



1 de diciembre de 2014 | Vol. 15 | Núm. 12 | ISSN 1607 - 6079

ARTÍCULO

INGENIERÍA DE PROTEÍNAS PARA EL MEJORAMIENTO DE ENZIMAS

*Claudia Martínez Anaya (Instituto de Biología, UNAM) y
José Fernando García Guevara
(Instituto de Biotecnología, UNAM)*

INGENIERÍA DE PROTEÍNAS PARA EL MEJORAMIENTO DE ENZIMAS

Resumen

Las condiciones en las que suceden las reacciones catalizadas por enzimas que son usadas en la industria son muy diferentes a las condiciones de su entorno natural, afectando en ocasiones su función. Para contender con esta limitación existen métodos de ingeniería de proteínas que tienen como objetivo el mejoramiento de las propiedades de las enzimas. Estos métodos utilizan una combinación de técnicas computacionales y de evolución dirigida. En este artículo se explica en qué consisten estos métodos, y se ejemplifican tres casos de éxito de enzimas mejoradas.

“

Las enzimas o biocatalizadores son energéticamente eficientes y prácticamente todas llevan a cabo su función bajo condiciones moderadas: temperatura ambiente, presión atmosférica y en ambientes acuosos.

”

Palabras clave: Estructura proteica, enzimas, ingeniería de proteínas, diseño racional, evolución dirigida.

PROTEIN ENGINEERING FOR ENZYME IMPROVEMENT

Abstract

The conditions under which enzyme-catalyzed reactions occur during industrial processes are very different to those found in their natural environment, sometimes affecting their activities. To contend with this limitation different methods of protein engineering are directed at improving the properties of enzymes. These methods combine the techniques

of computational analysis and directed evolution. This article explains the basis of these methods, and three case studies of improved enzymes are presented.

Keywords: Protein structure, enzymes, protein engineering, rational design, directed evolution.

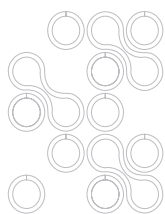
INGENIERÍA DE PROTEÍNAS PARA EL MEJORAMIENTO DE ENZIMAS

Introducción

La definición del diccionario de biocatalizador es: "aquella sustancia, en especial una enzima, que inicia o modifica la velocidad de una reacción química sin ser ella misma afectada por esta reacción". Efectivamente, las enzimas son macromoléculas (normalmente proteicas) que tienen la capacidad de aumentar la velocidad de las reacciones en millones de veces comparadas con el tiempo que les tomaría ocurrir espontáneamente. Estas enzimas o biocatalizadores son energéticamente eficientes y prácticamente todas llevan a cabo su función bajo condiciones moderadas: temperatura ambiente, presión atmosférica y en ambientes acuosos. Además, generan muy pocos productos de desecho. Sin embargo, para poder llevar a cabo procesos aplicados de manera rentable, las capacidades de las enzimas pueden mejorarse mediante métodos de ingeniería de proteínas (PEÑA-MONTES, 2008; SEGOVIA, 2010). Estos métodos tienen la finalidad de generar variantes enzimáticas que sean aún más activas o duraderas, o que puedan actuar sobre sustratos diferentes al original o funcionar en condiciones diferentes a las que trabajan de manera fisiológica dentro de su organismo de origen, tales como: temperaturas elevadas, presencia de compuestos inhibitorios o solventes diferentes al agua. La creciente presión para llevar a cabo procesos ambientalmente amigables es un motor para el uso de enzimas, por lo que su mejoramiento en el laboratorio es deseable.

Al emplear técnicas que intentan imitar la evolución natural en un proceso de generación de variabilidad-selección, la ingeniería de proteínas utiliza diferentes estrategias dependiendo del grado de conocimiento que se tenga sobre ellas. De una enzima es posible conocer la secuencia del gen que la codifica, que es indicativa de su secuencia de aminoácidos, según el código genético. A esta secuencia lineal de aminoácidos se le denomina estructura primaria. La posición que los diferentes aminoácidos adoptan puede resultar en la formación de hélices o láminas conocidas como alfa-hélices u hojas-beta plegadas. En su conjunto, a éstas se les conoce como estructura secundaria. La forma tridimensional del acomodo completo de todos los aminoácidos en el espacio se conoce como estructura terciaria (Figura 1), y se determina experimentalmente, entre otras maneras, mediante difracción de rayos X de arreglos cristalinos de las proteínas. Por esta razón es que comúnmente se puede llamar también a la estructura tridimensional de una proteína, estructura cristalina¹.

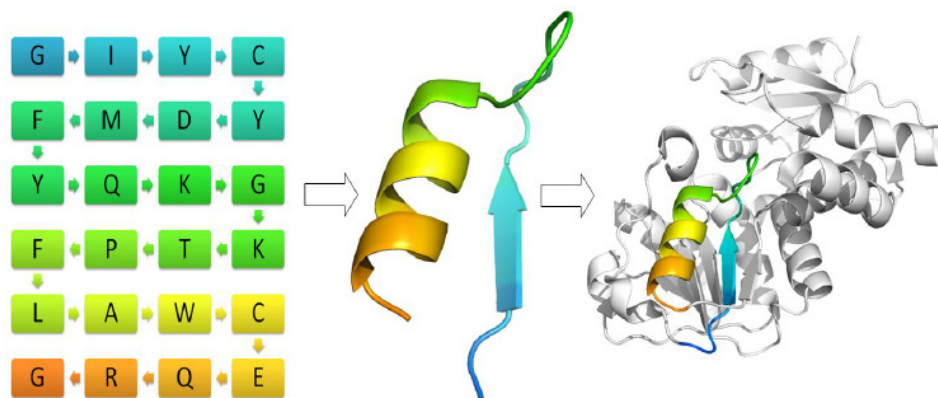
Aunque la determinación de la estructura tridimensional de las proteínas se ha vuelto más cotidiana, solamente un pequeño porcentaje de las enzimas con las que se trabaja en los laboratorios cuenta con datos estructurales. Por este motivo, es muy común tener datos de la secuencia primaria y secundaria de una proteína, pero no siempre se conoce la estructura terciaria. Esta información determina el método que se quiera aplicar para llevar a cabo el mejoramiento de una enzima. En general, los métodos de ingeniería de proteínas pueden dividirse ampliamente en dos tipos: a) métodos de evolución dirigida, y b) métodos de diseño racional. En muchos casos exitosos la combinación de ambos enfoques ha permitido mejores resultados.



[1] Se puede encontrar más información al respecto de la estructura de las proteínas y de los métodos para su análisis en: <http://laguna.fmedic.unam.mx/~evazquez/0403/proteinas.html>

* Un ejemplo ilustrativo de la evolución dirigida de proteínas se puede ver en el siguiente video animado: <http://youtu.be/r7mSweNf54c>

Figura 1. Se muestra la secuencia lineal (indicada por las flechas pequeñas) que corresponde a la estructura primaria de una proteína. Cada bloque indica el nombre del aminoácido con el código de una sola letra, ejemplo: G = glicina. El color de esta secuencia de aminoácidos corresponde al lugar que ocupa en el espacio en la formación de hélices u hojas (estructura secundaria). La estructura terciaria o tridimensional resulta del plegamiento de toda la cadena de aminoácidos, y en este ejemplo se indica con color el lugar dentro de toda la proteína del fragmento de en medio.



Evolución dirigida

Cuando la información estructural de una enzima es limitada, la producción adecuada de la proteína en el laboratorio es suficiente para llevar a cabo ensayos de evolución dirigida. En este caso, la teoría de la evolución conocida para las especies, en la que la acumulación de mutaciones genera variabilidad, es aplicada en la evolución dirigida de proteínas, aunque en escalas de tiempo miles de veces más cortas.

Para este propósito el gen de la enzima de interés es sometido a la técnica conocida como PCR (por sus iniciales en inglés: polymerase chain reaction), en la que el fragmento de DNA que codifica a la enzima es copiado y replicado millones de veces por una enzima polimerasa de DNA. Si esta polimerasa replica al DNA introduciendo errores que no son reparados, al final de la reacción de PCR se tendrán millones de moléculas portadoras de cambios (o mutaciones) con respecto a la original. Cuando se desconoce qué secuencia de aminoácidos es la que conferirá una mejor actividad a la enzima, es deseable que los cambios que se producen en la reacción de PCR ocurran al azar y con poca frecuencia (para evitar generar muchas variantes inactivas) (Figura 2).

Las técnicas actuales en el laboratorio permiten manipular o controlar la tasa de error durante la replicación de un gen, pudiendo llegar a tener de 2 a 8 mutaciones por cada 1000 pares de bases replicadas. Las poblaciones de moléculas variantes que llegan a obtenerse con este método deben ser posteriormente clonadas en un vector para su producción y la verificación del cambio enzimático buscado (Figura 3). Los genes de la o las variantes que presenten cambios positivos pueden ser sometidos a nuevos ciclos de PCR para intentar obtener aún mejores actividades. Cuando la mutagénesis ocurre a lo largo de toda la secuencia proteica, el número de variantes generadas es inmenso y en la práctica es posible el análisis de sólo una fracción infinitesimal de estas colecciones (también llamadas librerías).

Figura 2. Mediante PCR mutagénico se genera variabilidad a partir de la secuencia de un solo gen. Con esta técnica la molécula parental (indicada en gris) es replicada a través de muchos ciclos (aproximadamente 30), y cada vez son introducidos errores en diferentes posiciones (indicados con líneas rojas). El método de barajeo de DNA involucra el corte de varias secuencias genes similares (indicados por líneas dobles de diferente color) y su posterior re-ensamblaje al azar originando genes compuestos de diferentes fragmentos parentales (mostrados en líneas dobles de colores combinados). Después, las moléculas obtenidas por cualquiera de estas dos formas son clonadas en un fragmento de DNA llamado vector en el que se liga cada fragmento modificado para generar las librerías conteniendo miles de variantes.

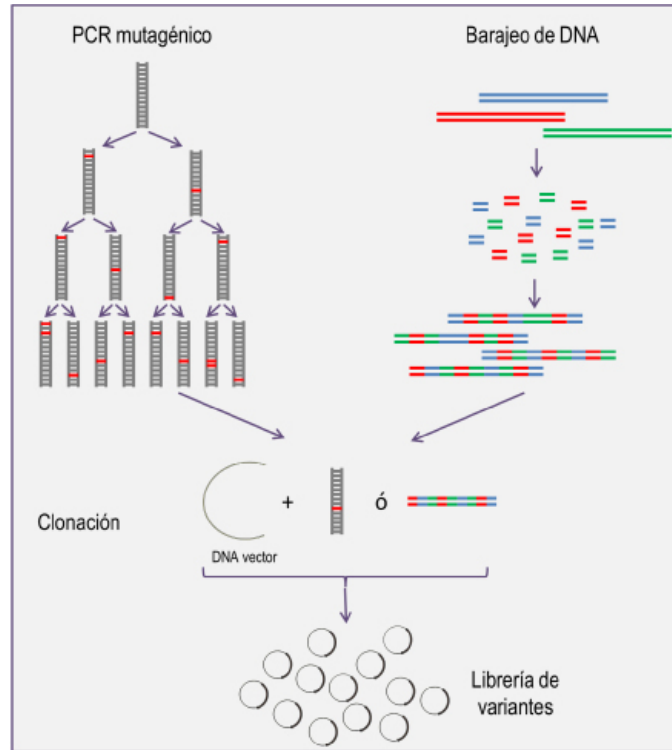
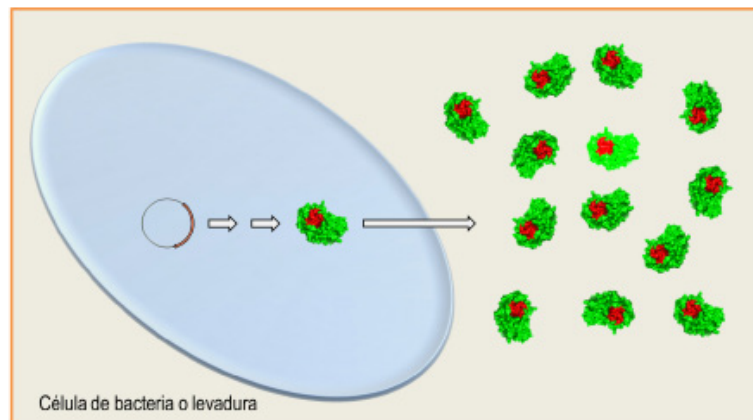


Figura 3. Las librerías de mutantes son introducidas dentro de células –que pueden ser de bacterias o levaduras, por ejemplo– para que, mediante los mecanismos celulares (indicados por las flechas pequeñas), sean producidas enzimas con pequeños cambios respecto a la proteína original (en este caso se ilustra una proteína de color pálido como aquella que por azar resultó con una mejor actividad). La flecha larga indica la producción en alta cantidad de la enzima y su exportación hacia afuera de la célula.



Otra manera en la que puede llevarse a cabo la evolución dirigida es mediante barajeo de DNA que genera diversidad mediante la combinación de secuencias de genes similares que existen naturalmente. Las colecciones de DNA recombinantes son generadas mediante la fragmentación al azar de los genes homólogos, seguida del ensamblaje de los fragmentos, lo que genera entrecruzamientos en las áreas en las que existe identidad en la secuencia.

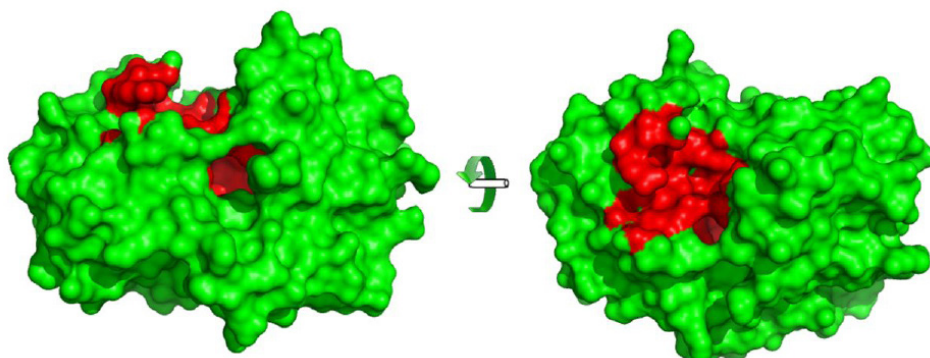
El éxito de las técnicas de evolución dirigida depende de contar con un método que permita la identificación de la variante con la actividad deseada de entre los millones de mutantes generados en una sola librería. Debido a que no todas las enzimas son sujeto para el desarrollo de sistemas de búsqueda sencilla dentro las librerías, en

muchas ocasiones el método para la selección de las variantes mejoradas es difícil de implementar, y esto constituye la principal desventaja de la evolución dirigida al azar. Por esta razón, los esfuerzos actuales se enfocan en la obtención de librerías inteligentes que produzcan un número de variantes mucho menor mediante la identificación de los mejores aminoácidos candidatos a la mutagénesis. Por ejemplo, las zonas de funcionalidad de una proteína pueden ser identificadas usando información evolutiva que permite la identificación de aquellos aminoácidos que son comunes entre genes de diferentes organismos, y que indican que su posición en la secuencia o en la estructura es importante, lo que delimita la mutagénesis a estas zonas blanco haciendo más sencilla la búsqueda de las variantes mejoradas.

Diseño racional

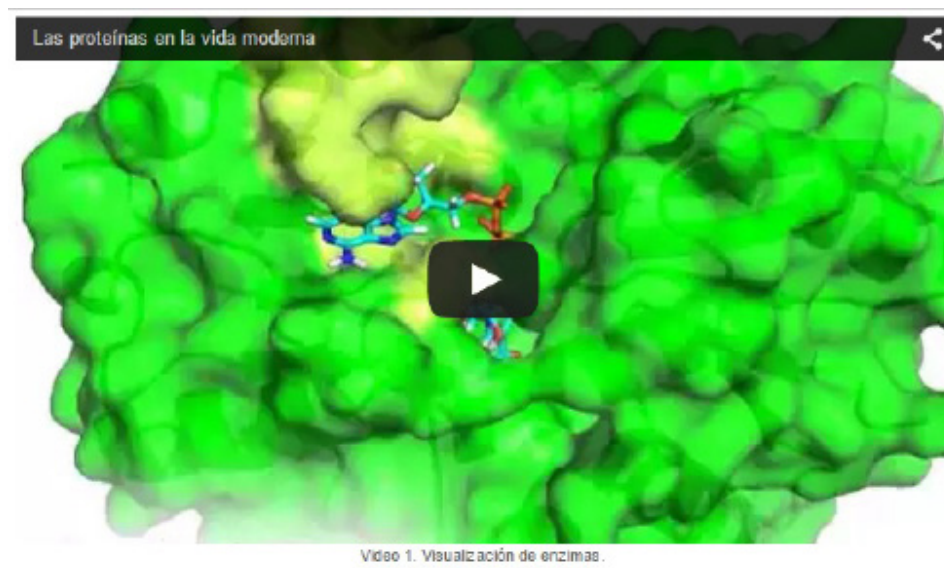
El diseño racional se refiere a la capacidad de predecir los cambios requeridos para lograr alguna función preconcebida de una enzima. Esta estrategia requiere de un conocimiento profundo de la estructura de una enzima y de su sitio activo, que es el lugar en donde se lleva a cabo la catálisis (Figura 4), o de los cambios moleculares que ocurren durante la reacción, debido a que la función está estrechamente relacionada a la estructura.

Figura 4. Estructura tridimensional de una proteína, en la que se ha marcado con color rojo el sitio activo. A la izquierda se muestra una vista lateral y a la derecha una vista de frente de la enzima.



Como se mencionó anteriormente, la determinación de las estructuras cristalográficas de muchas proteínas ha permitido que las bases de datos en donde se deposita esta información estén en constante crecimiento. Por otro lado, existen herramientas computacionales cada vez más poderosas que permiten la generación de modelos tridimensionales de proteínas para las cuales no hay datos cristalográficos, y que se basan en la información de proteínas homólogas cuya estructura ya ha sido determinada. Al contar con una estructura (real o modelada) los aminoácidos que forman el sitio activo de la enzima (también conocidos como residuos) pueden ser visualizados, así como otros residuos clave cercanos a éste, residuos que se encuentran en la frontera entre los diferentes fragmentos que componen una enzima, o aquellos que le permiten movilidad. La identificación de estos puntos clave permite proponer cambios que resulten en una actividad diferente a la original. El cambio necesario puede ser a nivel de sustitución de aminoácidos específicos, los cuales, por ejemplo, adicionan o eliminan carga eléctrica, o son más o menos voluminosos para acomodar a sustratos de cierto tamaño requerido,

como puede verse en el ejemplo del siguiente video. Otros cambios involucran la adición o eliminación de pequeñas regiones de una proteína para promover o restringir los cambios conformacionales que ocurren durante la catálisis.



Las proteínas en la vida moderna

<http://youtu.be/kg2hcJtpSRs>

Casos de éxito de proteínas mejoradas

El número reportado de enzimas mejoradas mediante ingeniería de proteínas es grande, y existen ejemplos para diferentes propósitos tales como: enzimas con usos terapéuticos, industriales o medioambientales. A continuación se describen tres ejemplos del mejoramiento de enzimas utilizando diferentes métodos.

Ejemplo 1

Para eliminar restos biológicos de la ropa, algunos detergentes incorporan en su fórmula enzimas como proteasas, lipasas o amilasas. Las enzimas añadidas deben ser estables en las condiciones de lavado que incluyen la presencia de surfactantes, agentes oxidantes y, en ocasiones, altas temperaturas. La subtilina es una proteína con actividad de proteasa –es decir, que destruye otras proteínas– que probó ser útil al ser incorporada en detergentes. Sin embargo, se encontró que el aminoácido metionina, cercano al sitio activo, es susceptible a ser oxidado y esto reduce la actividad catalítica de la enzima. Con base en este dato se planteó modificar dicho residuo por cada uno de los otros 19 aminoácidos para identificar un reemplazo de la metionina resistente al daño oxidativo (ESTELL et al., 1985). Los resultados del trabajo mostraron que, en general, los cambios por aminoácidos pequeños conservaban la actividad mientras que aminoácidos grandes o cargados eléctricamente provocaban una disminución considerable en la misma. El cambio de metionina por serina o alanina mantuvo actividades suficientemente buenas y generó resistencia al daño oxidativo. Este caso ejemplifica la solución a un problema de

inestabilidad de la enzima que fue abordado al generar una pequeña librería de variantes del sitio único responsable de esta inestabilidad.

Ejemplo 2

En este estudio se requería el cambio de especificidad de sustrato de una enzima, es decir, que ésta prefiriera actuar sobre un compuesto diferente al que transforma naturalmente. El objetivo fue invertir la especificidad de la enzima xilosa reductasa de una levadura llamada *Pichia stipitis*, que naturalmente utiliza NADP como cofactor, para que en su lugar utilizara NAD en la reacción. Este cambio disminuiría el desbalance oxidativo poco favorable en la producción de alcohol a partir de xilosa. Este trabajo tuvo un enfoque semi-racional usando una combinación de diseño racional y evolución dirigida. La estrategia utilizada consistió en el análisis de la estructura tridimensional de la proteína para identificar y hacer una lista de los aminoácidos que están involucrados en el reconocimiento del cofactor (KRAHULEC et al., 2009).

Al momento del estudio, la enzima xilosa reductasa no contaba con datos cristalográficos de estructura, pero fue posible predecir un modelo estructural por su parecido a proteínas homólogas con estructura conocida. En el modelo estructural se identificaron 6 aminoácidos involucrados en el reconocimiento del cofactor lo que los convertía en candidatos a ser modificados. Dos de los sitios candidatos se mutaron al azar simultáneamente en una primera ronda de mutación, y la variante con mejor actividad se eligió como molde de una segunda ronda de mutación donde se cambiaron otros dos sitios de la lista inicial. Una tercera ronda de mutaciones se realizó con la mejor mutante de la segunda ronda. La variante final pasó de preferir dos veces NADP a preferir 25 veces NAD. Esta estrategia se basó en el uso de la información estructural de la proteína lo que redujo considerablemente el número de variantes posibles, lo cual fue de gran ayuda debido a que el método de selección no permitía abarcar el análisis de muchas variantes.

Ejemplo 3

En este ejemplo se abordó el problema de la degradación de compuestos tóxicos introducidos al ambiente por procesos industriales y agrícolas mediante la estrategia de barajeo de DNA. Como sistema, se eligieron dos enzimas llamadas atzA y triA, que presentan solamente nueve aminoácidos de diferencia entre sí. Éstas inician la hidrólisis de los compuestos s-triazinas, una mediante la eliminación de cloro (desclorinación) y la otra mediante la eliminación del grupo amino (desaminación) de los compuestos atrazina y melamina, respectivamente. Las s-triazinas comenzaron a usarse recientemente como herbicidas, pesticidas o en plásticos, por lo cual su liberación en el ambiente ha sido extensa. Para el estudio se diseñó una familia de 15 compuestos distintos de s-triazinas. Posteriormente, se combinaron las secuencias de ambos genes para crear una colección de variantes para examinar su capacidad hidrolítica contra cada uno de los sustratos. De esta colección se identificaron enzimas que hidrolizaban cinco de ocho sustratos que no eran hidrolizados por las enzimas parentales. Mientras que para los otros siete sustratos, donde las enzimas parentales mostraban algo de actividad, la librería produjo variantes con actividades mejoradas que presentaron hasta 150 veces más actividad que la mejor

enzima parental. Además de esto, fue posible extraer, a partir de las variantes encontradas, algo de información acerca de la relación entre la secuencia y la función en estas enzimas (RAILLARD et al., 2001).

Conclusión

El uso de biocatalizadores podría satisfacer la creciente demanda de procesos industriales limpios. Sin embargo, la producción de enzimas a gran escala es todavía costosa. El uso de enzimas que resistan, por más tiempo, las diferentes condiciones operacionales, podría ayudar a aliviar este problema. La ingeniería de proteínas es una alternativa para la producción de actividades específicas tomando como base lo que la naturaleza ha creado. El conocimiento cada vez mayor sobre la genética y bioquímica de las proteínas, junto con el número creciente de secuencias de proyectos de secuenciación masiva y que se encuentran públicamente disponibles, y sobre la estructura de diferentes tipos de proteínas, han facultado el desarrollo de técnicas que en su conjunto permiten identificar residuos clave para la manipulación y confección de enzimas con características mejoradas. ❁

Glosario de términos

Aminoácido: Bloque de construcción de las proteínas que se une con otro en sus extremos para formar cadenas. Existen 20 aminoácidos diferentes en las proteínas encontradas en la naturaleza.

Proteína: Cadena compuesta por los 20 aminoácidos con un orden específico indicado por el gen que la codifica.

Secuencia: Es el orden en que se encuentran los aminoácidos (que forman las proteínas) o los nucleótidos (que forman el DNA) en una cadena lineal.

Residuo o residuo de aminoácido: Aminoácido que forma parte de la secuencia de una proteína.

DNA: También llamado ADN, es la molécula que guarda la información para especificar la secuencia de las proteínas (entre otras funciones) en los seres vivos.

Gen: Cadena de DNA que contiene la información que indica la secuencia específica de una proteína.

Polimerasa de DNA: Un tipo particular de enzima cuya función es replicar nuevas cadenas de DNA.

Mutación: Cambio en la secuencia de la cadena de DNA.

Mutagénesis: Generación de mutaciones en el DNA.

Homólogo: Proteína que tiene cierto parecido en su secuencia con otra.
Catálisis: Aceleración de una reacción química.

Lipasas, proteasas, amilasas: Proteínas que degradan grasas, proteínas y almidón, respectivamente.

Surfactantes: Agentes químicos que humedecen, emulsifican y dispersan compuestos cuando se disuelven en agua.

Agentes oxidantes: Son compuestos que aceptan electrones en una reacción de oxidación/reducción.

Serina, alanina, metionina: Nombres de tres aminoácidos de los 20 encontrados en las proteínas.

NADP, NAD: Son sustancias que actúan en conjunto con algunas proteínas para permitirles llevar a cabo reacciones de oxidación/reducción.

Cofactor: Moléculas que actúan en conjunto con algunas proteínas para permitirles llevar a cabo su función, como el NADP o el NAD.

Reductasa: Enzima que acelera reacciones de oxidación/reducción.

Hidrólisis: Reacción química que consiste en la ruptura de un enlace con la formación de una molécula de agua en el proceso.

Bibliografía

- [1] ESTELL, D.A., T.P. Graycar, y J.A. Wells, "Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation". *Journal of Biological Chemistry*, 1985. 260(11): p. 6518-6521.
- [2] KRAHULEC, S., M. Klimacek, and B. Nidetzky, "Engineering of a matched pair of xylose reductase and xylitol dehydrogenase for xylose fermentation by *Saccharomyces cerevisiae*". *Biotechnology Journal*, 2009. 4(5): p. 684-694.
- [3] PEÑA-MONTES, C. y A. Farrés Gonzalez-Saravia, "Evolución Dirigida en la generación de biocatalizadores: Biocatalizadores hechos a medida". *BioTecnología*, 2008. 12(2): p. 5-23
- [4] RAILLARD, S., *et al.*, "Novel enzyme activities and functional plasticity revealed by recombining highly homologous enzymes". *Chemistry & Biology*, 2001. 8(9): p. 891-898.
- [5] SEGOVIA, L. y M. Peimbert, "Ingeniería de Proteínas y Evolución Dirigida". *Mensaje Bioquímico*, 2010. XXXIV: p. 135-141.