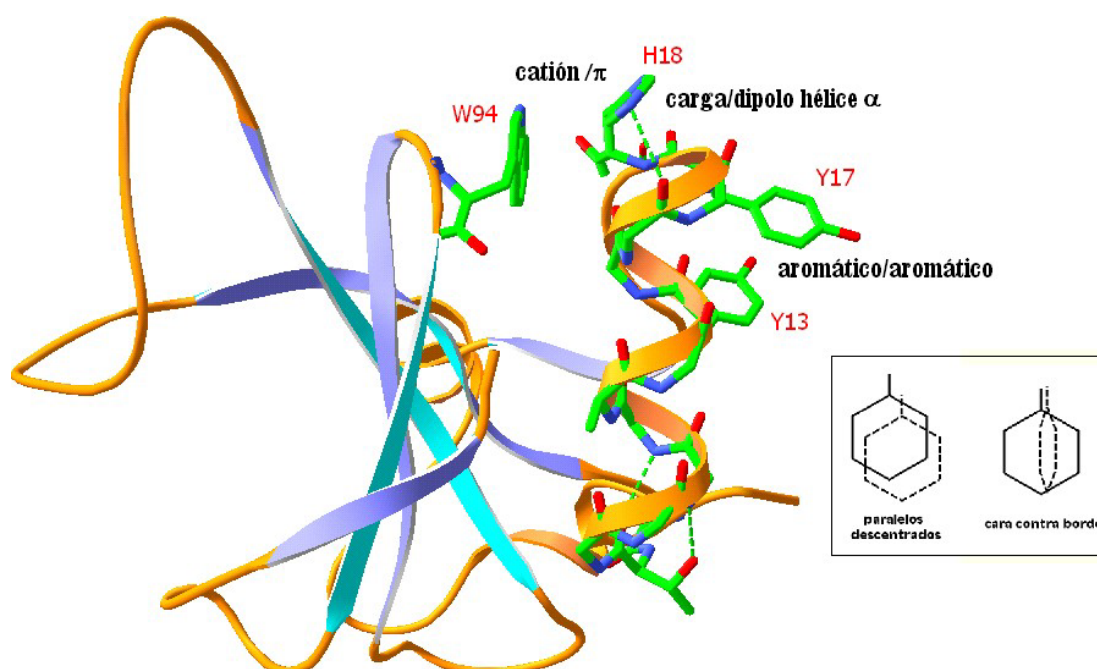


Guión de Prácticas de la asignatura

Estructura de Macromoléculas

del Grado de Biotecnología

Curso 2026-2027



Javier Sancho jsancho@unizar.es
Marta Martínez mmartine@unizar.es

Departamento de Bioquímica y Biología Molecular y Celular. Facultad de Ciencias &
Instituto de Biocomputación y Física de Sistemas Complejos (BIFI).
Universidad de Zaragoza

Práctica nº 1: Obtención de coordenadas de macromoléculas del PDB y visualización de su estructura tridimensional

1. INTRODUCCION

Las coordenadas XYZ de los átomos de una macromolécula son la información necesaria para representar su estructura. Las coordenadas de las macromoléculas cuya estructura tridimensional se ha resuelto por difracción de Rayos X o resonancia magnética nuclear se encuentran depositadas en el **Protein Data Bank (PDB)**. Al PDB se accede fácilmente por Internet: <http://www.rcsb.org/>. Una vez ahí se introduce el código PDB de la proteína deseada o se realiza una búsqueda por palabras (en inglés).

2. VISUALIZACION DE ESTRUCTURAS DE PROTEINAS CON EL PROGRAMA SPDBVIEWER

SPDBviewer (<http://us.expasy.org/spdbv/>) es un programa gratuito que permite representar la estructura tridimensional de moléculas utilizando coordenadas con formato pdb. Se puede utilizar, como una aplicación independiente al margen de Internet, abriéndolo y cargando las coordenadas de la molécula, o se puede utilizar como un programa auxiliar del navegador de internet que se activa automáticamente cada vez que hace falta. Para esta práctica y para la práctica 2, descargaremos a nuestro ordenador de trabajo la aplicación (versión 4.1 para MS Windows), la guía de usuario (User Guide.pdf) y el tutorial (Tutorial Material).

Una vez cargada una molécula en **SPDBviewer** (abriendo desde el menú de la aplicación el archivo .pdb correspondiente a la estructura que se desea visualizar) podemos verla de diversas maneras: como varillas, como bolas y varillas, como bolas, sólo los carbonos alfa de cada aminoácido, sólo unas cintas que representan el discurrir del polipéptido, etc.

La molécula se puede rotar, trasladar y ampliar o reducir convenientemente. Se puede ver los puentes de hidrógeno, calcular distancias, los vecinos de un residuo, las cavidades de la molécula, etc. Las posibilidades son muchas y hay que descubrirlas explorando el menú o leyendo el manual.

3. EJERCICIOS

- ⇒ **1nfn** ¿Qué proteína es? ¿A qué tipo estructural pertenece? ¿Cuál es el aminoácido C-terminal de la hélice N-terminal?
- ⇒ **1flv** ¿Qué proteína es? ¿A qué tipo estructural pertenece? ¿Qué aminoácidos aromáticos están en contacto con el grupo prostético FMN?
- ⇒ **1hti** ¿Qué proteína es? ¿A qué tipo estructural pertenece? ¿Cuántas metioninas contiene?
- ⇒ **1aqb** ¿Qué proteína es? ¿A qué tipo estructural pertenece? ¿Cuál es su cofactor? ¿Qué aminoácidos polares se encuentran en contacto con el cofactor?
- ⇒ ¿Cuántos mutantes de la lisozima del fago T4 se han cristalizado?
- ⇒ Visualizar algún complejo proteína/DNA
- ⇒ Visualizar alguna molécula de DNA
- ⇒ Visualizar algún carbohidrato
- ⇒ Comparar la estructura de la ribonucleasa barnasa resuelta por difracción de rayos X (**1a2p**) con la determinada por RMN (**1bnr**).

Práctica nº 2: Resolución estructural de una proteína mediante cristalografía de rayos X: cristalización y refinamiento

2.1 Cristalización de lisozima

1. INTRODUCCION

La característica principal de un cristal es su estructura interna periódica y ordenada en tres dimensiones. El hecho de que un sistema renuncie a sus grados de libertad parece una contradicción desde el punto de vista entrópico y, sin embargo, la cristalización de moléculas es un fenómeno espontáneo y reversible cuyos parámetros cinéticos y termodinámicos dependen de las propiedades químicas y físicas del solvente y del soluto involucrados. La cristalización se produce cuando, bajo ciertas condiciones (de supersaturación), el sistema se dirige hacia el estado de equilibrio entre dos fases, una soluble y una sólida. A pesar de que las moléculas individuales fijan su posición, disminuyendo así la entropía del sistema, establecen al mismo tiempo nuevas interacciones reduciendo la energía libre total del sistema. Las moléculas, por tanto, cristalizan minimizando su energía libre al disponerse periódicamente en unidades repetitivas y simétricas en el estado sólido, maximizando las interacciones intermoleculares favorables.

La solvatación es un factor vital para la cristalización de macromoléculas. En soluciones extremadamente concentradas, donde no hay agua suficiente para mantener la hidratación, las moléculas pueden o bien agregar como un precipitado amorfo o bien cristalizar (figura 1). La estrategia para inducir la cristalización consiste en llevar al sistema hacia el estado de mínima solubilidad lentamente y lograr así un grado limitado de supersaturación. Para lograr que las moléculas tengan oportunidad de formar el mayor número de interacciones favorables con los vecinos se modifican las propiedades del solvente con agentes precipitantes o se alteran algunas propiedades físicas como la temperatura.

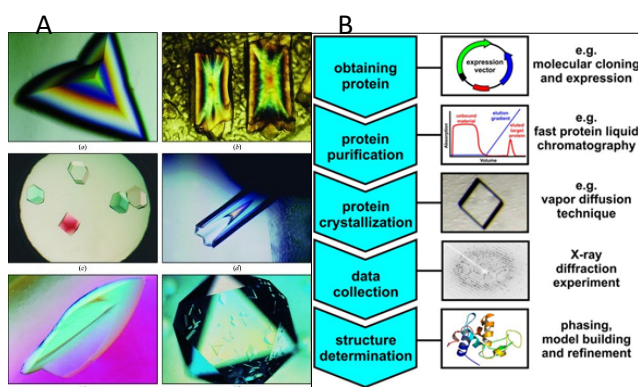


Figura 1. A) Microfotografías de cristales de proteínas y virus cultivados en el laboratorio de AM que muestran la variedad de hábitos comunes de los cristales macromoleculares. (a, b, f) Virus satélite del mosaico del tabaco, (c) Virus del moteado amarillo del Desmodium, (d) Canavalina hexagonal y (e) Anticuerpo intacto contra el linfoma canino. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3943105/figure/fig1/> **B) Pasos principales durante la cristalografía de rayos X para generar un modelo estructural de una proteína.** <https://link.springer.com/article/10.1007/s40828-018-0064-1>

2. FUNDAMENTOS FÍSICOS DE LA CRISTALIZACIÓN

La determinación de la estructura cristalográfica empieza con el crecimiento de un cristal adecuado. Para ello pueden aplicarse distintos métodos que se diferencian por las condiciones y solutos que emplean. Para decidir el método mejor para cristalizar una proteína determinada debemos entender los fundamentos de la cristalización.

La formación de un cristal de proteína macroscópico que contiene hasta 10^{15} moléculas, empieza con la aparición de agregados de proteína mediante contactos intermoleculares (*agregación*). Estos agregados prenucleares eventualmente alcanzan el tamaño nuclear crítico. Una vez formado un núcleo estable (*nucleación*), el crecimiento prosigue por la adición de moléculas a la red cristalina. Tanto la nucleación del cristal como el *crecimiento* ocurren en soluciones supersaturadas donde la concentración de proteína es mayor que su valor de

solubilidad en equilibrio. La región del *parámetro de solución* aconsejable para la cristalización viene recogida en el diagrama de fase (figura 2).

La supersaturación es una función de la concentración de la macromolécula y de diversos parámetros que afectan su solubilidad. En general, se logra a altas concentraciones de macromoléculas y a valores crecientes de parámetros de solución reductores de la solubilidad macromolecular. Los factores que pueden influir en la solubilidad proteica son por ejemplo, la inclusión de aditivos tales como alcoholes, polímeros hidrofílicos y detergentes que puedan disminuir la solubilidad proteica. Estos son *agentes precipitantes* ya que influyen en la solubilidad, pudiendo causar la precipitación macromolecular. Otro factor que influye sobre la solubilidad proteica es la presencia de sal, que varía generalmente como una curva asimétrica en forma de campana indicando que la solubilidad decrece tanto a altas como a bajas concentraciones de sal. Parámetros de la solución tales como pH o temperatura también pueden afectar dramáticamente a la solubilidad macromolecular.

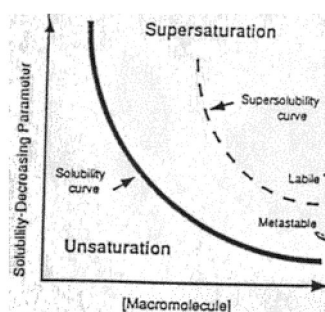


Figura 2. Diagrama de fase. La curva de solubilidad (sólida) divide el espacio de fase en regiones que promueven el proceso de cristalización (soluciones supersaturadas) de aquellas donde se disolverá el cristal (soluciones insaturadas). La curva de supersolubilidad (líneas entrecortadas) divide además la región supersaturada en condiciones de más alta supersaturación donde compiten la nucleación y crecimiento (fase lábil) y niveles menores donde sólo crecerán cristales (fase metastable).

Las condiciones de supersaturación óptimas para la nucleación y para el crecimiento del cristal son diferentes. Esto se muestra en el diagrama de fase donde la región de supersaturación es además dividida en regiones de más alta supersaturación (región lábil) donde tanto el crecimiento como la nucleación ocurren y la disminución de supersaturación (fase metastable) donde sólo se alcanza el crecimiento.

3. MÉTODOS DE CRISTALIZACIÓN

Métodos de difusión de vapor

Las condiciones dentro de una solución que contiene proteína son manipuladas mediante la difusión a través del aire (figura 3). Como con los otros métodos de cristalización, el objetivo inicial es formar núcleos. En este método, la nucleación ocurre cuando se aumenta la concentración de proteína por medio de la reducción del volumen de la solución. Esto ocurre porque el vapor de agua de la gota que contiene la proteína se deshidrata buscando el equilibrio con la solución de reserva más higroscópica. Como muestra la figura 3, una vez que se han alcanzado las condiciones de nucleación, la solución permanece altamente supersaturada, de modo que puede producirse simultáneamente la nucleación y el crecimiento de cristales. Para evitar que los valores de supersaturación excedan los requeridos para la nucleación, se han diseñado métodos para alcanzar muy lentamente el equilibrio. A pesar que dentro de un marco teórico pareciera que las condiciones no son óptimas para cristalizar una proteína, la difusión de vapor por gota colgante ("hanging drop") y gota posada ("sitting drop") son amplia y muy exitosamente usados para el crecimiento de cristales de proteína. El hecho que la mayoría de gotas de difusión de vapor no estén llenas de cristales microscópicos sugiere que el alto grado de supersaturación que motiva nucleación no es mantenido a lo largo del experimento de cristalización. Varios factores, incluso la disminución de concentración de macromolécula debido a la formación de núcleos o precipitados, la

incorporación en cristales crecientes o la desnaturalización macromolecular en la interfase solución-aire podría disminuir la concentración proteica a valores menos propicios para la nucleación y más adecuados para el crecimiento.

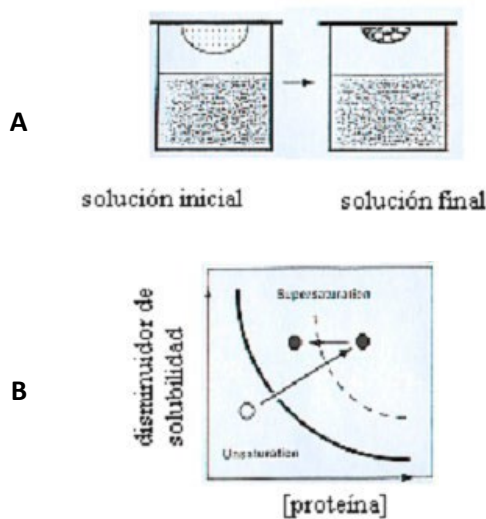


Figura 3. Diagrama esquemático de la cristalización por difusión de vapor. En este método de cristalización, las soluciones de proteína que contienen precipitante insaturado (indicado en el diagrama de fase por círculos abiertos) son suspendidos sobre un reservorio (A). El equilibrio, por medio de vapor de la gota y el reservorio, causa que la solución de la proteína alcance un nivel de supersaturación donde ocurren la nucleación y el crecimiento inicial (círculos sólidos en (B)). Como se ve en el texto, es probable que cambios en la concentración de la gota disminuyan la supersaturación a lo largo del tiempo del experimento (círculos sólidos de izquierda).

Otros métodos de cristalización (consultar)

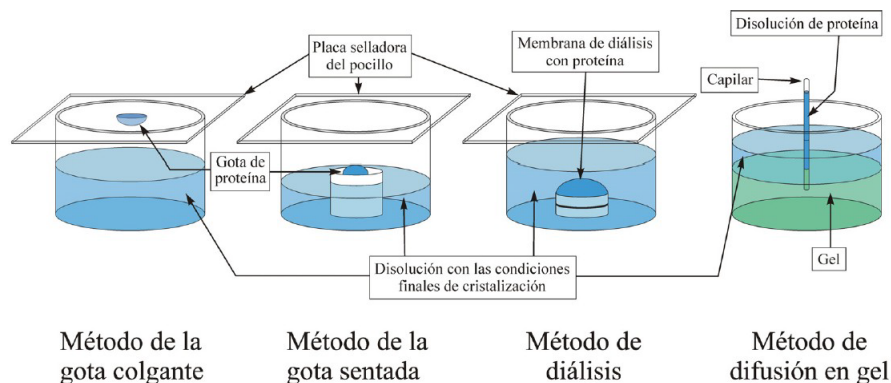
-Methods in Enzymology Volume 276. Macromolecular crystallography. Part A. Overview of protein crystallization methods. Patricia C. Weber. 13-22. Academic Press.

-Preparation and analysis of protein crystals. 1989. Alexander McPherson. 82-159. Krieger Publishin Company.

-How to grow protein crystals:

https://peakproteins.com/wp-content/uploads/2021/03/bio_2020_108.pdf?utm_source=chatgpt.com

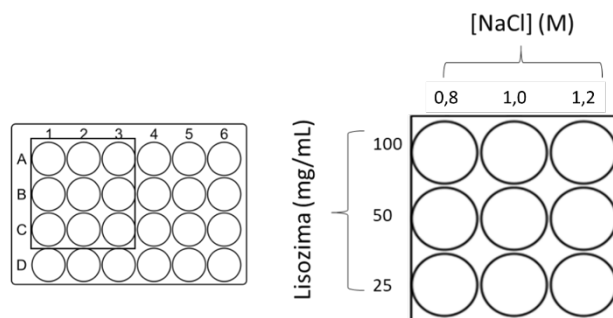
<https://www.sine2020.eu/about/the-road-to-the-ess/phase-diagram.html>



4. CRISTALIZACIÓN DE LISOZIMA

En esta práctica intentaremos reproducir condiciones de cristalización previamente establecidas de la lisozima de huevo. Para ellos utilizaremos la técnica de difusión de vapor en gota colgante (hanging drop). Realizaremos un estudio breve de la influencia de la concentración de la proteína y del precipitante en la aparición y tamaño de cristales. Para ello, variaremos ambos parámetros.

Las condiciones de cristalización de la lisozima contendrán lo siguiente: lisozima a concentraciones de **25, 50 y 100 mg/ml**, cloruro de sodio (NaCl) a **0,8; 1,0 y 1,2 M**, y tampón de **acetato de sodio (0,1 M, pH 4,6)**. Los ensayos se llevarán a cabo en una placa de 24 pocillos, de acuerdo con el siguiente esquema:



Paso 1. Hacer los cálculos respectivos para preparar las soluciones madre a partir de stocks de NaCl (2,5 M), buffer acetato sódico (AcNa 1 M, pH 4,5). El stock de lisoizima (100 mg/ml) se diluirá con buffer AcNa 0,1 M, pH 4,5 hasta la concentración deseada.

Paso 2. Con ayuda de una jeringa, se colocará grasa de animal alrededor de cada pocillo. Para evitar confusiones, marcar cada pocillo con la respectiva concentración de enzima y solución madre empleada.

Paso 3. Se pondrán 500 μ l de solución madre en cada reservorio. Luego, se prepararán gotas en cristales cubreobjetos (previamente silinizados) mezclando 3 μ l de solución proteica con 3 μ l de solución madre a la concentración indicada. Con mucho cuidado, voltear el cubreobjetos y colocar sobre el reservorio, presionando ligeramente para que se adhiera bien a la superficie.

Paso 4. Tapar y mantener las placas bajo condiciones de temperatura estable y en una superficie libre de movimiento. Esperar de 5 a 12 horas.

Paso 5. Se observará cada uno de los pocillos con lupa de 40 aumentos y con lámpara de luz fría.

Paso 6. Indicar:

- ⇒ ¿Qué condiciones son óptimas para el crecimiento de cristales de lisoizima de huevo?
- ⇒ ¿Cómo es un cristal de lisoizima?
- ⇒ ¿Qué ocurre en las demás condiciones probadas?

Nota: la lisoizima de huevo corresponde a la enzima liofilizada, de la compañía Sigma (código EC 3.2.1.17) contiene 47000 unidades/mg sólido y 95% de pureza.

2.1. *Ajuste de una secuencia de aminoácidos a un mapa de densidad electrónica*

1. INTRODUCCION

La difracción de un haz de rayos X por un cristal de proteína proporciona reflexiones de distinta intensidad y fase a partir de las cuales se puede calcular un mapa tridimensional de la densidad electrónica de la molécula proteica que ha difractado el haz. El cristalógrafo debe establecer un modelo tridimensional de la proteína que explique la distribución electrónica observada. En el proceso de generación del modelo (lo que al final vamos a llamar la estructura de la proteína) es frecuente realizar ajustes 'manuales' en el ordenador para que los aminoácidos de la proteína encajen bien en la densidad electrónica. Estos ajustes se realizan utilizando programas capaces de representar simultáneamente el modelo de la proteína y el mapa de densidad electrónica y capaces también de generar conformaciones alternativas de la proteína para conseguir un mejor encaje en la densidad electrónica.

Nuestra tarea consistirá en refinar un modelo de proteína haciendo las siguientes tareas:

- Corregir y modificar posiciones de aminoácidos en el modelo de manera que sea coherente con los datos de difracción de rayos X (los mapas de densidad electrónica en azul, verde y rojo).
- Añadir moléculas de agua (en su correspondiente densidad)
- Encontrar la posición de un ligando (en este caso un inhibidor) y ajustarlo en su densidad electrónica.

2. USO DEL PROGRAMA MOORHEN PARA ENCAJAR UNA SECUENCIA DE AMINOÁCIDOS Y UN LIGANDO EN MAPAS DE DENSIDAD ELECTRÓNICA

En esta práctica vamos a trabajar con **Moorhen**, que es la versión web de **Coot**.

Coot es un programa muy utilizado en biología estructural, ya que permite **construir, completar y validar modelos macromoleculares** a partir de datos de difracción de rayos X. Con él podemos visualizar mapas y modelos y realizar operaciones como:

- idealización de la geometría,
- refinamiento en espacio real,
- rotación y traslación manual de fragmentos,
- ajuste de cuerpos rígidos,
- búsqueda e incorporación de ligandos,
- solvatación (inclusión de moléculas de agua),
- mutaciones,
- elección de rotámeros
- análisis mediante gráficos de Ramachandran, entre otras.

En este ejercicio seguiremos una **versión adaptada del Tutorial 1 de Moorhen**, utilizando como ejemplo la estructura cristalográfica de una **quinasa dependiente de ciclina (OCR)** con código **2VTQ**.

Publicación: Wyatt, P.G. et al. (2008). *J. Med. Chem.* 51, 4986.

EMPEZAMOS:

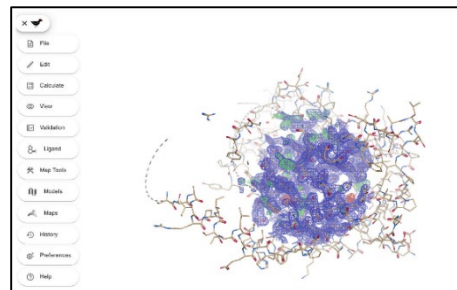
Abre una ventana del navegador web y teclea “moorhen.org”.

Se abre el programa Moorhen en la ventana del navegador

Vamos a cargar algunos datos:

File — Load Tutorial Data ... — Tutorial 1 - OK

Moorhen muestra un modelo de proteína, un mapa azul de tipo 2Fo-Fc y un mapa de tipo Fo-Fc en verde (positivo) y rojo (negativo).



Breve explicación de estos mapas:

En **crystalografía de proteínas**, los mapas de densidad electrónica se usan para **visualizar dónde están los átomos**. Hay varios tipos de mapas, pero los más comunes son:

1. Mapa 2Fo-Fc (azul)

- Es el **mapa principal** de densidad electrónica
- Se calcula combinando:
 - **Fo** → las amplitudes **observadas** experimentalmente.
 - **Fc** → las amplitudes **calculadas** a partir del modelo atómico.
- El “2Fo-Fc” significa que se da **más peso a los datos experimentales**.
- **Color azul** → indica **dónde debería estar la densidad electrónica** según los datos y el modelo.
- **Uso principal** → sirve para **colocar los átomos** y verificar que el modelo encaje bien en la densidad.

2. Mapa Fo-Fc (verde y rojo)

- Es un **mapa de diferencia** entre los datos experimentales (**Fo**) y el modelo calculado (**Fc**).
- Muestra **dónde el modelo no coincide con los datos**.
- **Verde (positivo)** → hay **densidad extra**: sugiere que **faltan átomos** o una conformación alternativa.
- **Rojo (negativo)** → hay **densidad sobrante**: indica que **hay átomos en el modelo que no deberían estar ahí**.
- **Uso principal** → detectar errores en el modelo y corregirlo.

ALUMNO:

Para mover el modelo estructural en la pantalla

- ✓ Utiliza el botón izquierdo del ratón y arrastra para girar la vista. (solo tienes que hacer clic y arrastrar)
- ✓ Utiliza la rueda de desplazamiento para acercar y alejar la imagen
- ✓ Utiliza el botón central del ratón (presionar rueda) para centrar en un átomo
- ✓ Para desplazar la vista, haz clic con el botón central del ratón y arrastrar

MODELOS (MODELS)

Haz clic en el botón “Models”. Verás botones para varios estilos de representación del modelo.

Haz clic en el botón «Sequences». Verás un visor de secuencias para desplazarse por la estructura. Hacer clic en algunas letras (que representan los residuos de la proteína). Observa que la malla de densidad del mapa se vuelve a dibujar alrededor del nuevo centro.

¿Ves el rectángulo gris sobre los números de secuencia? Mueve el cursor dentro del cuadro y utiliza el botón izquierdo del ratón para arrastrar y desplazarse por la secuencia.

Puedes cambiar el tamaño del rectángulo para mostrar más residuos (haciendo clic y arrastrando el borde del cuadro), pero si lo haces demasiado ancho, no se mostrará ninguna.

Haz clic en la pestaña “Ligands”

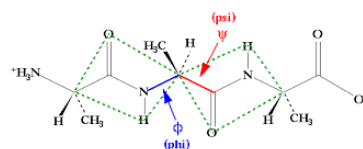
Observa que no hay «ningún ligando». Añadiremos uno más adelante.

Usar Control+rueda de ratón para poner el mapa 2Fo-Fc a un valor de 0,30 (1,0 rmsd).

VALIDACIÓN (VALIDATION)

Vamos a utilizar alguna de estas herramientas para evaluar la calidad y corrección del modelo molecular que estamos construyendo o refinando.

Vamos a ver el gráfico de Ramachandran que es una forma de visualizar las regiones energéticamente permitidas para los ángulos diedros (también llamados ángulos de torsión, ángulos phi (ψ) y psi(ϕ)) de los residuos de aminoácidos en la estructura de las proteínas.



Ángulos de torsión phi y psi del esqueleto peptídico de una proteína. Tomado de <https://ingprot.wordpress.com/diagrama-de->

Abrir la ventana “Validation”.

Seleccionar “Ramachandran Plot”. Moorhen muestra el gráfico de Ramachandran para la cadena «A» de esta proteína.

Podemos cambiar el tamaño del cuadro de diálogo para ver mejor el gráfico (arrastra lentamente el triángulo inferior derecho). Verás que hay varios puntos rojos interesantes. Hacer clic en el punto situado en la parte superior central.

Moorhen colocará el residuo A180 en el centro de la pantalla.

Echa un vistazo a la región... Mmmm... los átomos de oxígeno carbonílico están demasiado cerca unos de otros.

Para medir la distancia entre átomos, mantén pulsada la tecla M y haz clic en un átomo, luego (sin soltar la tecla «M») haz clic en otro átomo. Están a 2.26 Å, demasiado cerca. Para ocultar la distancia, pulsa la tecla «C».

Haz clic en «Models» y, a continuación, en el botón «Rama. Balls»

Moorhen muestra las bolas de Ramachandran, que utilizan colores para representar la probabilidad de esos ángulos phi y psi para ese residuo.

El **residuo A180** tiene una bola roja y, justo en dos posiciones anteriores, en A178, la LYS tiene una bola naranja. Quizás ambos sean mejorables.

Observa que hay una densidad verde cerca de la N de A180. Quizás sería mejor que esa densidad encajara con el carbonilo del péptido.

Para hacer esto vamos a emplear la función “Flipping a Peptide” (Voltear un péptido)

Haz clic con el botón central del ratón en el átomo N del residuo A 180.

Haz clic con el botón derecho del ratón en un átomo del péptido (átomo O de 179).

Moorhen muestra una cuadrícula con iconos para modelar. Al pasar el ratón por encima de los iconos de la barra de herramientas, verás una descripción emergente de ese icono.

Para voltear el péptido, debes utilizar el botón «Flip Peptide». Haz clic en él. Moorhen voltea el

péptido y la bola de Ramachandran para esos residuos se vuelve verde.

¡Bien! Avanzamos. Veamos si podemos hacer lo mismo con el **enlace peptídico entre A177 y A178**. Lleva A177 CA al centro de la pantalla. Haz clic con el botón derecho del ratón sobre A177 O y, a continuación, haz clic en el botón «Flip Peptide» (Voltear péptido). Moorhen voltear el péptido y la bola de Ramachandran para 177 se vuelve verde.

Más progreso. Muy bien.

Si volvemos a mirar el gráfico de Ramachandran se observa que los puntos rojos de los residuos problemáticos han desaparecido.

Consejo de navegación 1: Haz clic con el botón central del ratón sobre un átomo para colocar el átomo seleccionado en el centro de la pantalla

Consejo de navegación 2: Para ir a un residuo específico, utilizar “Edit”— “Go to”... y en la entrada «Atom selection» escribir la selección de residuos, por ejemplo, «//A/120» para referirse al residuo 120 de la cadena «A».

REFINAMIENTO DEL ESPACIO REAL (REAL SPACE REFINEMENT)

Seguimos comprobando si nuestro modelo se ajusta en los mapas de densidad electrónica calculados a partir de los datos experimentales de difracción de rayos X. Para mejorar la calidad del modelo atómico en regiones donde no se ajusta bien al mapa se hace un refinamiento del modelo que consiste en mejorar los parámetros del modelo ajustando las coordenadas atómicas, las ocupaciones y los parámetros térmicos para minimizar la diferencia entre los patrones de difracción calculados (F_c) y los observados (F_o).

Haz clic con el botón derecho del ratón en un átomo del residuo **A179**.

Haz clic en el botón «Refine Residues» (el icono es una diana).

Moorhen refina de manera automática y basándose en los mapas, los residuos a lo largo de la cadena. Ahora la cadena principal local encaja bastante bien en el mapa azul (pero es posible que las bolas de Rama no sean completamente verdes).

Actualizar los mapas

Para actualizar los colores de los mapas (que el verde pase a azul cuando se ha refinado correctamente)

File — connect molecule and map for updating... — OK.

Moorhen mostrará una «notificación» en la parte superior derecha informándote del factor R actual (discrepancia entre los datos experimentales y los que se calculan de nuestro modelo. Debería ser menor de 0.25)

Nota: El uso de mapas actualizados/conectados ralentizará un poco el proceso de construcción del modelo, pero ahora tenemos la ventaja de ver cómo baja el factor R a medida que hacemos cambios.

AÑADIR AGUAS (SOLVATACIÓN)

Añadir moléculas de agua a un modelo ayuda a mejorar la precisión y la integridad del modelo estructural, especialmente a mayor resolución, al tener en cuenta las moléculas de agua ordenadas que suelen estar presentes en los mapas de densidad electrónica. Estas moléculas de agua desempeñan un papel crucial en la estructura de las proteínas, ya que a menudo forman enlaces de hidrógeno con ellas y otras moléculas, lo que influye en la estabilidad y función de la proteína.

Podríamos añadir las aguas una por una, pero un método automatizado consiste en añadir muchas al mismo tiempo.

Calculate — Add waters... — OK

Esto añadirá alrededor de 100 aguas. Y, como antes, los mapas y los factores R se actualizarán. El mapa debería mejorar un poco para que el ligando sea más fácil de distinguir.

CAMBIAR EL TIPO DE RESIDUO (Mutación *in silico*)

Ahora navega hasta el **residuo A193**. ¿Qué te indican los mapas que está sucediendo aquí? ¿Cuál es el tipo de residuo? ¿Qué dice el modelo? ¿Qué dice el mapa?

Bien, primero llenemos la cadena lateral con los átomos del tipo de los átomos de la cadena principal:

«TYR». Utiliza el botón derecho del ratón para acceder al menú contextual y haz clic en el botón «Auto-fit Rotamer» (arriba a la izquierda).

¿Qué vemos? ¿Qué sugiere eso?

TIP: Sugiere que la secuencia del modelo no coincide con la secuencia de la proteína de la que se obtuvieron los datos. Bien, pues mutémosla.

En el menú contextual, haz clic en el botón «Mutate Residue» (es el símbolo de la radiación) y, a continuación, en el selector de tipo de residuos, actualmente «ALA (A)», elige «PHE (F)».

Moorhen actualiza el mapa para que desaparezca la densidad roja.

Navega hasta el **residuo A168**. ¿Qué vemos? ¿Qué debería ser en su lugar? Vale, mutémoslo: Utiliza el botón «Mutate Residue» (del menú contextual del botón derecho del ratón) para cambiar el tipo a «TYR (Y)».

BÚSQUEDA Y AJUSTE DE LIGANDO

Utiliza los picos del mapa de diferencias para navegar hasta el ligando.

Validation/Difference map peaks

Varios de los 5 picos principales verdes deberían corresponder ahora al ligando.

Ligand — Get monomer —

Escribe «LZA» en la entrada «Monomer Identifier» - OK

Moorhen importa el **ligando LZA**. (((2,6-difluorophenyl)carbonyl)amino)-N-piperidin-4-yl-1H-pyrazole-3-carboxamide/ C16 H17 F2 N5 O2

Bien, perfecto. Ahora vamos a ocultarlo:

En la ventana Model, la parte inferior debería ser ahora del ligando recién importado («#3 Mol LZA»).

Haga clic en el icono del ojo para ocultar el ligando.

Moorhen cambia el icono a un ojo tachado y el ligando desaparece.

Ligand — Find ligand...

Marque la opción “Search everywhere” y haz clic en “Find”.

Moorhen presenta un nuevo cuadro de diálogo que indica que hay 6 posibles ubicaciones de ligando.

Para visualizar el ligando, haz clic en el icono del «círculo enmarcado» y el ligando aparecerá en la densidad. Debería encajar razonablemente bien, pero no exactamente, ya que el algoritmo no utilizó la variación conformacional.

Para refinar el ligando, haz clic en el icono «objetivo» de ese cuadro de diálogo y para fusionarlo con la cadena peptídica, utiliza el icono «flecha» de ese cuadro de diálogo.

Moorhen actualiza los mapas para que cambien los colores de los mapas diferenciales.

CREA UNA IMAGEN BONITA

En Maps y Models, oculta los mapas, bonds, Rama. Balls... utilizando el icono del ojo.

Haz clic en Ligandos para mostrar el ligando en el modelo.

Haz clic en RIBON para mostrar el modelo en modo Cinta.

Para navegar hasta el ligando: Ligand — Center on ligand... Haz clic en «mol-1», en «Chain A» y en la etiqueta «/1/A/1(LZA)».

View — Scene settings — Background Colour: cámbialo si quieres.

Ajusta los controles deslizantes para recortar y difuminar para que el ligando sea más visible. OK.

View-Activate scene preset-Figure Making

Utiliza la tecla «S» del teclado para activar la captura de pantalla «en la aplicación».

Exportar su molécula y en la ventana Modelo, haz clic en el icono «descargar»

Práctica nº 3: Medida de la estabilidad conformacional de una proteína mediante desnaturalización química

1. INTRODUCCION

La conformación plegada, nativa, de las proteínas es la más estable sólo en un determinado intervalo de valores de temperatura, pH, fuerza iónica, etc.

Una forma sencilla de determinar la estabilidad conformacional de una proteína es realizar una desnaturalización térmica registrando el valor de alguna propiedad de la proteína en función de la temperatura y ajustando los datos obtenidos a un modelo realista del equilibrio de desnaturalización térmica. Otra forma, análoga, es mediante desnaturalización química, utilizando una sustancia que desestabilice la conformación nativa. Este segundo método, que es extratermodinámico, recibe el nombre de método de extrapolación lineal (LEM). En esta práctica vamos a utilizarlo para determinar la estabilidad de la proteína lisozima empleando cloruro de guanidinio como desnaturalizante.

2. REACTIVOS

- Tampón acético/acetato sódico 45 mM, pH 3.0
- Lisozima 10mg/ml en tampón acético/acetato sódico 45 mM, pH 3.0
- Cloruro de guanidinio 7M en tampón acético/acetato sódico 45 mM, pH 3.0

3. PROCEDIMIENTO

- Preparar, en viales Eppendorf, 950 µl de cloruro de guanidinio 0.0, 0.5, 1.0, 1.5, 2.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 3.6, 4.0, 4.5, 5.0, 5.5, 6.0 y 6.5 M, por dilución en tampón acético/acetato sódico 45 mM, pH 3.0
- Añadir a cada vial, 50 µl de solución de lisozima
- Incubar a temperatura ambiente durante 20 minutos
- Medir la absorbancia de las disoluciones a 301 nm, empezando por la de menor concentración de cloruro de guanidinio.

4. DETERMINACION DE LA ΔG DE DESPLEGAMIENTO

A: Procedimiento manual:

- Representar la curva de desnaturalización
- Trazar con regla las líneas base de la absorbancia de los estados plegado y desplegado
- Calcular, en la zona de transición, las constantes de equilibrio a distintas concentraciones de cloruro de guanidinio
- Calcular las correspondientes ΔG a esas concentraciones
- Representar las ΔG en función de la concentración de cloruro de guanidinio
- Calcular la ΔG de la lisozima en ausencia de cloruro de guanidinio, por extrapolación lineal a concentración 0 M de desnaturalizante

B: Ajuste computacional:

El modelo utilizado para realizar el ajuste manual asume lo siguiente:

- a) La absorbancia observada (A) es la que corresponde a las de los estados nativo (A_N) y desplegado (A_D) ponderadas por sus fracciones molares:
- $$A = A_N X_N + A_D X_D$$

- b) La absorbancia de los estados nativo y desplegado no son siempre igual sino que dependen linealmente de la concentración de agente desnaturante presente en la disolución:

$$A_N = A_{N0} + m_N * D$$

$$A_D = A_{D0} + m_D * D$$

donde A_{N0} e A_{D0} son las absorbancias de los estados nativo y desnaturado, respectivamente a concentración 0 M de desnaturante mientras que m_N y m_D cuantifican la dependencia lineal de esas absorbancias con la concentración de agente desnaturante.

- c) Cada molécula de proteína sólo puede adoptar dos conformaciones: o bien nativa o bien desplegada relacionadas por la constante de equilibrio de desplegamiento, la cual determina el valor de la diferencia de energía libre estándar:

$$K_{\text{desp}} = X_D / X_N$$

$$\Delta G = -RT \ln K_{\text{desp}}$$

- d) La energía libre estándar depende linealmente de la concentración de agente desnaturante según:

$$\Delta G = \Delta G_w - m * D$$

donde D es la concentración de agente desnaturante, m es una constante de proporcionalidad entre la estabilidad y la concentración de agente desnaturante, y ΔG_w es la estabilidad de la proteína en ausencia de desnaturante.

Combinando las ecuaciones anteriores, la dependencia de la absorbancia observada (A) de la concentración de agente desnaturante (D) se puede expresar como:

$$A = \frac{(A_{N0} + m_N D) + (A_{D0} + m_D D) e^{-(\Delta G_w - mD) / RT}}{1 + e^{-(\Delta G_w - mD) / RT}}$$

Se recomienda deducir la ecuación anterior a partir de las asunciones del modelo físico indicado.

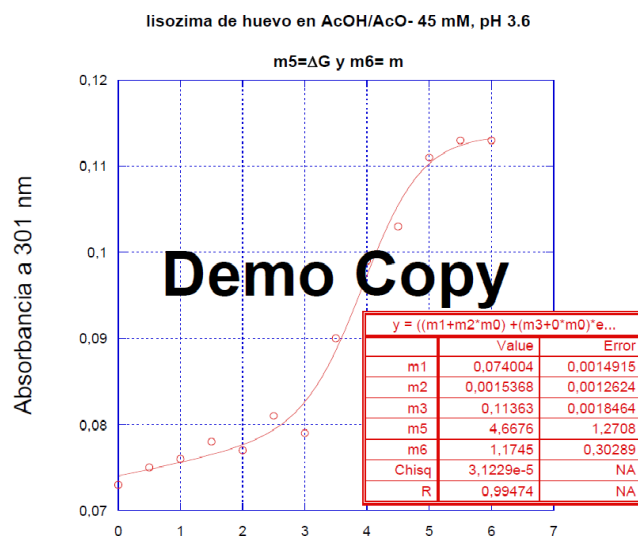
Una vez conocida la ecuación, se puede realizar un ajuste no lineal para estimar los valores de los parámetros ajustables A_{N0} , A_{D0} , m_N , m_D (no interesantes pero necesarios pues son parte del modelo), m (una medida de la cantidad de área expuesta por la proteína al desplegarse, dato de interés para un especialista), y ΔG_w (la energía libre de desplegamiento de la proteína en el tampón de trabajo elegido y sin desnaturante: (la magnitud termodinámica que se desea conocer).

Para poder realizar el ajuste hay que utilizar aplicaciones, raramente gratuitas, que permiten escribir las ecuaciones específicas del modelo que se ha planteado para ajustar los datos experimentales. Utilizaremos la aplicación Kaleidagraph.

- Descargar versión demo de la aplicación KaleidaGraph (<http://www.synergy.com/demowinint.htm>)
- Introducir en MENU CurveFit la ecuación General equivalente a la deducida:

$$((m1+m2*m0)+(m3+m4*m0)*\exp((-m5+m6*m0)/0.56))/(1+\exp((-m5+m6*m0)/0.56));m1=10;m2=0.001;m3=3;m4=0.001;m5=6;m6=3$$

- Ajustar la curva de desnaturalización a esa ecuación para obtener el valor de ΔG .
Ejemplo:



Práctica nº 4: Acaba de secuenciar un gen: ¿y ahora qué?

1. INTRODUCCION

Tras secuenciar un gen, pueden sentirse deseos irrefrenables de conocer la estructura tridimensional de la proteína correspondiente.

Para lograr semejante propósito, lo habitual es establecer una colaboración con un grupo de cristalografía de proteínas, de criomicroscopía electrónica o de resonancia magnética nuclear. En cualquier caso, pueden transcurrir meses, o quizá años, hasta que la estructura quede resuelta. Afortunadamente, los modelos computacionales no experimentales son cada vez más precisos.

En todo caso conviene conocer algunas herramientas computacionales básicas y gratuitas que nos facilitan el trabajo con secuencias y estructuras de ácidos nucleicos y proteínas.

2. RESCRITURA DE SECUENCIAS DE ADN Y DE PROTEINAS

Las secuencias de aminoácidos de una proteína pueden escribirse utilizando código de tres letras (e.g. TyrGlySer, que significa N-tirosina-glicina-serina-C) o de una letra (e.g. YGS, con el mismo significado). Salvo el PDB que sigue consignando las secuencias en código de tres letras, las bases de datos usan código de una letra, que es imprescindible conocer. Existen en internet servidores altruistas que convierten secuencias en código de tres letras a código de una letra, y viceversa. Deberás buscar uno y utilizarlo.

e.g.

Three to one https://www.bioinformatics.org/sms2/three_to_one.html

Three-/one-letter Amino Acid Codes https://www.bioline.com/media/calculator/01_17.html

Three to one https://www.genscript.com/sms2/three_to_one.html

One to three https://www.bioinformatics.org/sms2/one_to_three.html

Las secuencias de nucleótidos de un ácido nucleico se escriben en código de una letra. Existen también servidores que traducen de ADN(ARN) a proteína y viceversa.

e.g.

Translate tool: <https://web.expasy.org/translate/>

DNA to protein converter: <https://olvttools.com/en/translation>

Reverse translate: https://www.bioinformatics.org/sms2/rev_trans.html

Reverse translate: https://www.genecorner.ugent.be/rev_trans.html

Cuando se traduce de ADN a proteína obtendremos 6 lecturas distintas y sólo una será correcta. Debes entender porqué. Cuando se traduce de proteína a ADN las posibilidades son casi infinitas (¿porqué?) y si el objetivo es diseñar un gen sintético conviene especificar la utilización de una frecuencia de codones característica del organismo en que se desea obtener la proteína.

3. BÚSQUEDA DE INFORMACION BASADA EN HOMOLOGIA DE SECUENCIA

Una herramienta muy útil es la búsqueda de secuencias de proteínas homólogas. Cuando se encuentra una proteína de secuencia altamente similar a la de nuestra proteína problema cuya estructura tridimensional es conocida, la estructura tridimensional (y las propiedades biológicas) de nuestra proteína serán parecidas a las de ella.

Hay servidores en Internet que establecen el grado de identidad de una secuencia cualquiera con las secuencias ya registradas en distintas bases de datos. Estos servidores identifican secuencias parecidas y proporcionan información sobre las mismas (código de la proteína, código de su estructura tridimensional si se conoce, etc).

Un ejemplo es el comparador de secuencias BLAST (<https://www.uniprot.org/blast>) alojado en la base de datos de secuencias de proteínas UniProt: (<https://www.uniprot.org/>).

Otro ejemplo es el comparador de secuencias BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) alojado en el NCBI (National Center for Biotechnology Information).

Dentro de estos servidores puede resultar necesario seleccionar en primer lugar una aplicación específica para alinear secuencias de aminoácidos o secuencias de nucleótidos.

4. PREDICCIÓN DE ESTRUCTURA TRIDIMENSIONAL

Los más atrevidos pueden hacer una predicción completa de la estructura tridimensional en Internet.

Modelado por homología:

Si existen proteínas con alta identidad de secuencia y estructura conocida, el modelo será razonable (aunque la conformación predicha para las cadenas laterales, especialmente las expuestas al disolvente, es muy difícil que sea correcta del todo).

Usar SWISS-MODEL: <https://www.expasy.org/resources/swiss-model>

Todavía puede hacer modelos competitivos

Modelado por técnicas de AI

Alphafold: <https://alphafoldserver.com/>

Apéndice. Secuencias propuestas para identificar homólogos que insinúen su función y realizar una predicción de estructura tridimensional cuando sea posible.

(1)

```
ATAGGCGATGAGCCCCTTGAGAACTACCTGGATACTGAGTACTTCGGCACCATCGGCATTGGAACCCCC
GCTCAGGACTTCACTGTCATCTTTGACACCGGCTCCTCAACCTGTGGGTGCCCTCTGTCTACTGCTCTA
GTCTCGCCTGCAGCGACCACAACCAGTTCAACCCTGATGACTCCTCCACCTTCGAGGCCACCAGCCAGG
AGCTGTCCATCACCTATGGCACCGGTAGCATGACAGGCATCCTTGGATACGACACCGTCCAGGTCGGA
GGCATCAGCGACACCAACCAGATCTTTGGCCTGAGCGAGACAGAGCCCGGCTCCTTCTGTACTACGCC
CCCTTCGACGGCATCTAGGTCTGGCCTACCCAGCATCTCCGCTCCGGGGCCACCCTGTCTTTGACA
ACCTGTGGGACCAGGGCCTGGTTTCCCAAGACCTCTTCTCCGTCTACCTGAGCTCCAATGACGATAGCG
GCAGCGTGGTGTGTTGGGCGGCATCGATTCTTCTACTACACAGGAAGCCTGAAGTGGTGCCTGTTT
CCGTCGAGGGTTACTGGCAGATTACCTTGGACAGCATCACCATGGACGGAGAGACCATCGCTTGACGC
GGTGGCTGCCAGGCCATTGTGGACACGGGCACCTCTCTGCTGACCGGTCCCACCAAGTGCAATTGCAATC
AACATCCAGAGCGACATCGGAGCCAGCGAGAACTCAGATGGCGAGATGGTGATCAGCTGCTCCTCCAT
CGACAGCCTGCCTGACATCGTCTTACCATCAACGGTGTCCAGTACCCTCTGAGCCCCAGTGCCTACATC
CTGCAGGACGATGACAGCTGCACCAGCGGCTTCGAGGGCATGGATGTCCCCACCTCCTCCGGAGAGCT
CTGGATCCTGGGTGACGTCTTCATCCGCCAGTACTACACCGTCTTTGACAGGGCCAACAACAAGGTCGG
CCTGGTCCCCGTGGC
```

(2)

```
MILKILNEIASIPSTKQRQAIFENKDDKLLKKVYRVYTSAGLQWYIKKYPKPGIATQSFAMLTISDMLMFIEFS
LATRKVTGNAAFEELTGYISDGKKDEVEVMRDLESGASYSIANKVWKGLIPLQPQMLATSYDERGINKQIKF
PAAQKLKDDGAFCEVRADELDEVRLDRAGNEFLGLDLLKDELIKMTTEARNIHPEGALIDGELVYDEQVKK
```

KPEGLDFAFDAHPEQSKVKDYTEVAESRSASNGIANKTLKGTISEKELQCMKFQVYDYVPLLEVYGLPAFKLK
YDVRWSKLEQLTSGYDCCKLCVILIDNQVVNALDEAKVIYNKYIDQGLEGILKNIDGLWENARSKNLYKFKEV
IDVDLKIVGIYPHRKDPTKAGGFILESECGKIKVNAGSGLKDKAGVKTHELDRTKIMENQNYLGKILECDCN
GWYKSDGRDYLKFLPLAIRLRDDKTKADTFEDVWGDFHELTGL

(3)

CTGAAAGGAGGGCAATCCTTAATGAATCAAGCGAGTTCGTCTGGAGNAAGTGTTGCCTATGCACTTGCT
CAAGGTTTGAAAGTGATTGCTTGTGTTGGTGAGACTCTTGAGGAGCGGGAAGCTGGATCGACCATGGA
TGTTGTNGCTGCCCAGACCAAGGCTATTGCCATTCGGGTNACAACTGGTCAAATNTTGTATAGCCTA
TGAACCANTGTGGGCCATTGGAACCGNNNGTCGNTAGCCCAGNCCAAGCTCAAGAAGTNCACGAT
GAGCTGAGGAATGGCTNGCCAAGANCGTGA

(4)

GLY TRP GLU ILE PRO GLU PRO TYR VAL TRP ASP GLU SER PHE ARG VAL PHE TYR GLU GLN LEU
ASP GLU GLU HIS LYS LYS ILE PHE LYS GLY ILE PHE ASP CYS ILE ARG ASP ASN SER ALA PRO ASN
LEU ALA THR LEU VAL LYS VAL THR THR ASN HIS PHE THR HIS GLU GLU ALA MET MET ASP ALA
ALA LYS TYR SER GLU VAL VAL PRO HIS LYS LYS MET HIS LYS ASP PHE LEU GLU LYS ILE GLY GLY
LEU SER ALA PRO VAL ASP ALA LYS ASN VAL ASP TYR CYS LYS GLU TRP LEU VAL ASN HIS ILE LYS
GLY THR ASP PHE LYS TYR LYS GLY LYS LEU