



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

REDISEÑO DEL DOMINO ROSSMANN DE LA ENZIMA SHIKIMATO
DESHIDROGENASA DE *E. COLI*: INTERCAMBIO DE LA ACTIVIDAD DE NADP
HACIA NAD

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

José Fernando García Guevara

TUTOR PRINCIPAL

Dr. Lorenzo P. Segovia Forcella
Instituto de Biotecnología, UNAM

MIEMBROS DEL COMITÉ TUTOR

Dr. Daniel Alejandro Fernández
Facultad de Medicina, UNAM

Dra. Mariana Peimbert
Universidad Autónoma Metropolitana, Unidad Cuajimalpa

Ciudad Universitaria, Cd. Mx. Enero, 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Este proyecto fue financiado con fondos de la Dirección General de Asuntos del Personal Académico (DGAPA): IN213511 (<http://dgapa.unam.mx/html/papiit/papit.html>); y fondos del Consejo Nacional de Ciencia y Tecnología (CONACyT): 132580 (<http://www.conacyt.gob.mx>).

Agradezco el apoyo financiero por parte del Programa de Apoyo a los Estudios del Posgrado (PAEP) y al CONACyT por la beca Número (CVU/Becario): 347975/240044 para cursar mis estudios.

Quiero agradecer a mi tutor y al comité tutor por todo su apoyo y guía durante esta etapa.

A los miembros del jurado —Dr. Adelfo Escalante, Dra. Nina Pastor, Dr. Enrique Rudiño, Dra. Gloria Saab y Dr. Alejandro Sosa— por su tiempo y aportaciones.

A todos los miembros del laboratorio.

Y sobre todo, gracias a mi familia y amigos.

Contenido

1 Resumen.....	1
2 Introducción.....	3
2.1 Diseño <i>in silico</i> de enzimas.....	3
2.2 Estrategias alternas que explotan el análisis de secuencias: Ingeniería de consensos.....	5
2.3 Inversión en la especificidad de la enzima Shikimato deshidrogenasa de <i>Escherichia coli</i>	6
2.3.1 La enzima EcoSDH: estructura, función y familia.....	7
2.3.2 Nuestro interés por intercambiar la actividad con NADP+ por NAD+ en EcoSDH.....	11
2.3.3 Propuesta experimental para intercambio de especificidad por el cofactor.....	12
3 Antecedentes.....	14
3.1 Dominio Rossmann consenso de la enzima Shikimato deshidrogenasa de <i>E. coli</i>	14
4 Hipótesis.....	17
5 Objetivos.....	17
6 Materiales y Métodos.....	18
6.1 Diseño <i>in silico</i> con <i>Rosetta Enzyme Design</i> para intercambiar la actividad de NADP+ hacia NAD+.....	18
6.2 Plásmidos y genes.....	18
6.3 Construcción de las mutaciones puntuales.....	19
6.4 Expresión y purificación de proteína.....	19
6.5 Medición de actividad y determinación de los parámetros cinéticos aparentes k_{CAT_app} y K_{M_app}	20
6.6 Complementación del crecimiento en medio mínimo.....	21
6.7 Modelos estructurales de las variantes construidas.....	21
7 Resultados.....	22

7.1 Rediseño del dominio Rossmann usando la aplicación <i>Rosetta Enzyme Design</i>	22
7.1.1 Modelado de los cofactores.....	22
7.1.2 Selección de los residuos para rediseñar.....	23
7.1.3 Selección de los mejores diseños.....	24
7.1.4 Construcción de la mutante N149D.....	26
7.2 Mutación N149D en el fondo del dominio Rossmann consenso.....	26
7.2.1 La mutación consenso S131A nos propone otra mutación para aumentar la actividad con NAD+.....	28
7.3 Complementación del crecimiento en medio mínimo de una cepa de <i>E. coli</i> carente del gen que codifica para la enzima SDH.....	29
8 Discusión	34
8.1 Diseño <i>in silico</i> con <i>Rosetta</i>	34
8.2 Ingeniería de consensos como estrategia complementaria.....	35
8.3 Intercambio en la actividad de NADP+ hacia NAD+ y producción de shikimato.....	37
9 Conclusiones	40
10 Perspectivas	41
11 Referencias	42
12 Anexo	50
Figura anexo 1	50
Figura anexo 2	52
Figura anexo 3	53
Figura anexo 4	54
Tabla anexo 1.....	55
Tabla anexo 2	55

1 Resumen

La ingeniería de proteínas estudia la relación entre la estructura y la función en las proteínas, con la intención de desarrollar las herramientas que nos permitan obtener proteínas con nuevas funciones. Una de las herramientas más prometedoras en la ingeniería de proteínas es el diseño *in silico* de enzimas. A pesar de que ha habido grandes avances en esta herramienta, aún se encuentra en fases tempranas de su desarrollo y todavía tiene una baja tasa de éxito. Para mejorar la actividad de los resultados del diseño por computadora es común utilizar estrategias experimentales complementarias. Normalmente se usa evolución dirigida, pero cuando no se cuenta con un sistema de selección o un sistema de tamizaje de actividad eficiente, usarla presenta algunas limitaciones. Propusimos que podíamos usar la ingeniería de consensos como herramienta complementaria para los resultados del diseño *in silico*; la estrategia normalmente se ha usado para aumentar la estabilidad térmica de proteínas, pero algunas observaciones apuntan a que podría usarse para aumentar la actividad con otros sustratos.

Una modificación frecuente en la ingeniería de proteínas es el intercambio de la especificidad por el cofactor en enzimas que utilizan NADP⁺, para que ahora utilicen NAD⁺, porque es más abundante y estable. El shikimato es un metabolito de interés industrial, por ser precursor en la síntesis química del oseltamivir-fosfato. Obtener una enzima shikimato deshidrogenasa que prefiera NAD⁺, en lugar de NADP⁺, podría beneficiar la producción del shikimato.

Usamos la aplicación de diseño de enzimas *Rosetta Enzyme Design* para modificar a la enzima shikimato deshidrogenasa de *Escherichia coli* para que ahora prefiera NAD⁺ en lugar de NADP⁺; posteriormente probamos el efecto de introducir las mutaciones consenso sobre el diseño. Usando *Rosetta Enzyme Design*, identificamos una mutación puntual que nos permitió disminuir la actividad de la enzima usando NADP⁺, pero encontramos sólo un ligero aumento en la actividad con NAD⁺. Al juntar el diseño *in silico* con las mutaciones consenso logramos aumentar la actividad de nuestro diseño. Las mutaciones consenso cercanas al sitio de unión 2' del cofactor son suficientes para el aumento que observamos en la actividad. Al final construimos una variante que pierde más de mil veces la actividad con NADP⁺ y aumenta un orden de magnitud su actividad con NAD⁺, comparada con la enzima silvestre. Tal vez necesitemos mejorar todavía más la actividad con NAD⁺ de la variante que obtuvimos, antes de usarla en estudios para intentar mejorar producción de shikimato.

Proponemos a la ingeniería de consenso como una herramienta auxiliar para mejorar la actividad con otros sustratos. Entre las ventajas de la ingeniería de consensos están que aprovecha la información de secuencia disponible en las bases de datos.

Summary

Protein engineering studies the relationship between structure and function in proteins, and develops tools for constructing proteins with new functions. One of the most promising tools in protein engineering is *in silico* enzyme design. This tool has shown significant advances, but still is in an early stage of development presenting a low success rate. A common practice to enhance *in silico* enzyme design results is to employ complementary tools. The most common and successful tool is directed evolution, however if there is not an efficient selection system for the desired function, then the implementation of directed evolution becomes burdensome. We proposed we could use consensus engineering as a complementary tool for the results of *in silico* enzyme design; consensus engineering has been previously employed to increase the thermostability in some proteins, but some observations suggest it also could be used to increase the enzymatic activity with different substrates.

We employed the *Rosetta Enzyme Design* application to design a cofactor specificity switch in the Shikimate dehydrogenase enzyme from *Escherichia coli*; then we tested the effect of introducing the mutations proposed by consensus engineering on top of the *in silico* design results. We identified a point mutation using *Rosetta Enzyme Design* that lead to a decreased enzyme activity using NADP⁺ and a slight increase using NAD⁺. When we added the mutations proposed by consensus engineering, we increased the enzyme activity. We found that the consensus mutations closest to the 2' adenine moiety produced the activity increase.

The enzyme activity using NADP⁺ in our final variant decreased over a thousand fold, meanwhile enzyme activity using NAD⁺ increased a tenfold. Our findings suggest consensus engineering could also be used as a complementary tool for increasing or modifying enzyme activity during design. Consensus engineering is straightforward to implement and now days we rely on huge sequence databases which aid to implement this technique.

2 Introducción

La secuencia de aminoácidos de cada proteína determina sus características estructurales y funcionales (Anfinsen, 1972). Lo anterior significa que podemos modificar la función de las proteínas, si realizamos cambios en su secuencia primaria. Gran parte del esfuerzo de la Ingeniería de proteínas se ha enfocado en estudiar cómo rediseñar enzimas naturales, para que utilicen sustratos distintos. Muchas pistas para rediseñar interacciones con otros sustratos, se han obtenido a través de estudios estructurales de la unión entre una enzima y su sustrato (W. Yang & Lai, 2017).

Las enzimas pueden acelerar reacciones químicas varios órdenes de magnitud al estabilizar los estados de transición y reducir la energía de activación. Esto permite que las reacciones químicas ocurran en escalas de tiempo que permitan la vida. Los residuos catalíticos deben estar colocados alrededor de los sustratos en la orientación correcta para estabilizar los estados de transición (Tyack, Furnham, Sillitoe, Orengo, & Thornton, 2017). Dentro de los factores principales que dominan la interacción de una enzima con el sustrato encontramos: las preferencias de los ángulos diedro, las interacciones hidrofóbicas, los puentes de hidrógeno, las interacciones de van der Waals, las interacciones electrostáticas, y la interacción con el solvente. Las cadenas laterales y la cadena principal se encuentran orientados de forma precisa para unir sus sustratos, existen ciertos ángulos diedro que se prefieren dependiendo del aminoácido del que se trate. Según la naturaleza del sustrato que una la enzima, podremos observar interacciones con puentes de hidrógeno o interacciones hidrofóbicas. Los átomos en un complejo proteína-sustrato se encuentran empacados de forma ajustada, esto resalta la importancia de las interacciones de van der Wals. Algunos aminoácidos poseen carga, con la que pueden atraer o repeler a otras moléculas. Un sustrato pasa de interactuar con las moléculas del solvente hacia un ambiente distinto cuando se une a la enzima, pudiendo a su vez desplazar moléculas del solvente del sitio activo de la enzima (Dill & MacCallum, 2012). La mecánica molecular ha usado campos de fuerza (modelos moleculares de energías potenciales) para tratar de describir estas fuerzas físicas, que a su vez usamos para realizar simulaciones por computadora. Sin embargo hasta la fecha las simulaciones por computadora aún no modelan con precisión la estabilidad de las proteínas, sus propiedades termodinámicas, o las interacciones y afinidades de unión con otras moléculas (Dill & MacCallum, 2012).

2.1 *Diseño in silico de enzimas*

Con los avances en los modelos moleculares, la bioinformática estructural y la capacidad de cómputo, la ingeniería de proteínas ha desarrollado herramientas que nos permiten rediseñar enzimas para que

usen otros sustratos. Entre las herramientas más prometedoras encontramos al rediseño de enzimas por computadora basado en la estructura. El objetivo de estas herramientas es regresar secuencias que adopten una estructura definida, la cual pudiera unir al sustrato deseado en la orientación correcta para catalizar una reacción. Para ello el algoritmo busca las secuencias que minimicen una función de energía y que cumplan con los criterios establecidos (Gainza, Nisonoff, & Donald, 2016; Liu & Chen, 2016; W. Yang & Lai, 2017). Una estrategia eficiente para el diseño de enzimas es usar como punto de partida estructuras de enzimas que ocurren naturalmente, aprovechar la actividad original de la enzima y solo modelar los residuos que interactúan con el sustrato deseado (W. Yang & Lai, 2017).

La función de energía toma algunos de sus términos de los campos de fuerza de la mecánica molecular. Aunque también contiene términos basados en estadística, que derivan de datos de secuencia y estructura de proteínas naturales, por ejemplo la distribución de los ángulos de Ramachandran, la distribución de rotámeros, o la exposición al solvente; estos términos basados en estadística son relevantes porque capturan información que todavía no se puede capturar con los modelos físicos (Liu & Chen, 2016).

El diseño por computadora aún se encuentra en etapas tempranas del desarrollo. Existen algunos casos exitosos, sin embargo la tasa de éxito aún es muy baja; así como también es baja la actividad en los diseños que sí funcionan (Jiang et al., 2008; Röthlisberger et al., 2008; Siegel et al., 2010). El éxito de los programas de diseño depende de dos aspectos, el primero es modelo biofísico para representar los eventos que ocurren en la naturaleza y el segundo es un algoritmo que construya la mejor secuencia de acuerdo al modelo (Gainza et al., 2016). Los retos que hay que enfrentar son muchos. Del lado del modelo biofísico podemos mencionar algunos ejemplos: un sustrato puede tomar diferentes conformaciones, pero durante el diseño la interacción enzima-sustrato se predefine, lo que limita la búsqueda; tampoco se considera la dinámica intrínseca de la proteína y sus cambios conformacionales; también, los residuos remotos pueden alterar las propiedades en la unión pero normalmente solo se modelan los residuos de la primera capa de unión; el solvente se trata implícitamente, lo que impide modelar con precisión el efecto de las interacciones polares del soluto y el efecto de la desolvatación; finalmente, los modelos de la energía de unión proteína-ligando no tienen la suficiente precisión para distinguir las interacciones óptimas, los ligandos son muy sensibles a la precisión de los modelos de la energía de unión debido a pequeño tamaño (Gainza et al., 2016; W. Yang & Lai, 2017). Del lado del algoritmo de búsqueda: mientras más preciso se convierte un modelo, mayor es la complejidad

computacional de la búsqueda de la mejor secuencia, y la capacidad de computación se vuelve una limitante (Gainza et al., 2016).

Mientras los algoritmos en el diseño por computadora se siguen desarrollando y mejorando, una alternativa a nuestro alcance es emplear estrategias experimentales alternas para mejorar la actividad sobre los diseños construidos (Giger et al., 2013). Usualmente la estrategia empleada es la evolución dirigida, la cual consiste en generar nuevas variantes de la proteína insertando cambios al azar en la secuencia. Posteriormente se buscan aquellas que cumplan con nuestros criterios, usando un método de selección que permita revisar el mayor número posible de variantes en poco tiempo (Kuchner & Arnold, 1997). Sin embargo, aplicar la evolución dirigida resulta poco práctica si no se cuenta con sistema de selección eficiente, por ello nos interesa explorar estrategias alternativas para complementar los resultados de los diseños por computadora.

2.2 Estrategias alternas que explotan el análisis de secuencias: Ingeniería de consensos

El análisis de secuencias en enzimas homólogas nos permite obtener información valiosa acerca de la proteína, que podemos aprovechar en la ingeniería de proteínas. La ingeniería de consensos es un claro ejemplo ya que ha permitido generar proteínas con mayor estabilidad térmica (cuadro 1).

Observaciones previas sugieren que las secuencias consenso podrían ser usadas también para aumentar la actividad con sustratos distintos (Aerts, Verhaeghe, Joosten, Vriend, & Soetaert, 2013; Loening, Fenn, Wu, & Gambhir, 2006); por eso, nos preguntamos si podíamos usar la ingeniería de consensos como herramienta auxiliar, para complementar los resultados del diseño racional y ampliar nuestra caja de herramientas en la ingeniería de proteínas.

Cuadro 1. Ingeniería de consensos. La ingeniería de consensos es una técnica que se ha empleado para aumentar la estabilidad térmica de algunas proteínas (M Lehmann, Pasamontes, Lassen, & Wyss, 2000; Martin Lehmann, Pasamontes, Lassen, & Wyss, 2000; Steipe, Schiller, Pluckthun, & Steinbacher, 1994; Vázquez-Figueroa, Chapparro-Riggers, & Bommarium, 2007). La técnica consiste en introducir mutaciones en una secuencia usando el aminoácido más frecuente en un alineamiento de secuencias de proteínas homólogas. Para construir una secuencia consenso comenzamos con una secuencia de referencia. Después buscamos secuencias homólogas (pueden ser tan pocas como tres secuencias en total) y realizamos un alineamiento múltiple. Para cada sitio del alineamiento seleccionamos el residuo que se observa con mayor frecuencia. Cuando no se puede establecer claramente cuál es el residuo más frecuente normalmente se deja el mismo residuo que en la secuencia de referencia. Al final terminamos con una nueva secuencia a la que denominamos secuencia consenso. Una parte esencial de la técnica es iniciar con un buen alineamiento, es decir, estar seguro de que exista una correspondencia entre los residuos alineados. Normalmente esta parte es subjetiva por lo que se prefiere contar con secuencias con porcentaje de identidad altos para asegurar un buen alineamiento, pero también con suficiente divergencia entre ellas para evitar sesgos filogenéticos (Jackel, Bloom, Kast, Arnold, & Hilvert, 2010; Martin Lehmann, Kostrewa, et al., 2000; Martin Lehmann, Pasamontes, et al., 2000; Martin Lehmann & Wyss, 2001).

2.3 Inversión en la especificidad de la enzima Shikimato deshidrogenasa de *Escherichia coli*

Dentro de los temas más aplicables en el rediseño de enzimas encontramos el intercambio en la especificidad por la coenzima en enzimas que normalmente usan NADP⁺, para que ahora utilicen NAD⁺. Este intercambio ha mostrado ser de importancia biotecnológica para la producción de algunos compuestos de interés, además del interés intrínseco que representa para estudiar el reconocimiento molecular (Khoury et al., 2009; Scrutton, Berry, & Perham, 1990). En nuestro proyecto propusimos rediseñar al dominio Rossmann de la enzima Shikimato deshidrogenasa de *Escherichia coli* (EcoSDH) para que ahora prefiera NAD⁺ en lugar de NADP⁺.

El shikimato es un compuesto con un anillo de seis carbonos altamente funcionalizado con tres centros anoméricos. El shikimato es un precursor para la producción de compuestos con aplicaciones farmacéuticas como antipiréticos, antioxidantes, anticoagulantes, antitrombóticas, antiinflamatorias, y analgésicas (Martínez, Bolívar, & Escalante, 2015; Rodríguez et al., 2014). Específicamente el shikimato

es relevante en la industria farmacéutica por ser el precursor en la síntesis química del oseltamivir fosfato, cuya demanda ha ido en aumento debido a brotes recientes de influenza (Rawat, Tripathi, Jahan, & Saxena, 2013). El oseltamivir fosfato se utiliza como antiviral contra algunos virus estacionales de influenza, los cuales representan un grave riesgo para la población en caso de una pandemia (Martínez et al., 2015).

La ingeniería de vías metabólicas en microorganismos recombinantes es una de las alternativas más llamativas para producir shikimato. Entre las razones están la alta productividad y rendimiento, la amabilidad con el ambiente, los bajos costos de la materia prima y la selectividad de los productos. *Escherichia coli* es uno de los microorganismos más estudiados en la producción de shikimato (Ghosh, Chisti, & Banerjee, 2012; Krämer et al., 2003; Martínez et al., 2015; Rawat et al., 2013). La enzima EcoSDH participa en la vía de síntesis de shikimato y utiliza NADP⁺ como cofactor favorito (Michel et al., 2003). Proponemos que una enzima SDH que prefiera NAD⁺ podría enriquecer los estudios para mejorar la producción de shikimato. En las siguientes secciones de la introducción presentamos la función y la estructura de la enzima EcoSDH; detallamos por qué nos interesa obtener una variante de la EcoSDH que ahora prefiera usar NAD⁺; y describimos nuestra propuesta para emplear a la ingeniería de consensos como estrategia complementaria en el diseño de enzimas.

2.3.1 La enzima EcoSDH: estructura, función y familia

La enzima EcoSDH cataliza la reacción reversible desde 3-dehidroshikimato hacia shikimato usando una molécula de NADPH y produciendo NADP⁺ (Figura 1). El shikimato es un intermediario en la vía común de síntesis de aminoácidos aromáticos (Chaudhuri & Coggins, 1985). Se ha reportado que la enzima EcoSDH sólo puede usar NADP⁺ como cofactor y shikimato como sustrato, mientras que no tiene actividad con NAD⁺ y ni siquiera lo reconoce (Michel et al., 2003).

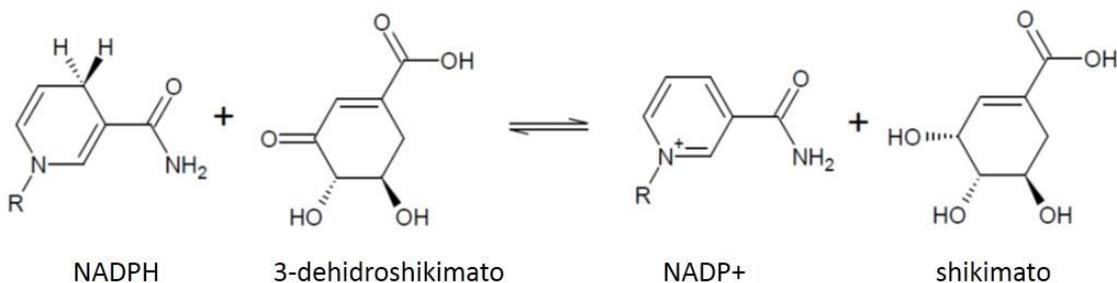
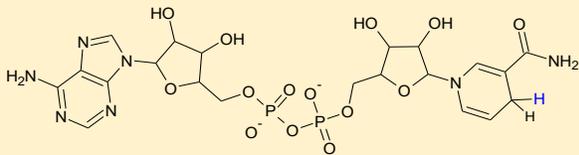


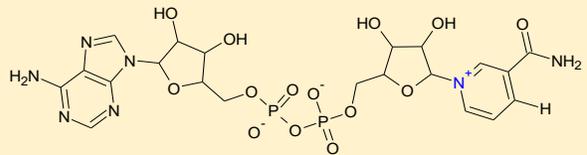
Figura 1. Reacción catalizada por la enzima EcoSDH (Michel et al., 2003). Construimos las figuras con el programa ACD/ChemSketch(Freeware version)2016 (Advanced Chemistry Development Inc.)

Cuadro 2. *NAD* y *NADP*. Las moléculas *NAD* y *NADP* están formadas por los nucleótidos adenina y nicotinamida unidos con un enlace difosfato por los extremos 5'; la única diferencia entre ellas es que el grupo 2' de la adenosina es un hidroxilo en *NAD* y un fosfato en *NADP*. La forma reducida de estas moléculas es *NADH* y *NADPH*, la forma oxidada de las moléculas es *NAD⁺* y *NADP⁺* (Figura cuadro 2); cuando participa como cofactor en reacciones óxido-reducción, la fracción de nicotinamida es la que dona o recibe el anión de hidrógeno (Pollak, Dölle, & Ziegler, 2007). El *NAD* está normalmente asociado al catabolismo, y el *NADP* está asociado al anabolismo. La forma reducida de estas moléculas muestra un máximo a 340 nm en su espectro de absorbancia, en contraste con la forma oxidada que a esa longitud de onda su espectro de absorbancia es plano, lo que permite seguir la velocidad de la reacción (McComb, Bond, Burnett, Keech, & Bowers, 1976).

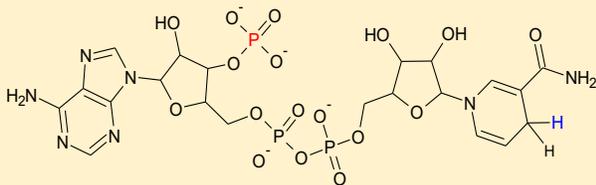
A)



B)



C)



D)

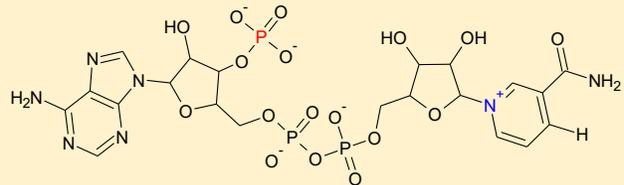


Figura cuadro 2. Los esquemas de arriba muestran la estructura química de los cofactores *NAD* y *NADP* en sus formas reducida y oxidada. A) *NADH*, B) *NAD⁺*, C) *NADPH*, y D) *NADP⁺*. Del lado izquierdo de cada figura se observa la fracción de adenosina y del lado derecho, la fracción de nicotinamida. En rojo se señala el grupo fosfato 2' en *NADP*. En azul se resalta si la molécula se encuentra en estado reducido u oxidado. Dibujamos las figuras con el programa *ACD/ChemSketch(Freeware version)2016* (Advanced Chemistry Development Inc.)

La enzima *EcoSDH* es un monómero en solución (Maclean et al., 2000). Tiene dos dominios estructurales: el primer dominio es conocido como *dominio catalítico* y une al shikimato; el segundo dominio es conocido como *dominio Rossmann* y este une a la molécula de *NADP⁺*. En la secuencia de aminoácidos

el dominio Rossmann se encuentra insertado dentro del dominio catalítico, pero en la estructura tridimensional forman dos dominios separados (Figura 2) (Michel et al., 2003).

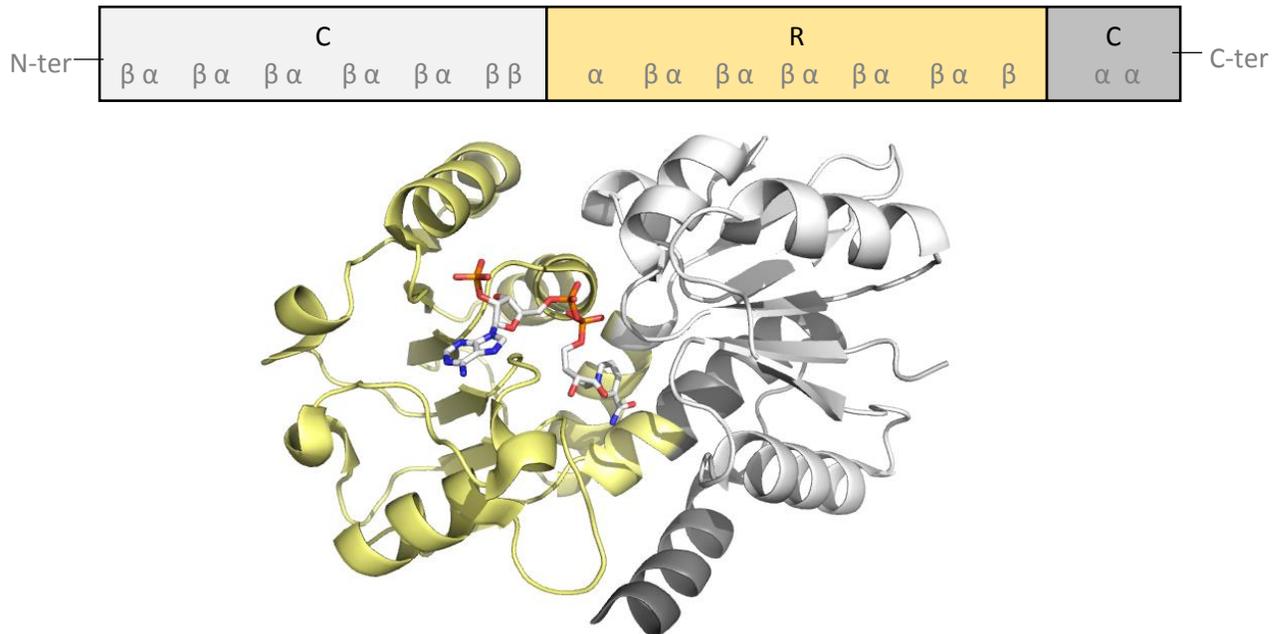


Figura 2. Dominios de la enzima EcoSDH. La enzima EcoSDH está formada por dos dominios, el extremo amino de la proteína comienza con el dominio catalítico (blanco), luego, antes de que el dominio catalítico quede completo, está insertado el dominio Rossmann (amarillo claro). Después del dominio Rossmann un par de hélices alfa en el extremo carboxilo (gris) completan al dominio catalítico. En palillos se muestra la estructura del NADP+ unido al dominio Rossmann. La imagen está elaborada a partir de la cadena A de la estructura 1NYT del PDB (Michel et al., 2003).

El mecanismo catalítico en EcoSDH no ha sido descrito con profundidad. Se sabe que la transferencia del anión de hidrógeno ocurre desde la cara A de la fracción de nicotinamida del cofactor (Haslam & Turner, 1971). La catálisis depende del pH, teniendo un pH óptimo entre 5 y 7 en la dirección de síntesis de shikimato y entre 8 y 10 en la dirección contraria (Michel et al., 2003). Se propone que la reacción que cataliza EcoSDH tiene un mecanismo secuencial, como también ocurre en otras deshidrogenasas. En las enzimas SDH de *Mycobacterium tuberculosis* y *Pisum sativum* se observa un mecanismo secuencial ordenado; pero en la enzima de *M. tuberculosis* se une primero el dehidroshikimato, mientras que en la de *P. sativum* se une primero el NADPH (Balinsky & Dennis, 1971; Fonseca et al., 2007). También se propone que una vez unidos los sustratos, ocurre un cambio conformacional en la enzima: el *linker* que

une a los dominios actúa como bisagra y los dominios giran ligeramente uno respecto al otro para acercar al sustrato con el cofactor (Michel et al., 2003).

Cuadro 3. *El dominio Rossmann.* El plegamiento Rossmann está formado por una lámina beta paralela con orden 3-2-1-4-5-6, conectadas por hélices alfa, la lámina beta queda en forma de emparedado en medio de las hélices alfa. La zona donde se lleva a cabo la función se encuentra siempre sobre el extremo carboxilo de la lámina beta central (Figura cuadro 2) (Rao & Rossmann, 1973). El dominio Rossmann está presente gran variedad proteínas con diversas propiedades y está principalmente representado en proteínas que unen dinucleótidos (como NAD⁺, NADP⁺ o FAD) para reacciones de óxido-reducción (Ojha, Meng, & Babbitt, 2007).

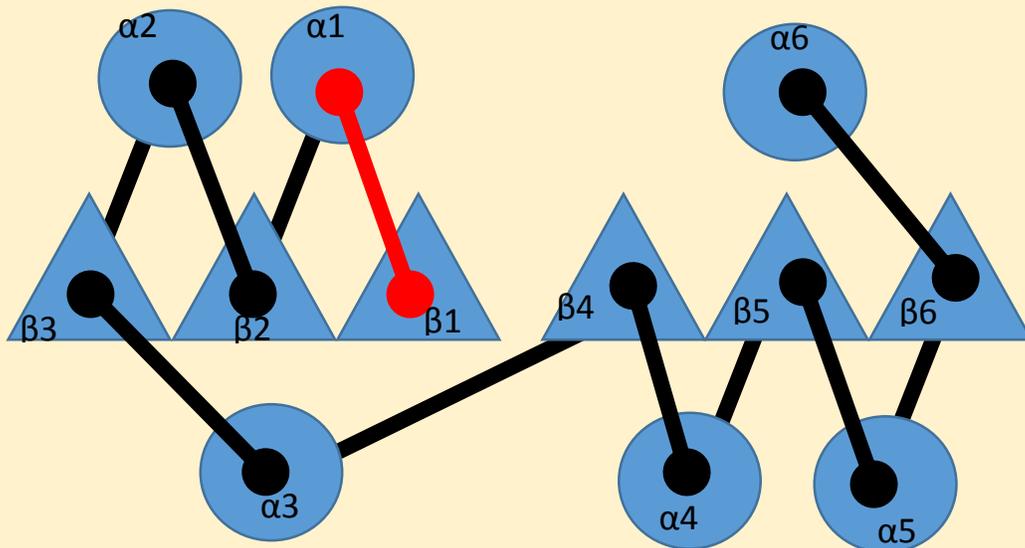


Figura cuadro 3. Esquema de la topología del dominio Rossmann. El esquema de arriba muestra la topología del dominio Rossmann. Los triángulos representan líneas beta y los óvalos hélices alfa, las líneas representan las asas. El dominio Rossmann canónico está formado por una lámina beta central de líneas paralelas con orden 3-2-1-4-5-6 y conectadas por hélices alfa (Rao & Rossmann, 1973). En rojo se muestra el asa de glicinas conservado característico en los dominios Rossmann que unen dinucleótidos (Bellamacina, 1996).

La enzima EcoSDH pertenece a la familia de proteínas *quinato/shikimato deshidrogenasa* (Q/SDH). La familia posee cinco grupos caracterizados de enzimas –AroE, YdiB, SdhL, Rifi, Ael1– que poseen estructuras similares, pero tienen diferentes funciones como: la síntesis de aminoácidos, el uso de

quinato o la producción de metabolitos secundarios (Peek, Garcia, Lee, & Christendat, 2013; Singh, Stavrinides, Christendat, & Guttman, 2008). Dentro la familia hay variantes que prefieren NAD⁺ como cofactor. Desafortunadamente ninguna de ellas es específica por shikimato, ya que todas ellas también usan quinato como sustrato, además tienen poca actividad con shikimato (Peek & Christendat, 2015).

2.3.2 Nuestro interés por intercambiar la actividad con NADP por NAD en EcoSDH.

La producción de shikimato en condiciones limitantes de oxígeno muestra un bajo consumo de glucosa y una producción alta de ácidos orgánicos como acetato y succinato (Kogure, Kubota, Suda, Hiraga, & Inui, 2016). Este evento se atribuye a que el NADH generado en la glicólisis, no puede ser oxidado de vuelta a NAD⁺ durante la síntesis de shikimato (Hasegawa et al., 2012). Por cada molécula de fosfoenol piruvato (PEP) generada a partir de glucosa se produce una molécula de NADH, mientras que por cada molécula producida de shikimato se consume una de NADPH, lo que genera un desbalance; si contáramos con una enzima SDH que prefiera NAD en lugar de NADP podríamos utilizar el NADH generado durante la glicólisis para la producción de shikimato. Existen enzimas en la naturaleza que prefieren NAD sobre NADP, pero todas prefieren como sustrato al quinato, lo que puede desviar el flujo de carbono en la producción de shikimato (Martínez et al., 2015). Por ello planteamos intercambiar la afinidad de NADP hacia NAD del dominio Rossmann en la enzima EcoSDH para obtener una enzima SDH que prefiera NAD como cofactor y use exclusivamente shikimato como sustrato (Figura 3).

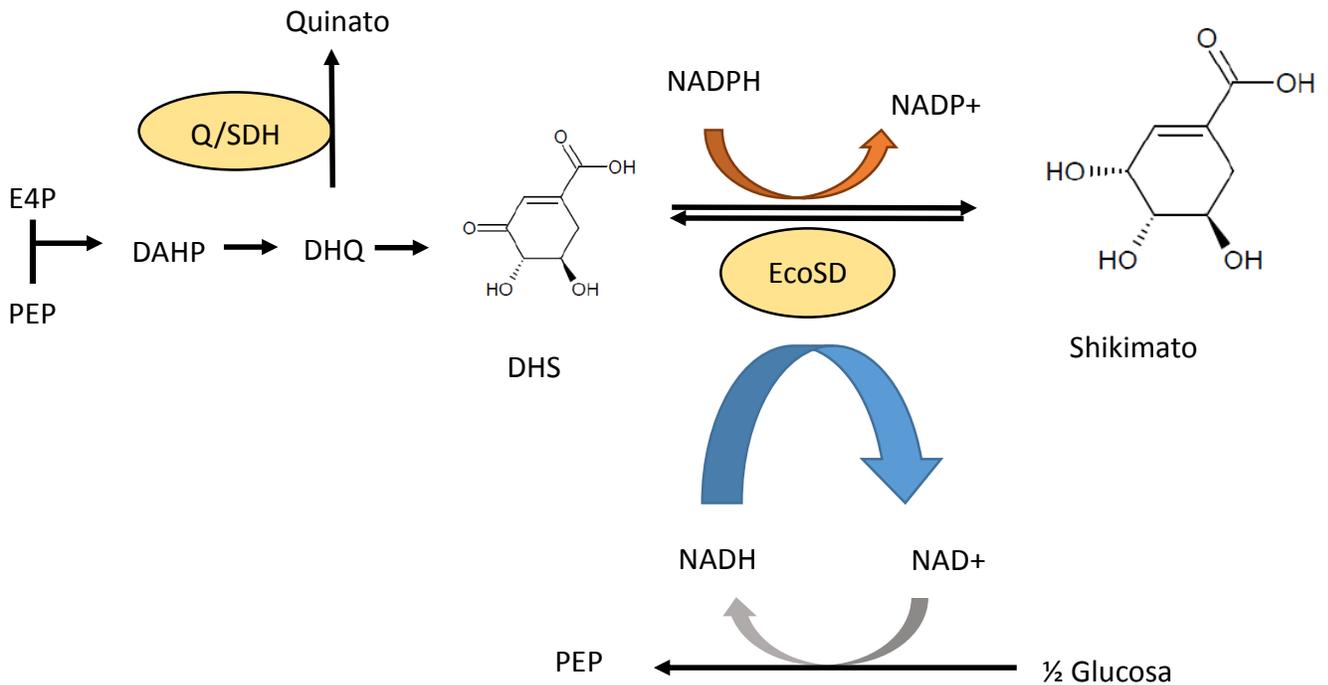


Figura 3. Esquema producción de shikimato. El shikimato se sintetiza en cuatro pasos partiendo de los precursores fosfoenol piruvato (PEP) y eritrosa 4-fosfato (E4P). La enzima SDH cataliza el cuarto paso de la vía y usa NADPH para producir NADP⁺ (flecha naranja). En la glicólisis por cada molécula de PEP se genera una de NADH (flecha gris). Proponemos cambiar la preferencia de NADP hacia NAD en la enzima SDH para usar el NADH producido en la glicólisis y regenerar NAD⁺ (flecha azul). Las enzimas homólogas conocidas que prefieren NAD también prefieren al quinato como sustrato, lo que puede desviar el flujo de carbono en la producción de (Martínez et al., 2015; Rodríguez et al., 2014).

2.3.3 Propuesta experimental para intercambio de especificidad por el cofactor.

Para que la enzima EcoSDH pueda usar NAD en lugar de NADP planeamos emplear la aplicación de *Rosetta Enzyme Design*, esta herramienta ha mostrado ser exitosa para diseñar nuevas actividades y modificar proteínas para unir ligandos distintos (Jiang et al., 2008; Röthlisberger et al., 2008; Siegel et al., 2010). *Rosetta* permite identificar secuencias compatibles con un plegamiento, ha mostrado tener éxito en problemas como: el rediseño de plegamientos existentes (Dantas, Kuhlman, Callender, Wong, & Baker, 2003), el diseño de nuevos plegamientos (Kuhlman et al., 2003), y el diseño de nuevas actividades enzimáticas (Bjelic et al., 2013; Jiang et al., 2012; Röthlisberger et al., 2008; Siegel et al., 2010). Las principales características del programa son: una función de energía que evalúa las secuencias propuestas y un método de búsqueda eficiente para buscar en el espacio de posibles secuencias. La función de energía es dominada por interacciones entre átomos (usando un potencial de Lennard-Jones), puentes de hidrógenos y un modelo de solvatación implícito. El método de búsqueda usa algoritmos de optimización para reducir el espacio de búsqueda y encontrar las secuencias con mejor energía (Dantas et al., 2003). Para realizar diseño de enzimas la aplicación *Rosetta Enzyme Design* recibe la estructura de la proteína donde va a realizar cambios junto con el ligando deseado, entonces comienza a realizar mutaciones en los sitios se le indique; después acepta o rechaza los cambios siguiendo un criterio *Metropolis* (cuando una mutación produce una mejor energía entonces acepta la mutación, cuando la mutación empeora la energía hay una baja probabilidad de aceptarla y una alta probabilidad de rechazarla). Al final regresa una nueva secuencia que, de acuerdo a la función de energía de *Rosetta*, podría reconocer con afinidad al ligando deseado (Röthlisberger et al., 2008).

Las actividades obtenidas por diseño con *Rosetta* suelen ser bajas, debido a que los modelos empleados por el programa requieren simplificaciones que aceleran los cálculos realizados (por ejemplo en las posibles conformaciones que pueden tomar los átomos en el sistema o en la energía de interacción entre la proteína y el sustrato). Sin dichas simplificaciones, el tiempo de cómputo aumentaría tanto que

haría extremadamente lenta una búsqueda amplia del espacio de posibles secuencias. Los diseños generados generalmente se complementan con ensayos de evolución dirigida, para aumentar su actividad (Röthlisberger et al., 2008). El éxito de la evolución dirigida está estrechamente relacionado con la cantidad de variantes probadas, sin embargo, no contamos con un sistema eficiente para seleccionar variantes con mejor actividad sobre NAD⁺. Normalmente los sistemas que han usado en otros reportes de inversión de especificidad son de tamizaje de actividad, que consumen muchos recursos, lo que dificulta probar muchas variantes (Liang, Zhang, & Lin, 2007). Por esta razón, nos preguntamos si podríamos usar una estrategia complementaria diferente, para aumentar la actividad sobre el resultado del diseño *in silico*. Como estrategia complementaria propusimos la Ingeniería de consensos, ya que se ha observado que las secuencias consenso aumentan la actividad con otros sustratos, pero todos los estudios de ingeniería de consensos se han enfocado en aumentar la estabilidad térmica de las proteínas (Aerts et al., 2013; Loening et al., 2006).

Planteamos primero usar el diseño *in silico* para diseñar un cambio en la especificidad por el cofactor, después nos preguntamos si podíamos complementar al diseño *in silico* usando la ingeniería de consensos. Elegimos la herramienta *Rosetta Enzyme Design* para intercambiar la actividad del cofactor y que ahora use NAD⁺ en lugar de NADP⁺. Las mutaciones propuestas por el diseño *in silico* las probamos en el fondo de un dominio Rossmann consenso, y vimos el efecto que tiene la ingeniería de consensos sobre la actividad de NAD⁺.

3 Antecedentes

3.1 Dominio Rossmann consenso de la enzima SDH de *E. coli*

En nuestro grupo usamos a la enzima EcoSDH como modelo de estudio. Iris Bravo, integrante de nuestro laboratorio, aplicó la ingeniería de consensos sobre el dominio Rossmann de EcoSDH. Con este trabajo tenía en la mira conseguir un dominio Rossmann con mayor estabilidad térmica, y eventualmente poder intercambiarlo con el dominio Rossmann de otras enzimas de la familia SDH. La enzima que construyó — a la que llamaremos conSDH de aquí en adelante— aumentó su estabilidad térmica y mantuvo su actividad original (Bravo, 2015).

Para construir la enzima consenso se comenzó con una búsqueda de secuencias homólogas a la enzima EcoSDH usando BlastX en la base de datos nr (Pruitt, Tatusova, & Maglott, 2007). Se recuperaron las secuencias con un *e-value* inferior a 1×10^{-5} , y para prevenir un sesgo filogenético se eliminaron aquellas con un porcentaje de identidad superior al 80% entre cualquiera de los posibles pares de secuencias. Después se procedió a realizar un árbol filogenético con las secuencias restantes (aproximadamente 1200 secuencias); de este árbol se seleccionó un grupo de las 47 secuencias más cercanas a EcoSDH, la mayor parte de éstas pertenecientes al orden de las enterobacterias (Bravo, 2015). A continuación, se construyó un alineamiento del dominio Rossmann; el porcentaje de identidad entre cualquiera de las posibles parejas de secuencias se encuentra entre el 40% y el 80%. Con este alineamiento se construyó la secuencia consenso del dominio Rossmann. Un residuo se consideraba consenso si era el más frecuente en la posición del alineamiento, y su frecuencia era superior al 30%; cuando en una posición no se podía encontrar un residuo consenso entonces se mantuvo la secuencia de la enzima silvestre EcoSDH (Bravo, 2015).

El dominio Rossmann consenso de la enzima SDH presentó quince mutaciones respecto a la secuencia EcoSDH silvestre: V104I, R117K, L120D, S131A, L135I, L141Y, D142G, V145I, V152F, E156Q, D195E, I208T, K219L, S231A y A243G (figura 4A). La mitad de las mutaciones son conservativas, mientras que la otra mitad cambian el tipo de aminoácido. Las mutaciones están distribuidas en todo el dominio Rossmann: podemos observar mutaciones tanto en la superficie, como en el interior de la proteína, en hélices alfa, láminas beta y asas (Bravo, 2015).

La secuencia nucleotídica del dominio Rossmann consenso se envió a sintetizar comercialmente. Después se clonó junto con el dominio catalítico de la enzima EcoSDH silvestre para producir una enzima quimérica, a la que llamamos conSDH. Esta construcción está formada por el dominio catalítico silvestre

de EcoSDH y el dominio Rossmann consenso descrito arriba. La enzima conSDH presenta una estructura secundaria similar a EcoSDH (figura 4B), y también mantiene una actividad comparable usando shikimato y NAD⁺. Al medir la estabilidad térmica de la enzima se encontró un ligero aumento de 2° C en la T_{m_app} , además se observó la aparición de un intermediario en la desnaturalización térmica, que posiblemente se deba a que el domino Rossmann se desnaturaliza después del dominio catalítico (Bravo, 2015).

Un hallazgo adicional en conSDH captó nuestra atención: la enzima también tiene algo de actividad con NAD⁺ como cofactor; esto contrasta con la falta de actividad con NAD⁺ en EcoSDH de acuerdo a lo reportado (Michel et al., 2003). Por esta razón, verificamos que la enzima EcoSDH fuera inactiva con NAD⁺, pero encontramos que no es así, EcoSDH sí tiene actividad con NAD⁺. La actividad de conSDH usando NAD⁺ es mayor que la actividad de EcoSDH, es decir, las mutaciones consenso aumentan la actividad con NAD⁺ (Tabla 1 sección Resultados) (García-Guevara *et al*, 2017).

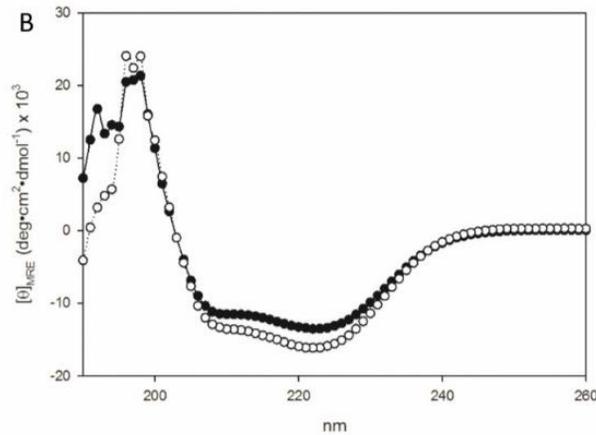
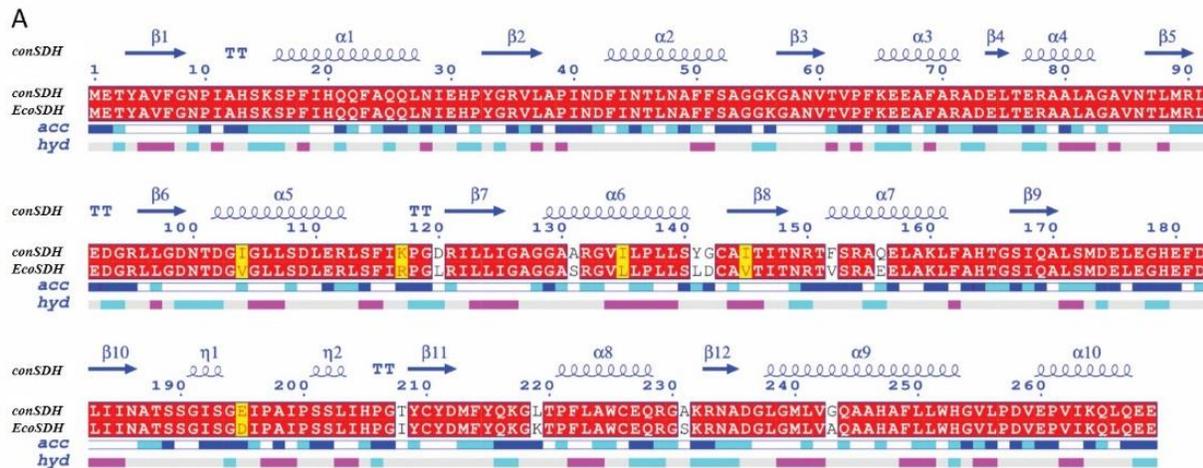


Figura 4. Estructura primaria y secundaria de la enzima conSDH. A) El esquema muestra la secuencia y la estructura secundaria de la enzima conSDH y la secuencia de EcoSDH. Los residuos que no presentaron mutaciones están coloreados con rojo, las mutaciones conservativas con amarillo y las mutaciones no conservativas en blanco; *acc* representa la accesibilidad al solvente, en azul oscuro se indican los residuos expuestos al solvente, en azul cian los residuos medianamente expuestos y en blanco los residuos enterrados; *hyd* indica la hidrofobicidad, en rosa se señalan los residuos hidrofóbicos y en azul cian los residuos hidrofílicos. Esta figura fue construida con el servidor ENDscript (Robert & Gouet, 2014). B) La figura muestra el espectro de dicroísmo circular de la enzima conSDH (círculos vacíos) y la EcoSDH (círculos negros), el espectro muestra la presencia de estructura alfa y beta muy similares en ambas enzimas. (García-Guevara et al., 2017).

4 Hipótesis

Las mutaciones consenso del dominio Rossmann aumentan la actividad con NAD⁺ cuando las agregamos al diseño *in silico*.

5 Objetivos

Intercambiar la especificidad por el cofactor de la enzima EcoSDH, de NADP⁺ hacia NAD⁺.

Objetivos particulares

- Realizar diseño *in silico* sobre EcoSDH con la aplicación *Rosetta Enzyme Design* para intercambiar la afinidad por el cofactor de NADP⁺ hacia NAD⁺.
- Construir las mutaciones propuestas por el diseño *in silico* y medir actividad con NAD⁺ y NADP⁺
- Introducir mutaciones consenso sobre proteína rediseñada con *Rosetta* y medir actividad con NAD⁺ y NADP⁺

6 Materiales y Métodos

6.1 Diseño *in silico* con Rosetta Enzyme Design para intercambiar la actividad de NADP⁺ hacia NAD⁺

Usamos la aplicación *Rosetta Enzyme Design* de *Rosetta 3* para el rediseño *in silico* de la enzima EcoSDH, para que ahora prefiera NAD⁺ sobre NADP⁺. Primero modelamos los cofactores NAD⁺ y NADP⁺ en el formato requerido para usar con la aplicación de *Rosetta*. Para NADP⁺ usamos la misma conformación del cofactor de la cadena A de la estructura cristalográfica 1NYT. Para la conformación de NAD⁺ solo reemplazamos el grupo fosfato 2' de la adenosina en NADP⁺ por un grupo hidroxilo, usando Pymol (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger). Después, usando la aplicación Ligprep con parámetros *default* del programa MAESTRO (Schrödinger Release 2012: Maestro, versión 9.3; LLC, New York, NY, 2012) añadimos los hidrógenos correspondientes sobre cada molécula a pH 7. Por último calculamos las cargas parciales de cada átomo usando el programa MOPAC (MOPAC_2012™, J.J.P. Stewart, Stewart Computational Chemistry; Colorado Springs, CO, USA) y el modelo PM6 (Stewart, 2007).

Los residuos del dominio Rossmann que seleccionamos para que *Rosetta* modificara fueron aquellos dentro de un radio de 10 Ångstroms alrededor del grupo fosfato 2' de NADP⁺. Realizamos dos series de diseños pidiendo al programa 1000 diseños en cada serie: en la primera serie usamos el cofactor deseado, NAD⁺; en la segunda usamos el cofactor original, NADP⁺, como control. De cada serie seleccionamos los diseños con mejor puntuación (menor energía total de *Rosetta*, menor energía de interacción con el cofactor y menor número de residuos polares enterrados no satisfechos), recuperamos la secuencia de los diseños seleccionados y la alineamos contra los respectivos controles. Comparamos los diseños que propuso *Rosetta* usando NADP⁺ contra los diseños que usaron NAD⁺, para buscar diferencias específicas. Seleccionamos las mutaciones que propusiera *Rosetta* usando NAD⁺ que no se encontraran cuando usamos NADP⁺.

6.2 Plásmidos y genes

Usamos un par de construcciones para expresar EcoSDH y conSDH en *E. coli*. Las construcciones contienen el gen que codifica para EcoSDH o conSDH, respectivamente, junto a una fusión traduccional en el extremo 3', que codifica para una cola de polihistidinas; el plásmido pT4 es el vector de expresión que contiene el promotor P_{trc} que produce la expresión constitutiva. El vector pT4 contiene un gen que confiere resistencia a kanamicina y un origen de replicación ColE1 (Osuna & Pe, 2002).

6.3 Construcción de las mutaciones puntuales

Usamos el kit *Quick Change Mutagenesis Multisite Lightning*, según las especificaciones de fabricante, para introducir las mutaciones puntuales sobre el dominio Rossmann de EcoSDH o conSDH (Agilent Technologies, Inc. 2012; California; USA). Cada reacción contenía 2.5 µL de *buffer* de reacción, 80 ng de DNA de doble cadena como templado, 100 ng del *primer* mutagénico, 1 µL de dNTP mix, 1 µL de la mezcla de enzimas y agua tetradestilada para un volumen final de 25 µL. Las reacciones en el termociclador comienzan con 2 minutos a 95° C; luego 30 ciclos de 20 segundos a 95° C, 30 segundos a 55 °C y 90 segundos a 65° C; y una extensión final de 5 minutos a 65° C. Después digerimos los 25 µL de la reacción con 1 µL de la enzima de restricción DpnI durante 5 minutos a 37° C para eliminar el templado de DNA. Electroporamos 1.5 µL de la reacción en la cepa XL10 Gold de *E. coli*. Confirmamos la presencia de las mutaciones mediante la secuenciación del plásmido. Los archivos con las secuencias de cada variante construida se incluyen en el material suplementario.

A continuación mostramos la lista de *primers* mutagénicos empleados:

Eco_N149D	5' GCGGTGACAATAACTGATCGGACGGTATCCCGCGC 3'
Cons_N149D	5' CATTACCAACCGTACCTTTAGCCGTGCG 3'
Eco_N149D V152F	5' CTGATCGGACGTTCTCCCGCGCGGAAG 3'
Eco_S131A	5' GCTGGTGGAGCAGCTCGCGGCGTACTAC 3'
Eco_S131A/L135A	5' CTCGCGGCGTAGCACTGCCACTCCTTTC 3'
Eco_D195E	5' GGCATCAGTGGTGAGATTCCGGCGATCC 3'
Eco_A243G	5' GAATGCTGGTGGGACAGGCGGCTCATG 3'

6.4 Expresión y purificación de proteína

Para expresar y purificar la proteína primero inoculamos 3 mL de medio LB con 25 µg/mL de kanamicina a 200 rpm y 36°C toda la noche con células JM101 de *E. coli* transformadas con el vector de expresión, luego tomamos de ahí 10 µL para inocular 50 mL de LB 25 µg/mL de kanamicina a 200 rpm y 36°C toda la noche. A partir de aquí mantuvimos las muestras a 4°C, recuperamos las células centrifugando los 50 mL del cultivo usando 4000 x g durante 5 minutos. Después re-suspendimos las células en 6 mL de *buffer* de lisis (*buffer* fosfatos 50 mM pH8, imidazol 20 mM y NaCl 300 mM) y añadimos fluoruro de fenilmetilsulfonilo (PMSF) como inhibidor de proteasas a una concentración final de 0.5 mM. Luego lisamos las células por sonicación durante cuatro minutos, usando pulsos de 10 segundos y dejando 15

segundos de descanso entre cada pulso, con una amplitud del 80% en un equipo Ultrasonic Processor Model GEX 130 (Sonics and Materials; Newton, CT, USA). Centrifugamos las muestras por 20 minutos a 10 000 x g y separamos la fracción soluble.

Realizamos un solo paso de purificación usando una columna de cromatografía por afinidad a níquel His-Trap de 5 mL (GE Healthcare; Uppsala, Suecia), según las instrucciones del proveedor. Para la purificación usamos *buffer* fosfatos 50 mM pH8, NaCl 300 mM con las siguientes concentraciones de imidazol: *buffer* de unión, 15mM; *buffer* de lavado, 35mM; y *buffer* de elución, 250mM. Después de purificar, dializamos las proteínas contra *buffer* Tris-HCl 100 mM pH9 para medir actividad. En el anexo al final del documento mostramos los geles PAGE-SDS que muestran las variantes purificadas (Figura anexo 1).

6.5 Medición de actividad y determinación de los parámetros cinéticos aparentes k_{CAT_app} y K_{M_app}

Para medir la actividad enzimática usamos shikimato como reactivo, y el respectivo cofactor (NAD⁺ o NADP⁺). Seguimos la producción de NAD(P)H mediante el incremento de absorción a 340 nm (coeficiente de extinción molar 6220 M⁻¹cm⁻¹) usando un espectrofotómetro Beckman Coulter DU730 (Beckman Coulter; Brea, CA, USA) y una celda de cuarzo con una ventana de luz de 10 mm (Hellma; Müllheim, Alemania). Comenzamos la reacción agregando 10 µL de la enzima purificada y seguimos la reacción por dos minutos. La reacción fue realizada en *buffer* Tris-HCl 100 mM pH9 a temperatura ambiente en un volumen final de 100 µL. Realizamos las mediciones por triplicado.

Para estimar la $K_{Mapp_shikimate}$ y $k_{CATapp_shikimate}$ en las variantes EcoSDH y conSDH medimos la velocidad inicial (moléculas de NADPH producido por cada molécula de enzima cada minuto) a diferentes concentraciones de shikimato (0.04, 0.1, 0.4, 1 y 4 mM) y una concentración a saturación (4 mM) de NADP⁺. Para estimar la K_{Mapp_NADP+} y k_{CATapp_NADP+} en las variantes EcoSDH y conSDH medimos la velocidad inicial a distintas concentraciones de NADP⁺ (0.04, 0.1, 0.4, 1 y 4mM) y una concentración a saturación de shikimato (4 mM). La concentración final de proteína empleada para EcoSDH fue de 3.52x10⁻⁹ M y para conSDH de 8.1x10⁻⁹ M (Tabla anexo 1).

Realizamos mediciones de velocidad inicial usando NAD⁺ como cofactor a diferentes concentraciones (1, 4, 8, 12 y 20 mM) y shikimato a 4 mM. La concentración final de proteína fue: EcoSDH, 1.84 x 10⁻⁷ M; conSDH 2.84 x 10⁻⁷ M; EcoSDH_{N149D}, 1.27 x 10⁻⁷ M; conSDH_{N149D}, 1.53 x 10⁻⁸ M; EcoSDH_{N149D_V152F}, 1.69 x 10⁻⁸ M; EcoSDH_{S131A_N149D_V152F}, 1.34 x 10⁻⁸ M; EcoSDH_{N149D_V152F_D195E}, 5.49 x 10⁻⁸ M; EcoSDH_{N149D_V152F_A243G}, 5.63 x 10⁻⁸ M; EcoSDH_{S131A_L135A_N149D_V152F}, 1.12 x 10⁻⁸ M.

Para verificar si el shikimato estaba a saturación realizamos otra serie de mediciones de velocidad inicial pero con una concentración fija de shikimato de 9 mM en las variantes conSDH, EcoSDH y EcoSDH_{S131A_L135A_N149D_V152F} (Figura anexo 2).

Realizamos mediciones de velocidad inicial usando NADP⁺ como cofactor a diferentes concentraciones (1, 4, 8, 12 y 20 mM) en todas las variantes excepto EcoSH y conSDH, usando una concentración fija de shikimato (4 mM). La concentración final de proteína fue: EcoSDH_{N149D}, 1.27×10^{-7} M; conSDH_{N149D}, 1.53×10^{-8} M; EcoSDH_{N149D_V152F}, 3.38×10^{-7} M; EcoSDH_{S131A_N149D_V152F}, 6.7×10^{-8} M, EcoSDH_{N149D_V152F_D195E}, 2.2×10^{-7} M; EcoSDH_{N149D_V152F_A243G}, 4.5×10^{-7} M; EcoSDH_{S131A_L135A_N149D_V152F}, 5.6×10^{-8} M.

Para estimar los parámetros cinéticos aparentes, ajustamos un modelo simple de Michaelis-Menten sobre los valores de velocidad inicial usando el método de mínimos cuadrados. Cuando los valores estimados de K_{M_app} eran superiores a 1 mM ajustamos también una línea recta sobre los valores de velocidad inicial de concentraciones menores o igual a 8mM de cofactor usando el método de mínimos cuadrados, a la pendiente de la recta la usamos como el valor aparente de $(k_{CAT}/K_M)_{app}$ (Tabla anexo 2).

6.6 Complementación del crecimiento en medio mínimo

Transformamos una cepa derivada de *E. coli* BW2511 (derivado de la colección *Keio*, sensible a kanamicina y Δ aroE) con el vector pT4:EcoSDH, una segunda con el vector pT4: EcoSDH_{N149D}; y otra con el vector pT4: EcoSDH_{S131A_L135A_N149D_V152F}. Como control negativo usamos la cepa con el plásmido pT4 vacío. Crecimos las variantes en 3 mL de LB con kanamicina (25 μ g/mL) por 10 horas a 37°C y 250 rpm. Centrifugamos las células y lavamos el pellet 3 veces en 1 mL de sales M9 (6g/L Na₂HPO₄, 0.5 g/L NaCl, 3 g/L KH₂PO₄, 1 g/L NH₄Cl), suspendimos las células en 1 mL de sales M9 e inoculamos 20 μ L en 2 mL de medio mínimo M9 (sales M9, glucosa al 0.4%) con kanamicina (25 μ g/mL) y suplementado con tiamina (10 μ g/mL). Incubamos los cultivos a 37° C a 250 r.p.m., monitoreamos el crecimiento siguiendo la densidad óptica a 600 nm.

6.7 Modelos estructurales de las variantes construidas

Usamos el servidor I-TASSER (J. Yang et al., 2015; J. Yang & Zhang, 2015; Y. Zhang, 2008) para generar los modelos estructurales de las variantes construidas y de conSDH. Manualmente introdujimos las mutaciones en la secuencia de aminoácidos de la enzima EcoSDH. Después le pedimos al servidor generar un modelo usando como templado la estructura del archivo 1NYT del PDB. El servidor I-TASSER regresa un archivo con extensión PDB. Usamos el programa PyMOL para generar las imágenes que presentamos.

7 Resultados

Encontramos que la enzima EcoSDH silvestre es naturalmente activa usando NAD⁺ como cofactor, contrario a un reporte anterior donde no detectan actividad (Michel et al., 2003). Es posible que ellos no detectaron actividad porque la concentración de proteína que usaron para medir actividad fue baja (0.017 nm), suficiente para detectar actividad con NADP⁺ pero no para detectarla con NAD⁺. Nosotros empleamos concentraciones más altas de proteína con los que pudimos detectar actividad y medir los parámetros cinéticos aparentes; la enzima presenta cerca de tres órdenes de magnitud menor actividad con NAD⁺ que con NADP⁺ (Tabla 1 y Figura 14, al final de la sección de resultados).

7.1 Rediseño del domino Rossmann usando la aplicación Rosetta Enzyme Design

Para intercambiar la actividad de NADP⁺ hacia NAD⁺ en el domino Rossmann de EcoSDH, propusimos realizar una serie de diseños usando NAD⁺ como cofactor. Posteriormente, comparar las secuencias resultantes contra otra serie de diseños donde usamos NADP⁺ como cofactor, para así proponer cambios específicos para NAD⁺. El proceso que seguimos consistió de los siguientes pasos: la selección de residuos para rediseñar, el modelado de los cofactores, la construcción y selección de los diseños *in silico*, y por último, la comparación de las secuencias entre los diseños con NAD⁺ contra los diseños con NADP⁺ (Figura 5).

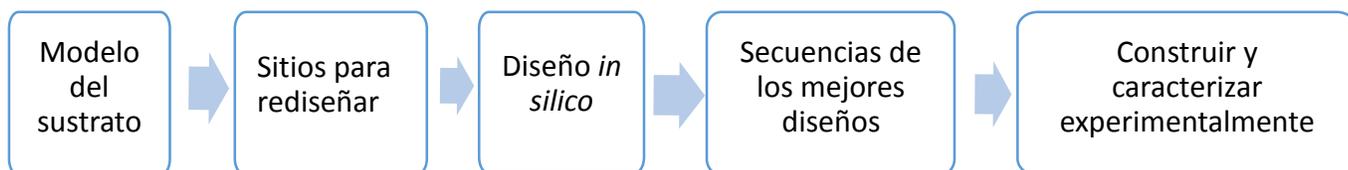
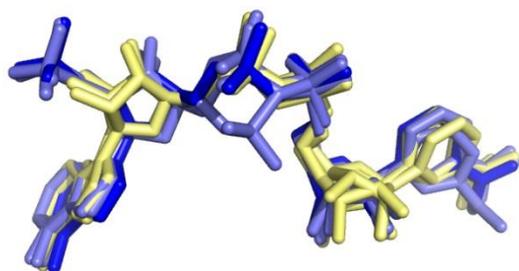


Figura 5. Esquema de pasos para el diseño *in silico* con Rosetta Enzyme Design. Comenzamos construyendo el modelo de los sustratos, en este caso los cofactores NAD⁺ y NADP⁺, luego seleccionamos una lista de residuos los cuales permitimos sean modificados por el programa. Iniciamos el diseño *in silico*, una serie de diseños usando NAD⁺ como cofactor y una segunda serie usando NADP⁺. Seleccionamos los diseños mejor calificados y comparamos las secuencias cuando usamos NAD⁺ contra las secuencias cuando usamos NADP⁺. Seleccionamos los residuos propuestos solo en los diseños con NAD⁺ para caracterizar experimentalmente.

7.1.1 Modelado de los cofactores

En el primer paso del diseño *in silico* modelamos los cofactores NAD⁺ y NADP⁺. Para elegir la conformación de NAD⁺ comparamos la estructura del cofactor en enzimas de la familia Q/SDH que usan

NAD⁺ contra las que usan NADP⁺. Observamos una conformación muy parecida entre NAD⁺ y NADP⁺, la desviación cuadrada media (Figura 6) entre todos los cofactores es menor que 1 Ångstrom – por debajo de la resolución de las estructuras cristalográficas– lo que nos sugiere que la conformación es la misma para ambos cofactores, entonces para modelar NAD⁺ tomamos la conformación de NADP⁺ de la estructura cristalográfica de EcoSDH (1NYT) y reemplazamos el fosfato 2' por un hidroxilo. La conformación de NADP⁺ fue la misma que en la estructura cristalográfica. Por último representamos los hidrógenos y calculamos las cargas parciales de cada átomo usando el programa MOPAC (MOPAC2009, James J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA).



RMSD	3phi	2ev9	1o9b	3jyq	3t4e	4k28
Å						
1nyt	0.524	0.352	0.44	0.894	0.415	0.804
3phi		0.415	0.398	0.628	0.389	0.615
2ev9			0.362	0.762	0.274	0.758
1o9b				0.665	0.209	0.816
3jyq					0.751	0.324
3t4e						0.792

Figura 6. Desviación cuadrada de la estructura entre cofactores de la familia Q/SDH. La figura de la izquierda muestra las estructuras cristalográficas alineadas de los cofactores en miembros de la familia Q/SDH. En color azul se muestra la estructura del cofactor NADP⁺ proveniente de la enzima SDH de *E. coli*, en color azul claro la estructura de NADP⁺ de las enzimas SDH de *Thermus thermophilus* y *Helicobacter pylori*. En color amarillo claro la estructura del cofactor NAD⁺ proveniente de las enzimas QSDH de *E. coli*, *Corynebacterium glutamicum*, *Salmonella enterica* y *Pseudomonas putida*. La tabla de la derecha muestra la desviación cuadrada media entre las posibles parejas de estructuras de cofactores en la familia QSDH. Alineamos los cofactores usando el comando *align* del programa pyMol. 1nyt, SDH de *E. coli*; 3phi, SDH de *H. pylori*; 2ev9, SDH de *T. thermophilus*; 1o9b, QSDH de *E. coli*; 3jyq, Q/SDH de *C. glutamicum*; 3t4e, Q/SDH de *S. enterica*; 4k28, Q/SDH de *P. putida*.

7.1.2 Selección de los residuos para rediseñar

Luego elegimos los residuos donde permitimos a la aplicación de *Rosetta* realizar modificaciones para mejorar la interacción con el cofactor deseado. *Rosetta Enzyme Design* toma un andamio sobre el cual construir el diseño y una lista de residuos que le indica cuáles residuos se van a modificar, cuáles se van

a mantener fijos, y cuáles van a rotar para lograr un buen empaquetamiento de los residuos modificados durante el diseño. Elegimos como andamio la cadena A de la estructura 1nyt. Decidimos modificar los residuos dentro de un radio de 10 Å alrededor del grupo fosfato 2' del cofactor. Además mantuvimos sin cambios al dominio catalítico (residuo 1-100 y 245- 271) y los residuos en contacto directo con la fracción de nicotinamida o el enlace difosfato del cofactor (102, 125-131, 187-189, 212-214, 237, 238, 240 y 241); y permitimos el re-empaquetamiento del resto de los residuos (Figura 7).

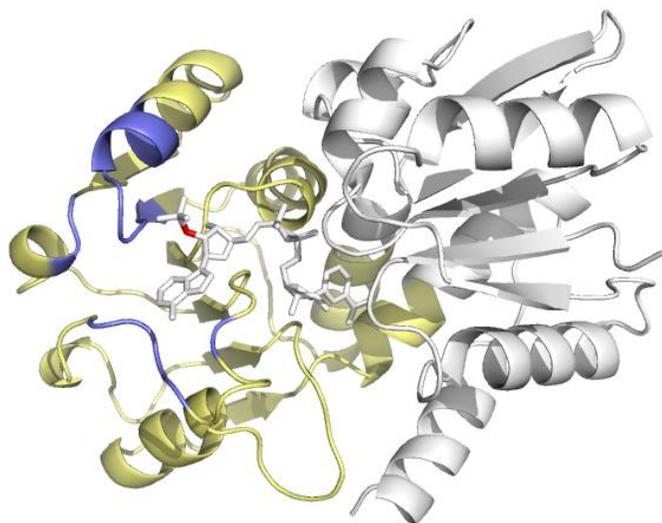


Figura 7. Residuos dentro de un radio de 10 Ångstroms del grupo fosfato 2' de NADP+. La figura muestra la cadena principal de de residuos (147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 172, 173, 190, 195, 196, 197, 198) que permitimos variar en el dominio Rossmann durante el diseño in silico con Rosetta Enzyme Design, mantuvimos sin cambios los residuos en el asa beta1–alfa1 del dominio Rossmann, por estar conservados entre Aroe y YdiB en contacto directo con el puente difosfato y la fracción de nicotinamida. En blanco se muestra el domino catalítico, en amarillo claro se muestra el dominio Rossmann, en azul mostamos los sitios que permitimos variar, el cofactor NADP+ se muestra en palillos blancos y con rojo marcamos el sitio 2' de la adenosina. Estructura 1nyt del PDB.

7.1.3 Selección de los mejores diseños

En el siguiente paso recuperamos las secuencias de los diseños donde usamos NAD⁺ como cofactor, y las comparamos con la secuencias de los diseños donde usamos NADP⁺, para identificar las mutaciones introducidas únicamente en los diseños con NAD⁺. Construimos 1000 diseños con NAD⁺ y otros 1000 diseños con NADP⁺ usando la aplicación. Seleccionamos los mejores diseños, usando los criterios recomendados en el manual de la aplicación: energía total, energía de unión a ligando, estadística de

empaquetamiento y residuos polares enterrados no satisfechos. Obtuvimos la secuencia de los diseños que cumplieron con los criterios de selección y alineamos los residuos que permitimos variar durante el diseño (Figura 8). Sugerimos introducir un residuo de aspartato en lugar de asparagina (N149D) después de comparar los diseños donde usamos NAD⁺ como cofactor contra los diseños donde usamos NADP⁺. Una gran parte de los sitios en los diseños conservan el residuo silvestre, a pesar de que se le permitió elegir entre cualquier residuo. En otros sitios los diseños sugieren mutaciones similares usando cualquiera de los cofactores, por ejemplo en la posición 151 los diseños sugieren introducir un residuo de aspartato con NAD⁺ y también con NADP⁺. La mutación N149D es la única que aparece más de una vez en los diseños con NAD⁺, pero ninguna vez en los diseños con NADP⁺.

Diseños modificando los residuos dentro de un radio de 10 Å del grupo 2' de la adenina del cofactor

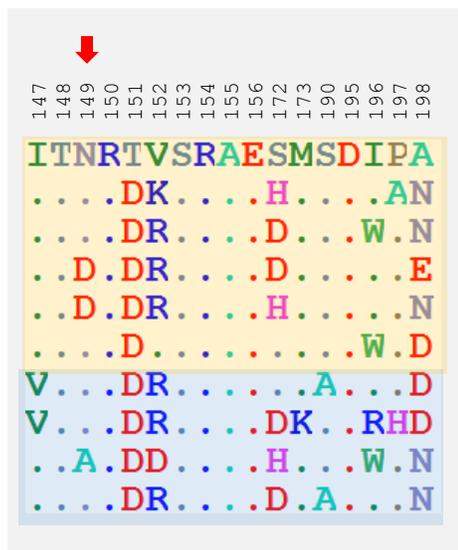


Figura 8. Alineamiento de los mejores diseños obtenidos con Rosetta Enzyme Design. Aquí mostramos el alineamiento de los diseños (con mejor puntaje) usando NAD⁺ comparados contra los diseños con NADP⁺. Sombreados con naranja mostramos los diseños donde usamos NAD⁺ como cofactor, y con azul los diseños con NADP⁺. El alineamiento solo muestra las posiciones que permitimos variar durante el diseño *in silico*: los sitios dentro de un radio de 10 Å del grupo 2' de la adenina del cofactor. Con flechas rojas marcamos los sitios donde observamos que un residuo se sugiere más de una vez en los diseños con NAD⁺ y que no se sugiere en los diseños con NADP⁺. Las posiciones en donde el diseño con Rosetta regresa el mismo residuo que el de la secuencia silvestre son marcadas con un punto. La secuencia en la parte superior corresponde a EcoSDH.

7.1.4 Construcción de la mutante N149D

La mutación N149D aumenta ligeramente (1.5 veces) la k_{CATapp} usando NAD⁺, sin cambios en la K_{Mapp} , mientras que la K_{Mapp} usando NADP⁺ aumenta dos órdenes de magnitud y la k_{CATapp} disminuye un orden de magnitud; la enzima EcoSDH_{N149D} prefiere NAD⁺ cinco veces sobre NADP⁺ (Tabla 1 y Figura 14). La posición 149 está al final de la segunda lámina beta del dominio Rossmann, la carga negativa del aspartato repele al fosfato 2' de NADP⁺ y puede formar puentes de hidrógeno con los hidroxilos 2' y 3' de NAD⁺ (Figura 9).

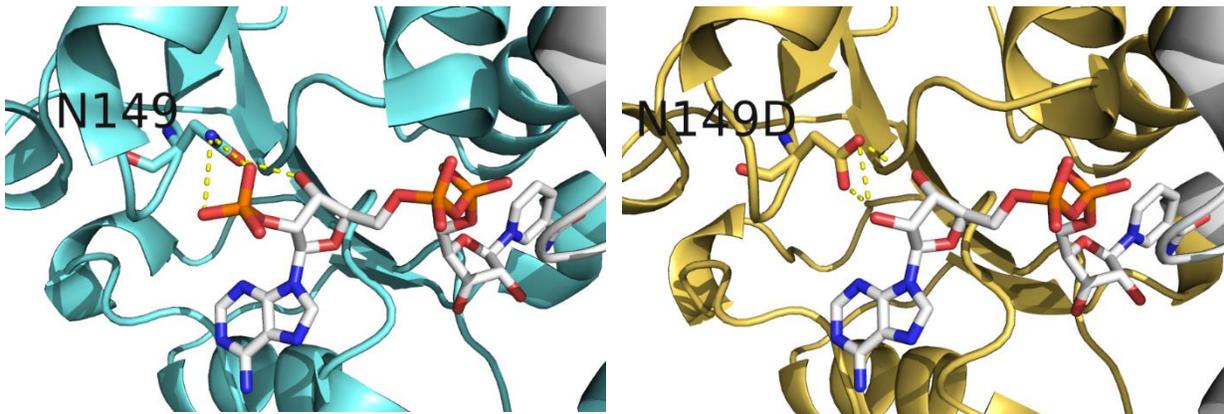


Figura 9. Modelo de la estructura EcoSDH_{N149D}. El panel de la izquierda muestra el residuo N149 en la enzima silvestre EcoSDH, la asparagina forma puentes de hidrógeno con el grupo fosfato 2' y el hidroxilo 3' de la fracción de adenina de NADP⁺. En el panel de la derecha mostramos la mutación N149D: el nuevo residuo introduce una carga negativa, lo que repele al fosfato 2' y disminuye la actividad con NADP⁺; al mismo tiempo podría formar puentes de hidrógeno con los grupos hidroxilo 2' y 3' de la fracción adenina de NAD⁺.

7.2 Mutación N149D en el fondo del dominio Rossmann consenso

Nos preguntamos si podíamos aprovechar las mutaciones consenso para mejorar la actividad en nuestro diseño. Para responder probamos la mutación propuesta por nuestro diseño *in silico* (N149D) en el contexto de un dominio Rossmann consenso. La enzima conSDH tiene 15 mutaciones respecto a la enzima silvestre EcoSDH y es tan activa como la enzima silvestre usando NADP⁺ y shikimato; cuando medimos la actividad con NAD⁺ encontramos que la enzima consenso es casi el doble de activa que EcoSDH silvestre usando NAD⁺ (Tabla 1). Probamos la mutación N149D sobre el fondo conSDH, y la variante conSDH_{N149D} prefiere NAD⁺ sobre NADP⁺, con una actividad cinco veces mayor usando NAD⁺ que EcoSDH_{N149D} (Tabla 1 y Figura 10).

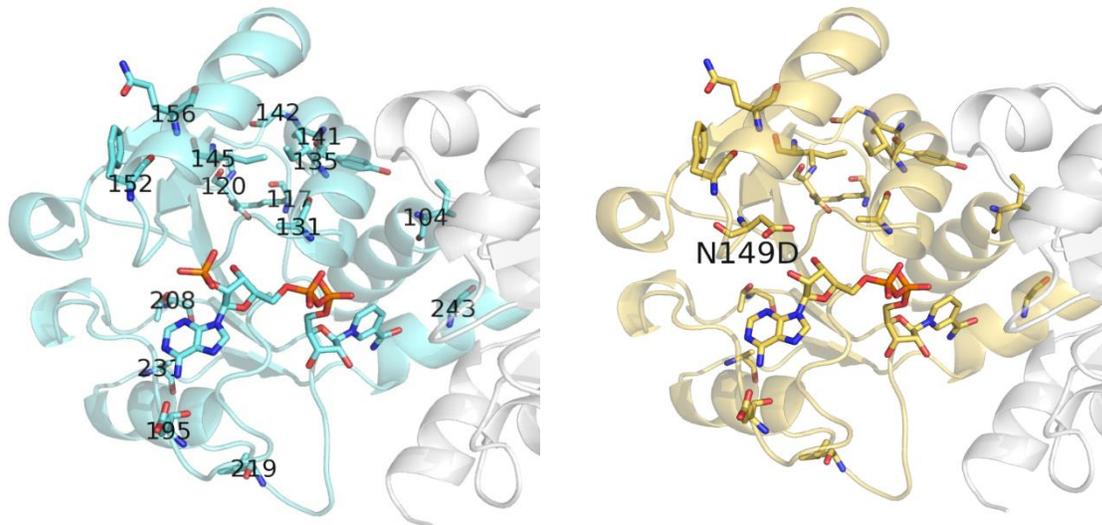


Figura 10. Mutaciones consenso con mutación N149D. La figura de la izquierda muestra el modelo de la estructura de la enzima consSDH y señala las 15 mutaciones introducidas respecto a la enzima silvestre de *E. coli*: V104I, R117K, L120D, S131A, L135I, L141Y, D142G, V145I, V152F, E156Q, D195E, I208T, K219L, S231A y A243G, en color azul mostramos el dominio Rossmann y el cofactor NADP+. La figura de la derecha muestra el modelo de la enzima consenso con la mutación N149D (consSDH_{N149D}) que obtuvimos del diseño con *Rosetta*, en color amarillo mostramos el dominio Rossmann y el cofactor NAD+.

Nuestro siguiente paso fue encontrar cuáles son las mutaciones consenso que aumentan la actividad con NAD+. Propusimos que las mutaciones más cercanas al cofactor son las responsables del aumento de actividad. Como primer candidato elegimos la mutación V152F debido a que es la más cercana al sitio de unión 2' del cofactor. La mutante EcoSDH_{N149D_V152F} aumenta la actividad usando NAD+, pero no recupera la actividad de consSDH_{N149D}. Seleccionamos como siguientes candidatos las mutaciones más cercanas al cofactor: S131A, D195E y A243G. Construimos las variantes EcoSDH_{N149D_V152F_S131A}, EcoSDH_{N149D_V152F_D195E} y EcoSDH_{N149D_V152F_A243G} (Figura 11). Los cambios en las posiciones 195 y 243 no tienen efecto sobre la actividad (Figura anexo 3), mientras que la mutación en la posición 131 aumenta actividad por arriba de la variante consSDH_{N149D} (Tabla 1 y Figura 14). Las mutaciones consenso responsables del aumento en la actividad con NAD+ son V152F y S131A, que también son las más cercanas al sitio de unión 2' de la adenosina en NAD+ (Figura 11).

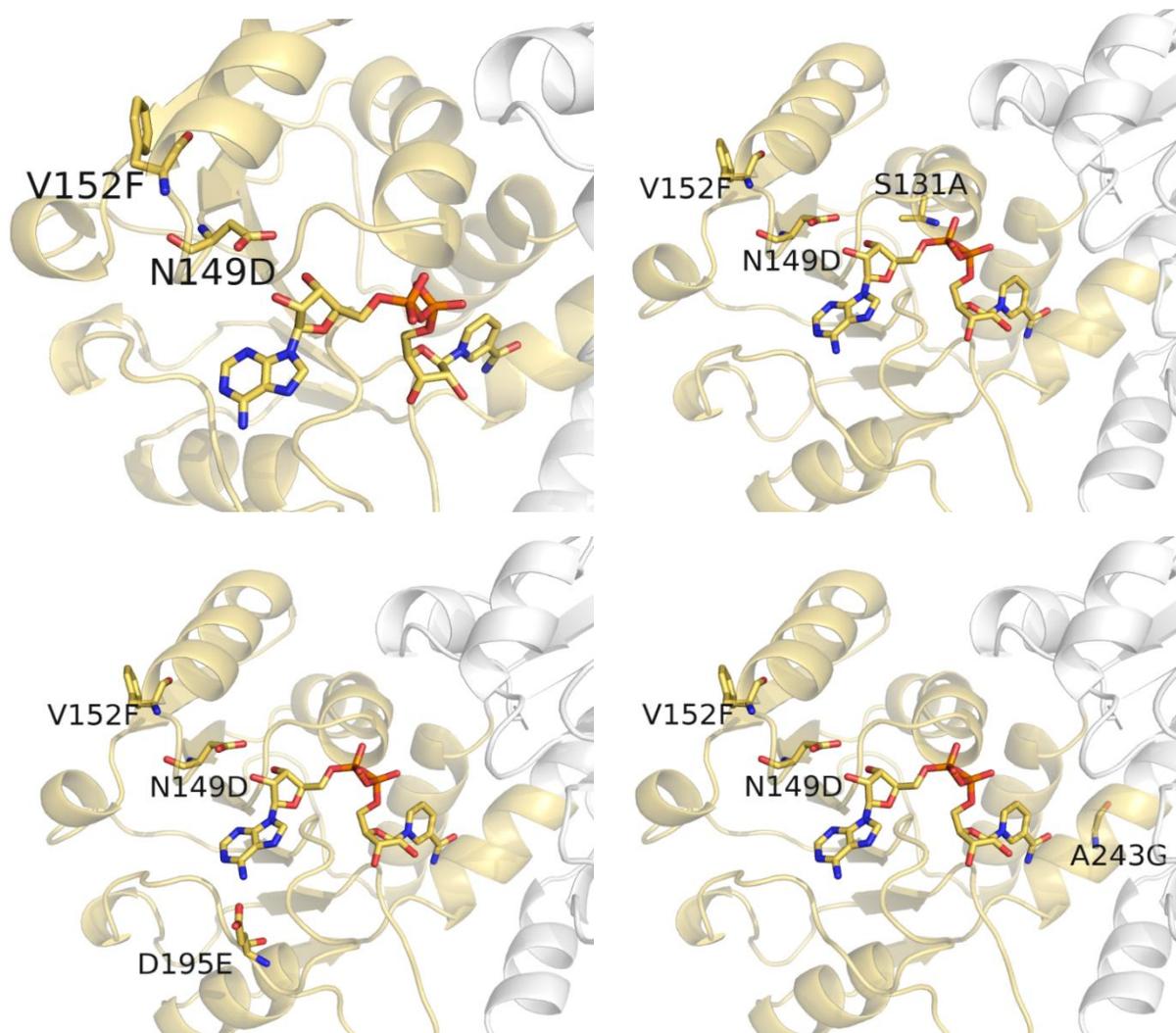


Figura 11. Búsqueda de mutaciones responsables del aumento en la actividad con NAD⁺ en la enzima consenso. La figura muestra los modelos de las mutaciones construidas para encontrar los residuos responsables del aumento de la actividad en la enzima consenso. Elegimos como candidatos las mutaciones-consenso más cercanas al sitio de unión al cofactor: A) EcoSDH_{N149D_V152F}, B) EcoSDH_{N149D_V152F_S131A}, C) EcoSDH_{N149D_V152F_D195E}, D) EcoSDH_{N149D_V152F_A243G}. En caricatura color amarillo mostramos al dominio Rossmann y en blanco al dominio catalítico; en palillos amarillos mostramos los residuos introducidos y el cofactor NAD⁺.

7.2.1 La mutación consenso S131A nos ayuda a proponer otra mutación para aumentar la actividad con NAD⁺

Anteriormente se había observado que en enzimas que prefieren NAD⁺, la cadena lateral de los residuos en la primera hélice alfa del dominio Rossmann, y que apuntan contra la lámina beta central, son más cortas que en las enzimas que prefieren NADP⁺ (Bellamacina, 1996; Rao & Rossmann, 1973). La mutación S131A cumple con esa observación. Por ello propusimos que la mutación L135A, que cumple

con esta descripción, podría aumentar la actividad con NAD⁺ (Figura 12). Construimos la mutante cuádruple EcoSDH_{S131A_L135A_N149D_V152F} y aumentó un poco la actividad; ésta es diez veces más activa que la enzima silvestre EcoSDH usando NAD⁺ y prefiere NAD⁺ sobre NADP⁺ cinco veces (Tabla 1 y Figura 14).

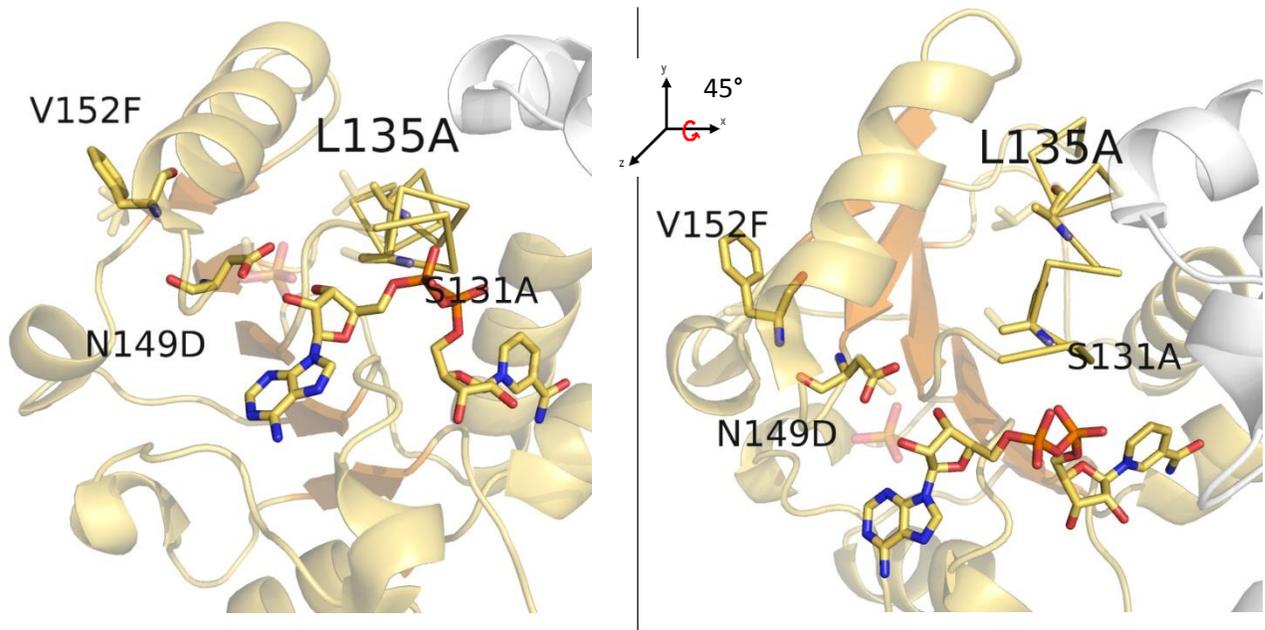


Figura 12. Modelo de la mutación L135A. La figura muestra el modelo de la mutante EcoSDH_{S131A_L135A_N149D_V152F}. El residuo L135 está sobre la primera hélice alfa del dominio Rossmann en EcoSDH y apunta contra la lámina beta central del dominio (color naranja), la mutación L135A reduce la longitud de la cadena lateral y aumenta la actividad con NAD⁺.

7.3 Complementación del crecimiento en medio mínimo de una variante de una cepa de *E. coli* carente del gen que codifica para la enzima SDH

Medimos la actividad *in vitro* de las enzimas en la dirección contraria a la que ocurre en la síntesis de shikimato, además encontramos valores muy altos en la $K_{Mapp_NAD^+}$, por lo que nos preguntamos si nuestras variantes eran activas *in vivo*. Realizamos ensayos de complementación en medio mínimo. Usamos una variante de la cepa BW2511 de la colección *Keio* (Baba et al., 2006) que tiene deletado el gen que codifica para EcoSDH (y sin el marcador de resistencia a kanamicina). Para realizar los ensayos de complementación elegimos las variantes EcoSDH_{N149D} (el diseño con *Rosetta*) y EcoSDH_{S131A_L135A_N149D_V152F} (la variante con mayor actividad con NAD⁺). Las variantes probadas complementan el crecimiento, aunque la fase de retraso en la cepa que expresaba la variante

EcoSDH_{S131A_L135A_N149D_V152F} se extiende por algunas horas más, comparada con la EcoSDH silvestre (Figura 13) (García-Guevara et al., 2017).

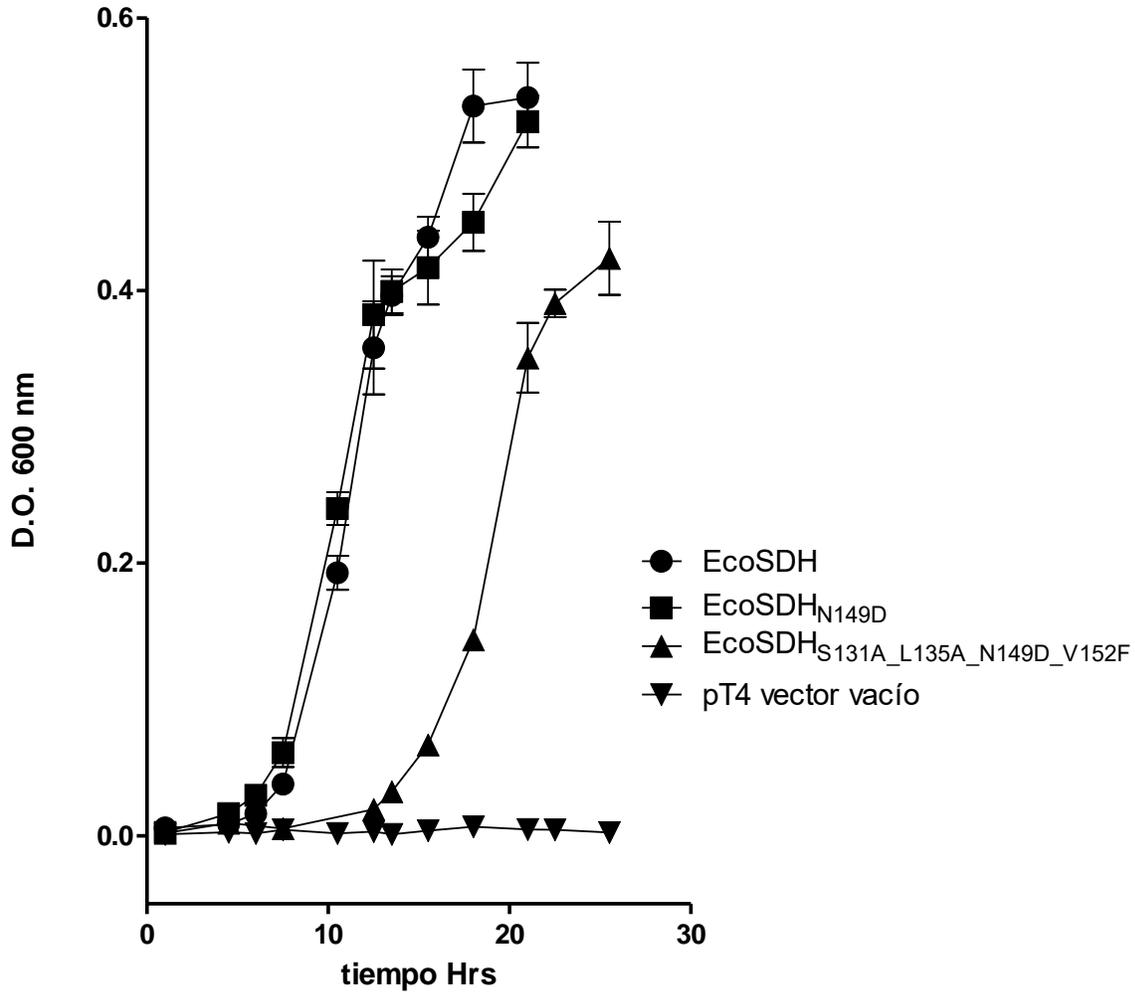


Figura 13. Complementación de crecimiento en medio mínimo con glucosa como fuente de carbono. La gráfica de arriba muestra la cinética de crecimiento en medio mínimo de la cepa de *E. coli* BW2511 carente del gen silvestre que codifica para la enzima SDH; la cepa es auxótrofa para aminoácidos aromáticos. En círculos se muestra la cepa transformada con el vector pT4 que contiene el gen para EcoSDH silvestre; en cuadrados se muestra la cepa transformada con el vector que lleva la variante EcoSDH_{N149D}; en triángulos hacia arriba se muestra la cepa transformada con el vector que lleva la variante EcoSDH_{S131A_L135A_N149D_V152F}; en triángulos hacia abajo se muestra la cepa transformada con el vector pT4 vacío. (García-Guevara et al., 2017).

Tabla 1. Parámetros cinéticos aparentes estimados para las variantes presentadas (García-Guevara et al., 2017).

		K_{Mapp} (mM)	K_{CATapp} (1/min)
EcoSDH	NAD+	10 ± 1.18	1463 ± 79.7
	NADP+	$5.8 \pm 0.69 \times 10^{-2}$	11170 ± 200
EcoSDH_{N149D}	NAD+	7.8 ± 0.82	1613 ± 70.5
	NADP+	10.2 ± 4.8	346 ± 78
EcoSDH_{N149D_V152F}	NAD+	13.8 ± 2.69	5226 ± 534
	NADP+	29.6 ± 6.2	958 ± 137
EcoSDH_{N149D_V152F_S131A}	NAD+	11 ± 1.34	13212 ± 773
	NADP+	11.6 ± 1.7	2450 ± 177
EcoSDH_{N149D_V152F_S131A_L135A}	NAD+	12 ± 1.62	19038 ± 1281
	NADP+	10.3 ± 1.9	3261 ± 286
conSDH	NAD+	10.5 ± 1.07	3178 ± 154
	NADP+	$1.05 \pm 0.1 \times 10^{-2}$	10510 ± 101
conSDH_{N149D}	NAD+	8.6 ± 0.94	6978 ± 331
	NADP+	12.7 ± 2.3	1290 ± 122

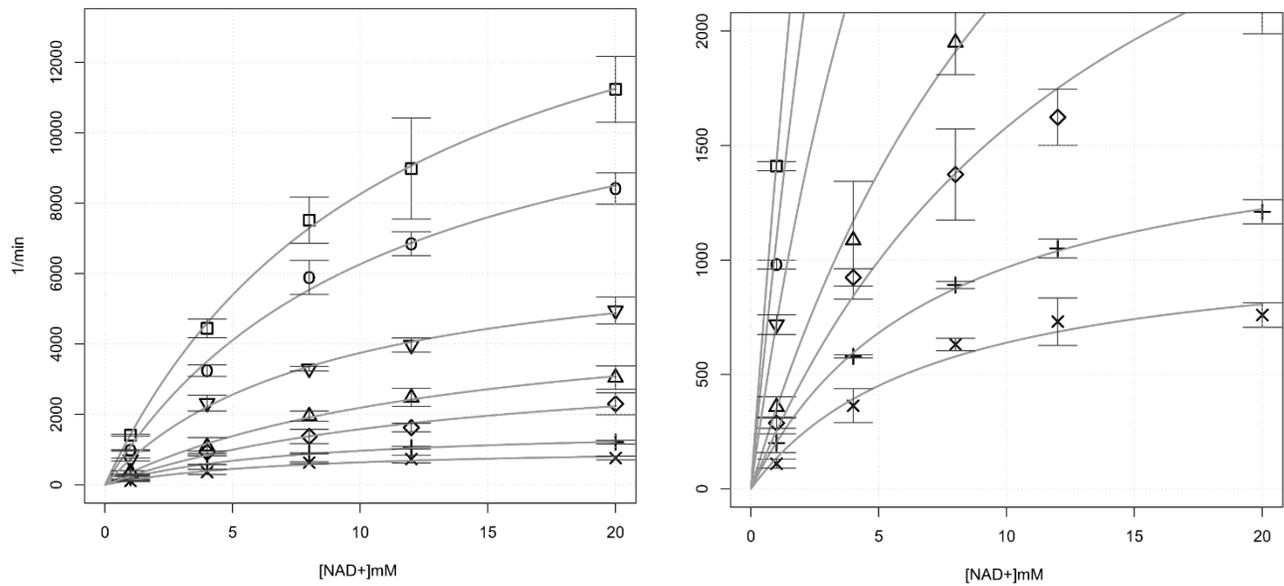


Figura 14 a. Perfil de actividad con NAD⁺ de las variantes reportadas. La figura de la izquierda muestra la velocidad inicial de las variantes purificadas a diferentes concentraciones de NAD⁺ (1, 4, 8, 12, y 20 mM), con una concentración fija de shikimato de 4mM. EcoSDH, letras x; EcoSDH_{N149D}, signos de más; EcoSDH_{N149D_V152F}, triángulos hacia arriba; EcoSDH_{N149D_V152F_S131A}, círculos; EcoSDH_{N149D_V152F_S131A_L135A}, cuadrados; conSDH, diamantes; conSDH_{N149D}, triángulos hacia abajo. Las barras de error muestran el intervalo de confianza del 95% de la media para cada punto. Las curvas grises muestran el ajuste de un modelo simple de Michaelis-Menten. La imagen de la izquierda muestra un acercamiento en la escala del eje Y para mayor claridad (García-Guevara et al., 2017).

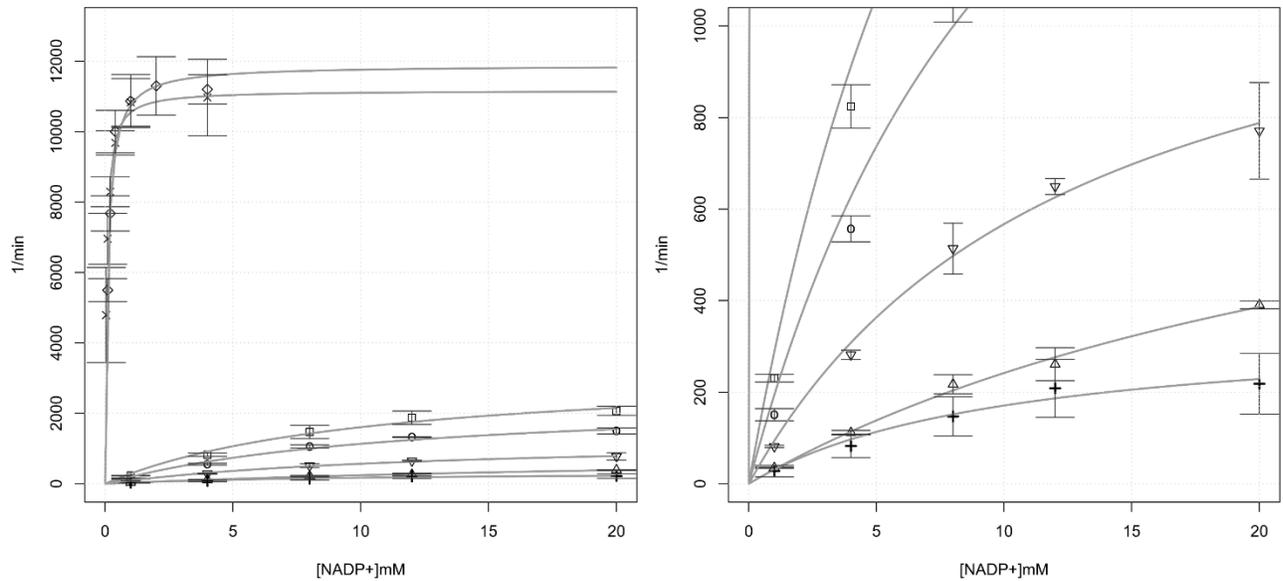


Figura 14 b. Perfil de actividad con NADP+ de las variantes reportadas. La figura de la izquierda muestra la velocidad inicial de las variantes purificadas, a diferentes concentraciones de NADP+ (1, 4, 8, 12, y 20 mM) con una concentración fija de shikimato de 4mM. EcoSDH, letras x; EcoSDH_{N149D}, signos de más; EcoSDH_{N149D_V152F}, triángulos hacia arriba; EcoSDH_{N149D_V152F_S131A}, círculos; EcoSDH_{N149D_V152F_S131A_L135A}, cuadrados; conSDH, diamantes; conSDH_{N149D}, triángulos hacia abajo. Las concentraciones para medir velocidad inicial en las enzimas EcoSDH y conSDH fueron: 0.04, 0.1, 0.4, 1, y 4mM. Las barras de error muestran el intervalo de confianza del 95% de la media para cada punto. Las curvas grises muestran el ajuste de un modelo simple de Michaelis-Menten. La imagen de la izquierda muestra un acercamiento en la escala del eje Y para mayor claridad (García-Guevara et al., 2017).

8 Discusión

Utilizando la aplicación *Rosetta Enzyme Design* propusimos la mutación N149D en el dominio Rossmann de la enzima EcoSDH. La mutación N149D intercambia la especificidad por el cofactor, casi exclusivamente por la disminución de la actividad con el cofactor original (NADP+). Usamos la misma mutación en la enzima conSDH y encontramos mayor actividad con NAD+ que en la enzima EcoSDH_{N149D}. Después identificamos que las mutaciones consenso más cercanas al sitio de unión 2' de la ribosa de la adenina causan este aumento de actividad. Por último probamos que nuestra variante tiene actividad *in vivo*, a pesar de que no pudimos disminuir la K_{Mapp_NAD+} . Proponemos a la ingeniería de consensos como una estrategia complementaria en el diseño de enzimas.

8.1 Diseño *in silico* con *Rosetta*

Nuestro diseño *in silico* con *Rosetta* propuso la mutación N149D. En la naturaleza los motivos Rossmann que prefieren NAD+ suelen tener un residuo cargado negativamente en la misma posición, y es interesante ver que recuperamos dicha información sin usar secuencias homólogas (Bellamacina, 1996). La mutación N149D introduce una carga negativa que repele al grupo fosfato 2' del cofactor NADP+, disminuye drásticamente la actividad con NADP+, y mantiene la actividad con NAD+ casi sin cambios. En otros estudios para intercambiar la actividad del cofactor con frecuencia se usa una mutación equivalente, pero el efecto es caso específico: a veces aumenta la actividad con NAD+ y a veces la disminuye (Rane & Calvo, 1997; Rosell et al., 2003). Los efectos reales de las mutaciones aún están fuera de nuestro alcance predictivo.

El diseño *in silico* de enzimas es una técnica llamativa, porque permite probar variantes en una computadora y elegir solo los cambios que permitan a la enzima blanco realizar la función deseada; sin embargo todavía se encuentra en sus primeras etapas de desarrollo y presenta muchas deficiencias. Los programas que usan modelos más precisos requieren una capacidad de cómputo excesiva, mientras que los programas que usan modelos más fáciles de calcular comprometen la veracidad de las predicciones. *Rosetta* es un programa que posee un modelo rápido de calcular, por ello compromete la veracidad de sus predicciones. Entre las limitaciones de *Rosetta* observamos que durante el proceso del diseño la cadena principal de la enzima se mantiene rígida, lo cual no es la mejor aproximación a la realidad y puede perder mucho al realizar esta simplificación. En reportes recientes han tratado de incorporar movimiento en la cadena principal y esto ha mostrado algunas mejoras en los resultados (Loffler et al., 2017). Esperamos que el desarrollo del diseño *in silico* vaya de la mano con el aumento en nuestra

capacidad de cómputo y el progreso de modelos que aceleren los cálculos mientras realizan predicciones veraces. Todavía estamos lejos de alcanzar los resultados deseados usando el diseño *in silico*, pero usar los programas disponibles nos permite examinar su eficacia y buscar áreas de oportunidad para la mejora.

Ya que los resultados del diseño *in silico* no siempre son tan eficientes como nos gustaría, usualmente se complementan con estrategias que introducen información alternativa (Giger et al., 2013). La evolución dirigida (que consiste en realizar cambios al azar en la secuencia y después seleccionar entre las variantes generadas aquellas que cumplan con nuestros criterios) podría considerarse como la estrategia conocida más efectiva para aumentar la actividad en una proteína (Kuchner & Arnold, 1997). Aunque en ocasiones presenta algunos puntos limitantes, debido a que el éxito de esta técnica depende en gran medida del número de variantes que podamos probar. El número de posibles variantes aumenta exponencialmente con el número de residuos en la proteína. Si se cuenta con un sistema de selección adecuado, entonces es posible probar un gran número de variantes y maximizar las probabilidades de éxito. Pero si no se cuenta con un sistema de selección eficiente, entonces se requiere de una gran cantidad de tiempo y recursos para tamizar la actividad de las variantes. Entonces, un enfoque que se le da a este problema es reducir el número de variantes por probar, como dirigir los cambios a zonas particulares de la proteína con conocimiento previo de la estructura; pero aun así el número de posibles variantes continúa siendo muy grande (Flores & Ellington, 2005; Liang et al., 2007). Nosotros proponemos que usando información disponible en bases de datos, podríamos extender nuestra caja de herramientas para mejorar la actividad en enzimas, sin requerir tantos recursos.

8.2 Ingeniería de consensos como estrategia complementaria

Decidimos probar el efecto de la mutación N149D en el contexto de la enzima consenso, porque se ha observado que en ocasiones las mutaciones consenso aumentan la actividad con sustratos alternativos (Aerts et al., 2013; Cerdobbel et al., 2011; Flores & Ellington, 2005). Encontramos que la variante conSDH_{N149D} tienen casi cinco veces más actividad que EcoSDH_{N149D}

La enzima conSDH difiere en 15 sitios comparada con la enzima EcoSDH; propusimos que las mutaciones más cercanas al sitio 2' del cofactor eran las responsables en el aumento de la actividad. Introdujimos las mutaciones consenso por cercanía al cofactor sobre la variante EcoSDH_{N149D}: las dos mutaciones más cercanas al sitio 2' recuperaron la actividad de la variante conSDH_{N149D}. La mutación V152F introduce un residuo más voluminoso en el asa beta-2 alfa-2, que posiblemente genera una conformación favorable

para la actividad. Los diseños con *Rosetta Enzyme Design* proponen para esta posición colocar un residuo con carga positiva, sin embargo decidimos no probarlo porque: sugería el cambio tanto en los diseños con NAD⁺ como en los diseños con NADP⁺; este residuo no está en contacto directo con el cofactor y durante el diseño la cadena principal se mantuvo fija; además se sabe que *Rosetta* tiende a colocar residuos polares y cargados en la superficie de las proteínas cuando realiza diseños (Dantas et al., 2003). Sin embargo, sin realizar las mutantes no podemos concluir nada acerca del efecto sobre la actividad de las mutaciones que propone *Rosetta*.

La mutación S131A apoya a observaciones anteriores en dominios Rossmann que prefieren NAD⁺, en los que las cadenas laterales de la primera hélice alfa que apuntan contra la lámina beta central son más cortas (Bellamacina, 1996; Rao & Rossmann, 1973; Scrutton et al., 1990). En los dominios Rossmann, el extremo amino de la hélice alfa interactúa con enlace di-fosfato del cofactor, ya sea que se trate de NAD o de NADP (Bellamacina, 1996). El extremo amino en una hélice alfa tiene un dipolo neto equivalente a casi media carga positiva, debido al arreglo regular de sus puentes de hidrógeno (Hol, van Duijnen, & Berendsen, 1978). Se propone que en los dominios Rossmann que prefieren NAD⁺, las cadenas laterales cortas en la cara interna de la primera hélice alfa permiten un empacamiento cercano contra la lámina beta y un arreglo regular en los puentes de hidrógeno de la hélice. Mientras que en dominios Rossmann que prefieren NADP⁺, las cadenas laterales más largas laterales separan a la hélice alfa de la lámina beta, y permiten acomodar mejor el espacio extra que necesita el grupo fosfato adicional del NADP⁺ (Figura anexo 4) (Bellamacina, 1996; Rao & Rossmann, 1973). Al reducir la longitud de la cadena lateral del residuo L135 por una alanina encontramos un aumento en la actividad con NAD⁺. Nuestra variante EcoSDH_{S13A_L135A_N149D_V152F} es la que mayor actividad presenta con NAD⁺ y sería interesante contar con la estructura experimental para ver si esto ocurre realmente.

Nosotros empleamos mutaciones consenso para aumentar la actividad con NAD⁺ en EcoSDH, pero el aumento en la actividad no es muy grande. Además hacen falta más estudios relacionados con la promiscuidad que pueden generar las mutaciones consenso en otras enzimas. La ingeniería de consensos es un enfoque fácil de implementar, además, actualmente contamos con grandes bases de datos de secuencias que podemos aprovechar. Proponemos que si los residuos consenso cercanos al sitio activo aumentan la actividad con otros sustratos, entonces podemos aprovechar ese aumento de actividad en el rediseño. La ingeniería de consensos usa residuos que ya han sido probados en la naturaleza y no requiere modelos moleculares precisos. Desafortunadamente estamos limitados por espectro de sustratos a los que las mutaciones consenso brinden promiscuidad, si es que la brindan. No

hay razón alguna para suponer que el espectro de sustratos, donde las mutaciones consenso brinden promiscuidad, se ajuste a nuestras necesidades en el diseño. Más bien debería ser al revés: probar los sustratos donde haya más actividad y buscar si podemos aplicar el descubrimiento a alguna aplicación que tengamos.

El enfoque de ingeniería de consensos se ha empleado principalmente para aumentar termo-estabilidad de las proteínas y hay muy poca información acerca del efecto que tiene para ampliar la especificidad por el sustrato. La ingeniería de consensos en esencia captura la información de conservación, podríamos pensar que si un residuo en particular es más frecuente entre otros posibles residuos esto podría reflejar que dicho residuo es ventajoso. Desde un punto de vista evolutivo la promiscuidad se puede considerar como una ventaja en el caso de cambios imprevistos en el ambiente. Desde un punto de vista tecnológico, explotar la promiscuidad de las enzimas para desarrollar nuevas funciones ha sido un enfoque aprovechado en años recientes (Shao & Arnold, 1996; Tyzack et al., 2017). Aunque no estamos seguros que los residuos consenso sean sinónimo de promiscuidad, es un efecto que también se ha encontrado en secuencias ancestrales; cuya construcción se basa en cierta medida en la conservación entre secuencias (Eick, Colucci, Harms, Ortlund, & Thornton, 2012; Risso, Gavira, Mejía-Carmona, Gaucher, & Sanchez-Ruiz, 2013; Risso, Gavira, & Sanchez-Ruiz, 2013).

8.3 Intercambio en la actividad de NADP+ hacia NAD+ y producción de shikimato

Un tema recurrente en la ingeniería de proteínas es el intercambio de actividad en enzimas que usan NADP+ para que ahora usen NAD+. Comparando nuestros resultados con otros estudios, nosotros logramos un aumento considerable en la k_{CAT_app} con NAD+, pero en otros estudios usualmente logran disminuir la K_{M_app} . Desafortunadamente nuestros valores de K_{M_app} para NAD+ se mantuvieron sin cambios significativos (Tabla 1). Debido a los valores de la $K_{M_app_NAD+}$, nos preguntamos si la variante EcoSDH_{S13A_L135A_N149D_V152F} presentaba actividad *in vivo*. Para responder, probamos la capacidad de la variante para complementar el crecimiento de la cepa de *E. coli* carente del gen que codifica para la SDH, esta variante es incapaz de crecer en medio mínimo, ya que no puede sintetizar aminoácidos aromáticos (Baba et al., 2006). Para crecer, necesita ser suplementada con estos aminoácidos o adquirir un gen capaz de conferir la actividad shikimato deshidrogenasa. Observamos que mientras la cepa carente del gen que codifica la SDH no crece, las variantes EcoSDH_{N149D} y EcoSDH_{N149D_V152F_S131A_L135A} sí complementan el crecimiento; lo que nos muestra que son activas *in vivo*, a pesar de que no pudimos disminuir los valores de K_{M_app} por NAD+. Los valores de la concentración intracelular de NAD+ estimados para *E. coli* en crecimiento exponencial son de 2.6 mM, si bien son menores que la $K_{M_app_NAD+}$, se

encuentran en un rango cercano (Bennett, Kimball, & Gao, 2009). Cabe además mencionar que, la actividad *in vivo* ocurre en la dirección opuesta a la actividad que medimos *in vitro*.

Este, junto con otros estudios para intercambiar la actividad hacia NAD⁺ muestran que: 1) los resultados dependen de cada caso (Banta, Swanson, Wu, Jarnagin, & Anderson, 2002; Bastian et al., 2011; Chen, Greer, & Dean, 1995; Dambe, Kuhn, Brossette, Giffhorn, & Scheidig, 2006; Eppink, Overkamp, Schreuder, & Van Berkel, 1999; Kristan, Stojan, Adamski, & Lanišnik Rižner, 2007; Pick, Ott, Howe, Schmid, & Sieber, 2014; Rane & Calvo, 1997; Rosell et al., 2003; Scrutton et al., 1990; Yaoi, Miyazaki, Oshima, Komukai, & Go, 1996; R. Zhang, Xu, Sun, Zhang, & Xiao, 2009); 2) nos falta profundizar más en nuestro conocimiento en la relación entre la secuencia y la función de las proteínas.

Nosotros propusimos obtener una enzima SDH que prefiera NAD⁺ como cofactor, porque podría ser de utilidad para producir shikimato en condiciones limitantes de oxígeno a partir de glucosa. En esas condiciones el principal equivalente reductor es NADH, mientras que la enzima silvestre usa NADPH; esto pudiera generar un desbalance en el uso de cofactores que limite la producción de shikimato (Kogure et al., 2016). Estudios previos han mostrado que un desbalance semejante en el uso de cofactores ocurre en la producción de otros metabolitos de interés. Una posible solución es construir una vía que sea dependiente de NAD⁺ (una vía fermentativa) mediante el intercambio de la actividad hacia NAD⁺ en las enzimas que participan. Esto ha ayudado a generar un mayor título y rendimiento en la producción del metabolito de interés, porque alivia el desbalance en el uso de cofactores y aprovecha la fuerza motora para regenerar NAD⁺ (Bastian et al., 2011; Hasegawa et al., 2012; Shen et al., 2011). Pero todavía queda la pregunta si la actividad en nuestra variante es suficiente para intentar producir shikimato a gran escala o haga falta mejorar la actividad de la enzima.

Además del nivel de actividad en la variante SDH que prefiere NAD⁺, existen otros puntos importantes que considerar para tratar de producir shikimato usando esta estrategia. El primer punto sería eliminar las vías naturales para la regeneración de NAD⁺. La propuesta sería eliminar las enzimas alcohol deshidrogenasa, lactato deshidrogenasa y succinato deshidrogenasa, como se mostró en anteriormente (Shen et al., 2011). El segundo punto es incluir las estrategias previas que se han usado en la ingeniería de vías metabólicas para la producción de shikimato en condiciones aerobias, las cuales incluyen: eliminar vías metabólicas que compitan con la formación del producto, incrementar la disponibilidad de los precursores PEP y E4P, y remover la regulación transcripcional o la regulación alostérica. Podemos resaltar los siguientes ejemplos: eliminar el sistema PTS que consume gran cantidad de PEP para introducir glucosa en la célula; aumentar la expresión de las enzimas que participan en la vía no

oxidativa de las pentosas fosfato para producir E4P; o remover el control alostérico de la enzima que cataliza la síntesis de DAHP (Escalante et al., 2010; Knop et al., 2001; Liao, Lin, Chien, & Hsu, 2001; Martínez et al., 2015; Rodríguez et al., 2014). Para abordar el problema de contar con ATP suficiente para producir los precursores del shikimato, la propuesta sería considerar el uso de condiciones micro-aerobias, donde una pequeña parte del NADH sea reducido a NAD⁺ produciendo el ATP necesario. Por último habría que considerar usar una cepa productora sin genes que codifiquen para enzimas QSDH (Martínez et al., 2015; Rodríguez et al., 2014).

9 Conclusiones

Intercambiamos la actividad en la enzima EcoSDH de NADP+ hacia NAD+ con ayuda de la aplicación de *Rosetta Enzyme Design* y la ingeniería de consensos. El intercambio de la actividad se debe principalmente a la disminución de la actividad con NADP+, mientras que la ingeniería de consenso nos ayudó a incrementar ligeramente la actividad en nuestro diseño.

El diseño *in silico* con *Rosetta Enzyme Design* nos permitió reducir selectivamente la actividad con NADP+, sin afectar la actividad con NAD+. La herramienta logró recuperar información de soluciones presentes en la naturaleza, aunque todavía presenta limitaciones lo cual se hizo evidente con los resultados obtenidos.

Encontramos que los residuos consensos más cercanos al sitio 2' de la adenosina del cofactor nos permitieron incrementar ligeramente la actividad de la enzima. La ingeniería de consensos es un enfoque relativamente sencillo, proponemos que podríamos estudiar con más profundidad su potencial como estrategia complementaria en el diseño de proteínas, para ampliar la actividad con sustratos distintos.

En el cambio de especificidad por los cofactores, observamos generalidades previamente reportadas: normalmente existen cargas positivas rodeando al grupo fosfato del NADP+; y para que se prefiera al NAD+, se introducen cargas negativas que repelen al NADP+ y pueden formar puentes de hidrógeno con hidroxilo 2'. Pero las mutaciones requeridas y su efectividad dependen de cada caso.

Logramos un incremento en la k_{CATapp_NAD+} pero la K_{Mapp_NAD+} en nuestras variantes permaneció sin cambios, a pesar de esto sí encontramos actividad *in vivo*. Podríamos intentar usar nuestras variantes para estudiar la producción de shikimato, aunque tal vez todavía sea necesario incrementar su actividad con NAD+.

10 Perspectivas

Dentro de nuestras perspectivas está encontrar la estructura tridimensional de la enzima conSDH y de nuestras mutantes. Nos llaman particularmente la atención la variante EcoSDH_{S131A_L135A_N149D_V152F}, que acorta la longitud de las cadenas laterales de los residuos internos en la primera hélice alfa en el dominio Rossmann. Sería interesante observar si en realidad existe un giro en el ángulo de dicha hélice alfa cuando introducimos esas mutaciones.

Otro tema interesante de explorar es el efecto de las mutaciones consenso sobre la promiscuidad en la actividad. Encontramos que las mutaciones consenso cercanas al sitio de unión 2' del cofactor aumentan la promiscuidad de la enzima SDH de *E. coli* usando el cofactor NAD⁺. Nos preguntamos si esto ocurre también con el dominio catalítico. Proponemos construir una secuencia consenso del dominio catalítico de la SDH, con el mismo conjunto de secuencias que usamos para el dominio Rossmann consenso. Para después probar su actividad y estabilidad térmica con sustratos análogos al shikimato, como quinato, que es un sustrato análogo que usan otros miembros de la familia. Esperamos que la enzima consenso sea más termoestable que la enzima de *E. coli* y que las mutaciones-consenso cercanas al sitio de unión del shikimato aumenten la promiscuidad con sustratos análogos.

Para estudiar el efecto de usar una enzima así en la producción de shikimato, es posible usar las variantes que obtuvimos, pero tal vez convenga primero mejorar la actividad con NAD⁺. Nuestra enzima tiene actividad *in vivo*, pero los parámetros cinéticos que estimamos nos sugieren que aún podemos mejorar la K_{Mapp} . Para tratar de mejorar la actividad con NAD⁺ proponemos abordar el problema usando evolución dirigida, tratando de usar una estrategia reciente que consiste en construir bibliotecas mutantes en asa beta 4 alfa 4, que une la cara A de la fracción de adenina. Nosotros ya introdujimos algunas variantes en esa zona, pero no encontramos mejoras. Lo que se propone es que aumentando el número de variantes hay mayores probabilidades de encontrar alguna con mayor actividad. Proponemos estudiar la producción de shikimato en condiciones limitantes de oxígeno, para la producción además requerimos considerar otros aspectos en la cepa que usemos: aumentar la disponibilidad de los precursores del shikimato, eliminar las vías que normalmente regeneran el NAD⁺ necesario en la vía de glicólisis y asegurar la expresión de las enzimas que participan en la síntesis de shikimato.

11 Referencias

- Aerts, D., Verhaeghe, T., Joosten, H., Vriend, G., & Soetaert, W. (2013). Consensus Engineering of Sucrose Phosphorylase : The Outcome Reflects the Sequence Input. *Biothechology and Bioengineering*, 110(10), 2563–2572. <http://doi.org/10.1002/bit.24940>
- Anfinsen, C. B. (1972). The Formation and Stabilization of Protein Structure. *Biochemical Journal*, 128(4), 737–749. <http://doi.org/4565129>
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., ... Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, 2, 2006.0008. <http://doi.org/10.1038/msb4100050>
- Balinsky, D., & Dennis, A. W. (1971). Kinetic and Isotope-Exchange Studies on Shikimate Dehydrogenase from Pisum sativum. *Biochemistry*, 10(10), 1947–1952. <http://doi.org/10.1021/bi00786a032>
- Banta, S., Swanson, B. a, Wu, S., Jarnagin, A., & Anderson, S. (2002). Alteration of the specificity of the cofactor-binding pocket of Corynebacterium 2,5-diketo-D-gluconic acid reductase A. *Protein Engineering Design and Selection*, 15(2), 131–140. <http://doi.org/10.1093/protein/15.2.131>
- Bastian, S., Liu, X., Meyerowitz, J. T., Snow, C. D., Chen, M. M. Y., & Arnold, F. H. (2011). Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in Escherichia coli. *Metabolic Engineering*, 13(3), 345–352. <http://doi.org/10.1016/j.ymben.2011.02.004>
- Bellamacina, C. R. (1996). The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *The FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, 10(11), 1257–1269.
- Bennett, B., Kimball, E., & Gao, M. (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature Chemical Biology*, 5(8), 593–599. <http://doi.org/10.1038/nchembio.186.Absolute>
- Bjelic, S., Nivón, L. G., Çelebi-Ölçüm, N., Kiss, G., Rosewall, C. F., Lovick, H. M., ... Baker, D. (2013). Computational design of enone-binding proteins with catalytic activity for the morita-baylis-hillman reaction. *ACS Chemical Biology*, 8(4), 749–757. <http://doi.org/10.1021/cb3006227>
- Cerdobbel, A., De Winter, K., Aerts, D., Kuipers, R., Joosten, H. J., Soetaert, W., & Desmet, T. (2011). Increasing the thermostability of sucrose phosphorylase by a combination of sequence- and structure-based mutagenesis. *Protein Engineering, Design and Selection*, 24(11), 829–834. <http://doi.org/10.1093/protein/gzr042>

- Chaudhuri, S., & Coggins, J. R. (1985). The purification of shikimate dehydrogenase from *Escherichia coli*. *The Biochemical Journal*, *226*(1), 217–23. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1144695&tool=pmcentrez&rendertype=abstract>
- Chen, R., Greer, a, & Dean, a M. (1995). A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(25), 11666–11670. <http://doi.org/10.1073/pnas.92.25.11666>
- Dambe, T. R., Kuhn, A. M., Brossette, T., Giffhorn, F., & Scheidig, A. J. (2006). Crystal Structure of NADP(H)-Dependent 1,5-Anhydro-D-fructose Reductase from *Sinorhizobium morelense* at 2.2 Å Resolution: Construction of a NADH-Accepting Mutant and Its Application in Rare Sugar Synthesis. *Biochemistry*, *45*, 10030–10042. <http://doi.org/10.1128/AEM.72.2.1248>
- Dantas, G., Kuhlman, B., Callender, D., Wong, M., & Baker, D. (2003). A large scale test of computational protein design: Folding and stability of nine completely redesigned globular proteins. *Journal of Molecular Biology*, *332*(2), 449–460. [http://doi.org/10.1016/S0022-2836\(03\)00888-X](http://doi.org/10.1016/S0022-2836(03)00888-X)
- Dill, K. A., & MacCallum, J. L. (2012). The Protein-Folding Problem, 50 Years On. *Science*, *338*(6110), 1042–1046. <http://doi.org/10.1126/science.1219021>
- Eick, G. N., Colucci, J. K., Harms, M. J., Ortlund, E. A., & Thornton, J. W. (2012). Evolution of Minimal Specificity and Promiscuity in Steroid Hormone Receptors. *PLoS Genetics*, *8*(11). <http://doi.org/10.1371/journal.pgen.1003072>
- Eppink, M. H., Overkamp, K. M., Schreuder, H. a, & Van Berkel, W. J. (1999). Switch of coenzyme specificity of p-hydroxybenzoate hydroxylase. *Journal of Molecular Biology*, *292*(1), 87–96. <http://doi.org/10.1006/jmbi.1999.3015>
- Escalante, A., Calderon, R., Valdivia, A., de Anda, R., Hernandez, G., Ramirez, O., & Bolívar, F. (2010). Metabolic engineering for the production of shikimic acid in an evolved *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. *Microbial Cell Factories*, *9*, 1–12. <http://doi.org/10.1186/1475-2859-9-21>
- Flores, H., & Ellington, A. D. (2005). A modified consensus approach to mutagenesis inverts the cofactor specificity of *Bacillus stearothermophilus* lactate dehydrogenase. *Protein Engineering, Design and Selection*, *18*(8), 369–377. <http://doi.org/10.1093/protein/gzi043>
- Fonseca, I. O., Silva, R. G., Fernandes, C. L., de Souza, O. N., Basso, L. A., & Santos, D. S. (2007). Kinetic and chemical mechanisms of shikimate dehydrogenase from *Mycobacterium tuberculosis*. *Archives of Biochemistry and*

- Biophysics*, 457(2), 123–133. <http://doi.org/10.1016/j.abb.2006.11.015>
- Gainza, P., Nisonoff, H. M., & Donald, B. R. (2016). Algorithms for protein design. *Current Opinion in Structural Biology*, 39, 16–26. <http://doi.org/10.1016/j.sbi.2016.03.006>
- García-Guevara, F., Bravo, I., Martínez-Anaya, C., & Segovia, L. (2017). Cofactor specificity switch in Shikimate dehydrogenase by rational design and consensus engineering. *Protein Engineering, Design and Selection*, (September), 1–9. <http://doi.org/10.1093/protein/gzx031>
- Ghosh, S., Chisti, Y., & Banerjee, U. C. (2012). Production of shikimic acid. *Biotechnology Advances*, 30(6), 1425–1431. <http://doi.org/10.1016/j.biotechadv.2012.03.001>
- Giger, L., Caner, S., Obexer, R., Kast, P., Baker, D., Ban, N., & Hilvert, D. (2013). Evolution of a designed retro-aldolase leads to complete active site remodeling. *Nature Chemical Biology*, 9(8), 494–498. <http://doi.org/10.1038/nchembio.1276>
- Hasegawa, S., Uematsu, K., Natsuma, Y., Suda, M., Hiraga, K., Jojima, T., ... Yukawa, H. (2012). Improvement of the redox balance increases L-valine production by *Corynebacterium glutamicum* under oxygen deprivation conditions. *Applied and Environmental Microbiology*, 78(3), 865–875. <http://doi.org/10.1128/AEM.07056-11>
- Haslam, E., & Turner, M. J. (1971). The shikimate pathway. Part II. Conformational analysis of (–)-quinic acid and its derivatives by proton magnetic resonance spectroscopy. *J. Chem. Soc., Section C*, 56(5), 1496–1500. [http://doi.org/10.1016/S0300-9084\(74\)80046-5](http://doi.org/10.1016/S0300-9084(74)80046-5)
- Hol, W. G., van Duijnen, P. T., & Berendsen, H. J. (1978). The alpha-helix dipole and the properties of proteins. *Nature*, 273, 443–446. <http://doi.org/10.1038/273443a0>
- Jackel, C., Bloom, J. D., Kast, P., Arnold, F. H., & Hilvert, D. (2010). Consensus protein design without phylogenetic bias. *Journal of Molecular Biology*, 399(4), 541–546. <http://doi.org/10.1016/j.jmb.2010.04.039>
- Jiang, L., Althoff, E. A., Clemente, F. R., Doyle, L., Zanghellini, A., Gallaher, J. L., ... Stoddard, B. L. (2012). De Novo Computational Design of Retro-Aldol Enzymes. *NIH Public Access*, 319(5868), 1387–1391. <http://doi.org/10.1126/science.1152692>
- Jiang, L., Althoff, E. a, Clemente, F. R., Doyle, L., Röthlisberger, D., Zanghellini, A., ... Baker, D. (2008). De novo computational design of retro-aldol enzymes. *Science (New York, N.Y.)*, 319(5868), 1387–1391. <http://doi.org/10.1126/science.1152692>
- Khoury, G. a., Fazelinia, H., Chin, J. W., Pantazes, R. J., Cirino, P. C., & Maranas, C. D. (2009). Computational design of *Candida boidinii* xylose reductase for altered cofactor specificity. *Protein Science*, 18(10), 2125–2138. <http://doi.org/10.1002/pro.227>

- Knop, D. R., Draths, K. M., Chandran, S. S., Barker, J. L., Von Daeniken, R., Weber, W., & Frost, J. W. (2001). Hydroaromatic equilibration during biosynthesis of shikimic acid. *Journal of the American Chemical Society*, 123(42), 10173–10182. <http://doi.org/10.1021/ja0109444>
- Kogure, T., Kubota, T., Suda, M., Hiraga, K., & Inui, M. (2016). Metabolic engineering of *Corynebacterium glutamicum* for shikimate overproduction by growth-arrested cell reaction. *Metabolic Engineering*, 38, 204–216. <http://doi.org/10.1016/j.ymben.2016.08.005>
- Krämer, M., Bongaerts, J., Bovenberg, R., Kremer, S., Müller, U., Orf, S., ... Raeven, L. (2003). Metabolic engineering for microbial production of shikimic acid. *Metabolic Engineering*, 5(4), 277–283. <http://doi.org/10.1016/j.ymben.2003.09.001>
- Kristan, K., Stojan, J., Adamski, J., & Lanišnik Rižner, T. (2007). Rational design of novel mutants of fungal 17 β -hydroxysteroid dehydrogenase. *Journal of Biotechnology*, 129(1), 123–130. <http://doi.org/10.1016/j.jbiotec.2006.11.025>
- Kuchner, O., & Arnold, F. H. (1997). Directed evolution of enzyme catalysts. *Trends in Biotechnology*, 15(12), 523–530. [http://doi.org/10.1016/S0167-7799\(97\)01138-4](http://doi.org/10.1016/S0167-7799(97)01138-4)
- Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L., & Baker, D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science (New York, N.Y.)*, 302(5649), 1364–1368. <http://doi.org/10.1126/science.1089427>
- Lehmann, M., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L., & van Loon, A. P. G. M. (2000). From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. *Protein Engineering Design and Selection*, 13(1), 49–57. <http://doi.org/10.1093/protein/13.1.49>
- Lehmann, M., Pasamontes, L., Lassen, S. F., & Wyss, M. (2000). The consensus concept for thermostability engineering of proteins. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*. [http://doi.org/10.1016/S0167-4838\(00\)00238-7](http://doi.org/10.1016/S0167-4838(00)00238-7)
- Lehmann, M., Pasamontes, L., Lassen, S. F., & Wyss, M. (2000). The consensus concept for thermostability engineering of proteins. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, 1543(2), 408–415. [http://doi.org/10.1016/S0167-4838\(00\)00238-7](http://doi.org/10.1016/S0167-4838(00)00238-7)
- Lehmann, M., & Wyss, M. (2001). Engineering proteins for thermostability: The use of sequence alignments versus rational design and directed evolution. *Current Opinion in Biotechnology*. [http://doi.org/10.1016/S0958-1669\(00\)00229-9](http://doi.org/10.1016/S0958-1669(00)00229-9)
- Liang, L., Zhang, J., & Lin, Z. (2007). Altering coenzyme specificity of *Pichia stipitis* xylose reductase by the semi-

- rational approach CASTing. *Microbial Cell Factories*, 6, 36. <http://doi.org/10.1186/1475-2859-6-36>
- Liao, H. F., Lin, L. L., Chien, H. R., & Hsu, W. H. (2001). Serine 187 is a crucial residue for allosteric regulation of *Corynebacterium glutamicum* 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase. *FEMS Microbiology Letters*, 194(1), 59–64. [http://doi.org/10.1016/S0378-1097\(00\)00507-3](http://doi.org/10.1016/S0378-1097(00)00507-3)
- Liu, H., & Chen, Q. (2016). Computational protein design for given backbone: Recent progresses in general method-related aspects. *Current Opinion in Structural Biology*, 39, 89–95. <http://doi.org/10.1016/j.sbi.2016.06.013>
- Loening, A. M., Fenn, T. D., Wu, A. M., & Gambhir, S. S. (2006). Consensus guided mutagenesis of Renilla luciferase yields enhanced stability and light output. *Protein Engineering, Design and Selection*, 19(9), 391–400. <http://doi.org/10.1093/protein/gzl023>
- Loffler, P., Schmitz, S., Hupfeld, E., Sterner, R., Merkl, R., & Hughes, M. (2017). Rosetta:MSF: a modular framework for multi-state computational protein design. *PLOS Computational Biology*, 13(6), e1005600. <http://doi.org/10.1371/journal.pcbi.1005600>
- Maclean, J., Samantha, A., Pollock, K., Chackrewarthy, S., John, R., & Adrian, J. (2000). Crystallization and preliminary X-ray analysis of shikimate dehydrogenase from *Escherichia coli*. *Acta Crystallographica, Biological Crystallography*, 512–515.
- Martínez, J. A., Bolívar, F., & Escalante, A. (2015). Shikimic Acid Production in *Escherichia coli*: From Classical Metabolic Engineering Strategies to Omics Applied to Improve Its Production. *Frontiers in Bioengineering and Biotechnology*, 3(September), 1–16. <http://doi.org/10.3389/fbioe.2015.00145>
- McComb, R. B., Bond, L. W., Burnett, R. W., Keech, R. C., & Bowers, G. N. (1976). Determination of the molar absorptivity of NADH. *Clinical Chemistry*, 22(2), 141–150.
- Michel, G., Roszak, A. W., Sauvé, V., Maclean, J., Matte, A., Coggins, J. R., ... Laphorn, A. J. (2003). Structures of shikimate dehydrogenase AroE and its paralog YdiB: A common structural framework for different activities. *Journal of Biological Chemistry*, 278(21), 19463–19472. <http://doi.org/10.1074/jbc.M300794200>
- Ojha, S., Meng, E. C., & Babbitt, P. C. (2007). Evolution of function in the “two dinucleotide binding domains” flavoproteins. *PLoS Computational Biology*, 3(7), 1268–1280. <http://doi.org/10.1371/journal.pcbi.0030121>
- Osuna, J., & Pe, A. (2002). Improving a circularly permuted TEM-1 β -lactamase by directed. *Protein Engineering, Design and Selection*, 15(6), 463–470.
- Peek, J., & Christendat, D. (2015). The shikimate dehydrogenase family: Functional diversity within a conserved structural and mechanistic framework. *Archives of Biochemistry and Biophysics*, 566, 85–99. <http://doi.org/10.1016/j.abb.2014.12.006>

- Peek, J., Garcia, C., Lee, J., & Christendat, D. (2013). Insights into the function of RifI2: Structural and biochemical investigation of a new shikimate dehydrogenase family protein. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1834(2), 516–523. <http://doi.org/10.1016/j.bbapap.2012.10.016>
- Pick, A., Ott, W., Howe, T., Schmid, J., & Sieber, V. (2014). Improving the NADH-cofactor specificity of the highly active AdhZ3 and AdhZ2 from *Escherichia coli* K-12. *Journal of Biotechnology*, 189, 157–165.
- Pollak, N., Dölle, C., & Ziegler, M. (2007). The power to reduce: pyridine nucleotides – small molecules with a multitude of functions. *Biochemical Journal*, 402(2), 205–218. <http://doi.org/10.1042/BJ20061638>
- Pruitt, K. D., Tatusova, T., & Maglott, D. R. (2007). NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, 35(SUPPL. 1), 61–65. <http://doi.org/10.1093/nar/gkl842>
- Rane, M. J., & Calvo, K. C. (1997). Reversal of the nucleotide specificity of ketol acid reductoisomerase by site-directed mutagenesis identifies the NADPH binding site. *Archives of Biochemistry and Biophysics*, 338(1), 83–89. <http://doi.org/10.1006/abbi.1996.9802>
- Rao, S. T., & Rossmann, M. G. (1973). Comparison of super-secondary structures in proteins. *Journal of Molecular Biology*, 76(2), 241–256. [http://doi.org/10.1016/0022-2836\(73\)90388-4](http://doi.org/10.1016/0022-2836(73)90388-4)
- Rawat, G., Tripathi, P., Jahan, F., & Saxena, R. K. (2013). A natural isolate producing shikimic acid: Isolation, identification, and culture condition optimization. *Applied Biochemistry and Biotechnology*, 169(8), 2290–2302. <http://doi.org/10.1007/s12010-013-0150-1>
- Risso, V. A., Gavira, J. A., Mejia-Carmona, D. F., Gaucher, E. A., & Sanchez-Ruiz, J. M. (2013). Hyperstability and substrate promiscuity in laboratory resurrections of precambrian β -lactamases. *Journal of the American Chemical Society*, 135(8), 2899–2902. <http://doi.org/10.1021/ja311630a>
- Risso, V. A., Gavira, J. A., & Sanchez-Ruiz, J. M. (2013). Thermostable and promiscuous Precambrian proteins. *Environmental Microbiology*, 16, 1485–1489. <http://doi.org/10.1111/1462-2920.12319>
- Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(April), 320–324. <http://doi.org/10.1093/nar/gku316>
- Rodriguez, A., Martinez, J. A., Flores, N., Escalante, A., Gosset, G., & Bolivar, F. (2014). Engineering *Escherichia coli* to overproduce aromatic amino acids and derived compounds. *Microbial Cell Factories*, 13(1), 126. <http://doi.org/10.1186/s12934-014-0126-z>
- Rosell, A., Valencia, E., Ochoa, W. F., Fita, I., Parés, X., & Farrés, J. (2003). Complete Reversal of Coenzyme Specificity by Concerted Mutation of Three Consecutive Residues in Alcohol Dehydrogenase. *Journal of*

- Biological Chemistry*, 278(42), 40573–40580. <http://doi.org/10.1074/jbc.M307384200>
- Röthlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., DeChancie, J., Betker, J., ... Baker, D. (2008). Kemp elimination catalysts by computational enzyme design. *Nature*, 453(7192), 190–195. <http://doi.org/10.1038/nature06879>
- Scrutton, N. S., Berry, a, & Perham, R. N. (1990). Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature*, 343(6253), 38–43. <http://doi.org/10.1038/343038a0>
- Shao, Z., & Arnold, F. H. (1996). Engineering new functions and altering existing functions. *Current Opinion in Structural Biology*, 6(4), 513–518. [http://doi.org/10.1016/S0959-440X\(96\)80117-5](http://doi.org/10.1016/S0959-440X(96)80117-5)
- Shen, C. R., Lan, E. I., Dekishima, Y., Baez, A., Cho, K. M., & Liao, J. C. (2011). Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Applied and Environmental Microbiology*, 77(9), 2905–2915. <http://doi.org/10.1128/AEM.03034-10>
- Siegel, J. B., Zanghellini, A., Lovick, H. M., Kiss, G., Lambert, A. R., Clair, J. L. S., ... Baker, D. (2010). Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder Reaction. *Science*, 329(5989), 309–313.
- Singh, S., Stavrinides, J., Christendat, D., & Guttman, D. S. (2008). A Phylogenomic Analysis of the Shikimate Dehydrogenases Reveals Broad-scale Functional Diversification and Identifies One Functionally Distinct Subclass. *Molecular Biology and Evolution*, 25(10), 2221–2232. <http://doi.org/10.1093/molbev/msn170>
- Steipe, B., Schiller, B., Pluckthun, A., & Steinbacher, S. (1994). Sequence statistics reliably predict stabilizing mutations in a protein domain. *Journal of Molecular Biology*, 188–192.
- Stewart, J. J. P. (2007). Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling*, 13(12), 1173–1213. <http://doi.org/10.1007/s00894-007-0233-4>
- Tyzack, J. D., Furnham, N., Sillitoe, I., Orengo, C. M., & Thornton, J. M. (2017). Understanding enzyme function evolution from a computational perspective. *Current Opinion in Structural Biology*, 47, 131–139. <http://doi.org/10.1016/j.sbi.2017.08.003>
- Vázquez-Figueroa, E., Chapparro-Riggers, J., & Bommarius, A. S. (2007). Development of a thermostable glucose dehydrogenase by a structure-guided consensus concept. *ChemBioChem*, 8(18), 2295–2301. <http://doi.org/10.1002/cbic.200700500>
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, 12(1), 7–8. <http://doi.org/10.1038/nmeth.3213>

- Yang, J., & Zhang, Y. (2015). I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, 43(W1), W174–W181. <http://doi.org/10.1093/nar/gkv342>
- Yang, W., & Lai, L. (2017). Computational design of ligand-binding proteins. *Current Opinion in Structural Biology*, 45, 67–68. <http://doi.org/10.1016/j.sbi.2016.11.021>
- Yaoi, T., Miyazaki, K., Oshima, T., Komukai, Y., & Go, M. (1996). Conversion of the coenzyme specificity of isocitrate dehydrogenase by module replacement. *Journal of Biochemistry*, 119(5), 1014–1018.
- Zhang, R., Xu, Y., Sun, Y., Zhang, W., & Xiao, R. (2009). Ser67Asp and His68Asp substitutions in *Candida parapsilosis* carbonyl reductase alter the coenzyme specificity and enantioselectivity of ketone reduction. *Applied and Environmental Microbiology*, 75(7), 2176–2183. <http://doi.org/10.1128/AEM.02519-08>
- Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9(1), 40. <http://doi.org/10.1186/1471-2105-9-40>

12 Anexo

Figura Anexo 1

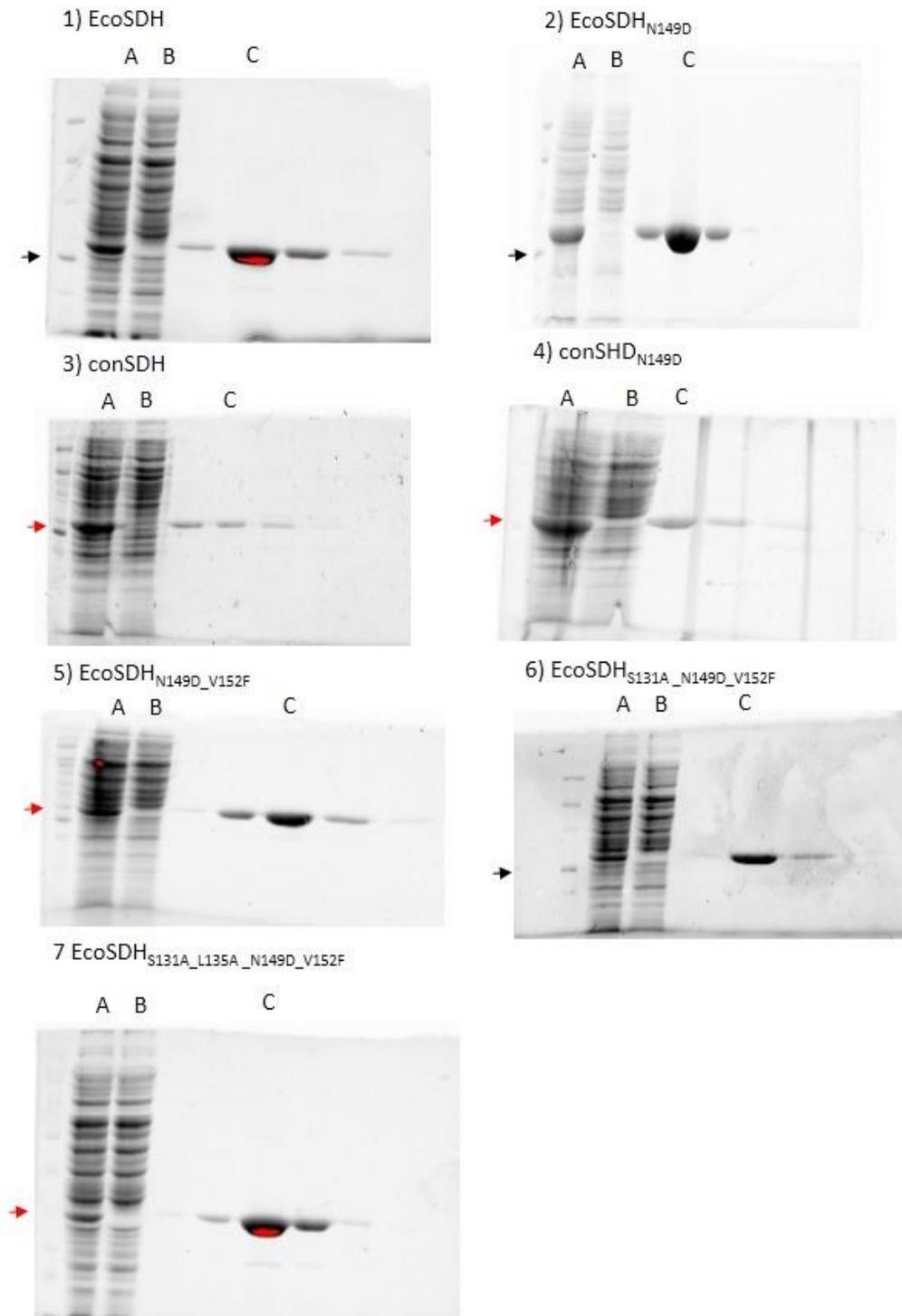


Figura anexo 1. Purificación de proteína de las variantes reportadas. La imagen de arriba fue obtenida a partir de geles PAGE-SDS teñidos con tricloroetanol. Para la purificación de la proteína utilizamos una columna de cromatografía por afinidad a níquel His-Trap de 5 mL (GE Healthcare; Uppsala, Suecia). Los carriles A muestran el extracto total de proteína; los carriles B muestran el extracto total después de pasar por la columna His-Trap; los carriles C muestran la fracción después de eluir con *buffer* con concentración de 250mM de imidazol. Las proteínas purificadas presentan una pureza superior al 95% juzgada por el análisis de la imagen con el software del equipo usado para tomar las fotografías: Gel Doc EZ Imager de BIORAD. En las imágenes 1), 2) y 6) el marcador de peso molecular usado fue *Unstained Protein MW Marker* de Jena Biosciences, catálogo #PS-101; la flecha con negro señala la marca con un peso de 28.5 KDa. En las imágenes 3), 4), 5) y 7) el marcador de peso molecular usado fue *PAGE Ruler unstained protein ladder* de *Thermo Scientific*, catálogo #26614. Las flechas con rojo señalan la marca con un peso de 30 KDa.

Figura anexo 2

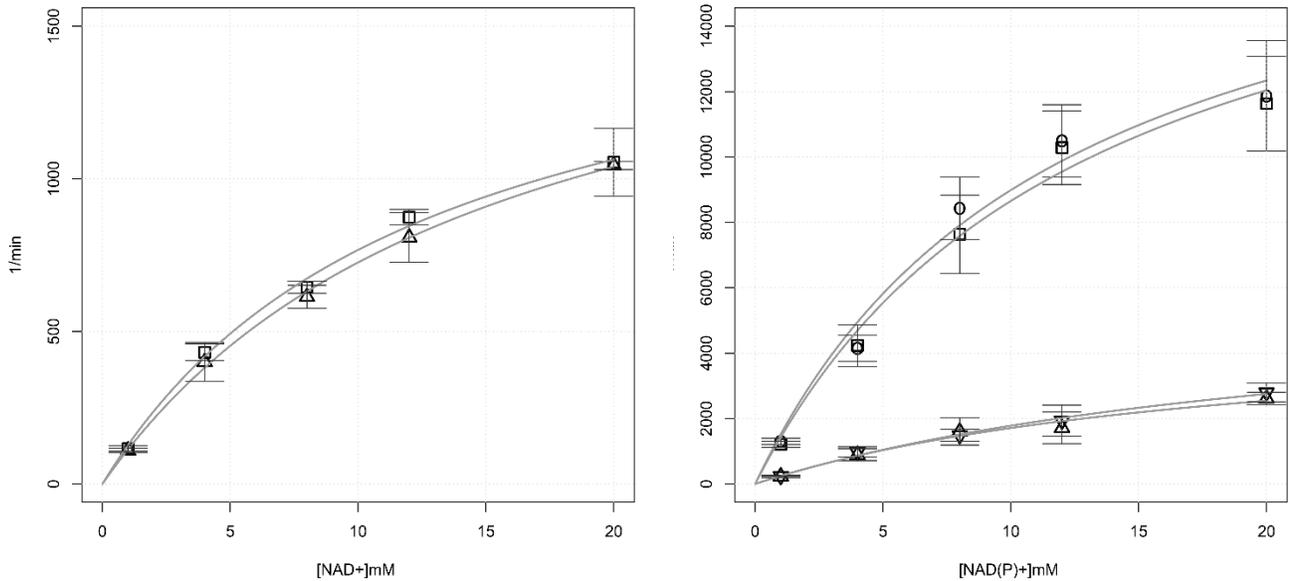


Figura anexo 2. Con 4mM de shikimato alcanzamos condiciones de saturación del sustrato. En la enzima EcoSDH, 4mM de shikimato es suficiente para alcanzar condiciones de saturación del sustrato usando NADP+. Nos preguntamos si esa concentración sigue siendo suficiente para alcanzar condiciones de saturación del sustrato cuando usa NAD+ o en la variante EcoSDH_{S131A_L135A_N149D_V152F} (la que tiene mayor actividad usando NAD+). Si 4mM todavía no alcanza condiciones de saturación entonces aumentar la concentración de shikimato produciría un aumento en la actividad, en cambio, si ya se encuentra en condiciones de saturación entonces no esperamos ver un aumento en la actividad. Encontramos mediciones similares usando 9mM de shikimato indicando que con 4 mM alcanzamos condiciones de saturación del sustrato. La figura de la izquierda muestra la velocidad inicial de EcoSDH a diferentes concentraciones de NAD+ (1, 4, 8, 12 y 20 mM) con shikimato a 4mM (círculos) y 9 mM (cuadrados). La figura de la derecha muestra la velocidad inicial de EcoSDH_{S131A_L135A_N149D_V152F} con diferentes concentraciones del cofactor (1, 4, 8, 12 y 20 mM): usando NAD+ y 4mM de shikimato, círculos; NAD+ y 9mM de shikimato, cuadrados; NADP+ y 4mM de shikimato, triángulos hacia arriba; NADP+ y 9mM de shikimato, triángulos hacia abajo. Las barras de error muestran los intervalos de confianza del 95% de la media y las curvas grises muestran un ajuste de un modelo simple de Michaelis-Menten.

Figura anexo 3

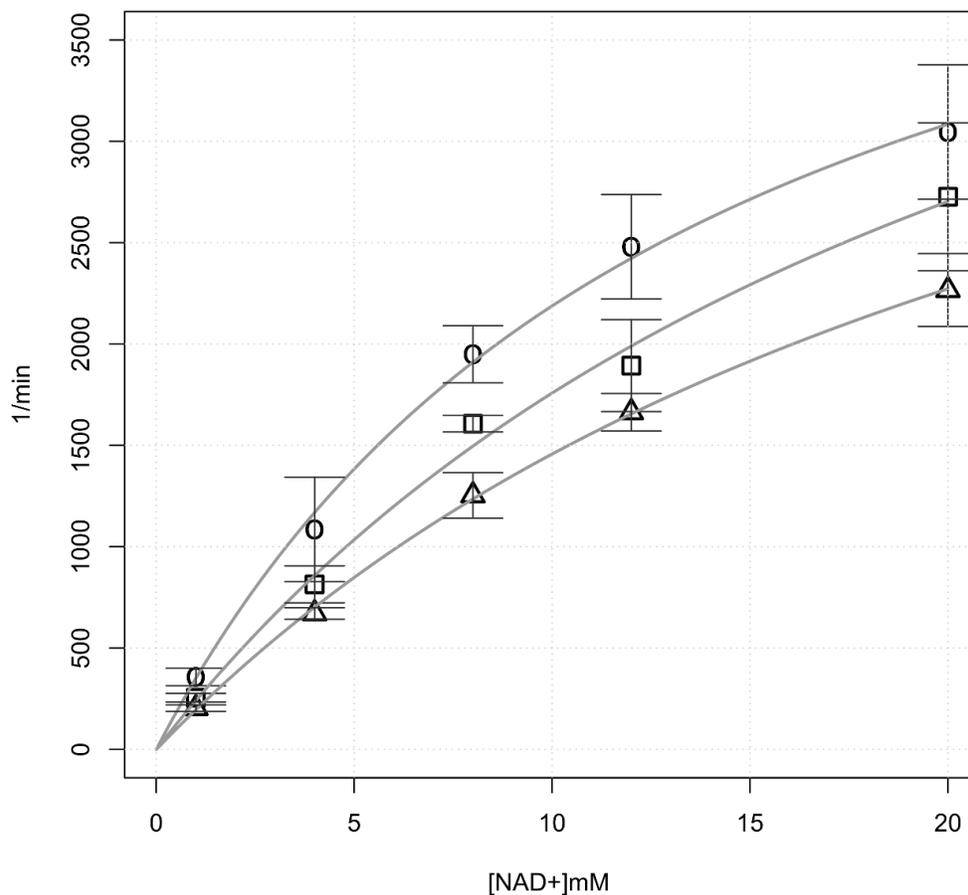


Figura anexo 3. Las mutaciones consenso D195E y A243F no aumentan la actividad con NAD⁺. Medimos la velocidad inicial a diferentes concentraciones de NAD⁺ (1, 4, 8, 12 y 20 mM) y una concentración fija de shikimato (4mM) en las variantes EcoSDH_{N149D_V152F_D195E} (cuadrados) y EcoSDH_{N149D_V152F_A243G} (triángulos) y las comparamos contra el perfil de velocidad inicial medido en la variante EcoSDH_{N149D_V152F} (círculos). Las mutaciones consenso D195E y A243F no aumentan la actividad con NAD⁺.

Figura anexo 4

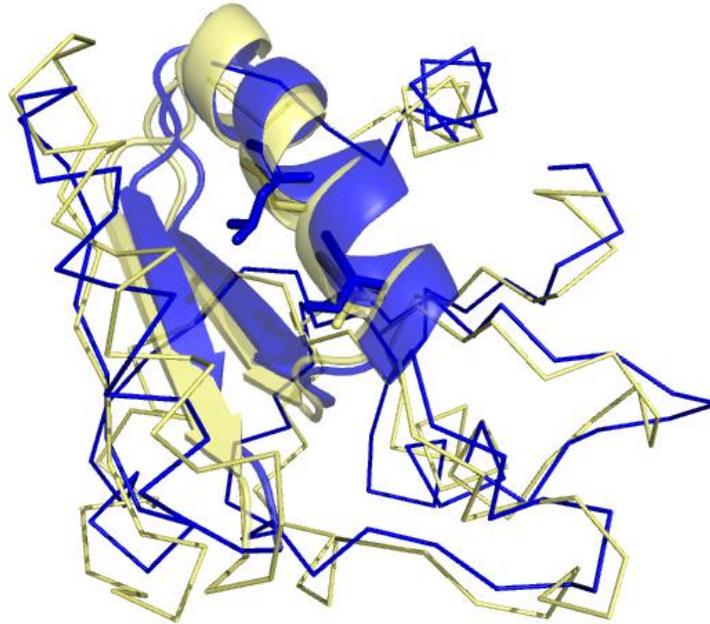


Figura anexo 4. Discusión, longitud de la cadena en dominios Rossmann que prefieren NAD⁺. La figura muestra el dominio Rossmann de EcoSDH (cadena A, 1nyt) en azul, alineado contra el dominio Rossmann de la enzima Q/SDH de *S. entérica*, en amarillo (cadena A, 3t4e). En EcoSDH, que une NADP⁺, las cadenas laterales S131 y L135 separan la primera hélice alfa de la lámina beta. En la enzima Q/SDH de *S. entérica*, que une NAD⁺, los residuos correspondientes A136 y G140 permiten un empaquetamiento más cercano de la primera hélice alfa contra la lámina beta.

Tabla anexo 1.

Parámetros cinéticos aparentes estimados para EcoSDH y conSDH usando shikimato y NADP+

	Sustrato	k_{CATapp} (s ⁻¹)	K_{Mapp} (mM)	k_{CATapp}/K_{Mapp} (s ⁻¹ M ⁻¹)
EcoSDH	Shikimato	190 ± 5 (237 ^a)	0.13 ± 0.013 (0.065 ^a)	1.4 x 10 ⁶ (3.6 x 10 ⁶ ^a)
	NADP+	186 ± 4.6 (237 ^a)	0.058 ± 0.007 (0.056 ^a)	3.2 x 10 ⁶ (4.2 x 10 ⁶ ^a)
conSDH	Shikimato	183 ± 5.3	0.061 ± 0.01	3.0 x 10 ⁶
	NADP+	198 ± 3.8	0.105 ± 0.01	1.8 x 10 ⁶

^a Valores en (Michel et al., 2003)

Tabla anexo 2.

(k_{CAT}/K_M) aparentes.

	NAD+ (k_{CAT}/K_M)_{app} (min ⁻¹ mM ⁻¹)	NADP+ (k_{CAT}/K_M)_{app} (min ⁻¹ mM ⁻¹)
EcoSDH	85 ± 2	1.92 x 10 ⁵ ^a
conSDH	186 ± 4.7	1.08 x 10 ⁵ ^a
EcoSDH_{N149D}	108 ± 4.5	19.3 ± 1.5
EcoSDH_{N149D_V152F}	250 ± 9.4	27.4 ± 0.6
EcoSDH_{N149D_V152F_S131A}	753 ± 18	133 ± 1.8
EcoSDH_{N149D_V152F_S131A_L135A}	1003 ± 20	188 ± 6.1
conSDH_{N149D}	449 ± 25	65 ± 1.7

^a Los valores para EcoSDH y conSDH usando NADP+ se calcularon dividiendo los valores de k_{CATapp_NADP+} and K_{Mapp_NADP+} de la tabla anterior y los valores para el restantes se calcularon de acuerdo a lo descrito en la sección de materiales y métodos.