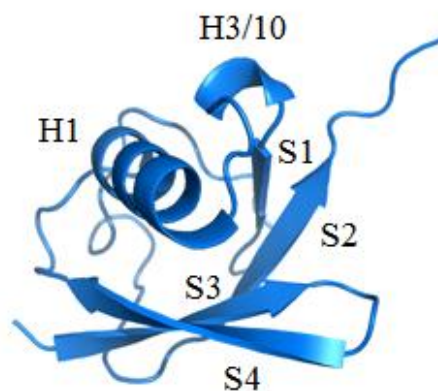


Figura 1. Secuencia correspondiente a la estructura de la ubiquitina. Se puede observar la hélice α (H3), la hélice 3_{10} , y las cuatro hebras β (S1, S2, S3, S4).

Esta proteína realiza una modificación covalente en proteínas diana, por medio de monoubiquitinación o poliubiquitinación, juega un papel clave en diversos eventos en el interior de la célula, entre los que destacan: el control de la progresión del ciclo celular,² la traducción de señales,³ el tráfico de proteínas,⁴ la degradación transcripcional,⁵ la endocitosis,⁶ etc. Debido al gran número de procesos en los que interviene, la ubiquitina es capaz de ser reconocida por una gran variedad de proteínas y dominios.⁷

Pero la función que más destaca de las mencionadas anteriormente es la de reciclaje. Es precisamente esta función la que es secuestrada por virus como VIH y Ébola para desarrollarse dentro del organismo. En este último, la proteína matriz Vp40 se adhiere a

la proteína TSG101-UEV, secuestrando las funciones de la ubiquitina a través de ésta (figura 2). También está íntimamente relacionada con otras enfermedades como el cáncer.

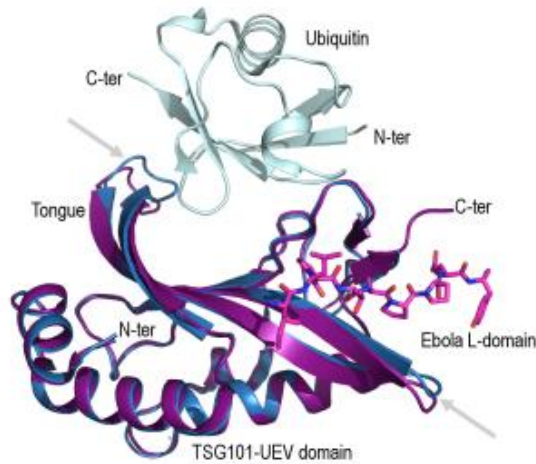


Figura 2. Superposición de las estructuras cristalográficas de la TSG101-UEV (púrpura), el dominio-L con la secuencia PTAP de la proteína matriz del Ébola VP40 (magenta) y la proteína Ubiquitina (azul cian).

La ubiquitinización es un proceso enzimático de modificación proteica post-traducciona l en el cual el ácido carboxílico de la glicina terminal (que se encuentra en el motivo de di-glicina de la ubiquitina activada) forma un enlace amida con el grupo amino épsilon de la lisina en la proteína modificada.¹ El proceso de marcar una proteína con ubiquitina consta de una serie de pasos que se describen a continuación:

1. Activación de la ubiquitina: la ubiquitina es activada en una reacción de dos pasos por una enzima activadora de ubiquitina en un proceso que requiere ATP como fuente de energía.
2. El paso inicial implica la producción de un producto intermedio adenilil-ubiquitina. El segundo paso consiste en transferir la ubiquitina al centro activo del E1, concretamente uniéndose al residuo de cisteína, con la liberación de AMP. Este paso se realiza mediante un acoplamiento tioéster entre el carbono terminal del grupo carboxilo de la ubiquitina y el grupo sulfhidrido o tiol (-SH) de la cisteína del E1.
3. Transferencia de la ubiquitina del centro activo del E1 a una enzima E2 de conjugación de ubiquitina a través de una reacción de transtioesterificación.

4. El paso final de la cascada de ubiquitinización crea un enlace isopeptídico entre una lisina de la proteína diana y la glicina del carbono terminal de la ubiquitina. En general, este paso requiere la actividad de una de los cientos de enzimas ligasas E3 de ubiquitina (a menudo llamada simplemente ligasa de ubiquitina). Las enzimas E3 funcionan como los centros de reconocimiento del sustrato del sistema y son capaces de interactuar tanto con el E2 como con el sustrato.

En la cascada de ubiquitinización, la enzima E1 puede unirse con docenas de enzimas E2, que pueden a su vez unirse con unos cientos de enzimas E3 de un modo jerárquico. En la figura 3 se muestra un esquema del proceso de ubiquitinización.

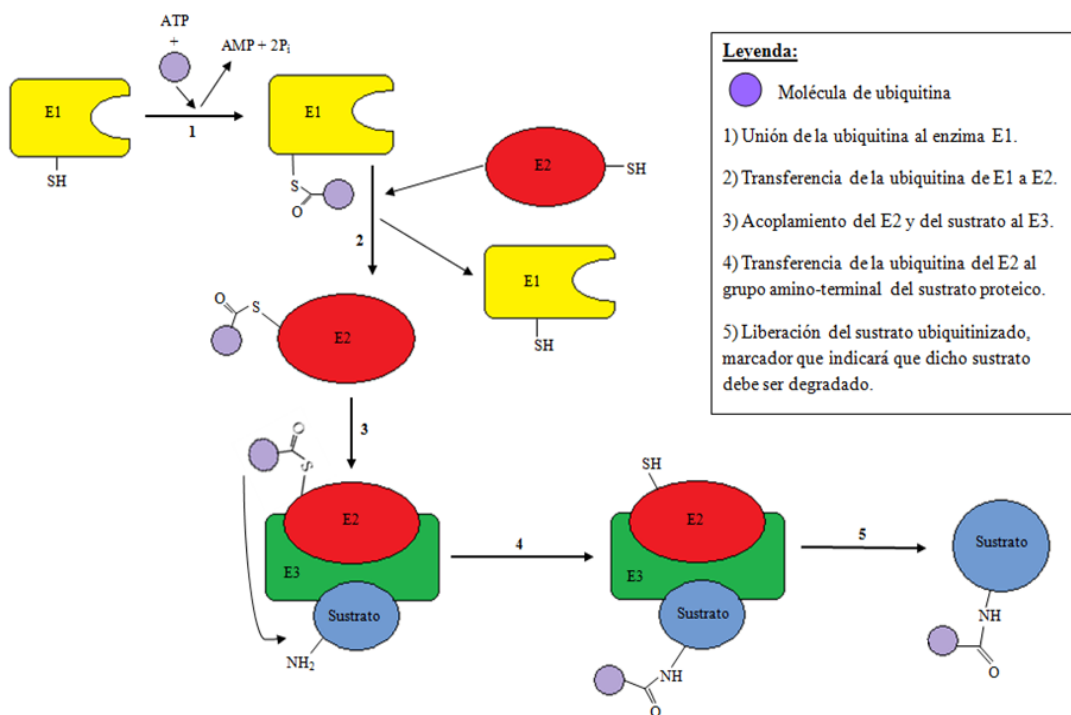


Figura 3. Esquema del proceso de ubiquitinización.

Una vez marcada la proteína que debe ser destruida, ésta es reconocida por el proteosoma. El proteosoma es un complejo proteico presente en todas las células eucariotas que se encarga de realizar la degradación de proteínas no necesarias o dañadas. Suele estar presente en el núcleo de las células y su citoplasma. Gracias a él, las células controlan la concentración de determinadas proteínas mediante la degradación de las mismas en un proceso llamado proteólisis. La proteína mono o poliubiquitinada permite al proteosoma identificar que ésta debe ser destruida.

1.1.2 Estructuras cristalinas

Actualmente existen 34 estructuras cristalográficas de la ubiquitina aislada depositadas en el Protein Data Bank (PDB).⁸ 16 de estas estructuras se tratan de ubiquitina humana wild-type (ubiquitina WT) y el resto corresponden a mutantes de la misma. Cabe destacar, que 11 de estas estructuras fueron obtenidas en presencia de cationes metálicos divalentes como Cd^{2+} (código PDB 3EEC), Hg^{2+} (código PDB 3EFU), Pt^{2+} (código PDB 3N32) o Zn^{2+} (códigos PDB 3N30, 4K7S, 4K7U, 4K7W, 4XOK y 4XOL).^{9,10,11} Estos metales se localizan principalmente en los contactos cristalinos, por lo que favorecen el crecimiento del cristal. En la figura 4 se muestran los contactos cristalinos mediados por un metal en algunas de las estructuras resueltas de ubiquitina.¹²

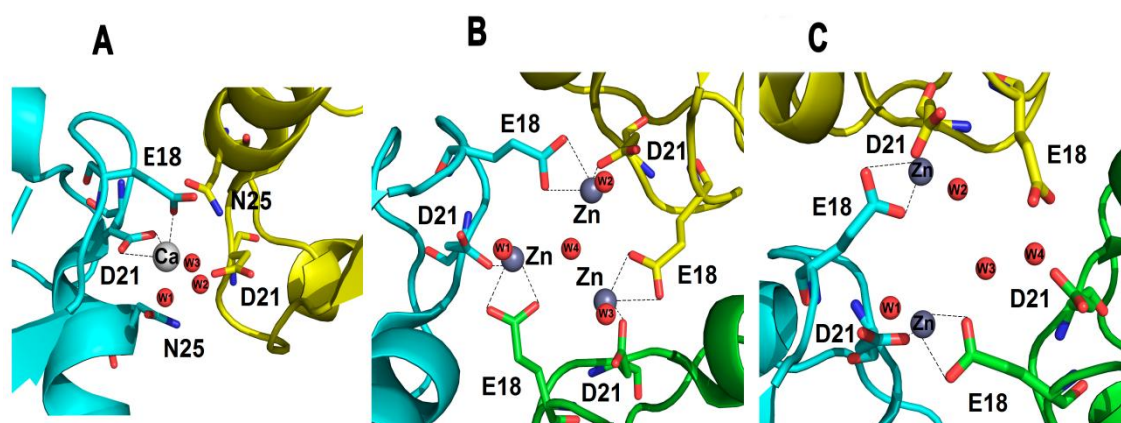


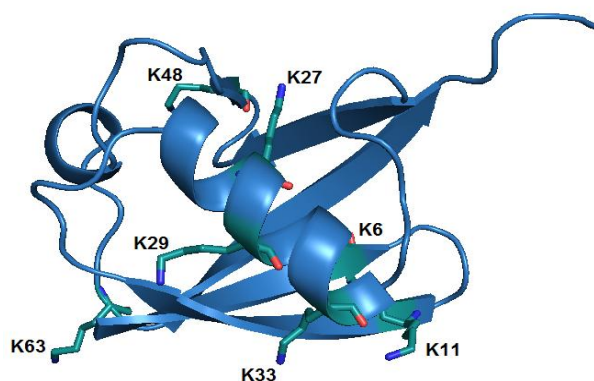
Figura 4. Empaquetamiento cristalino de la ubiquitina humana en presencia de (A) Ca^{2+} (código PDB 4pjh) y (B, C) Zn^{2+} (código PDB 4xol y 4xok). Las moléculas de ubiquitina en una posición especial de simetría se muestran en color azul, amarillo y verde. H_2O , Mg^{2+} , Ca^{2+} y Zn^{2+} se representan como esferas en color rojo, verde, blanco y gris, respectivamente. Los enlaces de hidrógeno se muestran con líneas negras discontinuas.

En la tabla 1 se muestran las estructuras de la ubiquitina disponibles en el Protein Data Bank.

Tabla 1. Estructuras de la ubiquitina en el PDB

Código PDB	Condiciones de cristalización	Grupo espacial/celda unidad	Referencia
5DK8	0.2 M cloruro de magnesio hexahidratado, 30% w/v PEG 4000, 0.1 M Tris pH 8.5, 288 K	P 1	12
4XOF	25% w/v PEG 1450, 50 mM HEPES pH 7.0, 293 K	P 2 ₁ 2 ₁ 2 ₁	11
4XOK	20 mM acetato amónico pH 4.3, 35% w/v PEG 1500, 25 mM acetato de zinc, 50 mM HEPES pH 7.0, 293 K	P 2 ₁ 2 ₁ 2 ₁	11
4XOL	20% w/v PEG 3350, 100 mM acetato de zinc, 100 mM MES pH 6.33, 293 K	P 4 ₃ 3 2	11
4K7S	25% w/v PEG 1450, 35 mM acetato de zinc, 50 mM HEPES pH 7.0, 293 K	P 2 ₁ 2 ₁ 2 ₁	Pendiente de publicación
4K7U	25% w/v PEG 1450, 70 mM acetato de zinc, 50 mM HEPES pH 7.0, 293 K	P 2 ₁ 2 ₁ 2 ₁	Pendiente de publicación
4K7W	25% w/v PEG 1450, 100 mM zinc acetate, 50 mM HEPES pH 7.0, 293 K	P 2 ₁ 2 ₁ 2 ₁	Pendiente de publicación
3ONS	50-56 % MPD, 8-18 % w/v glicerol, 27mM citrato sódico pH 4.0-4.2, 277 K	P 3 ₂ 2 1	13
3N30	25% w/v PEG 1450, 200 mM acetate de zinc, 50 mM HEPES pH 7, 293 K	P 4 ₃ 3 2	10
3N32	30% w/v PEG 1450, 0.1 M HEPES pH 7, 293K	P 2 ₁ 2 ₁ 2 ₁	10
3HM3	5 % w/v PEG 3000, 50 mM acetato de zinc, 0.1 M Sodium acetate pH 4.6, 293 K	P 2 ₁ 3	14
3EHV	25 % w/v PEG 1450, 25 mM acetato de zinc, 50 mM HEPES pH 7.0, 293 K	P 2 ₁ 2 ₁ 2 ₁	9
3EEC	25 % w/v PEG 1450, 200 mM acetato de cadmio, 50 mM HEPES pH 7.0, 298 K	P 4 ₃ 3 2	9
3EFU	25 % w/v PEG 1450, 2.92 mM acetato de mercurio, 50 mM HEPES pH 7.0, 293 K	P 2 ₁ 2 ₁ 2 ₁	9
1UBI	30 % w/v PEG 4000, 50 mM cacodilato/HCl pH 5.6	P 2 ₁ 2 ₁ 2 ₁	15
1UBQ	60 % w/v MPD, 50 mM HEPES pH 7.5	P 2 ₁ 2 ₁ 2 ₁	16

La dificultad para obtener cristales de ubiquitina ha sido atribuida a la presencia de siete residuos lisina mostrados en la figura 5 que se encuentran localizados en la superficie de la proteína (9,2 % de los aminoácidos de la ubiquitina). Estos residuos incrementan la entropía superficial, impidiendo las interacciones necesarias para empaquetar la estructura y dificultando por tanto la cristalización. Debido a esto, se han llevado a cabo una serie de experimentos cristalográficos mutando cada una de las lisinas por serinas.¹⁷ Estos experimentos han revelado que algunas de las mutaciones mejoran el proceso de cristalización, mientras que otras sin embargo, lo dificultan. Cabe destacar que uno de estos residuos de lisina (Lys29) está presente en el contacto cristalino en el que se encuentran los cationes metálicos divalentes.



MQIFV**K**TLTG**K**TITLEVEPSDTIENV**KAKIQDK**EGIPPDQQLIFAG**K**QLEDGRTLSDYNIQ**K**E
STLHLVLRGG

Figura 5. Estructura y secuencia de la ubiquitina. En color verde aguamarina se muestran los siete residuos de lisina. El aminoácido lisina se muestra con la letra K.

En este Trabajo Fin de Máster se describe una nueva estructura cristalina de la ubiquitina WT monomérica en el espacio triclínico P1 ($a=29,96 \text{ \AA}$, $b=30,18 \text{ \AA}$, $c=41,41 \text{ \AA}$, $\alpha=88,52^\circ$, $\beta=79,12^\circ$, $\gamma=67,37^\circ$)¹² cristalizada con cloruro de magnesio. El catión divalente Mg^{2+} se coordinada a seis moléculas de agua y facilitan la cristalización al favorecer los contactos cristalinos y el empaquetamiento.

1.2 Cristalización de proteínas

La cristalización de proteínas es un paso crucial durante el curso de la determinación de la estructura mediante la técnica de difracción de rayos X. Desafortunadamente, no existe actualmente ninguna forma de predecir bajo qué condiciones una determinada proteína formará cristales adecuados para difractar. Por ello, la cristalización de una

proteína se basa en un proceso de prueba-fallo hasta determinar las condiciones más adecuadas para obtener cristales de la misma. Las proteínas presentan formas irregulares y flexibles y las interacciones que facilitan la formación del cristal son en general de carácter débil. Esto dificulta el ordenamiento de sus moléculas de una forma regular en la red periódica cristalina. Por otra parte no debemos olvidar que las condiciones de cristalización deben ser compatibles con la estabilidad de la proteína. Por ejemplo, condiciones muy extremas de pH o temperaturas altas pueden dar lugar a la desnaturalización de la misma.¹⁹

1.2.1 Fundamento

Una manera práctica de representar el cambio de solubilidad de una proteína con el precipitante es el diagrama de fases¹⁸ que se representa en la figura 6. Se puede observar que cuanto más alta es la concentración de precipitante menor es la concentración de proteína soluble. También puede verse como a una determinada concentración de precipitante hay un límite de solubilidad de la proteína. Entre la línea de solubilidad y la de precipitación se encuentra la región metaestable que representa la zona de supersaturación, donde eventualmente pueden surgir núcleos de cristales que facilitan el posterior crecimiento de los mismos. Este proceso está controlado cinéticamente y es necesario superar una barrera de energía para que se forme un núcleo mínimo a partir del cual se pueda desarrollar el cristal.

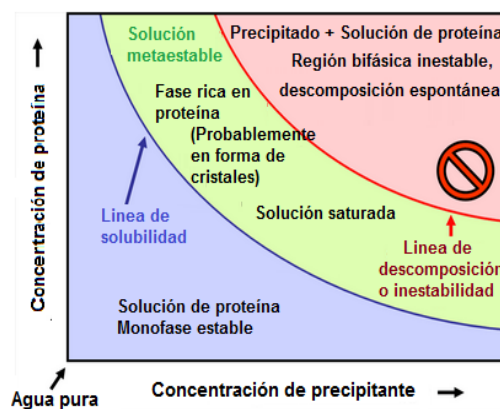


Figura 6. Diagrama de fases básico para una temperatura dada. La concentración de proteína está representada en el eje vertical y la concentración de precipitante en el eje horizontal. La línea de solubilidad separa la región que contiene una sola fase (de disolución de proteína) de una región con dos fases, que contiene proteína precipitada y una disolución saturada en equilibrio termodinámico.

Los métodos más comunes para alcanzar la sobresaturación de proteína son:¹⁹

- Adición de precipitante a la disolución de proteína.

- Eliminación de agua de la disolución de proteína-precipitante mediante difusión de vapor.
- Cambio de disolvente mediante diálisis.
- Difusión mediante interfase libre.
- Cambio de pH.

Efecto del pH en la solubilidad de la proteína. El pH de la disolución ejerce un efecto muy fuerte en la cristalización de proteínas. El punto de solubilidad mínima de la proteína se corresponde con el punto isoeléctrico (pI, es el pH en el cual la carga neta de la proteína es cero), por lo que generalmente las proteínas no cristalizan en su punto isoeléctrico. El cambio de pH es sin embargo un parámetro clave ya que produce cambios en la distribución local de cargas debido a los diferentes pKs de los residuos presentes en la superficie de la proteína.¹⁸ Ello permite en algunos casos crear interacciones específicas favorables para el empaquetamiento de las proteínas.

Efecto de la temperatura en la solubilidad de proteína. La solubilidad de una proteína puede aumentar o disminuir con la temperatura. Las estadísticas muestran que la mayoría de las proteínas cristalizan a temperatura ambiente, pero en muchos casos también se han utilizado temperaturas tan bajas como 4 °C. Las razones de la elección de esta temperatura son en gran parte históricas, basadas en el intento de evitar la degradación por proteasas, la cual es más común en las purificaciones de proteínas aisladas del tejido de origen que de las purificaciones actuales mediante sobreexpresión en *E.coli* u otro microorganismo. El uso de una temperatura reducida puede ser de ayuda para que el crecimiento cristalino sea lento. Por otra parte es importante señalar que un nivel de temperatura constante es conveniente para mantener estable cualquier cristal que se haya formado.

Uso de sales en cristalización. Las sales son precipitantes comunes usados en experimentos de cristalización. Las proteínas son poli-iones ya que poseen residuos cargados expuestos en la superficie. Los iones presentes en las disoluciones salinas afectan a la solubilidad de las proteínas de una forma compleja. Cantidades pequeñas de sal a menudo incrementan la solubilidad de proteína (fenómeno conocido como salting-in), mientras que si seguimos aumentando la concentración de sal ocurre el fenómeno contrario (salting-out) que reduce la solubilidad de esta.¹⁹

Uso de precipitantes orgánicos. Los precipitantes más usados en cristalización, además de las sales, son los polialcoholes orgánicos como el polietilenglicol (PEG), que según la longitud promedio de su cadena da lugar a su nomenclatura. Por ejemplo, si debido a la longitud de su cadena la masa molecular promedio es 200, se denomina PEG 200. Los PEGs son los precipitantes más exitosos para el crecimiento cristalino.¹⁹ La adición de PEGs siempre disminuye la solubilidad de proteína. Microscópicamente, el descenso en la solubilidad de proteína puede ser explicado por una competición de los polialcoholes por las moléculas de agua que están presentes en los alrededores de la proteína, así se fuerza a las proteínas a buscar interacciones entre ellas mismas más que con el disolvente polar de agua.²⁰

1.2.2 Técnicas de cristalización

En la figura 7 se muestran los métodos más comunes de cristalización de proteínas.^{18,19}

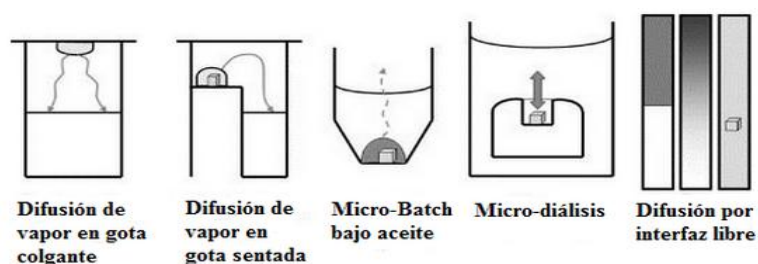


Figura 7. Esquema general de los procesos de cristalización más usualmente utilizados.

Difusión de vapor: se coloca una gota que contiene una mezcla (normalmente a partes iguales) de agente precipitante y de proteína en un pocillo en cuyo reservorio se encuentra el agente precipitante a mayor concentración. El pocillo se cierra herméticamente para evitar la evaporación de la gota y para permitir la difusión de vapor de agua desde la disolución de proteína en la gota hasta la disolución de agente precipitante en el reservorio. De esta forma se pretende alcanzar un equilibrio de saturación entre ambas disoluciones permitiendo que se formen cristales de proteína. Este tipo de técnica se puede clasificar en gota colgante o gota sentada. En el modo de gota colgante la mezcla entre proteína y agente precipitante se coloca en un cubre sobre la disolución precipitante del reservorio. En la gota sentada, la mezcla se coloca en un soporte que se encuentra en la parte superior del montaje, permitiendo la difusión de vapor de la gota hacia la disolución de precipitante en el reservorio. La bandeja entera

se cubre con cinta adhesiva transparente. Normalmente la gota colgante es el método favorito cuando se preparan las condiciones de cristalización de forma manual, pero es difícilmente automatizable y las gotas no pueden ser muy voluminosas. En aquellos casos en los que se usa un robot, el método utilizado es el de la gota sentada por difusión de vapor debido a limitaciones técnicas. Como desventaja, en algunas ocasiones los cristales de proteína pueden quedarse adheridos en la superficie del soporte.

Batch: esta técnica se suele emplear usando aceites que cubren la disolución de proteína-precipitante por lo que evitan la necesidad de sellado del pocillo. Sin embargo, la presencia de estos aceites hace que la manipulación de los cristales sea más complicada. En algunos casos el uso de aceites siliconados permiten el intercambio de disolvente por evaporación como se indica con la flecha punteada de la figura 7 y esta técnica aprovecha la ventaja de la eliminación progresiva y lenta del disolvente para alcanzar la sobresaturación.

Diálisis: esta técnica consiste en la difusión de moléculas de precipitante a través de una membrana semi-permeable hasta alcanzar el equilibrio de saturación, haciendo que la proteína cristalice.¹⁹ Esta técnica se suele llevar a cabo cuando el agente precipitante contiene moléculas pequeñas como sales o alcoholes, de manera que penetran fácilmente en la membrana haciendo que la proteína alcance lentamente el equilibrio con dicho agente. La diálisis presenta ciertas ventajas, como por ejemplo incluir la posibilidad de cambiar la disolución del reservorio varias veces. Esto permite que la condición de cristalización en la que se encuentra la proteína sea siempre la óptima, de manera que se pueden obtener cristales de gran tamaño. Como desventaja, es una técnica que presenta muchas complicaciones a la hora de miniaturizarla.

Difusión por interfaz libre: se basa en colocar una capa de la disolución precipitante sobre una solución de proteína concentrada en el interior de un capilar cuyos extremos han sido sellados con cera.¹⁸ El estrecho diámetro del capilar evita la mezcla de ambas disoluciones por convección, de esta manera el precipitante y la proteína difunden hasta que el sistema alcanza el equilibrio por un fenómeno denominado contra-difusión. Cuando inicialmente ambas soluciones entran en contacto y comienza la difusión, la proteína que se encuentra en los alrededores de la interfase alcanza la sobresaturación de manera que se empiezan a formar los primeros núcleos. A medida que la difusión

avanza, las dos disoluciones se van diluyendo una a otra, lo que permite que se vayan formando cristales cada vez más grandes de proteína a lo largo del eje del capilar. La ventaja de esta técnica es la posibilidad de obtener cristales de gran tamaño y además, realizar de forma automática un barrido de las concentraciones tanto de proteína como de agente precipitante. Sin embargo, necesita una cantidad relativamente grande de muestra y a veces es difícil que se diferencie la interfase entre las dos sustancias.

Cristalización en gel: se utiliza principalmente para minimizar los defectos de los cristales que ocurren durante el crecimiento de los mismos y para controlar la nucleación.¹⁹ El gel crea un medio en el que no hay convección inducida por la gravedad, por lo que el cristal queda atrapado en los poros de la matriz del gel y como consecuencia éstos no pueden moverse ni sedimentar. Cabe destacar que los geles suelen ser de agarosa y que son totalmente transparentes a los rayos X, por lo que su presencia inocua para mantener el cristal durante el proceso de congelación y difracción.

Contradifusión en gel: este método utiliza geles para mejorar los experimentos de contra-difusión. El gel de agarosa se utiliza para separar la proteína de la disolución de agente precipitante. Se rellena un capilar con la solución de proteína y se cubre uno de sus extremos con cera o arcilla mientras que el otro extremo está fijado al gel. Después, la matriz del gel se cubre con la disolución de agente precipitante. De esta manera se pueden evitar algunos de los inconvenientes del método de difusión por interfaz libre.

1.3 Cristalografía computacional

Actualmente existen diferentes métodos para resolver la estructura de una proteína, tales como difracción de rayos X, espectroscopía RMN o microscopía electrónica, siendo el método más popular sin duda alguna la cristalografía de rayos X. Aproximadamente un 87 % de las estructuras depositadas en el PDB han sido resueltas mediante esta técnica.⁸ Sin embargo, para poder llevarla a cabo es necesario obtener cristales de buena calidad y tal y como se ha visto en el apartado 1.2, esto no es nada sencillo ya que no se puede conocer a priori cuáles serán las condiciones en las que una determinada proteína cristalizará y mucho menos, si los cristales serán buenos para difractar. De hecho, un análisis reciente muestra que aproximadamente sólo el 4.6 % de proteínas objetivo dan lugar a cristales de calidad para difractar.²¹ Además, las

estimaciones muestran que más del 60 % de los cristales medidos son intentos fallidos.²²

Debido a todo lo explicado hasta este punto, se puede decir que la obtención de cristales de alta calidad es el cuello de botella a la hora de determinar una estructura. Es por esta causa por lo que ha surgido la cristalografía computacional para intentar predecir si una proteína va a tender a cristalizar o no según su secuencia.²³ Para ello, se utiliza la información estructural que ofrece el PDB y se correlaciona con secuencias primarias que pueden o no dar estructuras cristalinas y que se encuentran en bases de datos como TargetDB y UniRef50. TargetDB contiene datos procedentes de 20 centros SG²⁴ en Estados Unidos, Canadá, Alemania, Israel, Japón, Francia y Reino Unido. Estos sistemas tienen como objetivo por tanto predecir ciertas propiedades de la macromolécula para de esta manera ofrecer información acerca de la cristalizabilidad de la proteína. Entre los parámetros que suelen tener en cuenta estos sistemas, destaca el punto isoeléctrico calculado a partir de la secuencia, ya que permite sugerir rangos de pHs óptimos para cristalizar la proteína.^{25,26} También destacan los parámetros que se relacionan de forma directa con la predicción del grado de desorden de la proteína en el medio, tales como el número de residuos desordenados, la entropía de la cadena lateral, los residuos que se exponen al medio, el número de fenilalaninas y la cantidad de glicinas enterradas.²⁷

A la hora de trabajar con estas bases de datos lo único que hay que hacer es introducir la secuencia de la proteína de interés y la base de datos devolverá la probabilidad de cristalización y en algunos casos, una posible condición para que ésta cristalice. También existen otras bases de datos algo más completas que predicen además si la proteína es fácil de clonar y purificar y también si los cristales difractarán en una buena resolución.

La precisión de estos métodos de predicción se mide con la siguiente ecuación: $\frac{TP+TN}{TP+TN+FP+FN'}$ donde TP y TN son los números de verdaderos positivos y negativos, y FP y FN son los números de falsos positivos y negativos, respectivamente. De esta manera, para una misma proporción de resultados positivos y negativos la precisión mínima es del 50 %. También se utiliza el coeficiente de correlación de Matthews²¹

(MCC): $\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP) \times (TP+FN) \times (TN+FP) \times (TN+FN)}}$ para una mayor exactitud.

Las bases de datos con las que se han trabajado en este Trabajo Fin de Máster se describen a continuación en la siguiente tabla 2:

Tabla 2

Base de datos	Observaciones	Referencia
XtalPred	Predice la cristalizabilidad: hidrofobicidad, pI, número de C, M, W, Y y F y regiones desordenadas	22
CrystalP2	Predice la cristalizabilidad: hidrofobicidad, pI y la cantidad de pares y tripletes de aminoácidos separados entre sí por al menos 4 o 2 huecos, respectivamente	28
Pxs	Genera un análisis estadístico de las propiedades físicas que controlan la cristalización de la proteína a partir de su estructura, utilizando el factor Pxs, predice la calidad de difracción	27
Crysalis	Predice si la proteína es propensa a ser clonada, purificada, cristalizada y si sus cristales serán de calidad para difractar	29
<u>SVMCrys</u>	Clasifica a la proteína en “sensible” a ser cristalizada o “reacia” a ser cristalizada. No aporta más información puesto que la página no funciona correctamente	30
<u>ParCrys</u>	Predice la cristalizabilidad: número de S, C, M, G, F e Y	31
<u>PPCPred</u>	Realiza una predicción basada en la secuencia de la proteína acerca de la cristalizabilidad, purificación y rendimiento	32
<u>OB-Score</u>	Además del pI y de la hidrofobicidad, también tiene en cuenta la entropía superficial de la proteína debida a los residuos de la cadena principal. Especialmente indicada para proteínas de membrana	33
<u>MetaPPCP</u>	Predice si la proteína es cristizable o no cristizable con un 80 % de seguridad	24

Las bases de datos que aparecen subrayadas son aquellas que se han intentado utilizar pero que no funcionan correctamente por distintos motivos:

- SVMCrys: esta base de datos requiere el uso de unos ejecutables, pero al descargarlos la página no responde ni envía el resultado de la predicción. No ofrece medios para ponerse en contacto con el administrador de la página, por tanto no se ha podido utilizar.
- ParCrys: la base de datos no está operativa, al intentar acceder a ella se muestra un mensaje de error que indica que el servicio está temporalmente fuera de servicio. Tampoco hay ninguna manera de contactar con el administrador, por tanto parece que la página web está totalmente abandonada.
- PPCPred: nunca llega el resultado, en su lugar envía un mensaje de error diciendo que te pongas en contacto con el administrador de la página a través de

una dirección de correo electrónico facilitada en el texto. El administrador no responde por lo que no debe estar ofreciendo ningún tipo de soporte a la página.

- OB-Score: Ocurre lo mismo que en el caso de la base de datos SVMCrys.
- MetaPPCP: Nunca envía el resultado de la predicción ni tampoco llega mensaje de error. Como ocurre con ParCrys, al enviar un correo notificando del error al supuesto administrador de la página no se obtiene ningún tipo de respuesta.

Como se ha podido observar, no todas las bases de datos que se ofertan funcionan de forma correcta, por tanto se puede deducir que en su momento fueron diseñadas para trabajar con ellas asiduamente en un laboratorio o para un determinado trabajo de investigación pero que, al concluir dichas tareas, estas bases de datos han sido abandonadas ya que los administradores no ofrecen ningún tipo de soporte.

2. Objetivos

1. Entender y aplicar las competencias de un laboratorio de investigación del área de química física en estudios estructurales de proteínas.
2. Realizar un estudio bibliográfico de la ubiquitina WT, en especial de las estructuras disponibles.
3. Realizar un estudio bibliográfico de la cristalografía computacional y de las bases de datos disponibles, comprobando su funcionalidad.
4. Llevar a cabo la clonación, expresión y purificación de la ubiquitina.
5. Llevar a cabo una caracterización de la proteína mediante técnicas espectroscópicas y electroforéticas.
6. Determinar las condiciones de cristalización de la ubiquitina mediante técnicas de cribado y realizar un estudio de la optimización de dichas en presencia del catión magnesio Mg^{2+} .
7. Determinar la estructura de la ubiquitina mediante la difracción por rayos X de los cristales obtenidos.

4. Materiales y métodos

4.1 Materiales

4.1.1 Instrumentación

A continuación, se describen los instrumentos más utilizados en el laboratorio durante el desarrollo de este Trabajo Fin de Máster.

- Agitador Orbital Agimatic-S (J.P. Selecta): se ha utilizado para preparar todas las disoluciones.
- Micropipetas (Rainin): se utiliza para tomar volúmenes pequeños de 2 μL , 10 μL , 20 μL , 100 μL , 200 μL y 1000 μL .
- Baño termostático 12B Tectron S473-100 (Julabo): se utiliza en el proceso de transformación para que las células sufran el llamado choque térmico.
- Estufa bacteriológica Incudigit serie 2000 (J.P Selecta): se utiliza tras la transformación para la incubación de las placas.
- Armario incubador Medilow-LG (J.P Selecta): se han utilizado tanto para el crecimiento de los cultivos de *E. Coli* a una temperatura controlada cómo para el crecimiento de los cristales.
- Bandejas (Crystalgen): para la cristalización de proteínas se realiza en bandejas.
- Centrífuga Zentrifugen Rotanda 460R, con un rotor 5645 (Hettich): se utilizó durante el proceso de purificación.
- Ultracentrifuga XL-90 con un rotor 45TI (Beckman).
- Sonicador (GmbH)-UP 200S Ultraschall-prozessor (Grepütfe Sicherheit dr.Hielscher): se utilizó para el proceso de lisis celular.
- Filtros Ultra-Milipore de 3 kDa ó 10 kDa (Amicon): para cambiar el tampón y concentrar las proteínas.
- Espectrofotómetro Lambda 25 (Perkin Elmer) de UV-vis: utilizado para medir la absorbancia. Se utilizan cubetas de cuarzo con 1 cm de paso de luz.
- Microcentrífuga Sorvall Legend Micro 21 (Thermo Scientific): utilizada para tratar las proteínas.
- Microscopio M205 C (Leica): para observar los cristales.
- Equipo de DLS Zetasizer Nano ZS (Malvern): se ha utilizado para llevar a cabo las medidas de SLS.

- Columna (Bio-Rad): es utilizada para llevar a cabo la cromatografía de afinidad.
- Membrana para diálisis (Spectra/Por) Dialysis Membrane (MWCO: 3500): para dializar las proteínas.

4.1.2 Reactivos

Todos los reactivos utilizados en esta memoria son de grado para análisis. Fueron suministrados por las marcas: Sigma-Aldrich, Panreac, Conda, Cultimed, Fluka, Merck, Riedel-de Hæn y Acros Organics. Se describe a continuación la preparación de las disoluciones de uso más común.

Medio de cultivo LB (Lisogeny Broth). Es un medio nutricionalmente rico para las bacterias. Está compuesto por 10 g de triptona, 10 g de cloruro sódico y 5 g de levadura (para 1 L de agua destilada). El medio debe estar en un pH entre 7 y 8, por lo que se utiliza hidróxido sódico 1 M. Antes de utilizar el medio LB debe autoclavarse para asegurar su esterilización.

Células *E.coli* BL21-DE3 (Novagen). Almacenadas a -80 °C.

Plásmidos con los genes clonados de las proteínas que se vayan a purificar para introducir el gen que se desea. Estos plásmidos son proporcionados por la profesora Irene Luque del departamento de Química-Física de la Universidad de Granada.

Placas de Petri de Agar. Para su preparación se añade 1 g de Agar al medio LB por cada 50 mL del mismo. Una vez preparado el medio LB con agar se debe autoclavar para esterilizarlo. Cada placa de Petri contiene 25 mL de esta disolución, aproximadamente. Para incorporar el antibiótico (en nuestro caso kanamicina o ampicilina), se añade 1 µL de antibiótico por mL de medio. El antibiótico debe agregarse después del autoclavado.

Agente inductor IPTG. Se utiliza en el proceso de inducción. Se prepara en agua MiliQ a una concentración de 40 mg/mL.

Tampón de equilibrado de columna (TEC). Fosfato sódico 50 mM, cloruro sódico 300 mM pH 8. Se utiliza tras la lisis celular y la centrifugación para redissolver el precipitado (restos celulares).

Columna Ni²⁺-NTA Agarosa (Quiagen). Para llevar a cabo la cromatografía de afinidad.

Los reactivos utilizados en la electroforesis se almacenan en la cámara fría a 4 °C, exceptuando el SDS 10% y el tampón de desarrollo. Las disoluciones utilizadas son:

- Disolución A: acrilamida al 29,2 % y Bis-acrilamida al 0,8 %.
- Disolución B o gel B: tampón del gel separador. Tris 1,5 M a pH 8,8.
- Disolución C o gel C: tampón del gel apilador. Tris 0,5 M a pH 6,8.
- Disolución D: SDS 10 %.
- Disolución E: APS 10 % (persulfato amónico). Se prepara en el momento en el que se va a realizar la electroforesis porque se descompone. Se añaden 0,1 g de persulfato amónico en 1 mL de agua MiliQ.
- TEMED.
- Tampón de desarrollo concentrado: 100 mL de la disolución (Tris 0,125 M, glicina 1 M y SDS 0,5% a pH 8,3) en 1 L de agua destilada.
- Tampón de carga: 1 mL de disolución C, 0,8 mL de glicerol, 1,6 mL de SDS 10%, 0,4 mL de 2-mercaptoetanol (elimina los puentes disulfuro S-S de las proteínas), 0,2 mL de azul de bromofenol al 0,05 %. Debido al azul de bromofenol, esta disolución es de color azul y es la que nos va a permitir ver el desplazamiento de la proteína a lo largo del gel. Se añade 1 µL de esta disolución por µL de proteína.
- Las disoluciones empleadas para revelar el gel consisten en una disolución de fijación y secado compuesta de metanol al 40 % y ácido acético al 7 % y una disolución de tinción Comassie Brilliant Blue R-250 al 0,1 % en disolución fijadora.

4.2 Métodos

4.2.1 Clonación, expresión y purificación de la ubiquitina

Transformación

- Se añaden 2 µL del plásmido pET-101 Directional TOPO (Invitrogen) a un tubo Eppendorf que contiene la *E.coli* BL21-DE3. Todo este proceso se lleva a cabo en una cabina de flujo laminar para asegurar la esterilización del proceso.

- Se coloca el Eppendorf en un bloque térmico a 0 °C durante 5 minutos. A continuación se calienta en el baño termostatzado durante 40 segundos a 42 °C tras los cuales se vuelve a introducir el Eppendorf en el bloque térmico a 0 °C durante 2 minutos.
- Se añaden 500 µL de medio LB al Eppendorf y se incuba en una estufa a 37 °C durante 1 hora con agitación.
- Se realiza el sembrado de la *E.coli* sobre la placa de Petri que contiene el medio LB con el antibiótico al que son resistentes nuestras bacterias (ampicilina), con ello nos aseguramos de que las bacterias que crecen en la placa son las que contienen el plásmido de interés. Se deja incubar la placa en la estufa a 37 °C durante toda la noche.

Expresión

- Se preparan 10 mL de medio LB y se reparten en matraces de 2 L añadiendo aproximadamente 800 mL de medio en cada matraz (todos los matraces con el medio han sido previamente autoclavados).
- Se añade 1 mL de ampicilina a cada matraz, siempre cerca de un mechero Bunsen y bajo campana para asegurarse de que hay un medio estéril mientras se realiza el proceso.
- Se inocula una colonia de las que han crecido en la placa de Petri en cada matraz, también cerca del mechero bunsen para evitar contaminación.
- Se incuba el cultivo a 37 °C con agitación hasta que la densidad óptica del medio de cultivo esté entre 0,6 y 0,8 a la longitud de onda de 600 nm.
- Una vez alcanzada dicha absorbancia se procede con la inducción. Para ello se añade 1 mL de agente inductor IPTG (isopropil-β-D-1-tiogalactopiranosido) al medio de cultivo.
- Las células se incuban durante toda la noche.

Recolecta de células

- Para recoger las células y separarlas del resto de la disolución se utiliza una centrífuga. Se reparte el medio de cultivo en 6 botellas de centrifugación adaptadas a la centrífuga. Hay que igualar el peso del volumen de los botes para que al colocarlos en la centrífuga ésta esté correctamente equilibrada. Se centrifuga durante 10 minutos, a 5000 r.p.m a 4 °C.

- Tras centrifugar, las células se quedarán depositadas en el fondo de la botella, de modo que se desecha el líquido sobrenadante y se resuspende el precipitado en un volumen de tampón acorde con la cantidad de células obtenidas (50 mL por cada 10 g de precipitado).
- A continuación, se lleva a cabo la lisis celular empleando un sonicador. Con este procedimiento se rompe la membrana celular liberando al medio el contenido intracelular. El proceso consiste en intercalar 40 segundos de sonicado (configurando el aparato de manera que tenga una amplitud del 60 % y un tiempo de ciclo de 0,5 segundos) y 40 segundos de reposo. Esto se lleva a cabo en un baño de hielo para evitar el calentamiento de las células que pueden dar lugar a que se desnaturalice la proteína. El proceso se repite varias veces hasta que se observa un cambio en la densidad y el color de la disolución.
- Se pasa la disolución a unos tubos de ultracentrífuga, y se centrifuga durante 30 minutos a 4 °C y a 13000 r.p.m. Depositados en el fondo quedan los restos celulares, mientras que en la disolución deben quedar las proteínas de interés.

Columna de afinidad de Ni²⁺-NTA Agarosa

El uso de ingeniería genética ha permitido manipular los sistemas de expresión para facilitar la purificación de las proteínas de interés. En muchos casos la proteína se expresa con una “etiqueta” que presenta afinidad por algún grupo que se encuentra presente en una columna de cromatografía. Estos grupos se pueden introducir en los soportes cromatográficos por derivatización. Ello permite que la proteína de interés se pueda aislar de otras presentes en la célula de *E. coli* por su interacción específica a través de esta etiqueta con los grupos funcionales presentes en el gel cromatográfico. En concreto, el fundamento de la cromatografía de afinidad utilizada en nuestro caso es la siguiente: la columna está formada por una resina de agarosa derivatizada con ácido nitriloacético (NTA), que forma un quelato de iones Ni²⁺. Estos iones quedan adheridos a la columna a través de cuatro enlaces. Los dos enlaces restantes mantienen unida a la proteína ya que posee colas de histidina (figura 8). El imidazol compite por la interacción con el níquel presente en la columna y a determinadas concentraciones da lugar al desplazamiento de la proteína unida a la columna.

- Se empaquetan 10 mL de gel de Ni²⁺-NTA Agarosa en una columna de plástico de 20 mL. A continuación, se equilibra añadiendo 50 mL de TEC.

- Una vez equilibrada la columna se introduce en la misma la disolución de proteína obtenida a partir del sobrenadante de la centrifugación.
- La proteína de interés quedará unida a la columna ya que está expresada en un vector que introduce en el extremo amino-terminal de la proteína una secuencia de 6 histidinas que presentan una alta afinidad por el níquel. Otras proteínas presentes en la disolución carecen de esta “etiqueta” de histidinas y presentaran una baja o nula interacción con la columna. Para eliminar estas últimas se va añadiendo un gradiente escalonado de imidazol en el mismo tampón que se utilizó para equilibrar la columna (concentración de imidazol de 20-250 mM).
- La proteína se eluye, de forma general, a una concentración de imidazol de 250 mM. Una vez se inicie la elución con esta disolución se recogerá la proteína en tubos de ensayo.

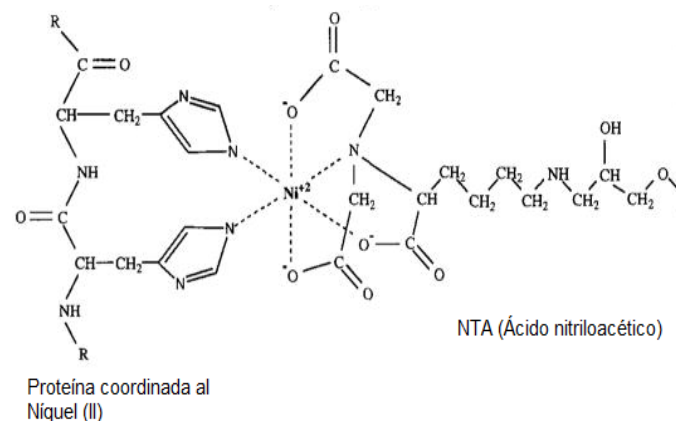


Figura 8. Representación gráfica de una columna de Ni^{2+} -NTA Agarosa. El Ni^{2+} adherido a la columna forma enlaces con las colas de histidina de la proteína.

Tras purificar la proteína, se pasan con una pipeta pasteur todas las eluciones obtenidas a una membrana para diálisis Spectra/Por Dialysis Membrane (MWCO: 3500) que ha sido previamente hidratada. La proteína se dializa en 5 L de TEC. Esto se hace para eliminar el imidazol y que así la proteína quede lo más pura posible.

Medida de la concentración

Tras la purificación es necesario conocer cuál es la concentración de la proteína obtenida. Para ello se utiliza espectrometría de ultravioleta visible (UV-vis). Se introduce en el espectrofotómetro la cubeta con la proteína y se mide la absorbancia a una longitud de onda de 280 nm (correspondiente a la absorbancia de los aminoácidos

aromáticos triptófano y tirosina). Una vez conocida la absorbancia se puede calcular la concentración fácilmente por medio de la ley de Lambert Beer.

Para determinar el coeficiente de extinción molar se utiliza la secuencia de aminoácidos de la proteína con el dominio correspondiente. Para ello se emplea la herramienta online proporcionada por ExPASy denominada Protparam (<http://web.expasy.org/protparam/>). Se introduce la secuencia de la proteína y la herramienta proporcionará un informe completo de los aminoácidos, incluyendo el cálculo de su composición atómica, peso molecular y coeficiente de extinción molar. En el caso de la ubiquitina, el coeficiente de extinción molar obtenido es de $1490 \text{ M}^{-1}\text{cm}^{-1}$ y el peso molecular es de 8650.8 Da.

4.2.2 Electroforesis en SDS-PAGE

Tras purificar la proteína, es indispensable comprobar el grado de pureza de la misma. Para ello se lleva a cabo una electroforesis en gel.

Dentro de las electroforesis en gel, la más empleada para el análisis de proteínas es el SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis). El uso de SDS provoca una normalización en la relación carga/masa ya que las proteínas se desnaturalizan, perdiendo su conformación tridimensional formando un complejo SDS-proteína en el cual la carga de la proteína queda enmascarada por la de las múltiples moléculas de SDS, obteniendo un fraccionamiento que obedece sólo longitud de la cadena (tamaño).

Las disoluciones utilizadas para la electroforesis aparecen detalladas en el apartado 3.1.2. El gel de poli(acrilamida) se trata de un gel sintético, termoestable, transparente, resistente y relativamente inerte químicamente hablando. Puede resistir gradientes de alto voltaje, ser viable para varios procedimientos de tinción y destinción y puede ser digerido para extraer fracciones separadas o desecado para su autorradiografía y registro permanente. Este gel se forma por polimerización de la acrilamida y la N,N'-metileno-bisacrilamida. La polimerización se inicia con la formación de radicales libres del monómero. Se trata de radicales libres de oxígeno procedentes de la descomposición del persulfato amónico. Las aminas terciarias como el TEMED se emplean como catalizadores de esta reacción de descomposición. La electroforesis se lleva a cabo en un medio discontinuo:

- Gel apilador: gel de poro grande que no limita la movilidad de las proteínas y cuyo papel es producir una concentración de la banda de proteína. Sobre el gel de apilamiento o espaciador se coloca el peine de 15 calles antes de que éste gelifique para crear la forma de los surcos donde se va a depositar la muestra.
- Gel separador: gel de poro pequeño que limita la movilidad de las proteínas en función de su tamaño.

En nuestro caso, los geles se han preparado al 16 % de acrilamida debido al tamaño de las proteínas a separar. Para preparar 10 mL de gel separador, se mezclan 2,1 mL de agua MiliQ, 5,3 mL de acrilamida, 2,5 mL de disolución B, 0,1 mL de SDS 10 %, 0,05 mL de APS 10 % y 5 μ L de TEMED. Una vez preparada la mezcla, se colocan las dos placas enfrentadas en el soporte y se rellena el espacio entre las placas hasta aproximadamente 3/4 de su altura. Se cubre cuidadosamente el espacio restante de agua destilada con la intención de generar un frente firme del gel y no se forme un menisco. Una vez polimerizado el gel separador se retira el agua de la superficie con ayuda de tiras de papel de filtro.

Para preparar 5 mL de gel apilador, se mezclan 2,7 mL de agua MiliQ, 1,0 mL de acrilamida, 1,25 mL de disolución C, 0,05 mL de SDS 10 %, 0,025 mL persulfato amónico 10 % y 5 μ L de TEMED. Se vierte la disolución encima del gel separador y se coloca el peine con el número de calles de nuestro interés.

Se coloca el gel polimerizado en la cubeta de electroforesis y la muestra de proteína se diluye en tampón de carga (1 μ L de tampón de carga por μ L de proteína). Se calienta a 80 °C durante 3 minutos y se introduce la dilución en el pocillo de la calle que se desee. Se rellena la cubeta con tampón de desarrollo diluido y se establece una intensidad de corriente de 16 mA por lámina, aumentando el amperaje a 30 mA por lámina cuando el frente de azul bromofenol alcance el gel separador.

Una vez finalizada la electroforesis, las bandas de las proteínas separadas se fijan con una disolución de metanol al 40 % y ácido acético al 7 % durante 30 minutos, se visualizan aplicando una disolución de teñir (Comassie Brilliant Blue) con una leve agitación durante 30 minutos. Por último, para la destinción empleamos la misma disolución que para fijar los geles, es aconsejable agitación para que la destinción sea más rápida. Finalizado este proceso, se obtiene el gel completamente revelado.

4.2.3 Cristalización de proteínas

4.2.3.1 Cribado de condiciones de cristalización

A la hora de intentar cristalizar una proteína por primera vez es necesario realizar un cribado de las posibles condiciones de cristalización. Para ello, se utilizan las disoluciones de cribado de Molecular Dimensions y Sigma Aldrich, concretamente:

- Structure Screen I y II de Molecular Dimensions (50 condiciones diferentes cada uno).
- ProPlex I y II de Molecular Dimensions (48 condiciones diferentes cada uno).
- Crystallization Kit for Protein-Protein Complexes de Molecular Dimensions (50 condiciones diferentes).
- Crystallization Cryo Kit for Proteins de Sigma-Aldrich (50 condiciones diferentes).

Para llevar a cabo esta tarea se utiliza una bandeja de cristalización de Crystalgen, que consta de 96 reservorios para colocar la disolución de cristalización y de dos pocillos para colocar las gotas por cada reservorio. La técnica de cristalización que se utiliza con esta bandeja es la de difusión de vapor por gota sentada. Se suele utilizar un volumen de gota de proteína de 1 μL probando diferentes concentraciones, generalmente entre 5-40 mg/mL. En el caso de la ubiquitina hemos visto que la concentración óptima para estos experimentos es 20 mg/mL. El volumen de agente precipitante (disolución de cristalización) en cada reservorio ha sido de 100 μL , añadiendo a la gota de proteína 1 μL del mismo.

Desde la puesta de la bandeja, se debe ir controlando la misma al día siguiente y después una vez a la semana con la ayuda de un microscopio para ver si hay evolución en las gotas. En este caso se ha utilizado un microscopio de la marca Leica, modelo M205 C.

4.2.3.2 Optimización de las condiciones de cristalización

Una vez determinadas las condiciones en las que cristaliza la proteína, o bien aquellas en las que se observa un precipitado microcristalino susceptible de ser mejorado, hay que intentar optimizar los cristales mediante un barrido de las

concentraciones del agente precipitante, variando también el pH y la temperatura. En algunos casos se ha usado también distintos aditivos.

Normalmente, una vez conocida la concentración de agente precipitante en la que cristaliza la proteína o con el que aparecen microcristales, se hace un barrido con concentraciones de dicho agente por encima y por debajo de la concentración de cristalización en el orden de ± 0.1 M. Lo mismo ocurre para el tampón. Cuando se conoce el pH al cual la proteína cristaliza o con el que aparece un precipitado microcristalino se lleva a cabo un barrido probando tampones con pHs por encima y por debajo del pH de interés en un orden de ± 1 unidades de pH. En algunos casos se puede llevar a cabo un barrido más fino y cambiando la composición de la disolución tampón. Para llevar a cabo este proceso de barrido más fino, normalmente se utiliza otro tipo de bandeja, con 48 reservorios y sólo un pocillo para colocar las gotas por cada reservorio (figura 9).

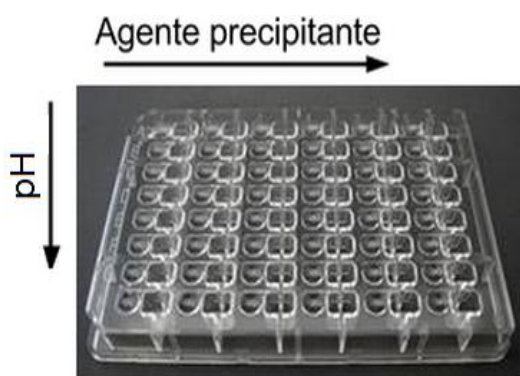


Figura 9. Modelo de bandeja utilizada para llevar a cabo la optimización de las condiciones. las flechas muestran como se suele variar el pH y el agente precipitante.

En este caso, las gotas de proteína se pueden poner de 2 o 2,5 μL , y en los reservorios se añaden 200 μL de agente precipitante. Como lo que se pretende es hacer un barrido de todas las condiciones posibles para optimizar las condiciones de cristalización, también se suele variar la relación proteína-agente precipitante manteniendo constante el volumen de proteína en toda la bandeja y variando el volumen de agente precipitante.

4.2.4 Capturado de las imágenes de los cristales

El microscopio se encuentra conectado a un ordenador. Leica Microsystems proporciona un programa llamado Leica LAS EZ que permite capturar en la pantalla del

ordenador todo lo que se vea por el microscopio. Esto es muy útil ya que nos permite fotografiar los cristales para poder comparar su crecimiento a lo largo del tiempo. También permite tomar anotaciones rápidas tales como medidas sobre la fotografía del cristal.

4.2.5 Difracción de los cristales, procesado, resolución de la estructura y refinamiento

Para poder obtener la estructura de la proteína es necesario difractar el cristal. Esto se lleva a cabo en un difractómetro o en instalaciones de luz sincrotrón. Estas últimas son las más utilizadas para medir los cristales ya que una radiación más potente permite realizar la medida en menos tiempo y ofrece mejores resultados. Sin embargo, un inconveniente del uso de radiación sincrotrón es el deterioro de los cristales por la radiación. Para evitar o retrasar este proceso se suelen utilizar bajas temperaturas. Habitualmente la recogida de datos se lleva a cabo a 100 K, pero debido al alto contenido de agua en los cristales de proteína es necesario añadir alguna sustancia que facilite su congelación para que el hielo no destruya el cristal. A estas sustancias se las conoce como crioprotectores. Los más comunes son el glicerol, el etilenglicol y los polietilenglicoles (PEGs). De esta manera los cristales se transportan en un dewar congelados bajo nitrógeno líquido hasta su destino.

En este caso, los cristales de ubiquitina fueron difractados en el “European Synchrotron Radiation Facility” (ESRF) en Grenoble, Francia. Los datos fueron obtenidos a una longitud de onda de 0.965 Å del haz de luz ID30A1³⁴. Los cristales difractaron a una resolución de 1.32 Å.

Las imágenes de difracción se procesaron utilizando el paquete autoPROC³⁵ y AIMLESS³⁶ y la resolución de la estructura se llevó a cabo utilizando PHENIX suite.³⁷ El programa Phaser³⁸ se utilizó para llevar a cabo la determinación de la fase por reemplazamiento molecular utilizando las coordenadas de la estructura cristalográfica del mutante de la ubiquitina K33S (código PDB: 4pjh).¹⁷ Una vez se obtuvo la solución de reemplazamiento molecular, se llevó a cabo un modelado manual de la estructura utilizando el programa Coot³⁹ que se alterna con ciclos de refinamiento utilizando phenix.refine en PHENIX.⁴⁰ El análisis de la estructura se realizó con los programas LSQKAB,⁴¹ CONTACT y AREAIMOL del programa CCP4 suite⁴² y MolProbity.⁴³

5. Resultados y discusión

5.1 Purificación de la ubiquitina

Los pasos específicos para la purificación de la ubiquitina se describen en el apartado 4.2.1, por lo que a continuación se muestran los resultados obtenidos. Durante todo el proceso de purificación se tomaron pequeñas alícuotas de todas las fracciones que salen de la columna cromatográfica para así comprobar donde está nuestra proteína en todo momento. Para ello se realiza una electroforesis en SDS-PAGE (figura 10).

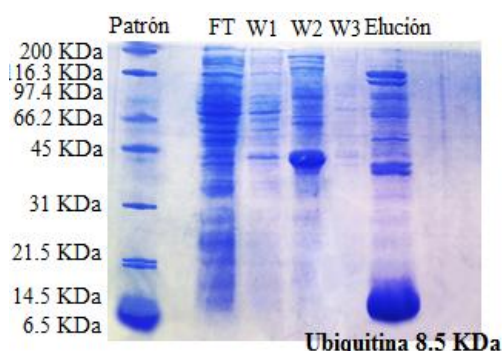


Figura 10. SDS-PAGE de la purificación de ubiquitina. FT (fracción del sobrenadante que sale de la columna de afinidad), W1 (lavado con 50 mL de TEC), W2 (lavado con 100 mL de TEC + imidazol 20 mM), Elución (eluciones de TEC + imidazol 250 mM).

Como se puede observar en el gel, hay una banda bastante intensa sobre la ubiquitina entre 31 y 45 KDa, con lo cual la proteína no está totalmente pura. Como la diferencia de peso molecular entre ambas es bastante grande, se llevó a cabo un paso por columna de exclusión molecular de la proteína, de manera que en primer lugar sale la proteína de mayor tamaño y finalmente la ubiquitina. Tras concluir con este paso, se volvió a llevar a cabo una electroforesis en SDS-PAGE (figura 11) para comprobar la pureza de la proteína.

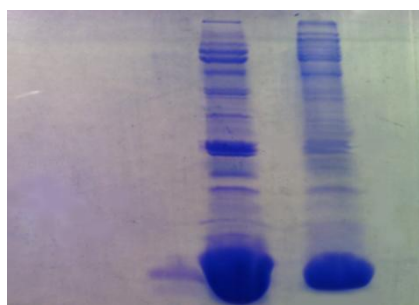


Figura 11. SDS-PAGE tras el paso por columna de exclusión molecular de la ubiquitina. Las bandas de la izquierda corresponden a la ubiquitina antes del paso por columna. La banda de la derecha corresponde a la ubiquitina tras pasar por la columna.

Se puede observar que tras realizar el paso por columna la proteína es más pura y que además la banda de la proteína interferente entre 31 y 45 KDa prácticamente ha desaparecido.

Tras dializar se midió la absorbancia obtenida en el espectrofotómetro para así calcular la cantidad de proteína purificada. Se obtuvieron 20 mL de proteína a 4 mg/mL, lo que supone una cantidad final de la misma en la disolución de 80 mg. Se concentró la proteína utilizando un Centricón Ultra-Milipore (Amicon) con un corte de tamaño molecular de 3 KDa hasta 20 mg/mL en un volumen final de 8 mL.

Para su posterior uso, se hicieron alícuotas de 1 mL en tubos Eppendorf. Las alícuotas de 1 mL se congelan empleando nitrógeno líquido y se conservan en el congelador a -80 °C hasta su uso. En estas condiciones, la proteína es estable durante al menos un año.

5.2 Cristalización de la ubiquitina

Para esta proteína, se utilizaron los kits de cristalización descritos en el apartado 4.2.3.1 La condición de agente precipitante en la que se observó la aparición del cristal tras tres meses fue la compuesta por PEG 4000 al 30 %, cloruro de magnesio 0,2 M y Tris 0,1 M a pH 8.5 del kit Structure Screen I de Molecular Dimensions.

El cristal de la figura 12 se obtuvo mediante la técnica de difusión de vapor en gota sentada a 15 °C con un volumen total en el reservorio de 100 μ L de agente precipitante. Para ello se pusieron gotas de 3 μ L formadas por 1,5 μ L de la disolución de ubiquitina (19 mg/mL en Tris 20 mM, pH 8) y 1,5 μ L de la disolución del reservorio (agente precipitante). Las fotografías se realizaron tal y como se explica en el apartado 4.2.4.

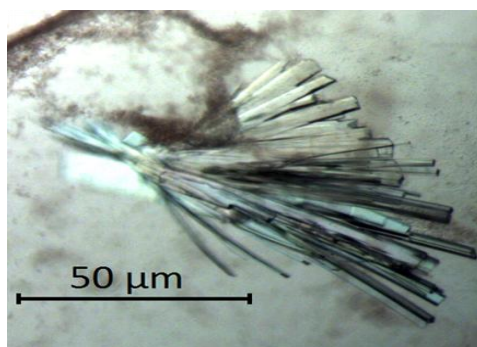


Figura 12. Cristal de ubiquitina obtenido con dimensiones 0,10 x 0,02 x 0,02 mm. Este cristal se difractó en instalaciones de luz Sincrotrón para determinar la estructura de la ubiquitina.

La condición de cristalización contiene cloruro de magnesio, por lo que tal y como se ha explicado anteriormente, el catión Mg^{2+} es un metal divalente que tiene un papel fundamental a la hora de cristalizar la proteína (apartado 1.1.2). Debido a esto y para intentar obtener nuevos cristales en otras condiciones y mejorar su calidad, se realizó un nuevo experimento añadiendo cloruro de magnesio a una concentración de 10 mM directamente sobre la ubiquitina a 20 mg/mL en tampón Tris pH 8 a 20 mM. Esta mezcla fue utilizada para poner nuevas bandejas utilizando los mismos kits de cristalización descritos en el apartado 4.2.3.1.

El procedimiento se llevó a cabo de la misma forma a la explicada en este mismo apartado, observando la aparición de cristales (figura 13) tras un mes de la puesta de la bandeja en la siguiente condición: 2,4-metilpentanodiol (MPD) al 50 %, fosfato amónico 0,2 M con tampón Tris pH 8.5 a 0,1 M.



Figura 13. Cristales obtenidos tras un mes de la puesta. 50 % MPD, 0,2 M $NH_4H_2PO_4$, 0,1 M Tris pH 8.5 y 10 mM $MgCl_2$.

En bibliografía¹⁶ existe una estructura resuelta de ubiquitina con MPD, pero en ausencia de metales divalentes. En nuestro caso, los cristales se han obtenido bajo dicha condición, pero además en presencia de cloruro de magnesio. Como se puede observar en la imagen, los cristales no son de muy buena calidad para difractar, ya que no tienen una estructura bien definida, no se observan caras planas y además han crecido en forma de aguja aunque en menor tiempo al obtenido en primer lugar (un mes frente a tres meses). Sin embargo, cuando esto ocurre se puede intentar optimizar la condición de cristalización tal y como se describe en el apartado 4.2.3.2 para intentar obtener mejores cristales. Para ello se realizó un cribado de las condiciones variando el porcentaje de MPD en la disolución (tabla 3) y el pH del medio. La técnica utilizada fue difusión de vapor en gota sentada a una temperatura de 15 °C.

Tabla 3. Representación esquemática de la bandeja. Se muestran cada una de las condiciones de cristalización presentes en cada reservorio.

10% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Citrato sódico pH 4.0	10% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 5.0	10% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.0	10% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	10% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Tris pH 8.0	10% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M bicina pH 9.0
20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Citrato sódico pH 4.0	20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 5.0	20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.0	20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Tris pH 8.0	20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M bicina pH 9.0
30% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Citrato sódico pH 4.0	30% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 5.0	40% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.0	40% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	40% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Tris pH 8.0	40% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M bicina pH 9.0
50% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Citrato sódico pH 4.0	50% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 5.0	60% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.0	60% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	60% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Tris pH 8.0	60% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M bicina pH 9.0

Tras cuatro días aparecieron los cristales que se muestran en la figura 14 visiblemente de mejor calidad y con una forma más definida bajo la condición: MPD al 60 %, fosfato amónico 0,2 M y Hepes pH 7,0 a 0,1 M como tampón. Cabe recordar que a la disolución de ubiquitina contiene cloruro de magnesio 10 mM.

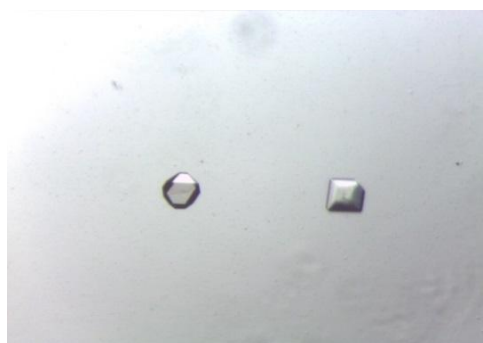


Figura 14. Cristales de forma más definida 60 % MPD, 0,2 M NH₄H₂PO₄, 0,1 M Hepes pH 7.0 y 10 mM MgCl₂.

Como se puede observar, esta condición de cristalización es muy similar a la de los cristales obtenidos inicialmente (figura 13), sin embargo el pH en este caso es de 7.0 y es posiblemente lo que mejora la forma y el crecimiento de los mismos.

También se ha optimizado el tiempo de aparición de los cristales. Una vez encontrada la mejor condición se prepara una nueva bandeja (tabla 4) realizando un

cribado muy fino de las condiciones del agente MPD y modificando el pH levemente en torno a 7.0. La técnica de cristalización utilizada es la misma que en los casos anteriores, difusión de vapor mediante gota sentada.

Tabla 4. Representación esquemática de la bandeja. Se muestran cada una de las condiciones de cristalización presentes en cada reservorio.

20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 6.0	30% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 6.0	40% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 6.0	50% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 6.0	60% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 6.0	70% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Acetato sódico pH 6.0
20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.5	30% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.5	40% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.5	50% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.5	60% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.5	70% MPD 0,2 M NH ₄ H ₂ PO ₄ 0,1M MES pH 6.5
20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	30% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	40% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	50% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	60% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0	70% MPD 0,2 M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.0
20% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.5	30% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.5	40% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.5	50% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.5	60% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.5	70% MPD 0,2M NH ₄ H ₂ PO ₄ 0,1M Hepes pH 7.5

A partir de esta bandeja se obtuvieron cristales (figura 15) en varias condiciones y un día siguiente a la puesta de la misma.



Figura 15. Cristales obtenidos. Imagen izquierda: 40 % MPD, 0,2 M NH₄H₂PO₄, 0,1 M Hepes pH 7.0 y 10 mM MgCl₂. Imagen derecha: 60 % MPD, 0,2 M NH₄H₂PO₄, 0,1 M MES pH 6.0 y 10 mM MgCl₂. Como se puede observar, el cristal tiene mayor tamaño y una estructura tetragonal.

Estos cristales obtenidos se difractaron en las instalaciones de luz sincrotrón situadas en Barcelona (Cataluña, España) ALBA. Desgraciadamente, los cristales difractados no eran de ubiquitina, sino que lo que se formó en la gota era una sal. En algunas ocasiones

es bastante complicado diferenciar si los cristales que se han formado son de sal o de proteína, y para ello nuestro grupo de investigación BIO238 ha puesto a punto un método que consiste en la tinción de los cristales con el colorante eosina mediante la técnica de soaking. Para llevarlo a cabo, la eosina se añade directamente sobre la gota que contiene el cristal ya formado. Si el cristal es de sal como en el caso de los obtenidos anteriormente, el colorante no difundirá hacia el cristal y por tanto éste no se teñirá. Este efecto se puede ver en la siguiente figura 16.

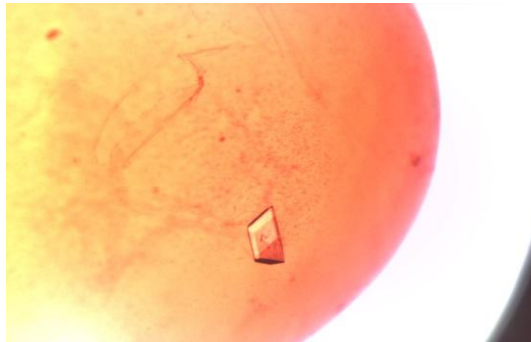


Figura 16. Cristal de sal. Como se puede observar, no hay diferencia de color entre el cristal y la disolución, por tanto, el cristal no se tiñe.

Sin embargo, si se lleva a cabo esta técnica sobre cristales de proteína, se observa una coloración intensa de los mismos tal y como se muestran en las figuras 17, 18 y 19.

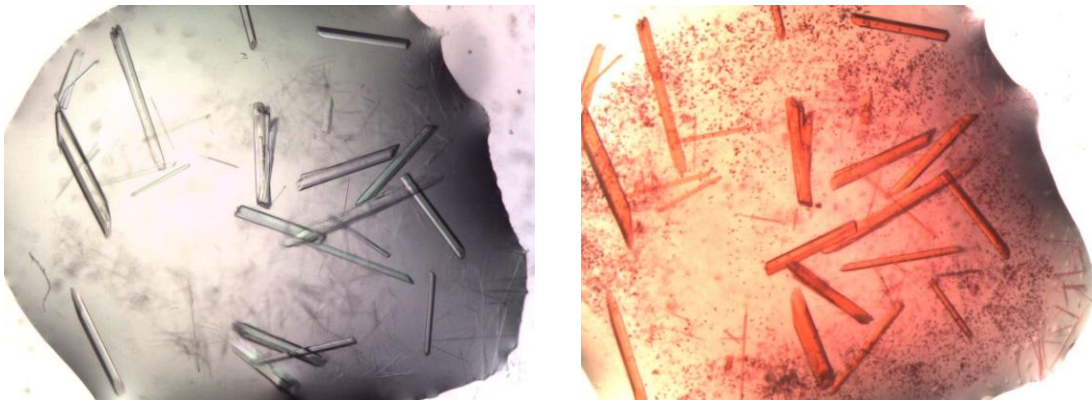


Figura 17. Cristales de Abl-SH3. En la imagen izquierda se muestran los cristales de Abl-SH3 antes de añadir la eosina. En la imagen derecha se muestran los mismos cristales tras 24 horas de la adición de eosina. Como se puede observar, los cristales tienen un color más intenso que la disolución, lo que quiere decir que la eosina ha difundido hacia el cristal, tiñéndolo.

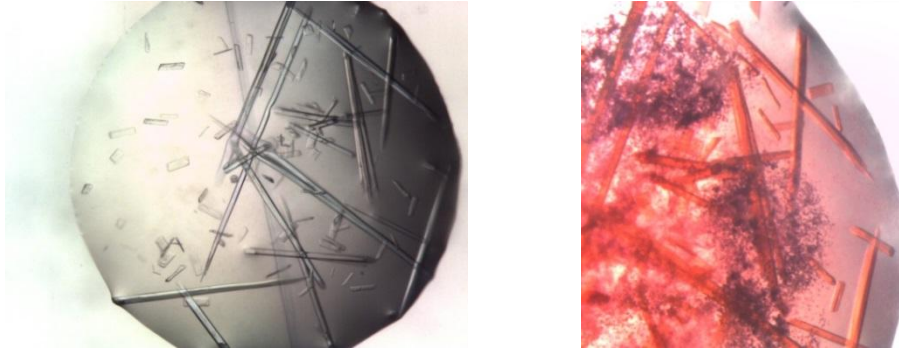


Figura 18. Cristales de Abl-SH3. En la imagen izquierda se muestran los cristales de Abl-SH3 antes de añadir la eosina. En la imagen derecha se muestran los mismos cristales tras 24 horas de la adición de eosina. Ocurre lo mismo que en el caso anterior.



Figura 19. Cristales de TSG101-UEV. En la imagen izquierda se muestran los cristales de TSG101-UEV antes de añadir la eosina. En la imagen derecha se muestran los mismos cristales tras 24 horas de la adición de eosina. Se puede observar que estos cristales también se tiñen.

De esta manera, se puede identificar fácilmente si el cristal obtenido es de sal o de proteína. Cabe destacar que la tinción de los cristales no afecta a su morfología ni tampoco a su calidad, ya que se han difractado cristales de lisozima teñidos con esta técnica y difractan a alta resolución.

5.3 Estructura de la ubiquitina

El cristal obtenido en el apartado 5.2.1 (figura 12) permitió resolver la estructura de la ubiquitina humana WT a alta resolución (1.32 Å) en una nueva forma cristalina (grupo espacial triclinico, P1). La novedad es el uso de cloruro de magnesio en la disolución precipitante, el cual favorece los contactos cristalinos. La unidad asimétrica está compuesta por dos monómeros de ubiquitina relacionados por un eje binario.

La estructura de la ubiquitina incluye una hélice α (residuos Ile23-Glu34), seis giros β (Thr7-Arg54, Glu18-Asp21, Ile44-Gly47, Phe45-Lys48, Glu51-Arg54 y Gln62-

Ser65), dos pequeños segmentos que forman una hélice 3_{10} entre los residuos Pro38-Gln40 y Ser57-Tyr59, una lámina β paralela formada por dos hebras en los residuos His1-Thr7 y Thr66-Leu71, y una lámina β antiparalela formada por tres hebras en los residuos Thr12-Val17, Gln41-Phe45 y Lys48-Gln49.

La red cristalina está formada por capas apiladas de moléculas de ubiquitina y por la presencia de dos interfaces, una mayor y otra menor, tal y como se muestra en las figuras 20 y 21.

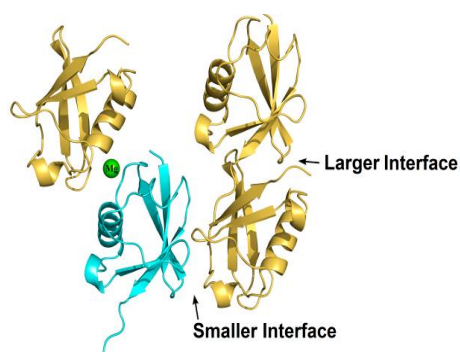


Figura 20. Se muestran las interfaces mayor y menor y la posición del catión de Mg (esfera de color verde) en una posición especial de simetría.

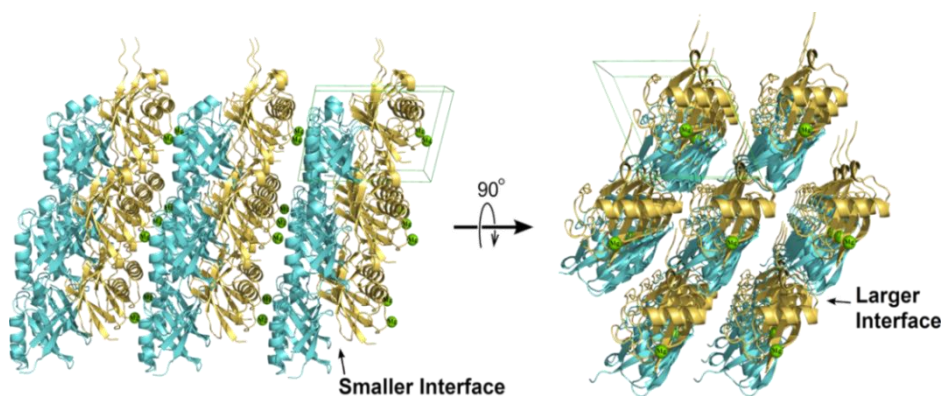


Figura 21. A la izquierda se muestra cómo la red cristalina está formada por capas apiladas de moléculas de ubiquitina humana. Las cadenas A y B se muestran en color celeste y dorado, respectivamente. El contorno de la celda unidad del grupo espacial triclinico aparece en verde, y los cationes de magnesio se muestran como esferas de color verde. La imagen de la derecha muestra la vista ortogonal del empaquetamiento.

La interfaz grande (704 \AA^2 de área superficial) está compuesta por la cadena A y por una molécula simétrica. Los contactos se estabilizan por once residuos en la cadena A y diez residuos de la molécula simétrica. Hay sólo dos enlaces de hidrógeno (Arg74 O-Gln49 NE2, Gly75 O-Gln49 N) y 45 contactos no enlazados. Estas interacciones son

principalmente hidrofóbicas y se estabilizan gracias a las cadenas laterales de los aminoácidos de las moléculas A o B de la unidad asimétrica y por aminoácidos parcialmente enterrados de la cadena simétrica (Val70, Leu8, Ile44, Leu73 y Leu 71). La interfaz menor (442 Å² de área superficial) muestra mayor carácter polar que la mayor e incluye cuatro enlaces de hidrógeno (Ala46A N-Ala46B O, Glu64A O-His68B NE2, Lys6A NZ-Asn60B O, Lys6A NZ-Gln62B OE1) y 54 contactos de van der Waals. Los residuos de histidina pertenecientes a la cola de expresión del vector pET-101D presente para optimizar el proceso de purificación, tal y como está explicado en el apartado 4.2.1) no están localizados ni en las interfaces ni cerca del catión de magnesio, por lo que no interfiere en el empaquetamiento del cristal.

El catión Mg²⁺ está coordinado con seis moléculas de agua formando el complejo [Mg(H₂O)₆]²⁺, el cual está unido a residuos que pertenecen a diferentes moléculas dentro de la unidad asimétrica. El contacto cristalino está mediado, por tanto, por el complejo [Mg(H₂O)₆]²⁺ y no se encuentra en ninguna de las principales interfaces de empaquetamiento. En su lugar, este ión metálico ocupa una posición entre las capas adyacentes del cristal, participando en contactos cristalinos muy importantes (figura 22). El catión magnesio tiende a coordinarse con moléculas de agua antes que con otros residuos u otros aniones en disolución. El número de coordinación usual del Mg²⁺ en estructuras cristalinas es de seis, formando hexa-aquo complejos.⁴⁴

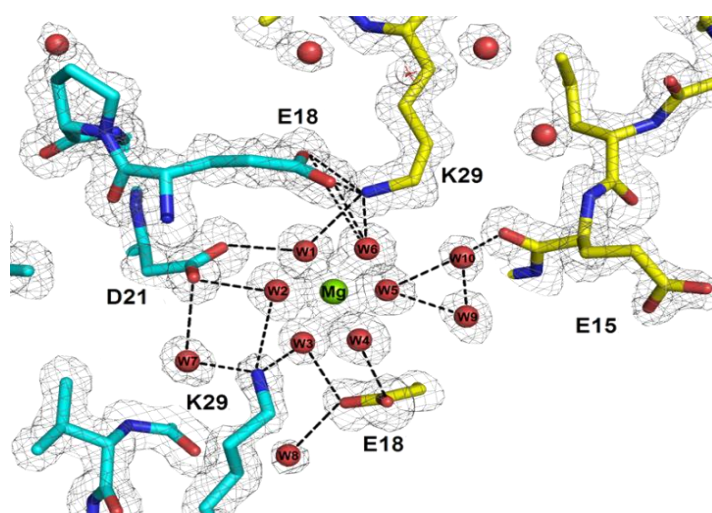


Figura 22. El catión magnesio se encuentra en una posición de simetría coordinando los residuos de las distintas cadenas en un contacto cristalino. Los residuos de la cadena A se muestran en color azul y los residuos de la molécula simétrica vecina son de color amarillo. El catión magnesio (verde) y las moléculas de agua (rojas) se muestran como esferas. Se puede observar que el catión magnesio estabiliza a la lisina 29 (K29) en el contacto cristalino.

Como se puede observar, el uso de cloruro de magnesio como aditivo en la condición de cristalización proporciona seis moléculas de agua ordenadas que pueden establecer interacciones con los residuos presentes en la superficie de la proteína, facilitando el empaquetamiento del cristal tal y como se muestra en la figura 23.

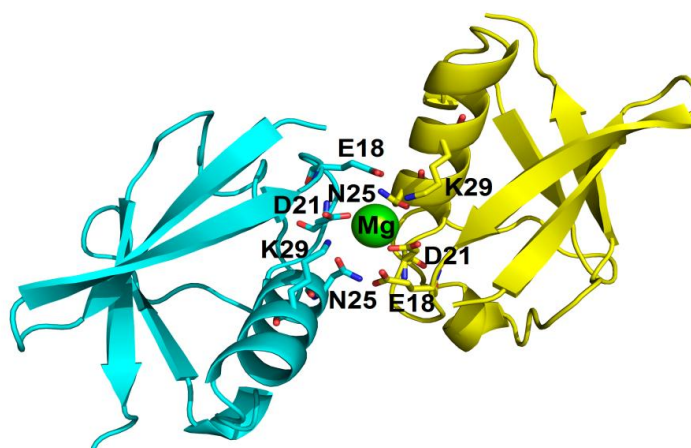


Figura 23. Empaquetamiento del cristal. El catión magnesio se representa como una esfera verde.

De esta manera, el empaquetamiento correcto de las moléculas de ubiquitina requiere la estabilización de algunos residuos que son muy flexibles y que además se encuentran en una posición crítica para permitir la formación de contactos cristalinos con sus moléculas vecinas. Esto es lo que ocurre en el caso de la Lys29, que está involucrada en el sitio de interacción del metal. LloI et al¹⁷ llevaron a cabo la mutación sistemática de todas las lisinas de la ubiquitina por serinas y observaron que la mutación de la Lys29 disminuía los hits de cristalización. Estos autores exponen por tanto que la Lys29 no está enterrada en el interior de la estructura de la proteína, sino que como ya se ha dicho, es un residuo superficial que interacciona con el átomo de oxígeno del esqueleto carbonílico del Glu16, así como con la cadena lateral del Asp21. Se atribuye el descenso en los hits de cristalización a la pérdida de interacciones intra-moleculares al mutar esa lisina por serina.

También se ha llevado a cabo una comparación de la coordinación del catión magnesio en nuestra estructura con otras donde están presentes iones metálicos. Esas estructuras han sido nombradas en el apartado 1.1.2 en presencia de Cd²⁺, Hg²⁺, Pt²⁺ y Zn²⁺. Esta es la primera estructura de ubiquitina monomérica WT cristalizada en presencia de magnesio y con grupo espacial P1. El Mg²⁺ se encuentra en una posición equivalente a la encontrada en estructuras obtenidas en el grupo espacial cúbico P4₃2 en presencia de Cd²⁺ y Zn²⁺ (códigos PDB 3EEC y 3N30, respectivamente). Merece la

pena mencionar que ambas estructuras fueron obtenidas a baja resolución (3 Å) haciendo difícil modelar el disolvente con precisión. De hecho, los cationes Cd^{2+} y Zn^{2+} no presentan su coordinación típica.⁴⁵

Las cadenas laterales grandes y flexibles son normalmente excluidas de los contactos cristalinos debido a la desfavorable pérdida de entropía durante el proceso de empaquetamiento.⁴⁶ La formación de nuevas interacciones fuertes alrededor del metal ocasionan un aumento de entalpía que puede compensar la pérdida de entropía. Así, este cristal obtenido a alta resolución da información acerca del rol del magnesio en la unión y cristalización de la ubiquitina humana a través de enlaces de hidrógeno mediados por el agua, entendiendo mejor por tanto cómo reduciendo la flexibilidad de algunos residuos el empaquetamiento en cristales se ve favorecido.

5.4 Análisis de las bases de datos

En este apartado se ha llevado a cabo un análisis de la efectividad de las bases de datos a la hora de predecir si una proteína es fácilmente cristizable o no según su estructura. Las bases de datos utilizadas se han descrito brevemente en el apartado 1.3.

Como se ha podido comprobar a lo largo de este Trabajo Fin de Máster, la ubiquitina es una proteína compleja de cristalizar y ha quedado demostrado experimentalmente. Esto se debe a que, tal y como se ha explicado en el apartado 1.1.2, la ubiquitina contiene siete residuos de lisina que incrementan mucho la entropía superficial de la misma. Por ello se utilizan cationes metálicos divalentes que, como se ha visto en el apartado 5.3, se colocan en el contacto cristalino en el que está presente la Lys29. La búsqueda realizada con estas bases de datos consiste en analizar el resultado que ofrece cada una de ellas de la ubiquitina WT y de cada una de las mutaciones de lisina por serina, que en teoría, deberían mejorar la cristalizabilidad de la misma.

XtalPred

Esta base de datos utiliza dos algoritmos distintos a la hora de analizar a la proteína: Expert Pool (EP) y Random Forest (RF). Expert Pool tiene en cuenta los siguientes parámetros: longitud de la proteína, punto isoeléctrico, desorden estructural, índice de inestabilidad y estructura secundaria de la proteína. Random Forest tiene en cuenta otros parámetros, tales como: entropía superficial de las cadenas laterales, hidrofobicidad y composición de los aminoácidos en la superficie.

A simple vista, es de esperar que RF sea mucho más preciso a la hora de predecir la cristalizabilidad de la ubiquitina que EP, ya que tiene en cuenta la entropía superficial de la misma. La ubiquitina se caracteriza por los siete residuos de lisina que incrementan en gran medida su entropía superficial.

En primer lugar se ha analizado la ubiquitina WT, sin ningún tipo de mutación, obteniéndose los resultados mostrados en la figura 24.

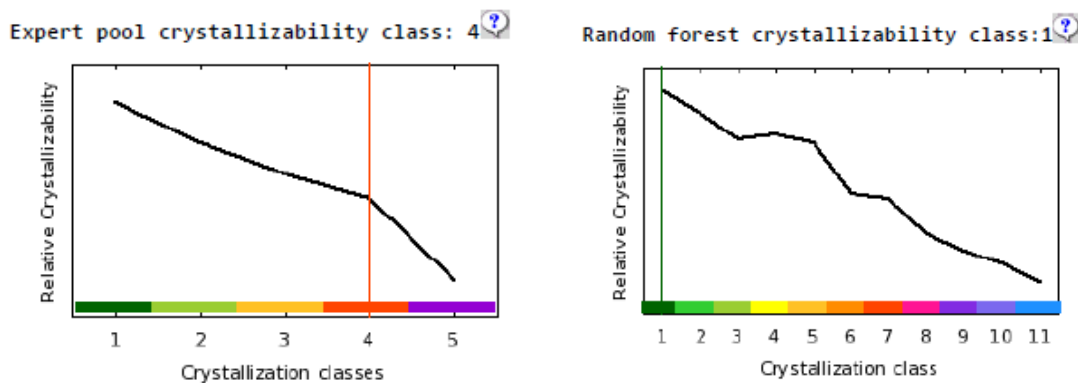


Figura 24. En ambos casos un valor de 1 indica que la proteína es fácilmente cristalizable. En EP un valor de 5 indica que la proteína es poco prometedora a la hora de cristalizar y en RF ocurre lo mismo.

Como se puede observar, EP da un valor de 4 para la ubiquitina WT, mientras que RF da un valor de 1, lo que indica que la proteína es muy sencilla de cristalizar cuando se ha demostrado experimentalmente que no es así.

Este mismo procedimiento se ha llevado a cabo para cada una de las mutaciones de la ubiquitina (K6S, K11S, K27S, K29S, K33S, K48S, K63S) y para el doble mutante K11S/K63S. Los resultados obtenidos fueron idénticos a los de la ubiquitina WT mostrados en la figura 24. Esto indica que la base de datos no es capaz de identificar las mutaciones ni el cambio que producen en la cristalizabilidad de la proteína a pesar de que experimentalmente Llol et al¹⁷ han demostrado que las mutaciones de cada una de las lisinas por serinas (figura 25) afecta a la facilidad con la que cristaliza la misma. Para ello han analizado los hits de cristalización de la ubiquitina WT y de las distintas mutaciones de las lisinas por serinas. Cabe destacar que en este experimento los hits de cristalización son todos los resultados positivos en los que se han obtenido cristales, sin tener en cuenta si éstos son de proteína o no.

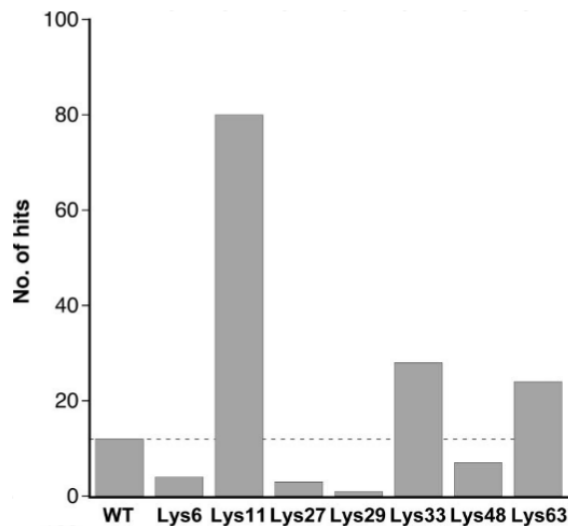


Figura 25. Hits de cristalización obtenidos para la ubiquitina WT y las siete mutaciones de lisinas por serinas. Se probaron 384 condiciones diferentes de screens comerciales.

También se ha hecho lo mismo utilizando la secuencia de nuestra ubiquitina, que contiene una cola de seis histidinas para facilitar el proceso de purificación durante el paso por columna de afinidad de Ni²⁺-NTA Agarosa descrito en el apartado 4.2.1. El resultado de la predicción se muestra en la figura 26.

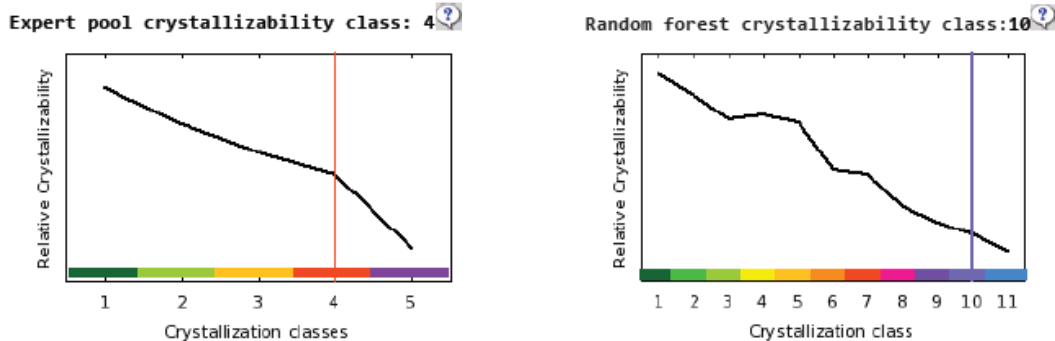


Figura 26. Resultado de la ubiquitina con la cola de seis histidinas.

En este caso, RF predice que la ubiquitina con la cola de seis histidinas es menos cristizable que la ubiquitina WT sin dicha cola. Las seis histidinas aumentan la solubilidad de la proteína, por lo que teóricamente deberían mejorar la cristalizabilidad de la misma y no empeorarla. Una proteína más soluble es más sencilla de cristalizar puesto que si ésta es muy insoluble alcanzará la zona de sobresaturación muy rápido (diagrama de fases mostrado en la figura 6) y no se formarán micro cristales, sino que aparecerá un precipitado.

Estudiando los resultados obtenidos, se puede observar que el algoritmo de EP predice que la ubiquitina no es nada sencilla de cristalizar, sin embargo, RF da valores bastante optimistas acerca de la cristalizabilidad de la misma, valores que por otra parte no son reales tal y como se ha comprobado experimentalmente. Tampoco es capaz de distinguir las mutaciones producidas entre lisinas y serinas. Esto puede deberse a que la base de datos sólo tiene en cuenta como aminoácidos “problemáticos” a las cisteínas, metioninas, triptófanos, tirosinas y fenilalaninas, es decir, a los aminoácidos hidrofóbicos. Si estos aminoácidos se encuentran en la superficie de la proteína disminuyen la solubilidad de la misma, por lo que es necesario tenerlos en cuenta para predecir si la proteína es sencilla o no de cristalizar. Sin embargo, los residuos de lisina son muy flexibles e incrementan mucho la entropía superficial de la ubiquitina, dificultando su cristalización. Como XtalPred no tiene en cuenta este aminoácido como “problemático” la predicción que da es errónea.

CrystalP2

Esta base de datos proporciona información de si la proteína es o no cristizable. Los parámetros que tiene en cuenta son la hidrofobicidad de la misma (basándose en los mismos aminoácidos que XtalPred), punto isoeléctrico y la cantidad de aminoácidos. Ofrece una hoja de resultados bastante sencilla y escueta con un tanto por ciento de cristalizabilidad de la proteína.

Para el caso de la ubiquitina WT se obtuvo el resultado que se muestra en la figura 27.

CRYSTALP2 RESULTS PAGE

Results for **CRYSTALP2** webservice.

Use this link to download the results as a CSV file: [RESULTS.CSV](#)

Results

The higher/lower the confidence the more probable that protein is crystallizable/non-crystallizable.

- Protein 1UBI_A_PDBID_CHAIN is non-crystallizable with 0.434 confidence.

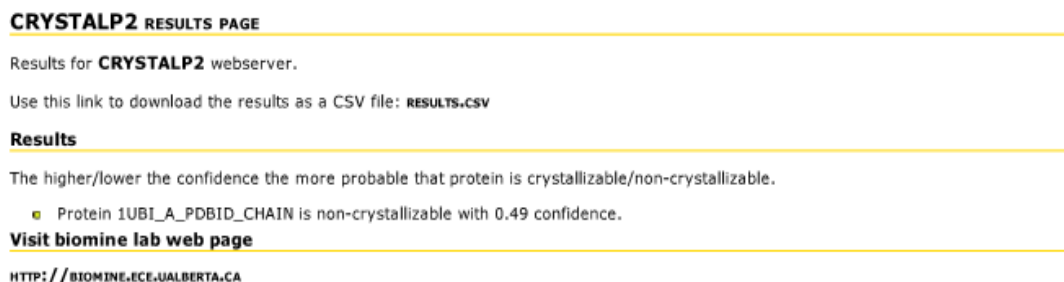
Visit biomine lab web page

[HTTP://BIOMINE.ICE.UALBERTA.CA](http://biomine.ice.ualberta.ca)

Figura 27. Resultado ofrecido por CrystalP2. La ubiquitina es no cristizable con un 43.4 % de confianza.

Para el resto de mutaciones de la ubiquitina (K6S, K11S, K27S, K29S, K33S, K48S, K63S) y para el doble mutante K11S/K63S el resultado es idéntico al obtenido en la figura 22. Por tanto, se observa el mismo problema, no distingue la mutación.

Para el caso de nuestra ubiquitina con la cola de seis histidinas, se observó un ligero cambio en el valor ofrecido por la base de datos, tal y como se muestra en la figura 28.



CRYSTALP2 RESULTS PAGE

Results for **CRYSTALP2** webserver.

Use this link to download the results as a CSV file: **RESULTS.CSV**

Results

The higher/lower the confidence the more probable that protein is crystallizable/non-crystallizable.

- Protein 1UBI_A_PDBID_CHAIN is non-crystallizable with 0.49 confidence.

Visit biomine lab web page

[HTTP://BIOMINE.ECE.UALBERTA.CA](http://biomine.ece.ualberta.ca)

Figura 28. Resultado ofrecido por CrystalP2. La ubiquitina con la cola de histidinas es no cristalizabile con un 49 % de confianza.

Se puede observar que los resultados obtenidos son similares a los que ofrece XtalPred, ya que tienen en cuenta parámetros muy parecidos y además ninguno de ellos tiene en cuenta a las lisinas y al aumento de entropía superficial a la hora de cristalizar la proteína. Tampoco tiene sentido que la ubiquitina sea más difícil de cristalizar con una cola de histidinas porque como se ha dicho anteriormente ésta aumenta la solubilidad de la misma. Sin embargo, esta base de datos se corresponde un poco más con la realidad puesto que predice que la ubiquitina es no cristalizabile con un mayor o menor tanto por ciento de confianza.

PxS

Esta base de datos es una herramienta que genera un análisis estadístico de las propiedades físicas que controlan la cristalización de una proteína a partir de la secuencia individual de la misma. Utiliza un parámetro denominado P_{xs} que se define como la probabilidad de obtener un cristal de proteína a partir de una disolución monodispersa de ésta. Este parámetro se puede calcular de la siguiente forma $P_{xs} = \frac{e^{\theta}}{1+e^{\theta}}$, donde θ es un factor $\theta = 1,85 + (14,3 \cdot F) + (8,14 \cdot G_{pb}) - (3,77 \cdot SCE_{pe}) - (3,2 \cdot Diso)$. F es la fracción de fenilalaninas, G_{pb} la fracción de glicinas enterradas, SCE_{pe} es la entropía que se predice de los residuos expuestos a la superficie y Diso es la fracción de residuos desordenados.

El procedimiento llevado a cabo para analizar esta base de datos es el mismo que en los casos anteriores. Los resultados obtenidos para la ubiquitina WT se muestran en la siguiente figura 29:

$$P_{xs} = 0.140$$

P_{xs} is the probability of obtaining a protein crystal structure from a monodisperse preparation of a soluble bacterial protein.

The metric is likely to be valid for well ordered domains from any source organism, but could overestimate the crystallization propensity of eukaryotic proteins containing disordered regions because these have different sequence properties compared to the mostly bacterial proteins used to develop the P_{xs} metric. See the relevant section in [Price et al.](#) and the corresponding Supplementary [Figs. 16-19](#).

$$P_{xs} = \exp(\Theta) / (1 + \exp(\Theta))$$

where

$$\Theta = 1.85 + (14.3 \times F) + (8.14 \times G_{pb}) - (3.77 \times SCE_{pe}) - (3.2 \times Diso)$$

For this protein

F - Fraction of phenylalanine: 0.0263
 G_{pb} - Fraction of glycine among buried residues: 0
 SCE_{pe} - Mean side chain entropy of predicted exposed residues: ?
 Diso - Fraction of predicted disordered residues: 0.0132

$$P_{C-XS-Hs} = 0.243$$

P_{C-XS-Hs} is the probability of obtaining a crystal structure from a human genomic protein sequence conflating expression, purification, and crystallization steps. It is a linear combination of several factors and is described in detail in the [Supplementary Notes and Supplementary Fig. 20](#)

Figura 29. La base de datos muestra de esta forma los resultados obtenidos. La ubiquitina WT tiene un P_{xs} de 0,140

Como se puede observar, en este caso la base de datos no es capaz de definir un valor de SCE_{pe} por lo que la predicción que ofrece podría no ser del todo precisa. Este valor es bastante importante para el caso de la ubiquitina porque hace referencia a la entropía superficial de la misma.

Para el resto de mutaciones (K6S, K11S, K27S, K29S, K33S, K48S, K63S) se obtiene el mismo resultado (figura 30) entre ellos pero distinto al obtenido para la ubiquitina WT.

$$P_{xs} = 0.155$$

P_{xs} is the probability of obtaining a protein crystal structure from a monodisperse preparation of a soluble bacterial protein.

The metric is likely to be valid for well ordered domains from any source organism, but could overestimate the crystallization propensity of eukaryotic proteins containing disordered regions because these have different sequence properties compared to the mostly bacterial proteins used to develop the P_{xs} metric. See the relevant section in [Price et al.](#) and the corresponding Supplementary [Figs. 16-19](#).

$$P_{xs} = \exp(\Theta) / (1 + \exp(\Theta))$$

where

$$\Theta = 1.85 + (14.3 \times F) + (8.14 \times G_{pb}) - (3.77 \times SCE_{pe}) - (3.2 \times Diso)$$

For this protein

F - Fraction of phenylalanine: 0.0263
 G_{pb} - Fraction of glycine among buried residues: 0
 SCE_{pe} - Mean side chain entropy of predicted exposed residues: 1.028
 Diso - Fraction of predicted disordered residues: 0.0132

$$P_{C-XS-Hs} = 0.196$$

P_{C-XS-Hs} is the probability of obtaining a crystal structure from a human genomic protein sequence conflating expression, purification, and crystallization steps. It is a linear combination of several factors and is described in detail in the [Supplementary Notes and Supplementary Fig. 20](#)

Figura 30. En este caso se puede observar como las mutaciones tienen un valor de P_{xs} de 0,155. También da un valor de SCE_{pe} de 1,028.

En el caso del doble mutante K11S/K63S también se observa una diferencia en la probabilidad de cristalización y en el factor SCE_{pe} (figura 31).

$$P_{xs} = 0.165$$

P_{xs} is the probability of obtaining a protein crystal structure from a monodisperse preparation of a soluble bacterial protein.

The metric is likely to be valid for well ordered domains from any source organism, but could overestimate the crystallization propensity of eukaryotic proteins containing disordered regions because these have different sequence properties compared to the mostly bacterial proteins used to develop the P_{xs} metric. See the relevant section in [Price et al.](#) and the corresponding Supplementary [Figs. 16-19](#).

$$P_{xs} = \exp(\theta) / (1 + \exp(\theta))$$

where

$$\theta = 1.85 + (14.3 \times F) + (8.14 \times G_{pb}) - (3.77 \times SCE_{pe}) - (3.2 \times Diso)$$

For this protein

$$F - \text{Fraction of phenylalanine: } 0.0267$$

$$G_{pb} - \text{Fraction of glycine among buried residues: } 0$$

$$SCE_{pe} - \text{Mean side chain entropy of predicted exposed residues: } 1.0198$$

$$Diso - \text{Fraction of predicted disordered residues: } 0$$

$$P_{C-XS-Hs} = 0.178$$

$P_{C-XS-Hs}$ is the probability of obtaining a crystal structure from a human genomic protein sequence conflating expression, purification, and crystallization steps. It is a linear combination of several factors and is described in detail in the [Supplementary Notes and Supplementary Fig. 20](#)

Figura 31. Como se puede observar, para el doble mutante el valor de P_{xs} es de 0,155 y el de de SCE_{pe} es de 1,0198.

Por último, también se ha llevado a cabo el análisis de la ubiquitina con la cola de seis histidinas obteniendo también unos resultados diferentes a los anteriores con un valor de P_{xs} de 0,111 y un valor de SCE_{pe} de 1,0907.

Analizando los resultados, se puede decir que esta base de datos se acerca un poco más a la realidad experimental. En ninguno de los casos la probabilidad de cristalización es alta por lo que se confirma que la ubiquitina es una proteína difícil de cristalizar. También se observa el hecho de que distingue la mutación de una lisina por serina, observándose un aumento en la probabilidad de cristalización. Sin embargo, a pesar de que considera la introducción de esta mutación, no es capaz de distinguir la posición en la que se produce y por eso todos los mutantes presentan el mismo valor de P_{xs} y de SCE_{pe} . Para el doble mutante se obtiene una probabilidad de cristalización aún mayor, por tanto, se puede deducir que para esta base de datos, a medida que se van sustituyendo las lisinas por serinas se produce un aumento en el valor de P_{xs} . En el caso de la ubiquitina con la cola de histidinas la probabilidad disminuye a pesar de que la cola de histidinas aumenta, como ya se ha dicho, la solubilidad de la misma.

En esta base de datos, el criterio que marca la diferencia entre las distintas secuencias de ubiquitina es el SCE_{pe} , que tiene en cuenta la entropía superficial de la proteína debido a los residuos expuestos a la superficie. En el caso de la ubiquitina, esos residuos son los de lisina y es por ese motivo por el cual este valor va disminuyendo a medida que se van eliminando lisinas de la proteína.

Crysalis

Crysalis es una base de datos desarrollada para llevar a cabo predicciones rápidas sobre ciertas características de una proteína. Entre estas características, puede predecir si la proteína objetivo va a superar con éxito cuatro procedimientos experimentales: clonación de su secuencia (CLF), purificación (PF), cristalización (CF) y determinación estructural (CRYs). De esta manera, Crysalis se puede utilizar para identificar el “cuello de botella” del procedimiento experimental cuando se quiere cristalizar una determinada proteína. Para llevar a cabo todo esto, la base de datos utiliza un algoritmo de dos etapas. En la primera etapa se lleva a cabo el método mRMR (mínima-redundancia y máxima-relevancia)⁴⁷ para reducir dentro de lo posible la información redundante e irrelevante para que, en la segunda etapa, se de una información completa y precisa de todos los parámetros expuestos anteriormente. De esta manera, se lleva a cabo un filtrado de la información disponible en la base de datos.

El procedimiento llevado a cabo fue el mismo que el utilizado en los apartados anteriores. En este caso, se obtuvieron los mismos resultados tanto para la ubiquitina WT como para los mutantes K6S, K11S, K27S, K29S, K33S, K48S, K63S y para el doble mutante K11S/K63S. A continuación, se muestra la predicción realizada por la base de datos (figura 32).

Your query sequence

```
>1UBI:A|PDBID|CHAIN|SEQUENCE_  
MQIFVKLTGKTTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRGG
```

Prediction results

Experimental procedure	CLF	MF	PF	CF	CRYs
Crysalis I (first-level predictor)	success, 0.572	success, 0.621	success, 0.648	failure, 0.402	success, 0.735
Crysalis II (second-level predictor)	success, 0.539	success, 0.644	success, 0.671	failure, 0.394	success, 0.587

CLF: Predicted propensity of sequence cloning:

MF: Predicted propensity of protein material production:

PF: Predicted propensity of protein purification:

CF: Predicted propensity of protein crystallization:

CRYs: Predicted propensity of protein diffraction quality crystals.

The prediction cutoff score was set as 0.5.

Figura 32. Resultado obtenido en la base de datos Crysalis. La ubiquitina no es propensa a cristalizar con un valor de 0.394 de confianza. El resto de parámetros son positivos y estiman que la proteína se puede clonar y purificar fácilmente y que además si se obtienen cristales, éstos serán de buena calidad para difractar.

Analizando los resultados se puede decir que Crysalis se acerca bastante a la realidad ya que estima que la ubiquitina no tiende a cristalizar. Otro parámetro que se puede demostrar experimentalmente es el hecho de que la proteína se puede purificar con éxito, hecho que es cierto ya que la ubiquitina es una proteína que se expresa muy bien en *E.coli*, obteniendo siempre un rendimiento bastante alto. A pesar de esto, ocurre lo mismo que en casos anteriores, ofrece los mismos valores para todas las mutaciones, por lo que no distingue entre lisinas y serinas y mucho menos entre la posición de la mutación.

Como podemos ver en el caso de las bases de datos para predecir si una proteína es o no cristizable, cuando la hemos aplicado a una proteína como la ubiquitina no funcionan correctamente. Experimentalmente se ha demostrado que la ubiquitina no cristaliza debido a sus problemas de nucleación (causados por la alta entropía superficial de los siete residuos de lisina) y además, las mutaciones afectan en gran medida y las bases de datos no son capaces de distinguirlas. Otro problema muy grave es que ninguna de estas herramientas tiene en cuenta el disolvente presente en el medio, cuando se ha estudiado experimentalmente que este factor es clave a la hora de cristalizar una proteína. Por ejemplo, en el caso de la ubiquitina, se ha comprobado que los cationes metálicos divalentes se colocan en un contacto cristalino y juegan un papel clave en su cristalización.¹²

Por otro lado, como se ha comentado brevemente en el apartado 1.3, existe el problema de que la mayoría de las bases de datos que aparecen en bibliografía (SVMCrys, ParCrys, PPCPred, OB-Score y MetaPPCP) ya no funcionan correctamente o están totalmente fuera de servicio, lo que indica que muchas de ellas han sido creadas para llevar a cabo algún experimento puntual en algún laboratorio, ya sea para realizar una tesis doctoral o para publicar algún artículo y que, alcanzados dichos objetivos, no tienen ningún tipo de soporte. Además, al pedir explicaciones o ayuda a la dirección de correo electrónico de los administradores o moderadores de las webs, no se obtiene ningún tipo de respuesta, lo que confirma el completo abandono de la base de datos.

Hay que tener en cuenta que la cristalización de una proteína no es tan “sencilla” como puede parecer, ya que hay que conseguir que la proteína se ordene en el espacio de manera que se establezcan el número de interacciones necesarias para formar una red tridimensional ordenada que da lugar al cristal. Esto depende de muchísimos factores,

tal y como se ha explicado en el apartado de métodos 4.2.3 y son estos factores los que difícilmente pueden tener en cuenta una base de datos a partir simplemente de la secuencia de aminoácidos de una proteína.

6. Conclusiones/Conclusions

1. Se ha purificado la proteína objeto de estudio y se ha logrado mejorar el rendimiento del protocolo. Se ha comprobado mediante electroforesis en SDS-PAGE que la pureza de la ubiquitina era la exigida en estudios cristalográficos, donde se requiere un grado de pureza muy alto.
 2. Se ha estudiado el efecto del magnesio en el proceso de cristalización de la ubiquitina observando claramente su influencia en un contacto cristalino gracias a la determinación de la estructura mediante la difracción de rayos X.
 3. Se han analizado las bases de datos descritas en bibliografía que predicen la cristalizabilidad de una proteína y se ha comprobado su efectividad con la proteína objeto de estudio, obteniendo resultados muy dispares que no siempre se pueden relacionar con la realidad experimental.
1. The protein under study has been purified, improving the performance of the protocol. It has been verified by SDS-PAGE that the purity of ubiquitin was the required one in crystallographic studies, where a very high degree of purity is required.
 2. The effect of magnesium has been studied in the crystallization process of ubiquitin, observing its influence in a crystal contact thanks to the structural determination by X-ray diffraction.
 3. The databases described in literature that predict the crystallizability of a protein have been analyze and its effectiveness has been proven with the protein under study, obtaining very different results that can not always be related to experimental reality.

7. Bibliografía

- ¹ Shields, S. B.; Piper, R. C. *Traffic* **2011**, *12*, 1306.
- ² King, R. W.; Deshaies, R. J.; Peters, J. M.; Kirschner, M. W. *Science* **1996**, *274*, 1652-1659.
- ³ Di Fiore, P. P.; Polo, S.; Hofmann, K. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 491-497.
- ⁴ Kostova, Z.; Wolf, D. H. *The EMBO Journal* **2003**, *22*, 2309-2317.
- ⁵ Hershko, A.; Ciechanover, A. *Biochem* **1998**, *67*, 452-479.
- ⁶ Hicke, L.; Dunn, R. *Annu. Rev. Cell Dev. Biol.* **2003**, *19*, 141-172.
- ⁷ Buchberger, A. *Trends Cell Biol.* **2002**, *12*, 216-221.
- ⁸ Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic acid Res.* **2000**, *42*, 292-296.
- ⁹ Falini, G.; Fermani, S.; Tosi, G.; Arnesano, F.; Natile, G. *Chem. Commun.* **2008**, 5960-5962.
- ¹⁰ Arnesano, F.; Belviso, B. D.; Caliandro, R.; Falini, G.; Fermani, S.; Natile, G.; Siliqi, D. *Chem. Eur.* **2011**, *17*, 1569-1578.
- ¹¹ Ma, P.; Xue, Y.; Coquelle, N.; Haller, J. D.; Yuwen, T.; Ayala, I.; Mikhailovskii, O.; Willbold, D.; Colletier, J.-P.; Skrynnikov, N. R.; Schanda, P. *Nature Commun.* **2015**, *6*, 8361.
- ¹² Cámara-Artigas, A.; Plaza-Garrido, M.; Martínez-Rodríguez, S.; Bacarizo, J. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2015**, *72*, 29-35.
- ¹³ Huang, K. Y.; Amodeo, G. A.; Tong, L.; McDermott, A. *Protein Sci.* **2011**, *20*, 630-639.
- ¹⁴ Datta, A. B.; Hura, G. L.; Wolberger, C. *J. Mol. Biol.* **2009**, *392*, 1117-1124
- ¹⁵ Ramage, R.; Green, J.; Muir, T. W.; Ogunjobi, O. M.; Love, S.; Shaw, K. *Biochem. J.* **1994**, *299*, 151-158

- ¹⁶ Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J.Mol.Biol.* **1987**, *194*, 531-544
- ¹⁷ Loll, P. J.; Xu, P.; Schmidt, J. T.; Melideo, S. L. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2014**, *70*, 1434-1442.
- ¹⁸ Rupp, B. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology* **2009**.
- ¹⁹ McPherson, A.; Gavira, J. A. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2014**, *70*, 2.
- ²⁰ Cooper D. R.; Boczet T.; Grelewska K., et al. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2007**, *63*, 636-645.
- ²¹ Kurgan, L.; Razib, A. A.; Aghakhani, S.; Dick, S.; Mizianty, M.; Jahandideh, S. *BMC Struct. Biol.* **2009**, *9*, 50.
- ²² Slabinski, L.; Jaroszewski, L.; Rychlewski, L.; Wilson, I. A.; Lesley, S. A.; Godzik, A. *Bioinformatics* **2007**, *23*, 3403-3405.
- ²³ Rodrigues, A.; Hubbard, R. E. *Bioinformatics* **2003**, *4*, 150-167.
- ²⁴ Kurgan, L.; Mizianty, M. J. *Nat. Sci.* **2009**, *1*, 93-106.
- ²⁵ Kantardjieff, K. A., Rupp, B. *Bioinformatics* **2004**, *20*, 2162-2168.
- ²⁶ Kantardjieff, K. A., et al. *Bioinformatics*, **2004**, *20*, 2171-2174.
- ²⁷ Price, W. N.; Chen, Y.; Handelman, S. K.; et al. *Nat. Biotechnol.* **2009**, *27*, 51-57.
- ²⁸ Chen K.; Kurgan L.; Rahbari M. *Biochem. Biophys. Res. Comm.* **2007**, *355*, 764-769
- ²⁹ Wang, H.; Feng, L.; Zhang, Z.; et al. *Sci. Rep.* **2016**, *6*.
- ³⁰ Kandaswamy K. K.; Pugalenti G.; Suganthan P.; Gangal R. *Prot. Pept. Lett.* **2010**, *17*, 423-430.
- ³¹ Overton, I.M.; Padovani, G.; Girolami, M. A.; Barton, G. J. *Bioinformatics* **2008**, *24*, 901-907
- ³² Mizianty M. J.; Kurgan L. *Bioinformatics* **2011**, *13*, 24-33.

- ³³ Overton I. M.; Barton G. J. *FEBS Lett.* **2006**, 580, 4005-4009.
- ³⁴ Svensson, O.; Malbet-Monaco, S.; Popov, A.; Nurizzo, D. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015**, 71, 1757-1767.
- ³⁵ Vorhein, C.; Flensburg, C.; Keller, P.; et al. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, 67, 293-302.
- ³⁶ Evans, P. R.; Murshudov, G. N. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, 69, 1204-1214.
- ³⁷ Adams, P. D., Alfonine, P. V., Bunkóczi, G., et al. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, 66, 213-221.
- ³⁸ Bunkóczi, G., Echols, N., McCoy, A. J., Oeffner, R. D., Adams P. D., Read, R. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, 69, 2276-2286.
- ³⁹ Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, 66, 486-501.
- ⁴⁰ Afonine, P. V.; Grosse-Kunstleve, R. W.; Caliandro, R.; et al. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, 68, 352-367.
- ⁴¹ Kabsch, W. *Acta Crystallogr., Sect. A* **1976**, 32, 922-923.
- ⁴² Winn, M. D. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, 67, 235-242.
- ⁴³ Chen, V. B.; Arendall, W. B.; Headd, J. J.; et al. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, 66, 12-21.
- ⁴⁴ Bock, C. W.; Kaufman, A.; Glusker, J. P. *Inorg. Chem.* **1994**, 33, 419-427.
- ⁴⁵ Harding, M. M.; Nowicki, M. W.; Walkinshaw, M. D. *Crystallogr. Rev.* **2010**, 16, 247-302.
- ⁴⁶ Derewenda, Z. S.; Vekilov, P. G. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, 62, 116-124.
- ⁴⁷ Ding, C.; Peng, H. *J. Bioinform. Comput. Biol.* **2005**, 3, 185-205.