



Tesina de Grado
Licenciatura en Bioquímica

Identificación, clonado, expresión y caracterización de nuevas imino reductasas y aminasas reductivas

Tutora: Dra. Paola Panizza

Cotutor: Dr. César Iglesias

Bach. Gonzalo Martín López Pérez

Laboratorio de Microbiología Molecular, Área
Microbiología, Departamento de Biociencias,
Facultad de Química, Universidad de la Republica

Montevideo, Uruguay

2020

Resumen	3
Agradecimientos	4
Introducción.....	5
Biocatálisis.....	5
Aminas quirales y la industria	6
Síntesis química de aminas quirales.....	7
Síntesis biocatalítica de aminas quirales.....	9
Imino reductasas.....	11
Estructura y mecanismo de las IRED	11
Aminación reductiva biocatalítica	12
Estructura y mecanismo propuesto para imino reductasas con actividad aminasa reductiva	14
Bioinformática y Biocatálisis	15
Búsqueda de proteínas	17
Modelado por homología	18
Molecular docking.....	20
Búsqueda de nuevas iminoreductasas en cepas nativas	21
Objetivos.....	22
Objetivos generales.....	22
Objetivos Específicos.....	22
Resultados	23
Clonado, expresión y caracterización de la IRED S encontrada en el genoma de <i>Streptomyces scabiei</i> ST129 en el vector pKK223-3.	23
Transformación en <i>E. coli</i>	24
Estudio computacional de la estructura tridimensional de la IRED R encontrada en el genoma de <i>Streptomyces sp.</i> ST1020 y su sitio activo.	28
Modelado por homología de IRED R	28
Estudio computacional de la actividad IRED y RedAm de la IRED R utilizando la metodología de Molecular Docking.	29
Actividad IRED.....	29
Actividad RedAm	33
Caracterización de la IRED R clonada en pET28b(+) y expresada en <i>E. coli</i>	35
Búsqueda de nuevas RedAm por el método basado en motivos conservados	37
Filtrado y comparación de los candidatos a RedAm encontrados con enzimas RedAm reportadas por literatura	39
Conclusiones y perspectivas.....	41
Materiales y métodos	42

Métodos microbiológicos.....	42
Protocolos Computacionales	43
Protocolos experimentales	46
Medios de cultivo y soluciones	51
Referencias	54

Las aminas quirales resultan de gran importancia dado que se encuentran presentes en una gran variedad de compuestos, entre ellos fármacos, agroquímicos, y productos naturales. Su presencia en la gran variedad de moléculas, hacen que estas sean integradas a bloques de construcción para el diseño de nuevos compuestos. Por estos motivos, hay un gran mercado alrededor de estas. Últimamente, teniendo en cuenta los factores ambientales, aunque también los económicos, hubo un gran desarrollo de rutas biocatalíticas, es decir, síntesis enzimáticas, que permiten obtener aminas enantioméricamente puras con métodos menos nocivos para el ambiente que los métodos clásicos.

En los últimos años, la capacidad de cómputo ha avanzado exponencialmente, permitiendo realizar grandes simulaciones en computadoras de un precio accesible. Gracias a esto, la bioinformática ha adquirido una gran importancia para el desarrollo de innumerables investigaciones en el área de las ciencias biológicas, y se han generado grandes bancos de datos de los cuales se puede obtener mucha información. A la hora de enfocarnos en la búsqueda de nuevos biocatalizadores capaces de sintetizar aminas, es muy acertado, por temas de costos, así también como de tiempos, el empleo de la bioinformática. Esta, nos permite buscar, modelar y realizar experimentos computacionales, que nos brindan grandes aproximaciones a lo experimental.

En trabajos previos de nuestro grupo se encontraron dos posibles iminoreductasas en genomas de *Streptomyces* nativos (IRED S y R) En este trabajo se clonó y expresó la IRED S en el vector pKK 223-3, de la cual no se logró ver actividad. Además, se realizó un estudio computacional de la posible actividad IRED y RedAm de la IRED R, para esto se realizó el modelado por homología de la proteína y posterior análisis mediante la metodología de docking molecular. Una vez obtenidos los resultados *in silico* se planeó un estudio *in vitro* para evaluar la eficiencia de este. Esta enzima se clonó y expresó de manera heteróloga en *Escherichia coli* y se evaluaron los resultados, los cuales fueron coherentes con los obtenidos de manera computacional. De esta manera se logró no solo caracterizar la IRED R encontrada por el grupo, sino que la actividad de esta fue respaldada por los resultados computacionales. Por último, dada la importancia que han tomado en este último tiempo las aminosas reductivas, decidimos realizar una búsqueda basada en motivos conservados del sitio activo obteniendo un grupo de enzimas muy diversas evolutivamente de las ya reportadas. Esto nos abre un abanico de posibilidades para el desarrollo de nuevos biocatalizadores de utilidad para la síntesis de aminas quirales. A su vez, la validación de las predicciones computacionales mediante resultados experimentales nos abre la vía de realizar screenings *in silico* a futuras enzimas que encontremos ahorrándonos los costos y el tiempo que conllevan los métodos tradicionales.

AGRADECIMIENTOS

Primero, en esta tesis, me gustaría agradecer a Sonia, por permitirme ingresar en su grupo de trabajo, asistiéndome día a día, y por estar siempre disponible para lo que necesitara más allá de lo académico.

También a Paola y Cesar, por tomarme como su alumno de licenciatura en bioquímica y por el apoyo que recibí de ambos durante la misma.

A Cecilia Porciúncula por aceptar la evaluación de mi tesis, en tiempos muy acotados.

De facultad de ciencias quería agradecer a todos mis compañeros de generación y en especial obviamente a los tortus, Rodri, Santi y Joaco.

A todo el Laboratorio de Microbiología Molecular (también conocido por nosotros como “el 12”) por bancarme desde 2018, y en especial a todos los sonios, Ariel, Diego, Agustín, Agustina y espero no olvidarme de nadie, que estuvieron siempre disponible ante consultas del día a día.

También a la gente del Laboratorio de Biocatálisis y Biotransformaciones, que les he ocupado la oficina, y también el laboratorio.

Institucionalmente quiero agradecer a Facultad de Química, al Área Microbiología y al Departamento de Biociencias, donde realice mi tesina de grado. También a Facultad de Ciencias, que fue la institución donde realice toda mi carrera de grado.

A la Agencia Nacional de Investigación e Innovación y al fondo L’Oreal por permitir la financiación de esta tesina.

Por último, le quiero agradecer a mi familia, que son quienes han estado apoyándome desde siempre, principalmente a mi madre, que nos banco siempre a pesar de ser sola, y a mi hermano, que fue la figura masculina que siempre tuve en mi casa.

A mi novia, Stephanie, que me ha apoyado desde el día 0, cursando la carrera junto conmigo, bancándome en todo momento, tanto los buenos, como los otros. Y a todos aquellos que me han apoyado, principalmente emocionalmente a lo largo de la carrera.

BIOCATÁLISIS

La biocatálisis se define como la síntesis química mediada por la ayuda de material biológico, utilizando como catalizador las enzimas, (ya sea que se utilicen aisladas, presentes en lisados celulares o en células enteras), que son capaces de convertir uno o más sustratos en productos. Existe una extensa variedad de enzimas utilizadas como biocatalizadores únicos o combinados en la preparación de distintas familias de compuestos¹. La gama de biocatalizadores está en constante crecimiento gracias al aumento de la cantidad de secuencias depositadas en bases de datos y al descubrimiento de nuevas enzimas en rutas biosintéticas.²

En cuanto a las aplicaciones sintéticas, las ventajas de las enzimas para suplir los catalizadores tradicionales son múltiples. Éstas poseen un alto poder catalítico, potenciando la velocidad de una reacción en el orden de hasta 10^{17} veces.³ Presentan gran especificidad por el sustrato, así como también una alta regio-, enantio- y estereo-selectividad en las reacciones que catalizan. Las enzimas trabajan en condiciones suaves de reacción, como por ejemplo pH cercano a la neutralidad y temperaturas moderadas. Si bien esto puede verse como una ventaja, limita la posibilidad de acelerar la reacción cambiando el pH, la concentración de sales o subiendo la temperatura, dado que esto afecta la actividad de muchas enzimas. Otra de las grandes desventajas es el hecho de que la mayoría de éstas tienen su máxima actividad en medios acuosos, por lo que se hace muy difícil utilizarlas en solventes orgánicos necesarios para que se lleven a cabo determinadas reacciones. Además, la actividad enzimática puede ser inhibida con diferentes compuestos o incluso por su propio producto.⁴

Sin embargo, a principio de los años 80 un trabajo pionero realizado con lipasas, demostró que estas son estables a altas temperaturas en solventes orgánicos. Esto reveló la gran capacidad que tienen algunas enzimas de catalizar reacciones en distintas temperaturas y solventes.⁵ El descubrimiento de esta característica de las lipasas fue lo que las hizo ser las primeras enzimas en captar la atención de la industria para su utilización en procesos de manufactura de productos masivos, ya sea como ingredientes o como catalizadores de reacción. Actualmente, la biocatálisis se enmarca en un área muy bien establecida de constante desarrollo tanto en la industria como en la academia. Esto último lo podemos notar gracias a dos avances paralelos evidentes en la industria en esta última época: la creciente demanda de científicos de las áreas biológicas, así como también la adaptación de los químicos orgánicos a la utilización de la biocatálisis en casos nuevos y/o difíciles de resolver de manera exclusivamente química.⁶

Las aminas son un grupo de compuestos que tienen la característica común de poseer átomos de nitrógeno con orbitales híbridos sp^3 que forman enlaces simples con otros elementos.⁷ Estos grupos funcionales los podemos encontrar formando parte de una gran diversidad de compuestos, como por ejemplo fármacos, productos naturales y alimentos (Fig. 1). Sin ir muy lejos, éstas son parte fundamental de nuestro cuerpo, las podemos encontrar en las bases nitrogenadas que forman las secuencias de ADN y ARN, como también en los aminoácidos, necesarios para la formación de péptidos y proteínas. Otro ejemplo muy claro de la importancia de las aminas lo podemos ver en diferentes neurotransmisores y en hormonas del grupo de catecolaminas. Estas hormonas son fundamentales para el correcto funcionamiento del cuerpo, jugando un papel central en la regulación de la temperatura y el pulso cardíaco, en las emociones e incluso en el desarrollo de enfermedades mentales.⁸⁻¹⁰

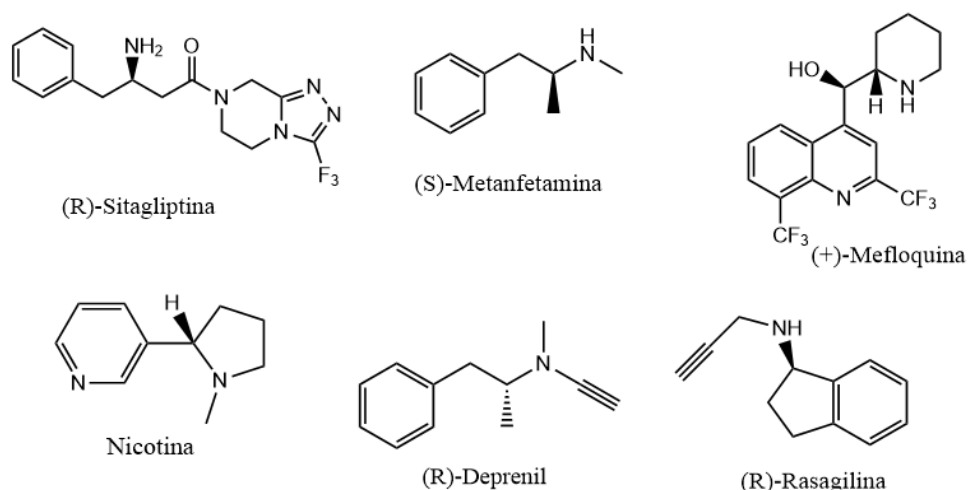


Fig 1. Estructura de diferentes moléculas con aminas quirales en su estructura utilizadas con fines farmacéuticos y recreativos.

Otro aspecto relevante es la gran importancia de la pureza óptica de las aminas. Los estudios sobre los efectos de la quiralidad de las moléculas han determinado que a pesar de que muchas tienen las mismas conexiones entre sus átomos, la disposición espacial de los mismos es muy importante. Esto ocurre porque los mismos receptores, enzimas y canales iónicos, dada su asimetría natural, tienen interacciones diferenciales con respecto a cada estereoisómero¹¹, haciendo que un fármaco ópticamente puro pueda mejorar altamente la efectividad como medicamento con respecto al racemato. Esto se debe a que en un fármaco administrado como racemato, sólo el 50 % de la dosis es efectiva, mientras que, en una droga ópticamente pura, el total de la dosis será farmacológicamente activa.

Un claro ejemplo lo da la empresa AstraZeneca, la cual poseía la patente del fármaco omeprazol, y en el año 2000 patentó un medicamento nuevo llamado esomeprazol, el cual era el enantiómero *S*-omeprazol puro, que mejoró ampliamente la efectividad en intensidad y tiempo con respecto al fármaco racémico.¹² O en casos más extremos aún, hay fármacos en los cuales un enantiómero cumple la función de medicamento mientras que el otro puede llegar a ser tóxico o causar reacciones adversas. Este es el caso de la talidomida, fármaco que comenzó a circular a principio de los 60 como racemato, y se recetaba para evitar las náuseas en los primeros meses de embarazo. Sin embargo, sólo el enantiómero *R* es el activo para esta función, mientras que al enantiómero *S* se le comprobó actividad teratogénica.¹³ Por estos motivos, a pesar de que a principios de los 90 sólo un 25% de las nuevas drogas eran enantioméricamente puras, a principios del siglo XXI este porcentaje pasó a ser el 75%.¹¹

Por los motivos expuestos en los párrafos anteriores, las aminas quirales son de gran importancia en la industria farmacéutica. Se encuentran presentes en un alto porcentaje de fármacos y compuestos bioactivos, por lo que su producción es vital en esta industria (Fig. 1). Las mismas, además de formar parte de diversos fármacos, se encuentran también en moléculas utilizadas muy comúnmente como “bloques de construcción”, que son utilizadas en librerías químicas para generar precursores de fármacos nuevos y ya existentes.¹⁴

SÍNTESIS QUÍMICA DE AMINAS QUIRALES

Industrialmente hay tres maneras de obtener aminas quirales por síntesis orgánica: cristalización mediante formación de sales diastereómericas, resolución cinética o síntesis asimétrica.

En el primer ejemplo, gracias a que los diastereómeros a diferencia de los enantiómeros son fisicoquímicamente diferentes, la cristalización de la mezcla racémica utilizando Ácido *D*-tartárico permite obtener una sal con alto exceso enantiomérico. Se puede mejorar dicha resolución con más de un proceso de cristalización.¹⁵ Otra estrategia para la resolución de los enantiómeros consiste en la resolución cinética de los mismos (Fig. 2). El principio de esta técnica es la velocidad diferencial con que reaccionan ambos enantiómeros ante otro reactivo o catalizador asimétrico, logrando convertir uno de ellos en otro producto mucho antes que reaccione una cantidad considerable del más lento. Este ejemplo, ampliamente usado, se puede ver en la resolución de alcoholes alílicos utilizando como catalizador el tartrato de alcóxido de titanio.¹⁶

La limitación más grande de estos dos métodos es el rendimiento teórico máximo del 50%. Sin embargo, existe la posibilidad de realizar una resolución cinética dinámica, en donde uno de los enantiómeros puede volver hacia atrás la reacción en forma de mezcla racémica. Para ello, la constante de racemización debe ser mayor que las otras. De esta forma se puede llegar a un rendimiento teórico del 100% con un exceso enantiomérico del 100%. Ejemplo de este método es la hidrogenación estereoselectiva de diferentes cetoésteres β -sustituídos utilizando como catalizador asimétrico el rutenio-BINAP, entre otros catalizadores organometálicos. Estos compuestos se racemizan naturalmente a través del equilibrio ceto-enólico, lo que permite realizar esta resolución cinética dinámica.¹⁷

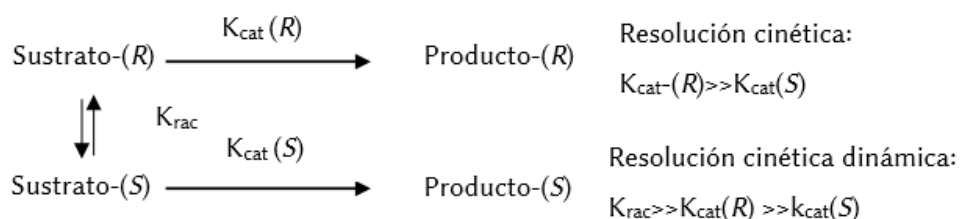


Fig 2. Principios básicos de una resolución cinética

Otro método de obtener aminas quirales por síntesis orgánica consiste en la síntesis asimétrica de estos compuestos. Los ejemplos de aminación reductiva asimétrica mediante el uso de síntesis química incluyen la catálisis organometálica utilizando diferentes complejos de metales de transición, como el rutenio, el rodio y el iridio, empleando ya sea H_2 ^{18,19} u otros agentes reductores.²⁰ Estas técnicas se han aplicado ampliamente y están bien establecidas en la industria.

Sin embargo, dada la necesidad de utilizar métodos menos nocivos con el medio, y el elevado costo asociado a los complejos de metales de transición utilizados como catalizadores para obtener las moléculas ópticamente puras, la biocatálisis entra en juego con un papel muy importante.

SÍNTESIS BIOCATALÍTICA DE AMINAS QUIRALES

Un gran número de aproximaciones han sido reportadas para la síntesis biocatalítica de aminas quirales.²¹ Una de las formas de obtener las mismas es mediante la resolución de los enantiómeros. En principio, se realizaron muchos intentos de síntesis centrándose principalmente en la obtención de las aminas racémicas mediante métodos químicos, y su posterior resolución mediada enzimáticamente utilizando lipasas, oxidasas, y transaminasas, obteniendo un rendimiento máximo de 50%.²² Recientemente, han surgido más enzimas con capacidad para resolver estos enantiómeros. Las investigaciones en esta área han planteado como estrategia mejorar los rendimientos de las reacciones, siendo una de las más eficaces la resolución cinética dinámica (RCD). Tomando como ejemplo las hidrolasas, se observa que la RCD se enfoca principalmente en dos estrategias. Una de ellas se basa en un mecanismo radical. Se realiza una abstracción homolítica reversible del átomo de hidrógeno α mediada por radicales tiilo, permitiendo la interconversión de los enantiómeros de amina a través de un intermediario radical centrado en el carbono α a la amina.²³ Un segundo enfoque es el de transformar asimétricamente el enantiómero no deseado al estado imina, el cual mediante una hidrogenación no selectiva vuelve a dar la mezcla racémica. Este tipo de hidrogenación también se utiliza como método dinámico para las monoaminoxidasas, entre otras. Un ejemplo de este método se observa en la figura 3, donde se muestra una RCD para la obtención de aminas quirales primarias, utilizando una lipasa para la resolución y un catalizador de níquel-Raney para la racemización, dando un rendimiento entre un 85 y 99%, con un ee de 98%.²²

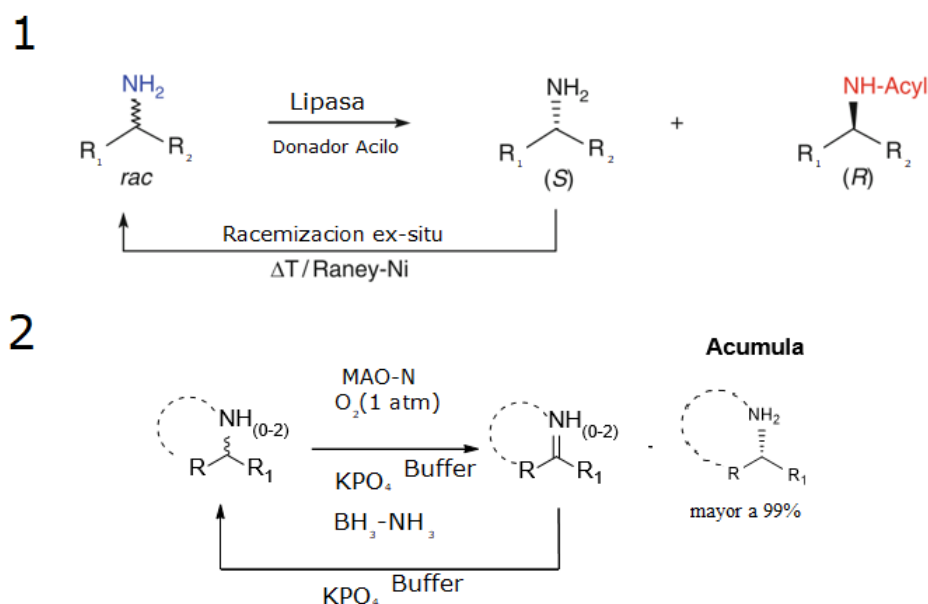


Fig 3. 1) RCD de aminas a partir de una lipasa y níquel Raney como catalizador.⁴ 2) RCD de Aminas a partir de una MAO

Más recientemente, el crecimiento exponencial de las bases de datos y los conocimientos generados sobre nuevas enzimas impulsó la búsqueda de nuevos biocatalizadores con características deseadas, por lo que la síntesis enzimática asimétrica de las aminas se estableció como una reacción de importancia para las industrias de síntesis. Entre las enzimas capaces de realizar esta reacción tenemos las transaminasas, amino deshidrogenasas, imino reductasas y aminosas reductivas, entre otras (Fig. 4).²⁴ Las transaminasas pueden realizar una resolución de enantiómeros al remover el grupo amina estereoselectivamente de uno de los productos de la mezcla racémica. Sin embargo, dicha reacción se puede realizar en ambos sentidos, por lo que esta enzima, al igual que las amino deshidrogenasas pueden realizar una síntesis asimétrica al agregar un grupo amino estereoselectivamente a una cetona, dando el producto de interés.

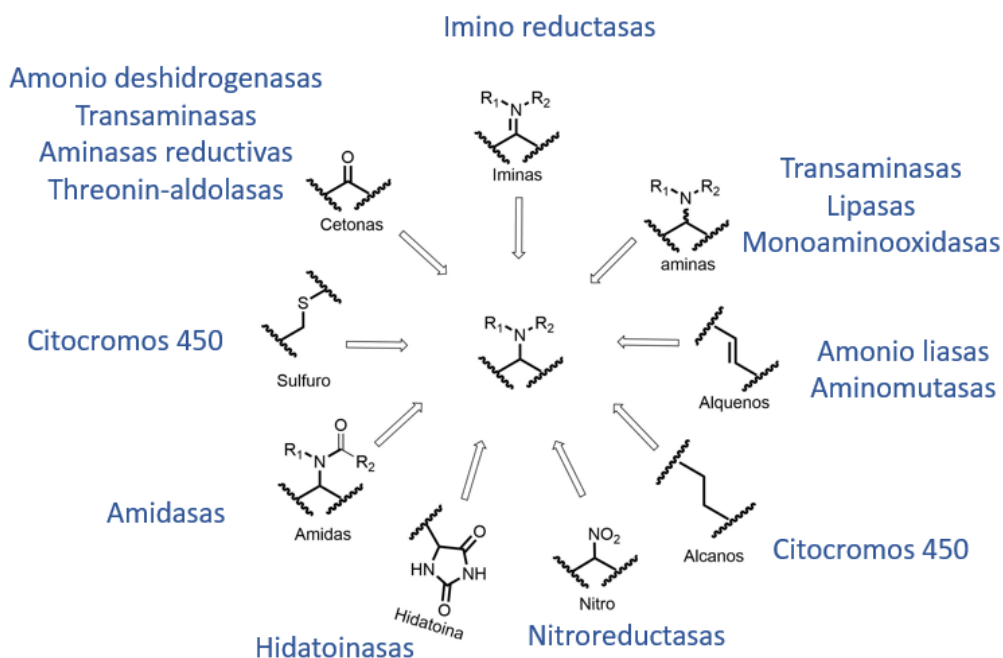


Fig 4. Enzimas capaces de sintetizar aminas enantiomericamente puras a partir de distintos sustratos.

IMINO REDUCTASAS

Las iminorreductasas han sido muy estudiadas para la síntesis de aminas quirales. Éstas son oxidorreductasas NAD(P)H-dependientes que catalizan la reducción asimétrica de iminas proquirales en sus correspondientes aminas.²⁵ Los estudios pioneros realizados por Mitsukura y Nagasawa en 2010 encontraron dos cepas de *Streptomyces* que contienen dos IREDs estereocomplementarias, capaces de reducir el sustrato modelo 2-metilpirrolina (Fig. 5).²⁶ Este descubrimiento, junto con la posterior purificación de estas enzimas y el descubrimiento de nuevas iminorreductasas promovieron el uso de éstas en biocatálisis. La ventaja en este tipo de enzima frente a las anteriormente descubiertas se da en la posibilidad de actuar también frente a iminas generando aminas secundarias y terciarias.

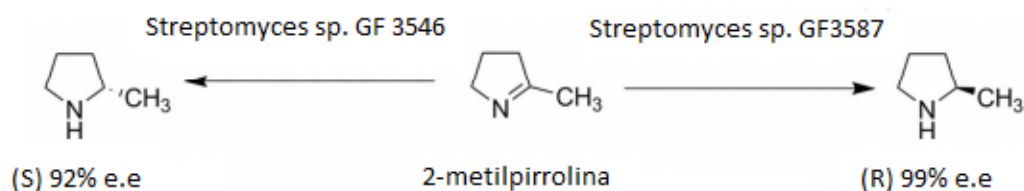


Fig 5. Reducción estereocomplementaria de 2-metilpirrolina por IRED utilizando distintas cepas de *Streptomyces* spp.²⁶

ESTRUCTURA Y MECANISMO DE LAS IRED

La primera estructura cristalina reportada de una IRED corresponde a una enzima proveniente de *Streptomyces kanamyceticus* (código Uniprot: Q1Eqe0). Esta enzima fue crucial para conocer la estructura y el sitio activo, así como para dilucidar el mecanismo de acción de estas. Las IREDs son proteínas diméricas y su sitio activo se forma con la participación de ambos monómeros. Éste se compone del extremo N-terminal de uno de ellos, el cual forma un pliegue de Rossman de unión a NAD(P)H, y un dominio helicoidal en el extremo C-terminal del otro monómero, unidos por una larga hélice alfa de 28 aminoácidos.²⁷ Para identificar los residuos funcionales relevantes se superpuso esta estructura con una β -hidroxiácido deshidrogenasa (β -HAD) estudiada y con muy alta similitud estructural. La β -HAD actúa catalizando la oxidación de alcoholes, y presenta una lisina en la posición 183 como residuo catalítico.²⁸ En dicha posición espacial se observó que la IRED posee un aspartato, cuya posición en la secuencia aminoacídica es la 187. La propuesta de mecanismo de las IRED plantea que el aspartato 187 protona el sustrato imina para la posterior transferencia del hidruro que se realiza entre el sustrato y el NAD(P)H (Fig. 6). Las mutaciones en el correspondiente aminoácido por alanina (D187A) y asparagina

(D187N) dieron como resultado una enzima inactiva. Otros estudios más recientes, realizan las mismas mutaciones en la (R)-IRED-Sr descubierta posteriormente dando una enzima no completamente inactiva, con una actividad mucho menor (entre 5% y 15% de la actividad inicial). En este último estudio, con únicamente información de (R)-IRED se procedió a analizar qué residuo se encontraba y cumplía la misma función en esa posición en las (S)-IRED hasta ese momento conocidas, encontrándose una tirosina. La mutación de tirosina por alanina (Y187A) y fenilalanina (Y187F) resulto nuevamente en una inactivación, por lo que se planteó el mismo mecanismo previamente descrito. A pesar de todo esto, el mecanismo sigue sin estar totalmente elucidado, dado que no se observó inactividad total en los mutantes de la (R)-IRED-Sr.²⁹

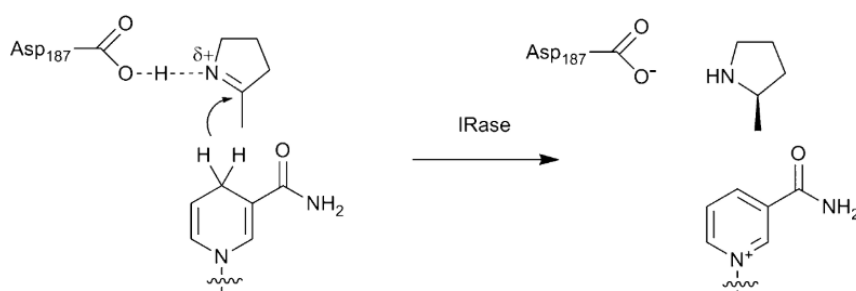


Fig 6: Mecanismo propuesto para la reducción de imina a amina en IREDs.

AMINACIÓN REDUCTIVA BIOCATALÍTICA

Otra aproximación, que ha comenzado a emplearse para la síntesis de aminas es la aminación reductiva, que permite acoplar directamente una amplia gama de cetonas a las aminas primarias y secundarias sin la necesidad de partir de una imina.^{25,30-33} Las octopina deshidrogenasas, que catalizan el acoplamiento de α -aminoácidos con α -cetoácidos, fueron foco en el 2014 para tratar de obtener esta actividad mediante ingeniería de proteínas, logrando obtener buenos resultados con algunos sustratos.³⁴

Otro tipo de proteínas también utilizado en aminación reductiva es el grupo de las aminoácido deshidrogenasas (AADH)³⁵. Sin embargo, su gran limitante es que actúan sobre alfa-cetoácidos y no cetonas. A su vez, por la parte de la amina donadora tienen también la limitación de solo aceptar amonio. Otro grupo de enzimas, las amino deshidrogenasas (AmDHs)^{36,37}, realizan la misma reacción que las anteriormente descritas, teniendo la ventaja de poder utilizar cetonas, aunque también presentan la limitación de aceptar sólo amonio. En un principio no se conocían proteínas con dicha actividad en la naturaleza, por lo que la primera enzima conocida de este tipo se obtuvo mediante diseño semi racional a partir de una leucina deshidrogenasa proveniente de *Bacillus stearothermophilus*

(BsLeuDh). La enzima, obtenida luego de 11 rondas de mutagénesis, dio un 92.5% de conversión y 98% e.e para la obtención del producto (R)-1,3-Dimetilbutilamina.³⁸

El gran potencial observado en las AmDHs diseñadas por ingeniería promovió la búsqueda de dicha actividad en la naturaleza. Esto dio como resultado una investigación en la cual se encontraron 6 amino deshidrogenasas nativas (nat-AmDH) con una identidad de secuencia entre 38 y 90 % entre sí, lo que habla de que éstas están relacionadas evolutivamente.³⁹ Recientemente, también se ha detectado dicha actividad en imino reductasas (IREDs).⁴⁰ Los primeros casos de estudio de esta actividad en dicho tipo de enzima dieron resultados muy bajos en cuanto a rendimiento, sin embargo, fueron los que abrieron paso al estudio de las mismas.⁴¹ Luego de este intento, el grupo de Hammer, quienes descubrieron la (R)-IREd-Sr, detectaron la actividad de aminación reductiva de ésta frente a benzaldehído con distintas aminas nucleófilas obteniendo resultados prometedores, sobre todo cuando se aumentaron la relación de equivalentes amina:cetona en una relación 50:1 llegando a 73% de conversión con metilamina.⁴² Posteriormente a este resultado, varios grupos de trabajo se volcaron en la búsqueda de actividad de aminación reductiva en imino reductasas, encontrando que muchas de estas enzimas son capaces de realizar esta reacción, abarcando un amplio espectro de cetonas y aminas.^{40,43,44} El 2017 fue un año de gran avance para el estudio de este tipo de enzimas, dado que el grupo de investigación liderado por el Prof. Turner descubrió esta actividad en enzimas pertenecientes a hongos del género *Asperigillus*. Estas enzimas tienen la capacidad de aceptar un amplio rango tanto de cetonas como de aminas, teniendo como principal ventaja frente a todos los tipos de enzima previamente encontrados la capacidad de realizar la reacción en concentraciones equimolares (Fig. 7).⁴⁵

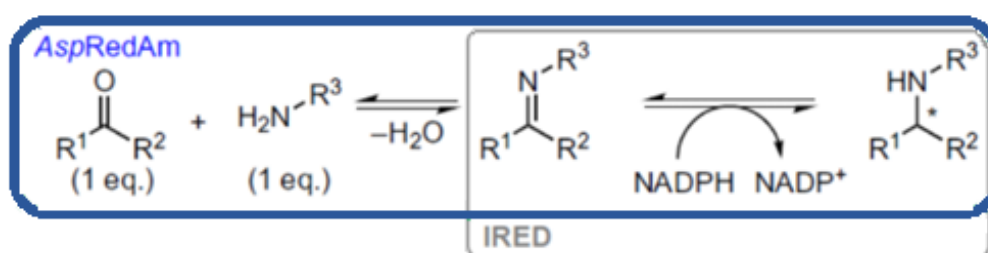


Fig 7. Reacción catalizada por aminosas reductivas e imino reductasas.

ESTRUCTURA Y MECANISMO PROPUESTO PARA IMINO REDUCTASAS CON ACTIVIDAD AMINASA REDUCTIVA

Mediante métodos de cristalización y modelado molecular, el grupo del Prof. N. Turner estudió la estructura de las RedAm, corroborando posteriormente mediante experimentos de mutagénesis los aminoácidos esenciales para que se lleve a cabo la reacción. Este estudio fue realizado sobre dos enzimas homólogas de la AspRedAm, provenientes de *Aspergillus terreus* (AtRedAm) y *Ajellomyces dermatitidis* (AdRedAm).⁴⁶ Estas enzimas presentan gran similitud estructural con las IREDs^{27,41,47,48}, encontrándose también en forma dimérica y formando el sitio activo entre un extremo N-terminal con pliegue de Rossman el extremo C-terminal del otro monómero con un dominio helicoidal. A través de la cristalización de la AtRedAm en distintas condiciones⁴⁹ se detectaron diferentes conformaciones de la proteína en presencia y ausencia de cofactores y sustratos. Las más distantes mostraron una rotación de aproximadamente 14° alrededor de los residuos 175 y 176, lo que indica que al igual que las IREDs presentan una forma abierta y una cerrada en su sitio activo. Los autores, presumen que la configuración cerrada le brinda un ambiente más hidrofóbico al ion iminio, permitiendo su estabilización. A su vez, dado que el espacio en el sitio activo es muy estrecho, favorecería la selectividad hacia uno de los enantiómeros. Este grupo logró cristalizar el complejo cuaternario (apoenzima, NADPH₄ (análogo inactivo del NADPH), cetona y amina) lo que permitió observar la interacción de los sustratos con la enzima. Además, demuestran que el residuo Y183 es necesario para la estabilización de la cetona por enlace de hidrógeno, y también lograron determinar que la densidad electrostática adyacente y la distancia del nitrógeno de la amina a los residuos D175 y N98 es consistente con que estos colaboran con la estabilización de la misma. Por último, describen la interacción de los sistemas pi de doble enlace de la alilamina con el bolsillo que forman los residuos L96 e I123.

Gracias a toda la información obtenida de la estructura cristalina, se ha logrado postular un mecanismo basado en el propuesto previamente para la enzima fenilalanina deshidrogenasa (PheDH).⁵⁰ El mecanismo plantea que al unirse el NADPH a la apoenzima ésta cambia su conformación al estado cerrado, y la cetona pasa a estabilizarse por enlace de hidrógeno directo o a través de una molécula de agua con la tirosina de la posición 183. Luego entra la amina al sitio activo siendo estabilizada por la asparagina 98 y el aspartato 175 de manera que ésta queda cargada positivamente. Posteriormente se da un ataque nucleofílico el cual permite la formación del ion iminio, el cual recibe el hidruro por parte del NADPH, pasando a NADP⁺ y liberando finalmente el producto de reacción (Fig. 8).

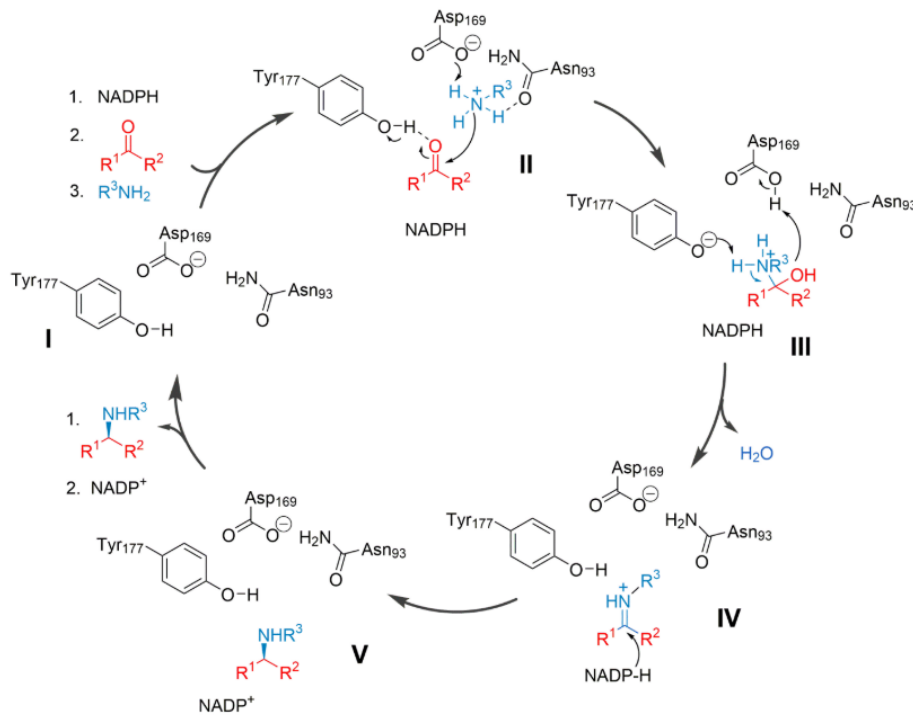


Fig 8. Mecanismo propuesto para la aminación reductiva a partir de las RedAm encontradas en hongos.⁴⁶

BIOINFORMÁTICA Y BIOCATÁLISIS

Es indispensable en la biología moderna el uso de herramientas cuantitativas, la mayoría de las investigaciones biológicas requieren la aplicación de algún tipo de herramienta matemática, estadística o computacional para ayudar a sintetizar datos e integrar varios tipos de información para responder a una pregunta biológica particular.⁵¹ A pesar de que existen muchas definiciones de bioinformática, tomaremos la definición propuesta por Luscombe y colaboradores: "la bioinformática implica la tecnología que utiliza computadoras para el almacenamiento, recuperación, manipulación y distribución de información relacionada con macromoléculas biológicas como el ADN, el ARN y las proteínas".⁵²

El surgimiento de la bioinformática fue facilitado por tres factores importantes a principios de la década de 1960. Una colección en expansión de secuencias de aminoácidos, imposible de resolver sin el poder numérico de las computadoras. La idea de que las macromoléculas transportan información unió también la computación con la biología molecular. Y, por último, el desarrollo de las computadoras de alta velocidad a partir de los programas de investigación de armas de la segunda guerra mundial permitió una mayor accesibilidad a éstas.⁵³ Si bien la bioinformática comenzó con la comparación de secuencias, ahora abarca una gran variedad de actividades. Es altamente interdisciplinaria,

requiriendo al menos conocimientos matemáticos, biológicos, físicos y químicos, y su implementación puede requerir además conocimientos de computación, ingeniería química, biotecnología, medicina, farmacología, etc.⁵⁴ Si realizamos una reseña temporal, probablemente el primer gran proyecto de bioinformática fue emprendido por Margaret Dayhoffin en 1965, desarrollando una base de datos de secuencias de proteínas llamada *Atlas of Protein Sequence and Structure*.⁵⁵ Posteriormente, se fueron desarrollando otros grandes proyectos. Por ejemplo, se estableció el *Protein Data Bank*,⁵⁶ se desarrolló el primer algoritmo de alineación de secuencias⁵⁷ y el primer algoritmo de predicción de la estructura de las proteínas,⁵⁸ entre otros. En la actualidad la gran capacidad de cómputo que adquirió el hardware informático, la fácil accesibilidad a éste, y el desarrollo de técnicas como lo son el modelado por homología,⁵⁹ el acoplamiento molecular y la dinámica molecular,⁶⁰ nos permite trabajar en la simulación de sistemas de miles de átomos como es el caso de los sistemas que involucran proteínas y ácidos nucleicos.^{61,62}

La bioinformática puede dividirse en dos grandes subáreas, una de ellas es el desarrollo de software bioinformático, el cual no solo engloba el desarrollo de software sino también de algoritmos, así como también el curado de las bases de datos. La otra es la aplicación de este, que se da en tres grandes campos, el análisis de la estructura, de la secuencia y de la función. Estos campos comúnmente se utilizan en conjunto para obtener respuestas a diferentes problemas.⁵¹ Dichas clasificaciones las podemos apreciar en la figura 9.



Clasificación de la bioinformática.⁵¹

En la década de los 90 la búsqueda de proteínas se basaba en la construcción de genotecas a partir de cepas o poblaciones de microorganismos.⁶³ Básicamente, el método consistía en la extracción de ADN de una o distintas poblaciones de microorganismos y su clonado en vectores para buscar una actividad deseada. Mediante secuenciación se identificaban los genes de interés de los clones positivos. Si bien, este método es muy efectivo, está restringido a enzimas que se pueden medir las actividades (por ejemplo, lipasas, oxidasas, etc.).⁶⁴ Además, el rastreo se realiza sin sobreexpresión de la proteína, lo que resulta en una sensibilidad limitada. Se estima que la biotecnología ha perdido hasta el 99% de los recursos microbianos existentes mediante el uso de técnicas de detección tradicionales.⁶⁵

Los métodos de secuenciado han avanzado notablemente en los últimos años, por lo que se cuenta con la secuencia de los genomas de numerosos microorganismos. La minería de estos permite la clonación directa de los genes objetivo y, por lo tanto, una expresión eficaz en una cepa huésped heteróloga adecuada. Además, contamos hoy en día con bases de datos abiertas conteniendo más de 200 millones de secuencias de diversos orígenes, lo que constituye una fuente de recursos muy interesante para la búsqueda y selección *in silico* de proteínas.

Existen diversas formas de obtener la información de los genomas obtenidos o de genomas disponibles en base de datos. Una muy interesante y útil, en el caso de la búsqueda de nuevas enzimas, es la estrategia basada en motivos conservados.⁶⁵ Los motivos conservados son patrones cortos y recurrentes en el ADN que se supone que tienen una función biológica.⁶⁶ En el caso de búsqueda de proteínas, se encuentran asociados a familias de éstas, pudiendo discriminar en el caso de encontrar enzimas de una misma familia que no posean dicha función/actividad y viceversa, aquellas que sí posean la función/actividad y están clasificadas en otra familia, o sin clasificar.

Para encontrar estos motivos, se parte de un alineamiento de secuencia múltiple con proteínas de las cuales se conoce su función, observando cuales son los aminoácidos conservados en ciertas posiciones. A partir del alineamiento múltiple de secuencias, se utilizan los Modelos Ocultos de Markov (HMM, por sus siglas en inglés),⁶⁷ el cual es un método estadístico que busca parámetros ocultos, a partir de parámetros observables. Este ofrece una gráfica con aminoácidos en cada posición en una escala de tamaños, de forma que las letras con mayor frecuencia en cada posición se pueden observar de mayor tamaño que aquellas que poseen menos frecuencia. Hoy en día son muy utilizados los programas donde se les da como archivo de entrada la alineación múltiple de secuencias, y el programa,

utilizando HMM da como salida una lista de motivos a utilizar. Un ejemplo de esto es la herramienta PRATT desarrollada por EMBL-EBI.⁶⁸

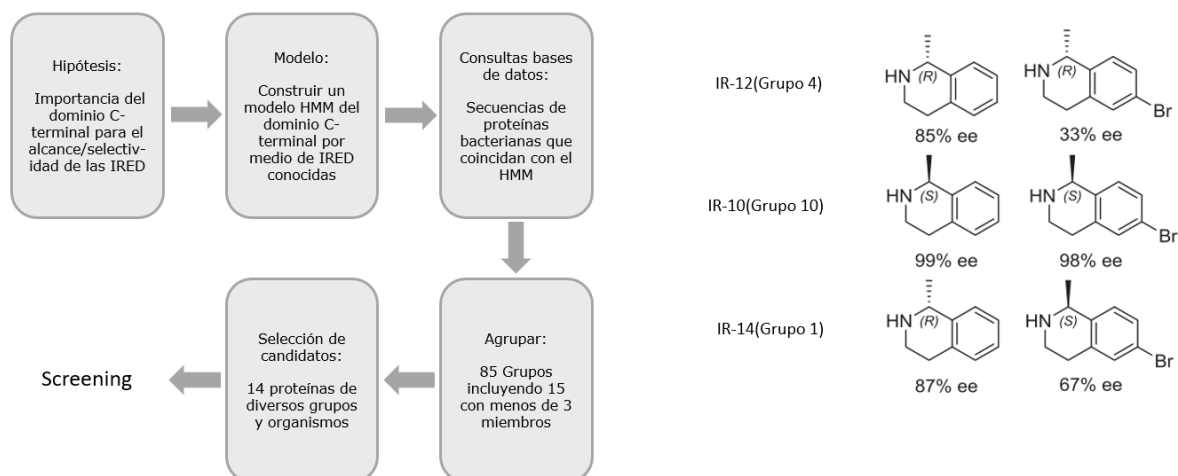


Fig 10. Ejemplos de IREDs descubiertas por Wetzl y colaboradores por motivos conservados en el dominio C-terminal de las mismas.

MODELADO POR HOMOLOGÍA

La estructura 3D de una proteína representa el patrón de plegado completo de la misma. Esta se utiliza de manera mucho más efectiva que la secuencia para comprender la función de la proteína, el mecanismo de acción y sus relaciones estructura-función.⁶⁹ Es importante el conocimiento de las estructuras tridimensionales para analizar la ubicación de los aminoácidos conservados, de los sitios activos y los aminoácidos presentes en ellos. Estos aminoácidos tendrán un efecto sobre la función de la proteína, y pueden explicar observaciones experimentales como especificidades de unión o propiedades antigénicas.⁷⁰

Existen dos métodos experimentales para obtener estructuras tridimensionales: Cristalografía de rayos X y espectroscopia de resonancia magnética nuclear (RMN). La cristalografía de rayos X es la técnica más comúnmente aplicada, lo que se demuestra por el número de estructuras resueltas: en abril de 2020 se depositaron en el banco de datos de proteínas alrededor de 145,000 estructuras cristalinas resueltas mediante rayos X, once veces más que el número de estructuras resueltas por RMN.⁷¹ Ambas técnicas son laboriosas, y requieren equipos costosos y procedimientos técnicos elaborados. Además, muchas proteínas no cristalizan, un requisito previo para la cristalografía de rayos X,⁷² o no son suficientemente solubles para los estudios de RMN.

Dado el costo y el tiempo involucrados en la determinación experimental de estructuras de proteínas, el desarrollo de métodos confiables para predecir la estructura tridimensional de una proteína a partir de su secuencia resulta de gran ayuda. Actualmente

los proyectos de secuenciación de genomas nos proporcionan día a día un número creciente información genética tanto de procariontas como de eucariotas. Se espera que la genómica estructural produzca una gran cantidad de estructuras de proteínas determinadas experimentalmente a largo plazo, resultando en una cobertura completa de todos los posibles plegados. Por lo tanto, refiriéndose a las estructuras de referencia adecuadas como base, será cada vez más factible generar modelos precisos para cualquier secuencia de proteína dada.

El modelado por homología es un método teórico para predecir la estructura 3D de una proteína utilizando información estructural de una proteína relacionada llamada modelo o *template*.⁵⁹ Los modelos se encuentran comúnmente mediante alineaciones simples que comparan la identidad de secuencia o también mediante HMM. Generalmente una mayor identidad de secuencia da como resultado una mayor calidad del modelo resultante. Como regla general, el porcentaje de similitud entre la proteína objetivo y la modelo debe ser mayor al 30%, siendo óptimo si es mayor a 40%.⁷³

El objetivo final, es crear una estructura 3D de proteínas que sea de calidad comparable a las estructuras de alta resolución realizadas por técnicas experimentales, como rayos X o RMN. Este es un proceso de varios pasos que comúnmente implica encontrar un modelo para realizar el alineamiento, el alineamiento en sí, la generación de la estructura principal de la proteína, el modelado de bucles y cadenas laterales, la optimización del modelo generado (mediante un refinado por medio de dinámica molecular) y la validación del modelo de homología para detectar errores estructurales e iteraciones en caso de que la estructura no sea lo suficientemente buena⁷⁴. Una vez culminado el modelo, se puede utilizar para diversos estudios posteriores, un ejemplo es el modelo generado de una proteasa del nuevo SARS-CoV2 para utilizar como blanco de fármacos (Fig. 11).

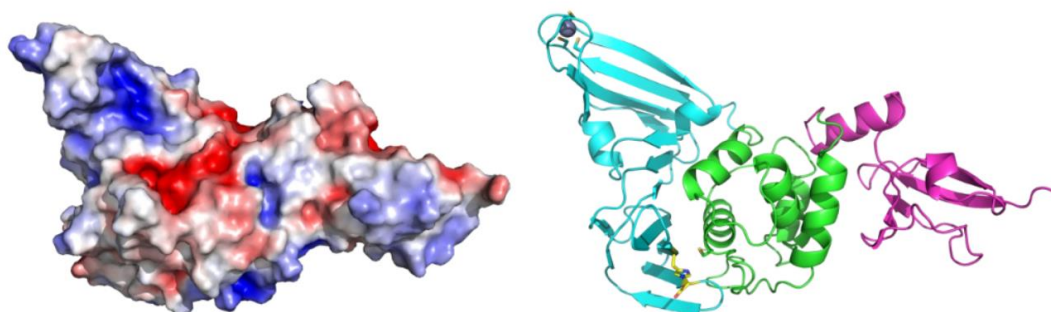


Fig 11. Dos visualizaciones diferentes del modelado por homología de una proteasa del nuevo SARS-CoV2 a partir de la proteasa 3CL del SARS.⁹⁸

En muchos campos científicos y tecnológicos existe la necesidad de diseñar o probar moléculas para objetivos específicos. Por ejemplo, en química, un investigador puede estar interesado en encontrar la proteína responsable de catalizar una reacción para acelerar el proceso; en medicina, los investigadores están interesados en encontrar nuevas moléculas o enfoques moleculares para curar enfermedades.⁷⁵ Dada esta necesidad, debido al creciente número de estructuras proteicas conocidas⁷⁶ y el desarrollo del modelado por homología, las técnicas computacionales capaces de predecir la interacción entre proteínas y proteínas, proteínas y ligandos, etc, son de gran importancia.

El docking molecular es una técnica computacional que ayuda a comprender las interacciones ligando-biomolécula. Puede utilizarse para el diseño y descubrimiento racional de fármacos, así como para el estudio mecanicista al colocar una molécula (ligando) en el sitio de unión de la biomolécula (receptor), donde se unirá principalmente de forma no covalente para formar un complejo estable de alta eficacia y especificidad.⁷⁷

La generación de una estructura receptor-ligando *in-silico* implica dos componentes principales: docking (o acoplamiento) y puntuación. El docking implica un muestreo conformacional y orientativo del ligando dentro de las limitaciones del sitio de unión al receptor. La función de puntuación selecciona la mejor pose (es decir, conformación, orientación y traslación del ligando) para una determinada molécula. Para tener éxito, el acoplamiento debe predecir con precisión la información disponible experimentalmente que se enfoca en dos puntos: la estructura del ligando (predicción de pose) y su propensión a la unión (predicción de afinidad).⁷⁸ En el área de biocatálisis, el docking molecular se utiliza para predecir posibles actividades enzimáticas de diversas moléculas, así como también su estereoselectividad dependiendo de la pose que adquiera la misma en el sitio activo. Un ejemplo, es el estudio de docking molecular realizado a la tolueno dioxigenasa con un sustrato específico, reportado de manera experimental previamente, para explicar la tendencia al exceso enantiomérico que presenta.⁷⁹ Otro ejemplo muy completo e interesante, que utiliza búsqueda, modelado y docking molecular, es el realizado por F. Spina y colaboradores. Realizaron una búsqueda de ene-reductasas en el hongo *Mucor circinelloides* logrando encontrar 10 genes de enzimas putativas de las mismas, las cuales luego se modelaron y estudiaron las diferencias estructurales de estas, viendo cual era más adecuada como biocatalizador mediante docking molecular.⁸⁰

Previo a la realización de esta tesis, nuestro equipo de trabajo se adentró en la búsqueda de imino reductasas a partir de genomas de *Streptomyces* nativos secuenciados por el grupo de la Dra. Maria Julia Pianzzola. Dada la novedad de este tipo de enzimas en ese momento, y la poca cantidad de imino reductasas caracterizadas, la búsqueda se realizó comparando la R y S IRED publicadas por el grupo de Nagasawa, con el fin de obtener zonas conservadas de las mismas.^{81,82} La búsqueda de estas zonas conservadas en los genomas nativos de *Streptomyces* permitió obtener 11 posibles candidatos, dos de los cuales fueron descartados luego de un BLAST con secuencias de referencia.⁸³

Las 9 potenciales IREDs de genomas nativos fueron analizadas mediante un algoritmo de identificación publicado por el grupo de Fademrecht.⁸⁴ Este autor establece los motivos de unión a NADPH y a sustrato de este tipo de enzimas, además de dos posiciones que indican la estereoselectividad de las mismas. En función de esto nuestro equipo de trabajo selecciono *in silico* 2 IRED que tienen el aminoácido que indica Fademrecht para cada tipo de estereoselectividad (IRED S e IRED R), con el objetivo de desarrollar biocatalizadores enantiocomplementarios.

OBJETIVOS

OBJETIVOS GENERALES

El objetivo general en esta tesis es la búsqueda, desarrollo y caracterización de nuevos biocatalizadores con capacidad de sintetizar aminas quirales. Más específicamente los tipos de enzima con los que se trabajará son las imino reductasas y aminosas reductivas, el primero partiendo de sustratos imina, mientras que el segundo parte de una cetona y una amina. Dado que se cuenta con dos hipotéticas IRED (IRED S e IRED R) encontradas previamente por el grupo de trabajo, se plantea también como objetivo la caracterización de éstas. Por otro lado, dado el reciente descubrimiento y auge de las aminosas reductivas, nos propusimos realizar la búsqueda de éstas en bases de datos utilizando métodos que permitan ampliar la diversidad de secuencias, pero conservando los aminoácidos del sitio activo esenciales para la actividad.

OBJETIVOS ESPECÍFICOS

- Clonado y expresión de la IRED S encontrada en el genoma de *Streptomyces scabiei* ST129 en el vector pKK223-3.
- Estudio computacional de la estructura tridimensional de la IRED R encontrada en el genoma de *Streptomyces sp.* ST1020 y su sitio activo.
- Estudio computacional de la actividad IRED y RedAm de la IRED R utilizando la metodología de Molecular Docking.
- Caracterización de la IRED R clonada en pET28b(+) y expresada en *E. coli*.
- Búsqueda de nuevas RedAm por el método basado en motivos conservados.
- Filtrado y comparación de los candidatos a RedAm encontrados con enzimas RedAm reportadas por literatura.

RESULTADOS

CLONADO, EXPRESIÓN Y CARACTERIZACIÓN DE LA IRED S ENCONTRADA EN EL GENOMA DE *STREPTOMYCES SCABIEI* ST129 EN EL VECTOR PKK223-3.

En este trabajo comenzamos a caracterizar las enzimas descubiertas previamente por el grupo de trabajo IRED S e IRED R. Como primer paso, se procedió al clonado de la enzima IRED S realizando la amplificación del gen ST129_05383 que codifica para ésta, a partir del ADN genómico de *S. scabiei* ST129 (Fig. 12). Esta amplificación se realizó por duplicado utilizando los primers FwRF y RvRF, los cuales poseen en sus extremos bases nucleotídicas para la posterior inserción en el vector pKK223-3.

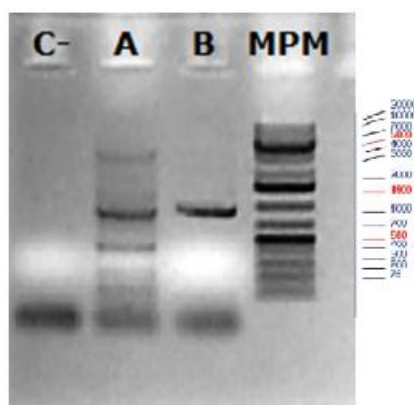


Fig 12. Gel de agarosa al 0.8% para verificar presencia del producto de PCR del gen *st129_05383* más la cola con los insertos para RF (megaprimer). 4 carriles. 1) Control negativo 2) Duplicado A 3) Duplicado B 4) Marcador de peso molecular *Gene Ruler 1 kb Plus DNA Ladder* (ThermoFisher)

El producto de PCR se purificó siguiendo el instructivo del kit *QIAquick PCR Purification Kit* (Qiagen). Posteriormente se realizó una estimación de la concentración en un gel de agarosa al 0.8% siguiendo los parámetros de intensidad de banda brindados por el marcador de peso molecular *GenRuler 1 Kb Plus DNA Ladder* (ThermoFisher), dando aproximadamente 110 ng/uL (Fig. 13).

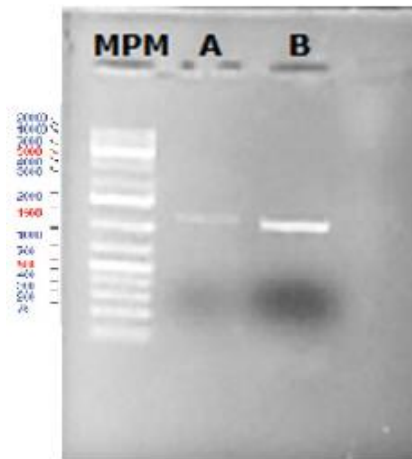


Fig 13. Gel de agarosa al 0.8% para la cuantificación de ADN del gen *ST129_05383* más la cola con los insertos para RF (megaprimer). 3 carriles. 1) Marcador de peso molecular Gene Ruler 1 kb Plus DNA Ladder 2) Duplicado A 3) Duplicado B

Para completar el clonado del inserto en el vector se realizó la RF PCR del megaprimer con el vector original. Luego de terminada la PCR, se eliminó el vector sin el inserto con la enzima de restricción DpnI, y se precipitó con ARNt.

TRANSFORMACIÓN EN *E. COLI*

El vector pKK223-3 posee el gen *lacZ* en la ubicación donde se realizó el clonado, solo aquellos plásmidos que hayan sido insertados con nuestro producto de PCR tendrán el gen *lacZ* interrumpido, y se visualizarán de color blanco, a diferencia de las que no, que se verán de color azul.⁸⁵ Para la transformación en *E. coli* se utilizó la cepa *JM109* y se realizó la preparación de las mismas siguiendo el protocolo para células electrocompetentes (BIO-RAD). Una vez obtenidas las células competentes, se procedió a realizar la electroporación de las mismas con el vector obtenido por RFC. Los posibles candidatos positivos se aislaron nuevamente en una Master Plate para su análisis.

MINIPREP PARA EVALUAR PRESENCIA DE PLÁSMIDOS CON INSERTO EN CANDIDATOS

Se realizaron minipreps siguiendo el protocolo descrito en *Materiales y métodos*, para los presuntos candidatos positivos obtenidos por la selección blanco/azul en busca de uno de ellos que presente el vector con el inserto (Fig. 14).

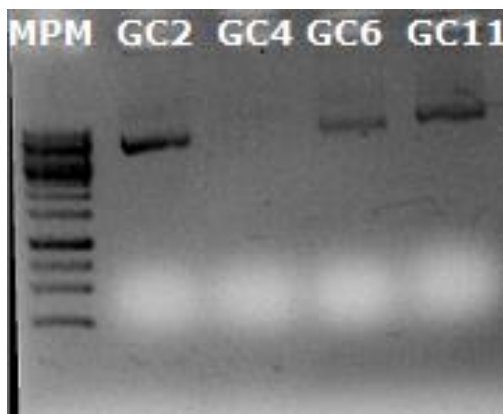


Fig 14. Gel de agarosa al 0.8% con corridas pertenecientes a las miniprep realizadas para observar si hay presencia de plásmido en las muestras GC2, GC4, GC6 y GC11

En la figura 14 se puede observar que excepto GC4, el resto de los candidatos presentaba el plásmido, aunque en este punto todavía no sabemos si este posee el inserto. Para determinar la existencia de inserto se realizó la caracterización de los plásmidos de los candidatos GC2, GC6 y GC11 mediante PCR y análisis con enzimas de restricción. La confirmación se realizó por PCR, por lo que si el plásmido contiene inserto debe obtenerse un producto de amplificación de alrededor de 950 pb. Para el análisis de restricción se utilizaron las enzimas BamHI y NcoI. En caso de no poseer el inserto, BamHI cortaría una vez dando una banda de 4600pb perteneciente el vector en estado lineal mientras que, si posee el inserto, cortaría una vez BamHI y otra NcoI, dando 2 bandas, una alrededor de los 800 pb y otra de 5300 pb (Fig. 15c)

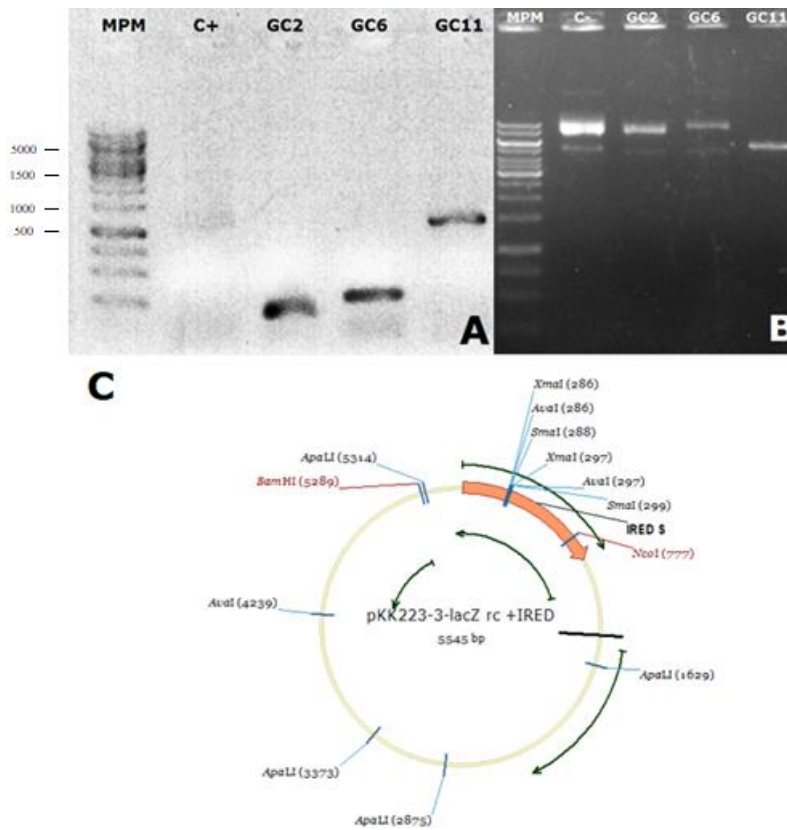


Fig 15. A) Gel de agarosa al 0.8% de la PCR del gen *st129_05383* para verificar la presencia del inserto B) Gel de agarosa al 0.8% del corte realizado por enzimas de restricción de los candidatos. C) Construcción del vector pKK223-3 con la IRED S con el mapa de enzimas de restricción.

En los carriles correspondientes a GC2 y GC6 se observan los tres estados del plásmido (superenrollado, lineal y circular). Esto indica que el corte con BamHI no se dio de forma correcta. En el carril correspondiente a GC11 se observa un corte con NcoI. Además, este plásmido es el único que dio amplificación por PCR, por lo que se concluye que es el único que contiene el inserto. Se volvió a realizar miniprep de GC11, obteniendo una concentración de ADN plasmídico purificado de 25 ng/uL para realizar verificación mediante secuenciación.

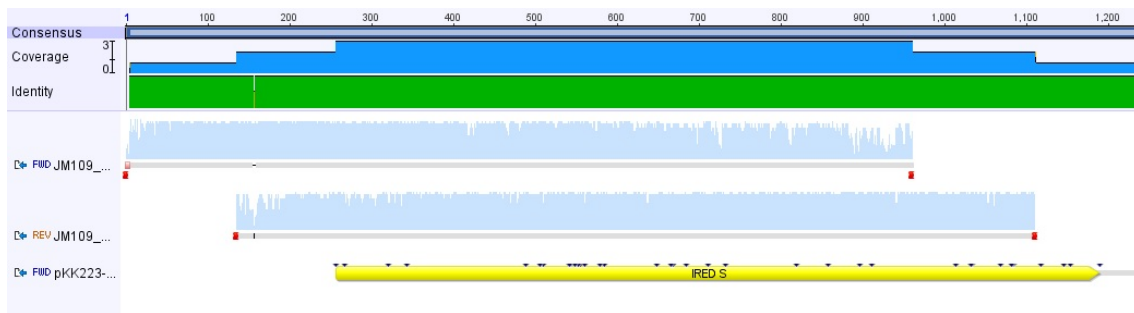


Fig 16. Resultados de la secuenciación alineados con la construcción del plásmido + inserto realizado Geneious.⁸⁶ En dicha figura se ve el secuenciado forward, el reverse, y por último la construcción teórica. En la región superior se ve la cobertura del ensamblado en azul, y la identidad en verde.

Una vez obtenidos los resultados de secuenciación, se ensamblaron las secuencias obtenidas, y se compararon con la construcción del vector con el inserto realizada *in silico* utilizando el programa VectorNTI.⁸⁷ Se puede ver que las secuencias obtenidas dan una identidad del 100% con el gen de la IRED S en la región que se logra analizar con las mismas, sin embargo, dada la baja calidad de secuenciado, no se pudo verificar la secuencia de los extremos, lo que no permitió confirmar el 100% de las bases del gen.

En paralelo con el análisis mediante secuenciado, se procedió a realizar biotransformaciones para analizar la actividad imino reductasa en la enzima. Se realizaron biotransformaciones con el sustrato 2-metilpirrolina y se procedió a hacer el análisis por cromatografía de capa fina y de gases. Por ninguno de los dos métodos se logró detectar presencia de producto, por lo que se decidió pasar la enzima al vector pET28b(+). Dado que el equipo se planteó este cambio de vector, se decidió no continuar con la confirmación del gen en el vector pKK223-3 mediante secuenciado, la cual se realizará sobre el nuevo vector una vez éste sea construido.

Modelado por homología de IRED R

El modelado por homología se ha convertido en una herramienta muy útil para la predicción de la estructura de la proteína cuando solo hay datos de secuencia disponibles. Para la caracterización computacional de la IRED R se realizó un estudio enfocado en su estructura tridimensional y su posible actividad. Para ello se llevó a cabo modelado por homología de las misma mediante el programa YASARA siguiendo el procedimiento descrito en *Materiales y métodos*.⁵⁹

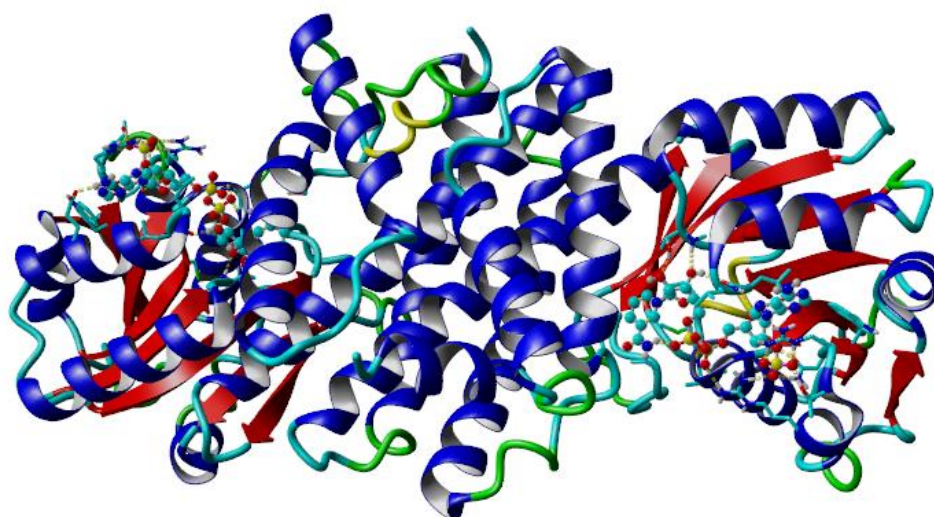


Fig 17. Modelado por Homología realizado para la IRED R.

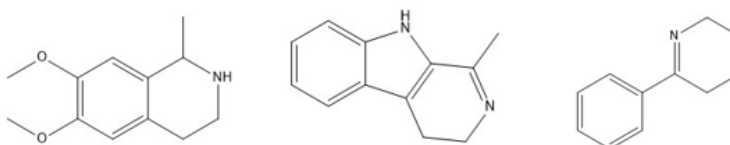
Posteriormente a la realización del modelo, se realizó un refinado de la estructura mediante una dinámica molecular de 500 ps de acuerdo al script md_refine brindado por el propio software YASARA el cual se basa en el artículo de Krieger y colaboradores.⁵⁹ Luego de aplicar el refinado, se procedió a seleccionar la estructura PDB con mejor Z-score (parámetro utilizado por YASARA para diferenciar la calidad de las estructuras) para posteriores análisis. Como se puede ver en la imagen, la IRED R, cumple con las condiciones estructurales vistas en las IRED y RedAm previamente descubiertas. Se puede ver que es un homodímero, el cual posee en su sitio activo el extremo N-terminal de uno de sus monómeros formando un pliegue de Rossman de unión a NAD(P)H y un dominio helicoidal en el extremo C-terminal del otro monómero unidos por una hélice alfa. Esta información nos da la pauta que a la hora de elegir posibles IRED o RedAm en los genomas de *Streptomyces* fue válido el método utilizado.

Molecular Docking de imino reductasas del grupo

Una vez obtenido el modelo de la estructura tridimensional de la IRED R se procedió a realizar estudios de docking. Como control positivo, y gracias a la colaboración del Dr. Schrittwieser y su equipo que nos brindó un set de 15 IREDs, elegimos la IRED D.⁸⁸ La elección se basó en 2 características, una de ellas fue que por literatura presentaba muy buena actividad IRED, la otra fue que luego de realizar un alineamiento, esta fue la que presentaba más aminoácidos conservados con aquellas enzimas que presentaron mayor actividad RedAm según el trabajo publicado por Frances y colaboradores.⁴³ La información que nos brindan los estudios de docking que realizamos se puede utilizar para sugerir la energía de unión, la energía libre y la estabilidad de los complejos ligando-receptor.

ACTIVIDAD IRED

Para analizar la actividad IRED se realizaron estudios de docking de tres sustratos: clorhidrato de 6,7-dimetoxi-1,2,3,4-tetrahidroisoquinolina(1); 1-Metil-1,2,3,4-tetrahidro-beta-carbolina(2) y 6-fenil-2,3,4,5-tetrahidropiridina(3). El método de docking utilizado fue VINA, se realizaron 50 corridas y se las agrupó con un RMSD máximo de 2.5 Å.⁸⁹



Parámetros	Enzima	6,7-dimetoxi-1,2,3,4-tetrahidroisoquinolina	1-metil-1,2,3,4-tetrahidro-beta-carbolina	6-fenil-2,3,4,5-tetrahidropiridina
Energía de unión(Kcal/mol)	IRED R	5.2	6.4	6.1
	IRED D	6.6	6.9	6.25
Distancia C-C(Å)	IRED R	4.0	4.4	3.6
	IRED D	4.0	3.7	3.8
Constante de disociación relativa	IRED R	11.2	2.2	1.4

Tabla 1. Parámetros para comparar la posible actividad IRED in silico utilizando la IRED D como control positivo.

Enzima	Aminoácidos que interaccionan con 6,7-Dimetoxi-1,2,3,4-tetrahidroisoquinolina												
IRED D		MET 244	-	-	ALA 247	LEU 75	ASP 76	THR 98	THR 99	MET 125	ASP 176	-	LEU 180
IRED R	VAL 230	MET 235	GLN 236	ASP 238	ALA 239	ASN 242	-	ASN 92	GLY 93	PRO 164	ASP 167	ILE 168	LEU 171

Enzima	Aminoácidos que interaccionan con 1-Metil-1,2,3,4-tetrahidro-beta-carbolina														
IRED D	VAL 230	SER 240	MET 235	GLN 236	ASP 238	ALA 239	-	ASN 242	ASN 92	GLY 93	-	ASP 167	LEU 171	MET 174	-
IRED R	TRP 214	SER 240	MET 244	GLN 245	-	ALA 247	GLY 248	THR 251	THR 99	ILE 124	MET 125	ASP 176	LEU 180	MET 183	TYR 184

Enzima	Aminoácidos que interaccionan con 6-Fenil-2,3,4,5-tetrahidropiridina													
IRED D	TRP 205	VAL 230	MET 235	GLN 236	ALA 239	ASN 92	ILE 117	ME T 118	ALA 119	ASP 167	LEU 171	MET 174	TYR 175	
IRED R	-	SER 240	MET 244	GLN 245	GLY 248	THR 99	-	ME T 125	-	ASP 176	LEU 180	MET 183	TYR 184	

Tabla 2. Aminoácidos que interaccionan con los sustratos en el sitio activo para comparar la posible actividad IRED in silico.

Como se puede observar en la tabla 1, los parámetros considerados para evaluar la posible actividad IRED en nuestra enzima de estudio fueron múltiples. La energía de unión del sustrato con la proteína (1): la distancia entre el carbono donador de hidruro perteneciente al NADPH y el carbono adyacente al nitrógeno que forma la imina (2): la constante de disociación relativa al control positivo (3). Además, la tabla 2 muestra los aminoácidos que interaccionan con cada sustrato en ambas enzimas brindados por el programa YASARA (4). La energía de unión se define como la energía mínima necesaria para separar el sustrato de la enzima, por lo que, a mayor energía se espera una mayor estabilidad del complejo. La distancia promedio entre el carbono donador del hidruro y el carbono que acepta el mismo oscila en 3.5 y 4 Å, por esto se espera que sólo sean válidos los valores incluidos en este rango. Por último, la constante de disociación es la relación entre el producto de la concentración de enzima y sustrato libres, respecto a la concentración de complejo enzima-sustrato, por lo que se espera que mientras más chica sea la misma, el equilibrio se encontrará más desplazado hacia el complejo.

6,7-DIMETOXI-1,2,3,4-TETRAHIDROISOQUINOLINA

Al comparar la energía de unión de este sustrato con nuestra enzima en estudio respecto al control positivo, encontramos una diferencia mayor al 20% a favor del control, lo cual indica una posible baja actividad, o una actividad menor al control con dicho sustrato. Si analizamos la distancia C-C podemos ver que el mejor resultado de docking en ambos casos muestra la misma distancia. La constante de disociación relativa al control es un orden de magnitud mayor, esta es la mayor diferencia encontrada para todos los estudios de docking realizados, por lo que más allá de la distancia C-C y los aminoácidos que interactúan en el sitio activo con el sustrato no es esperable observar actividad con este sustrato, o esta sería muy baja.

1-METIL-1,2,3,4-TETRAHIDRO-BETA-CARBOLINE

En este caso cuando analizamos la energía de unión vemos una diferencia menor al 10% entre ambas enzimas, de igual manera la constante de disociación se mantiene en el mismo orden con una diferencia de 2.2 veces. Los aminoácidos que interactúan con el sustrato en cada caso lo hacen de manera similar dado que 9 de 14 son los mismos aminoácidos o de similares características respecto al control positivo, algo también muy bueno. Sin embargo, la distancia C-C a de 4.4 Å, es bastante mayor a la descrita en literatura y a la obtenida en los estudios de docking realizados para el control positivo. Por este último motivo esperamos que la enzima en estudio presente muy baja actividad con este sustrato.

6-FENIL-2,3,4,5-TETRAHIDROPIRIDINA

En cuanto al sustrato 3, la energía de unión observada en ambos casos no es significativamente diferente. A la hora de evaluar la distancia C-C vemos que entra en el rango de distancia aceptable para que se dé la reacción efectivamente. La constante de disociación se mantiene en el mismo orden siendo 1.4 veces mayor que la del control positivo, y por último 10 de los 13 aminoácidos del sitio activo que interaccionan con el sustrato en el control positivo son comunes en ambas enzimas (Fig. 18). El análisis de todo el conjunto de datos obtenidos, nos da indicio que la IRED R presente actividad frente a este sustrato.

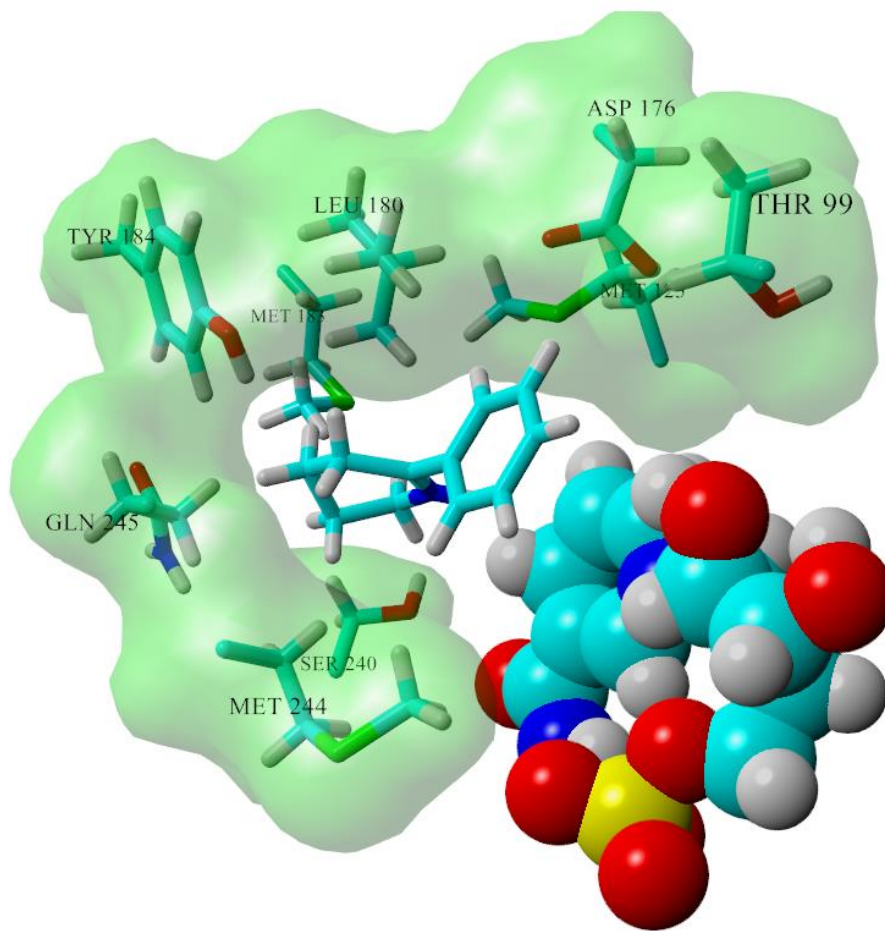


Fig 18. *Imagen del ligando 6-fenil-2,3,4,5-tetrahidropiridina con los aminoácidos que interaccionan en el sitio activo y una región del cofactor NADPH*

ACTIVIDAD REDAM

Luego de los diferentes dockings realizados para evaluar la actividad imino reductasa, se procedió a realizar estudios de docking en busca de actividad RedAm de acuerdo al mecanismo presentado por Turner y colaboradores.⁴⁶

Como sustrato para los docking, se utilizó propargilamina como dador de amino, y ciclohexanona y fenilacetona como cetonas. Dado que esta actividad (RedAm) se compone por la interacción en el sitio activo de dos sustratos, no nos fue posible comparar energías de unión o constantes de disociación. Por esto, decidimos evaluar *in silico* esta actividad en base a dos análisis correlacionados de docking molecular. En principio, realizamos el docking de la cetona, buscamos que esta tuviese una posición adecuada para formar enlace de hidrógeno con la tirosina reportada por literatura y fijamos la cetona. Luego con la cetona fija realizamos el docking con la amina y volvimos a ver si esta tenía alguna posible posición de acuerdo con la literatura interaccionando con un aspartato específico. Si ambos sustratos lograban estas posiciones consideramos que puede haber actividad con estos. Además, dado que experimentalmente detectamos actividad en algunas enzimas cuyos resultados de docking de los sustratos por separado no mostraron resultados favorables, y considerando que en literatura utilizan relaciones estequiométricas donde se favorece la formación espontánea de la imina (50:1 amina:cetona), también realizamos el docking con la imina formada por reacción de ambos sustratos para evaluar si esta se posiciona correctamente para que ocurra la transferencia del hidruro.

PROPARGILAMINA CON CICLOHEXANONA

El primer docking molecular realizado fue con la IRED R, al igual que con la IRED D, logramos encontrar una posición muy favorable en la que el oxígeno de la cetona queda a 2,4 Å de distancia del hidrogeno del grupo hidroxilo de la tirosina 184, distancia a la cual es muy probable que se genere un enlace de hidrogeno. Luego de esto, fijando la cetona, se realizó el docking de la propargilamina, la cual también dio una pose muy favorable, a una distancia de 2.9 Å para la formación del enlace de hidrogeno con el aspartato 176. Dado los resultados obtenidos, llegamos a la conclusión de que es probable que ocurra dicha reacción con una buena conversion.

Propargilamina con fenilacetona

En este docking, procediendo de la manera previamente explicada, no logramos encontrar una posición favorable en la cual la fenilacetona logre realizar un enlace de hidrogeno con la tirosina 184, dado que las distancias entre el oxígeno de la cetona y el hidrogeno del grupo hidroxilo de la tirosina superan los 5 Å, a diferencia de la IRED D, la cual tiene una distancia menor a 4 Å pudiendo formar un enlace de hidrogeno. Por esto y dado el carácter secuencial de la unión de los sustratos postulamos que dicha reacción no se da en la IRED R partiendo de estos 2 sustratos.

N-propargilciclohexilamina

Cuando estudiamos el intermediario imina en ambas enzimas, logramos ver que se encuentran poses favorables para que se dé la reacción en estas, por lo que, si estudiamos el conjunto de dockings realizados, con la cetona y la amina por separado, y con el intermediario ya formado podemos considerar que es factible que se logre obtener la actividad.

NORDEPRENIL

A diferencia del otro sustrato estudiado, cuando realizamos el docking con el intermediario imina que forma la fenilacetona con la propargilamina, en la IRED R no logramos encontrar ninguna pose favorable para que se dé la reacción, mientras que en la IRED D la pose con mayor energía de unión es favorable para que se dé la transferencia del hidruro, encontrándose una distancia de 3.8 Å C-C. Teniendo en cuenta estos resultados, esperamos que la IRED R no posea actividad RedAm con estos sustratos, ni con el intermediario directo. A diferencia de la IRED D que esperamos que posea dicha actividad. Con los datos obtenidos computacionalmente se planteó la estrategia experimental para poder comparar resultados.

CARACTERIZACIÓN DE LA IRED R CLONADA EN PET28B(+) Y EXPRESADA EN *E. COLI*

El gen de la IRED R se obtuvo mediante síntesis a través del servicio de la empresa GenScript.⁹⁰ Este gen se adquirió clonado en el vector pET28b(+), transformándose la cepa *E. coli* BL21 (DE3) para obtener la cepa productora de la IRED R. Para comprobar los resultados obtenidos *in silico* mediante el modelado por homología y docking molecular de la IRED R, se procedió a realizar biotransformaciones de acuerdo con cómo se indica en *Materiales y métodos*. Sabiendo que las IRED son NADPH dependiente, probamos dos métodos de regeneración de NADPH diferentes para que este no limite el rendimiento de la reacción. Uno de ellos fue el agregado de glucosa para que a partir del metabolismo de *E. coli* se regenere el NADPH que se consume por la reacción. El otro método se basa en la coexpresión de una Alcohol Deshidrogenasa (Ib-ADH, brindada también por el Dr. Schrittwieser), y el uso de isopropanol como co-sustrato.⁸⁸ La oxidación de isopropanol produce la reducción de NADP⁺ a NADPH, regenerando este cofactor. Los sustratos utilizados para la actividad IRED fueron los 3 utilizados en el diseño *in silico*, mientras que para la actividad RedAm se utilizó ciclohexanona y fenilacetona como sustratos cetona, y propargilamina como sustrato amina. Los resultados obtenidos se muestran en la siguiente tabla.

Enzima	Reg. NADPH	1	2	3	4	5
IRED R	Glucosa	0	0	1-2	99	0
	ADH + isopropanol	0	0	13		
IRED D	Glucosa	2	16	16	99	67
	ADH + isopropanol	-	68	99		

1. 6,7-dimetoxi-1-metil-3,4-dihidroisoquinolina
2. 1-methyl-3,4-dihydro-β-carboline
3. 6-fenil-2,3,4,5-tetrahidropiridina
4. Propargilamina + Ciclohexanona
5. Propargilamina + Fenilacetona

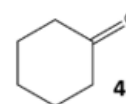
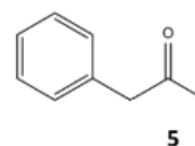
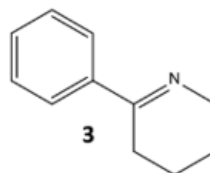
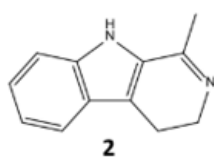
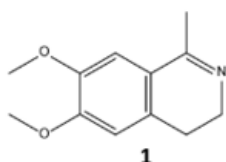


Tabla 2. Resultados obtenidos experimentalmente de la actividad IRED (sustratos 1,2 y 3) y RedAm (4 y 5) para la IRED R utilizando IRED D como Control positivo.

Efectivamente las enzimas se comportaron como esperábamos según los resultados previamente obtenidos en el estudio de docking molecular. La actividad IRED de la IRED R con los sustratos 1 y 2 fue nula, en estos casos los dockings me indicaban una energía del 21% y 7% menor respectivamente comparando con el control positivo. Con el sustrato 3 se ve baja actividad, principalmente con el sistema de regeneración de NADPH utilizando ADH. En este caso la energía fue tan solo de un 2% menor al control positivo, aunque se puede observar menos cantidad de interacciones con los aminoácidos del sitio activo. Por otro lado, la actividad RedAm frente a ciclohexanona con propargilamina fue muy buena, logrando transformar los sustratos en la amina secundaria, mientras que con fenilacetona no logramos observar actividad. Por esto, podemos concluir que la herramienta de docking molecular puede ser muy válida si se tienen los respectivos controles, y si se estudian los suficientes parámetros. Sin embargo, las diferencias de energía de unión en algunos casos pueden ser muy pequeñas, por lo que esto nos indican la necesidad de validar de manera experimental los resultados obtenidos por la misma.

FILTRADO Y COMPARACIÓN DE LOS CANDIDATOS A REDAM ENCONTRADOS CON ENZIMAS REDAM REPORTADAS POR LITERATURA

Con las secuencias obtenidas se realizó un árbol filogenético utilizando el programa MEGA X para conocer la distancia evolutiva entre nuestros posibles 28 candidatos, las 9 imino reductasas encontradas por el grupo, y las 13 obtenidas de literatura.⁹¹ El análisis de este árbol nos ayudará a seleccionar secuencias candidatas interesantes para ser estudiadas experimentalmente en el futuro.

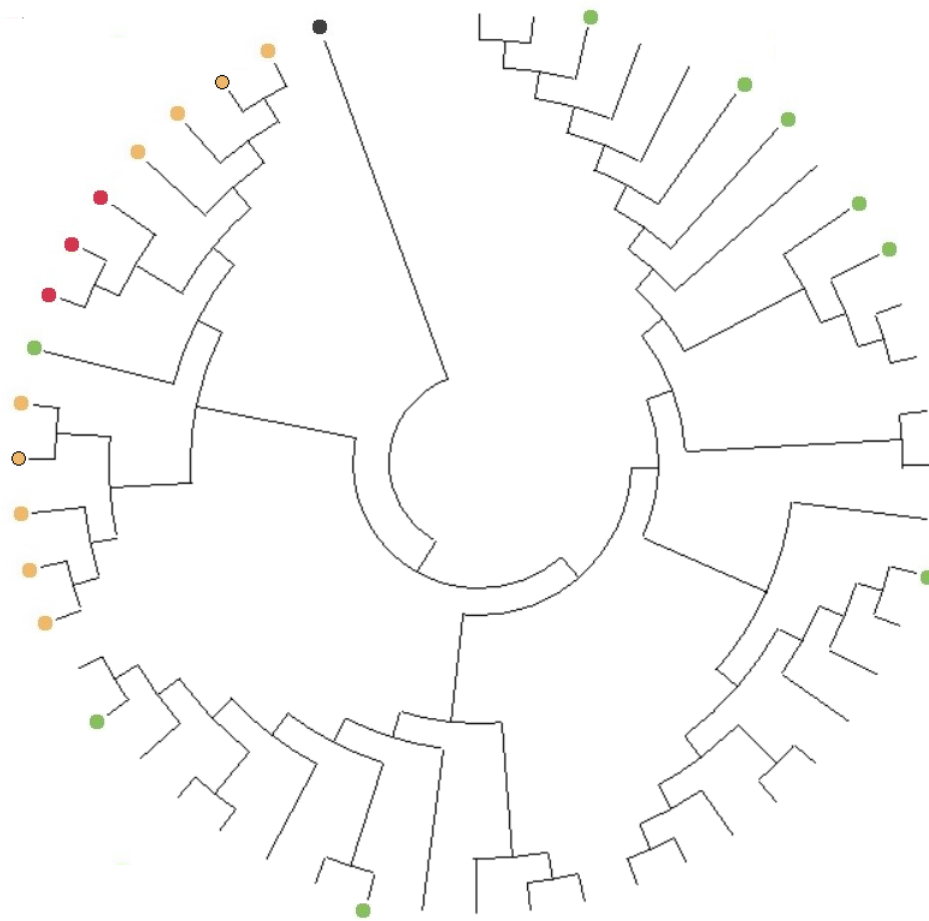


Fig 21. *Árbol filogenético de las RedAm reportadas por Turner en rojo, las RedAm reportadas por Frances en verde, las 9 IREDs encontradas por nuestro grupo de trabajo en amarillo, con un marco negro aquellas que fueron clonadas y, por último, los candidatos encontrados en la búsqueda se encuentran sin referencia. El Outgroup utilizado es una 3-hidroxitirato deshidrogenasa marcada en negro.*

La búsqueda de nuevas enzimas utilizando motivos conservados nos permitió aprovechar la información sobre las RedAm reportadas previamente para obtener una gran diversidad de posibles nuevas enzimas de este tipo, que se puede observar en el árbol filogenético. Algo importante que este árbol nos permite comparar son las IRED encontradas previamente por nuestro grupo de trabajo en los genomas de cepas de *Streptomyces* nativas. Éstas se encuentran estrechamente relacionadas, lo cual tiene mucho sentido dado que provienen del mismo género. A su vez tienen una gran similitud en cuanto a la totalidad de la secuencia con las reportadas por *Turner y colaboradores*,⁴⁵ aunque no así en su sitio activo, repercutiendo en su actividad RedAm como pudimos ver en los ensayos con las dos enzimas estereocomplementarias clonadas por nuestro grupo. En cambio, nuestro método de búsqueda de nuevas RedAm, nos permitió obtener una gran diversidad de secuencias, gracias a que nos basamos en motivos conservados y no en secuencias completas. Éstas ocupan un amplio rango evolutivo, pero todas ellas cumplen con la estructura común de las IRED-RedAm, principalmente el sitio de unión a NADPH y la similitud del sitio activo, manteniendo conservados aquellos residuos esenciales para la actividad RedAm.

CONCLUSIONES Y PERSPECTIVAS

Se efectuó el modelado por homología de la IRED R clonada proveniente de las nueve potenciales IRED encontradas en *Streptomyces spp.* Luego de obtener el modelo, se realizó el estudio de docking molecular para estimar su actividad con diferentes sustratos previo a la prueba en el laboratorio. El estudio computacional de IRED R logró predecir parte de la actividad obtenida tanto IRED como RedAm.

Se logró clonar exitosamente la IRED S en el vector pKK223-3 aunque no se observó actividad de esta en dicho vector. Se realizó el estudio de actividad IRED y RedAm de la IRED R clonada en pET28b(+) con diferentes sustratos obteniendo baja actividad IRED, pero buena actividad RedAm la cual actualmente está siendo aplicada a la síntesis de precursores de fármacos. Se realizó una búsqueda basada en motivos conservados de enzimas con actividad RedAm a partir de enzimas previamente reportada por literatura. Tras la construcción de un árbol filogenético se logró ver que obtuvimos un gran abanico de candidatos evolutivamente distantes, lo cual era uno de los objetivos de la búsqueda. A su vez se compararon las 9 enzimas encontradas por el equipo de trabajo provenientes de *Streptomyces spp.*, cuyo genoma fue secuenciado por el grupo de la Dra. Ma. Julia Pianzola. Estas últimas están todas estrechamente relacionadas entre sí, algo lógico proviniendo del mismo género.

Como perspectivas a futuro proponemos realizar un estudio *in silico* mediante el modelado por homología y docking molecular de todos los candidatos obtenidos en la búsqueda por patrones. Aquellos que muestren resultados de docking prometedores con algunos sustratos para actividad IRED o RedAm, se utilizarán para estudios de dinámica molecular de los mismos que permitan conocer la estabilidad del complejo enzima-sustrato. Finalmente, para aquella enzima con la cual tengamos mejores resultados se clonará en el vector de expresión en bacterias pET28b(+) y se realizarán pruebas de actividad *in vitro*.

MATERIALES Y MÉTODOS

MÉTODOS MICROBIOLÓGICOS

La triptona y el extracto de levadura utilizados para la preparación de medios de cultivo se compraron de Oxoid. El agar, la glucosa y la α -lactosa se adquirieron de Sigma. Las sales para los medios de cultivo y el SDS se obtuvieron de Sigma, Droguería Paysandú y USbiological. Y los antibióticos se adquirieron de Amresco y AppliChem.

Para la manipulación de las cepas se utilizaron las técnicas estándar de microbiología. El medio utilizado para cultivo de las cepas de *E. coli* en placa fue Luria Bertani (LB) con una adición de agar bacteriológico al 1.5%, los medios líquidos utilizados fueron LB y medio de autoinducción (AIM). A las cepas que contienen plásmidos con resistencia a antibióticos se les añadió ampicilina 200 $\mu\text{g}/\text{mL}$ y/o kanamicina 50 $\mu\text{g}/\text{mL}$. El mantenimiento a corto plazo de éstas se realizó en placas de LB selladas con Parafilm y almacenadas a 4°C. El mantenimiento a largo plazo se realizó en caldo LB y glicerol al 15% almacenadas en freezer a -70°C.

Cepa	Genotipo
<i>E. coli</i> JM109	<i>endA1, recA1, gyrA96, thi, hsdR17</i> (r_k^- , m_k^+), <i>relA1, supE44, Δ(lac-proAB)</i> , [F' <i>traD36, proAB, laqI^qZΔM15</i>]
<i>E. coli</i> BL21(DE3)	F- <i>ompT [lon] hsdS_B(r_B-m_B-;</i> una cepa de <i>E. coli</i> B) con DE3, un profago que posee el gen de la ARN polimerasa de T7
<i>E. coli</i> TOP10	[F- <i>mcrAΔ(mrr.hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 recAaraD139 Δ(ara-leu)7697 galUgaKrpsLendA1 nupG</i>]

Tabla 3. Cepas de *E. coli* utilizadas

A partir del top 10 de imino reductasas con mejor actividad aminsasa reductiva de Frances y colaboradores, y las 3 aminsas reductivas de hongos descubiertas por Turner *et al.*, se realizó una alineación de secuencias múltiples en MEGA X utilizando como método de alineamiento MUSCLE con su configuración preestablecida (Fig. 22).^{91,94} De allí se obtuvo el alineamiento múltiple que se utilizó como entrada para la aplicación PRATT. Dicha aplicación nos brindó como mejor patrón YxDGAI[ML]AxPx₂IG.

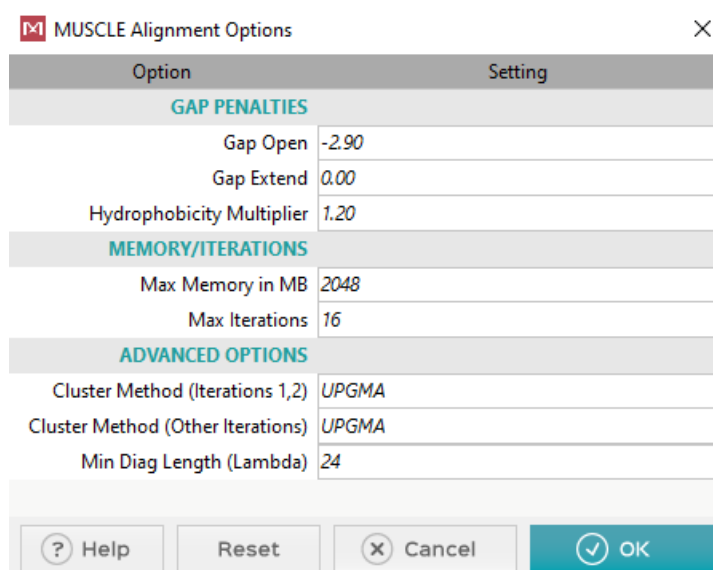


Fig 22. Parámetros utilizados para el alineamiento con MUSCLE

Se hizo una búsqueda basada en motivos conservados en ScanProsite, la cual toma como base de datos UniprotKB (SwissProt y TrEMBL).^{92,93} Como motivos conservados se utilizaron el obtenido por el software PRATT 2.1, y el obtenido por literatura de Fademretch y colaboradores.⁸⁴ Los resultados se filtraron siguiendo el criterio aplicado por Frances y colaboradores. para enzimas con actividad Redam basado en residuos homólogos en el sitio activo.⁴³

Ranking	IRED	93	169	177	210	239	240
45.9	IR48	S	D	W	W	A	H
40.9	IR47	T	D	W	M	V	H
39.9	IR50	S	D	W	W	P	H
39.8	IR44	S	D	W	W	A	H
38.4	IR77	S	D	W	W	P	H
37.4	IR56	S	D	W	W	P	H
34.7	IR61	S	D	W	W	P	H
23.5	IR66	S	D	W	L	A	H
22.6	IR91	S	Y	L	M	M	M
21.4	IR58	S	D	W	L	A	H

Fig 23. Ranking top 10 IREDs con actividad RedAm. Se muestran los residuos homólogos a los considerados esenciales en AspRedAm

Para conocer las distancias evolutivas de las secuencias se realizó un árbol filogenético en MEGA X, buscando primeramente el mejor modelo para proteínas teniendo en cuenta que será utilizado el método estadístico de máxima verosimilitud.⁹¹ Para el mismo se siguieron los siguientes parámetros.

Phylogeny Reconstruction	
Option	Setting
ANALYSIS	
Statistical Method →	Maximum Likelihood
PHYLOGENY TEST	
Test of Phylogeny →	Bootstrap method
No. of Bootstrap Replications →	1000
SUBSTITUTION MODEL	
Substitutions Type →	Amino acid
Model/Method →	WAG model
RATES AND PATTERNS	
Rates among Sites →	Gamma Distributed (G)
No of Discrete Gamma Categories →	5
DATA SUBSET TO USE	
Gaps/Missing Data Treatment →	Complete deletion
Site Coverage Cutoff (%) →	Not Applicable
TREE INFERENCE OPTIONS	
ML Heuristic Method →	Nearest-Neighbor-Interchange (NNI)
Initial Tree for ML →	Make initial tree automatically (Default - NJ/BioNJ)
Initial Tree File →	Not Applicable
Branch Swap Filter →	None
SYSTEM RESOURCE USAGE	
Number of Threads →	16

Fig 24. Parámetros utilizados para realizar el árbol filogenético

Luego de esto, con estas secuencias se procedió a eliminar todas aquellas que estuviesen repetidas y realizar el modelado por homología de las mismas, y de las poseídas por el grupo de trabajo. El modelado por homología se llevó a cabo en el programa YASARA siguiendo los siguientes parámetros.⁵⁹

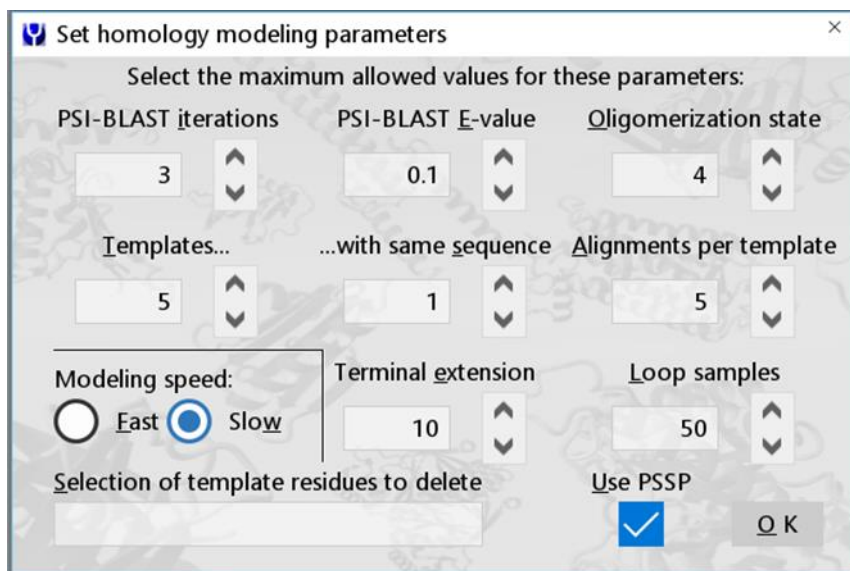


Fig 25. Parámetros utilizados para realizar el modelado por homología en YASARA.

Posterior a tener el modelo por homología hecho, se realizó un refinamiento de la estructura mediante una dinámica molecular de 500 ps en una caja de agua, tomando fotos a intervalos de 25 ps, a un pH=7.4, 25°C y el campo de fuerza YAMBER3 de acuerdo al script md_refine brindado por el propio software YASARA el cual se basa en el artículo de Krieger y colaboradores.⁵⁹ De cada modelo realizado se procedió a obtener la foto con mejor Z-score para posteriores análisis. Teniendo en cuenta las enzimas propias que ya posee el grupo de trabajo se procedió a realizar docking molecular de las mismas con diferentes sustratos que cuenta con stock el grupo de trabajo, El método de docking fue VINA, se realizaron 50 corridas y se las agrupó con un RMSD máximo de 2.5 Å.⁸⁹

RESTRICTION FREE CLONING

PCR del gen con Phusion.^{95,96}:

1. Se prepararon las reacciones de PCR según la siguiente tabla de componentes y volúmenes.

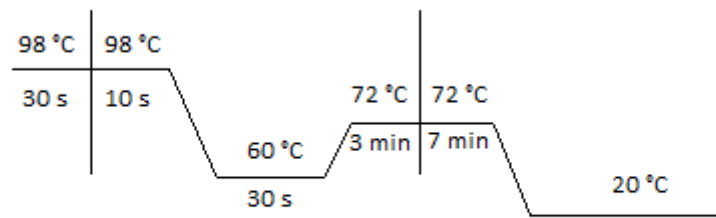
Componente	Reacción (μL)	Control negativo (μL)
H ₂ O MQ	50	54
Buffer GC 5X	20	20
Primer Fw	10	10
Primer Rv	10	10
DMSO	3	3
dNTPs	2	2
ADN	4	0
ADN pol Phusion	1	1

* Se utilizaron aproximadamente 10 ng de ADN genómico o vector por reacción de 100 uL.

Los primers utilizados fueron los siguientes:

Nombre	Secuencia
FwRF	GGATAACAATTTTCACACAGGAAACATGCCCGCACACATC
RvRF	GAAAATCTTCTCTCATCCGCCGACGAGGGAAACCCCGAC

2. Se estableció la temperatura y tiempo de cada ciclo (25 a 30 ciclos):



3. Se purificaron los productos en gel de agarosa, y luego se recuperó el ADN utilizando el kit Freeze and Squeeze (BIO-RAD).⁹⁷
4. Se consideró que el rendimiento en general no supera el 30%, por tanto, sólo se purifica si se tiene suficiente cantidad de ADN, sino repetir PCR y purificar todo junto.
5. Se verificó producto por electroforesis en agarosa (3 uL/reacción) y se cuantificó por comparación con un estándar de peso molecular.

RF PCR.

6. Se utilizaron 30 ng de ADN molde o vector a amplificar y 120 ng del megaprimer o amplicón obtenido en el punto 2. Dado que es una PCR sin primers, el producto de la primer PCR hace las veces de primer.
7. Se digirió con DpnI utilizando 100 µL de producto de RFPCR, 97µL de buffer y 3 µL de DpnI (20 U/µL), e incubando a 37°C, 5 hs. Inactivar DpnI a 80°C, 20 min.
8. Se precipitó con una solución de tRNA, utilizando 50 µL de muestra digerida con DPNI, 50 µL de solución stock de tRNA 0.2 µg/µL, 100 µL de agua Milli-Q y 500 µL de etanol absoluto.
9. Se transformó *E. coli* por electroporación. (las células tienen que tener una eficiencia de aproximadamente 10⁶).

PREPARACIÓN DE CÉLULAS COMPETENTES

1. Se inocularon 5 mL de medio LB con una colonia aislada en placa de células de *E. coli* BL21(DE3) (New England BioLabs Inc.). Se incubó el precultivo a 37°C, a 150 rpm durante 24 hs.

2. Se resuspendió el volumen de precultivo para una concentración 1/100 en LB (sin sal) y se incubó a 37°C, 150 rpm hasta una OD=0.4-0.5, Luego se colocó en hielo y se mantuvo en frío.
3. Se centrifugó a 4 °C, 4000 rpm, 15 min. Se descartó el sobrenadante y se resuspendió en 100 mL de Agua Mili Q.
4. Igual al paso 3, pero se resuspendió en 50 mL de Agua Mili Q.
5. Se volvió a realizar el paso 3, pero se resuspendió en 10 mL de Agua Mili Q.
6. Finalmente se siguió nuevamente el paso 3 y se resuspendió en 250 µL de Agua Mili Q.

TRANSFORMACIÓN DE E.COLI POR ELECTROPORACIÓN

7. Se transfirieron 40 µL de células competentes a eppendorf y se rotularon. Se agregaron 2 µL de ADN y se colocó la mezcla en celdas de electroporación de 2 mm pre-enfriadas.
8. Se electroporó con programa EC2 (2.5Kv durante 5 ms)
9. Se agregó 1 mL de SOC e incubó a 37°C, 1 hs.
10. Se sembraron 100 µL en placas (LB + AMP/KAN 50 µg/mL + X-gal 40 µL/placa + IPTG 4 µL/placa) y se incubaron a 37°C, 48 hs.

EXTRACCIÓN DE ADN PLASMÍDICO EN E. COLI

1. Se inocularon 5 mL de LB + amp con una colonia de *E. coli* que contiene el plásmido de interés. Se incubó a 37°C,
2. Se transfirió 1.5 mL de cultivo a un tubo eppendorf estéril.
3. Se centrifugó a 4°C, 4000 rpm, 10 min. Se descarto el sobrenadante.
4. Se agregó 1.5 mL de cultivo y se centrifugó a 4°C, 4000 rpm, 10 min. Posterior a esto se descartó el sobrenadante.
5. Se agregaron 300 µL de solución I y se homogeneizó
6. Se agregaron 300 µL de solución II y se homogeneizó.
7. Se agregaron 300 µL de solución III y se homogeneizó
8. Se centrifugó a 4°C, 12000 rpm, 5 min.
9. Se extrajeron 400 µL de sobrenadante y se igualó el volumen con una solución de fenol:cloroformo:alcohol isoamílico proporciones 25:24:1.
10. Se centrifugó a máxima velocidad, 5 min.
11. Se recuperaron 300 µL de la fase acuosa
12. Se precipitó agregando 2.5 volúmenes de etanol absoluto y 0.025 volúmenes de cloruro de sodio 5 molar (en este punto se puede guardar).
13. Se recolectó el precipitado por centrifugación a 4°C, a 13000 g, 5 min.
14. Se removió sobrenadante y se le agregaron 300 µL de etanol 70%

15. Se secó totalmente en estufa o termoblock y se resuspendió en 20 μ L de buffer TE.

LISADO QUÍMICO DE CÉLULAS PARA VER PERFIL DE EXPRESIÓN DE PROTEÍNAS

A partir de un cultivo de interés en medio de autoinducción o LB y antibiótico inducido a 28 °C, se sacaron muestras de 0.5 mL a 0 hs, 24 hs y 48 hs.

1. Se centrifugaron 0.5 mL de células a 4000 g en eppendorf por 15 min, se descartó sobrenadante. En este punto se puede parar y congelar las células.
2. Se resuspendieron en 0.5 mL de buffer de lisis.
3. Se puso en agua-hielo o bloque frío, se agregó 1 mg lisozima/mL de suspensión y se mezcló suavemente cada unos minutos, durante 30 min.
4. Se agregaron 0.05 mL de [Tritón X-100] 10 %/mL. Luego se mezcló suavemente cada unos minutos, durante 15 min.
5. Se centrifugó a 13000 g, 4 °C, 25 min y se separó el sobrenadante a otro eppendorf. Esto sería la fracción soluble.
6. Se disolvió la fracción insoluble en 1 mL y se diluyó en relación 1:5 muestra:buffer, se calentó a 100°C, 5 min.
7. Se sembraron en gel 20 μ L de la fracción soluble y 10 μ L de la insoluble.
8. Se corrió en la cuba de electroforesis a 100 mv, 2-2.5 hs.
9. Se tiñó con azul de Coomassie 2 hs.
10. Se destiñó con solución 7:2:1 H₂O:Isopropanol:Ác. acético glacial.

BIOTRANSFORMACIONES CON CÉLULAS ENTERAS EN SISTEMA DE CÉLULAS EN REPOSO

1. Se añadieron al medio de cultivo estéril los componentes necesarios (antibióticos, azúcares, etc.)
2. Se inoculó el medio con cultivo fresco con una colonia aislada de la cepa a estudiar y se incubó en shaker a 37°C, 150 rpm. durante toda la noche.
3. Se inoculó el volumen de precultivo para una concentración 1/100 en LB o AIM a 37°, 150 rpm, hasta una OD=0.4-0.5.
4. Se indujo de ser necesario (e.g., IPTG, AHT) y se dejó incubando durante un tiempo adecuado.
5. Se tomaron 500 µL de muestra para posterior análisis de expresión de proteínas.
6. Se rotuló y pesó tubos de centrifuga estériles.
7. Se transfirieron los cultivos a los tubos y se centrifugaron a 4000 g, 15 min. Se descartaron los sobrenadantes.
8. Se resuspendieron en 50 mL de buffer de lavado y se volvió a centrifugar a 4000 g, 15 min. Se descartaron los sobrenadantes.
9. Se pesaron los tubos y se calculó la cantidad de biomasa obtenida.
10. Se resuspendieron en buffer de reacción para dar una suspensión de aprox. 110 mg biomasa/mL.
Volumen de resuspensión = biomasa (mg)/110
11. Se largó cada reacción por duplicado.

Componente	Concentración final	Volumen
Biomasa	100 mg/mL	460 µL
Glucosa	2 % m/v	25 µL
Isopropanol	5 % v/v	25 µL
Sustrato 250 mM en DMSO	5 mM	10 µL

12. Se basificaron las reacciones con 100 µL de NaOH 10 M.
13. Se agregaron 750 µL de AcOEt.
14. Se agitó en vórtex 10 segundos y centrifugar.
15. Se separaron 500 µL a un eppendorf con una punta de espátula de Na₂CO₃ para secar el AcOEt.
16. Se agitó en vórtex 5 seg. y se dejó reposar 5 min. Luego de esto se centrifugó.
17. Se separaron 400 µL a un vial ámbar rotulado con tapa para inyector automático.
18. Se agregaron 50 µL de solución stock de estándar interno.

MEDIOS DE CULTIVO Y SOLUCIONES

- Luria Bertani(LB)

Triptona	10g
Cloruro de sodio	5g
Extracto de levadura	5g
Agua destilada	1 L

- Medio de Autoinducción (AIM)

Triptona	10g
Extracto de levadura	5g
Sulfato de amonio	3.3g
Fosfato de potasio monobásico	6.5g
Fosfato de potasio dibásico	7.1g
α -Lactosa	2.0g
Sulfato de magnesio	0.15g
Glucosa	0.5g
1000x Mezcla de elementos traza en H ₂ O (FeCl ₃ ·6H ₂ O, MnCl ₂ ·4H ₂ O 8.87mg, ZnSO ₄ ·H ₂ O 13mg, H ₃ BO ₃ 2mg, CoSO ₄ ·7H ₂ O 1.1mg, CuSO ₄ ·5H ₂ O 14mg, NiSO ₄ ·6H ₂ O 10mg, Na ₂ MoO ₄ 190.3mg).....	1mL

- SOB

Triptona	1g
Extracto de levadura.....	0.25g
Cloruro de sodio	0.025g
Agua destilada	50mL
Ajustar a pH 7.0	

- SOC

SOB estéril	5 mL
Cloruro de magnesio	50 μ L
Sulfato de magnesio	50 μ L
Glucosa 20 % estéril	100 μ L

Soluciones y reactivos

- Solución I

Glucosa	50 mM
Tris-CL	25 mM
EDTA	10 mM

- Solución II

NaOH	0.2 N
SDS	1% m/v

- Solución III

5 M KCH ₃ CO ₂ (Acetato de Potasio)	60 mL
Ác. acético glacial	11,5 mL
Agua destilada	28.5 mL

- 10x Buffer Tris Borato (TBE)

Tris base	108 g
Ácido bórico	55 g
EDTA 0.5 M	40 mL
Agua destilada.....	csp. 1 L

- 5x Buffer de carga (ADN)

Azul de bromofenol.....	0.25%
Glicerol.....	30%
Agua destilada	

- Solución stock de ampicilina 50 mg/mL

Ampicilina	0.5 g
Agua milli-Q csp.	10 mL

- Solución Stock de Kanamicina 5 mg/mL

Kanamicina 0.1g
 Agua milli-Q csp. 20 mL

- Fenol:Cloroformo :Alcohol Isoamílico:

Fenol 25 partes
 Cloroformo 24 partes
 Alcohol Isoamílico 1 parte

- Buffer de reacción (Tris-HCl 100mM pH 7.5):

Tris base 12.11g
 HCl hasta pH 7.5
 Agua destilada csp. 1L

- 2x Buffer de carga (Proteínas):

Buffer Tris-HCL 0.5 M pH 6.7 3.75 mL
 Glicerol 3 µL
 β-mercaptoetanol 0.3 µL
 SDS 0.6 g
 Azul de bromofenol 10 mg
 EDTA 1.5 µM
 Agua destilada csp. 10 mL

- Solución de tRNA 0.2 µg/µL para precipitación

Solución stock tRNA 10 µg/µL 20 µL
 Agua milli-Q csp. 1 mL

REFERENCIAS

1. Sheldon, R. A. & Pereira, P. C. Biocatalysis engineering: The big picture. *Chem. Soc. Rev.* **46**, 2678–2691 (2017).
2. Cosgrove, S. C. *et al.* Imine Reductases, Reductive Aminases, and Amine Oxidases for the Synthesis of Chiral Amines: Discovery, Characterization, and Synthetic Applications. *Methods in Enzymology* vol. 608 131–149 (2018).
3. Radzicka, A. & Wolfenden, R. A proficient enzyme. *Science (80-.)*. **267**, 90–93 (1995).
4. Faber, K. *Biotransformations in Organic Chemistry*. (Springer International Publishing, 2018). doi:10.1007/978-3-319-61590-5.
5. Zaks, A. & Klivanov, A. M. Enzymatic catalysis in organic media at 100°C. *Science (80-.)*. **224**, 1249–1251 (1984).
6. Netter, K. J. Xenobiotic conjugation chemistry. *Toxicology* **44**, 122–123 (1987).
7. Lawrence, S. a. *Quinoline, isoquinoline and quinolizine. Amines: Synthesis, Properties and Application* vol. 1 (2004).
8. Feldberg, B. W. & Myers, R. D. Effects on temperature of amines injected into the cerebral ventricles. A new concept of temperature regulation. *Natl. Inst. Med. Res.* **1703**, 226–237 (1964).
9. Schildkraut, J. J. & Kety, S. S. Biogenic Amines and Emotion Pharmacological studies suggest a relationship between brain biogenic amines and affective state. **156**, 21–30 (1967).
10. Davis, B. A. Biogenic amines and their metabolites in body fluids of normal, psychiatric and neurological subjects. *J. Chromatogr. A* **466**, 89–218 (1989).
11. Agranat, I., Caner, H. & Caldwell, J. Putting chirality to work: The strategy of chiral switches. *Nat. Rev. Drug Discov.* **1**, 753–768 (2002).
12. Lind, T. *et al.* Esomeprazole provides improved acid control vs. omeprazole in patients with symptoms of gastro-oesophageal reflux disease. *Aliment. Pharmacol. Ther.* **14**, 861–867 (2000).
13. Kim, J. H. & Scialli, A. R. Thalidomide: The tragedy of birth defects and the effective treatment of disease. *Toxicol. Sci.* **122**, 1–6 (2011).
14. Welsch, M. E., Snyder, S. A. & Stockwell, B. R. Privileged scaffolds for library design and drug discovery. *Curr. Opin. Chem. Biol.* **14**, 347–361 (2010).
15. Wang, Y. & Chen, A. Crystallization-Based Separation of Enantiomers. *Stereoselective Synth. Drugs Nat. Prod.* 1–20 (2013) doi:10.1002/9781118596784.ssd056.
16. Martin, V. S. *et al.* Kinetic Resolution of Racemic Allylic Alcohols by Enantioselective Epoxidation. A Route to Substances of Absolute Enantiomeric Purity. *J. Am. Chem. Soc.* **103**, 6237–6240 (1981).
17. Foundation, S., Noyori, R., Saito, T. & Kumobayashis, H. Stereoselective Hydrogenation via Dynamic Kinetic Resolution R. 9134–9135 (1989).
18. Fleury-Brégeot, N., de la Fuente, V., Castellón, S. & Claver, C. Highlights of transition metal-catalyzed asymmetric hydrogenation of imines. *ChemCatChem* **2**, 1346–1371 (2010).
19. Bartoszewicz, A., Ahlsten, N. & Martín-Matute, B. Enantioselective synthesis of alcohols and amines by iridium-catalyzed hydrogenation, transfer hydrogenation, and related processes. *Chem. - A Eur. J.* **19**, 7274–7302 (2013).
20. Chusov, D. & List, B. Reductive amination without an external hydrogen source. *Angew. Chemie - Int. Ed.* **53**, 5199–5201 (2014).
21. Höhne, M. & Bornscheuer, U. T. Biocatalytic routes to optically active amines. *ChemCatChem* **1**,

- 42–51 (2009).
22. Kim, M. J., Kim, W. H., Han, K., Yoon, K. C. & Park, J. Dynamic kinetic resolution of primary amines with a recyclable Pd nanocatalyst for racemization. *Org. Lett.* **9**, 1157–1159 (2007).
 23. Escoubet, S. *et al.* Thiyl radical mediated racemization of benzylic amines. *European J. Org. Chem.* 3242–3250 (2006) doi:10.1002/ejoc.200600120.
 24. Santos, A. I. Nuevas herramientas biocatalíticas para la síntesis asimétrica: Iminoreductasas, monoaminooxidasas y transaminasas. (Universidad de la República Facultad de Química-PEDECIBA Química, 2018).
 25. Schrittwieser, J. H., Velikogne, S. & Kroutil, W. Biocatalytic Imine Reduction and Reductive Amination of Ketones. *Adv. Synth. Catal.* **357**, 1655–1685 (2015).
 26. Mitsukura, K., Suzuki, M., Tada, K., Yoshida, T. & Nagasawa, T. Asymmetric synthesis of chiral cyclic amine from cyclic imine by bacterial whole-cell catalyst of enantioselective imine reductase. *Org. Biomol. Chem.* **8**, 4533–4535 (2010).
 27. Rodríguez-Mata, M. *et al.* Structure and activity of NADPH-dependent reductase Q1EQE0 from streptomyces kanamyceticus, which catalyses the R-selective reduction of an imine substrate. *ChemBioChem* **14**, 1372–1379 (2013).
 28. Zhang, L., Chooback, L. & Cook, P. F. Lysine 183 is the general base in the 6-phosphogluconate dehydrogenase-catalyzed reaction. *Biochemistry* **38**, 11231–11238 (1999).
 29. Scheller, P. N. *et al.* Enzyme Toolbox: Novel Enantiocomplementary Imine Reductases. *ChemBioChem* **15**, 2201–2204 (2014).
 30. Huang, H., Liu, X., Zhou, L., Chang, M. & Zhang, X. Direct Asymmetric Reductive Amination for the Synthesis of Chiral β -Arylamines. *Angew. Chemie - Int. Ed.* **55**, 5309–5312 (2016).
 31. Nugent, T. C. & El-Shazly, M. Chiral amine synthesis - Recent developments and trends for enamide reduction, reductive amination, and imine reduction. *Adv. Synth. Catal.* **352**, 753–819 (2010).
 32. Li, C., Villa-Marcos, B. & Xiao, J. Metal-bronsted acid cooperative catalysis for asymmetric reductive amination. *J. Am. Chem. Soc.* **131**, 6967–6969 (2009).
 33. Seiple, I. B. *et al.* A platform for the discovery of new macrolide antibiotics. *Nature* **533**, 338–345 (2016).
 34. Trotter, S. G. & Us, B. W. U. S. Engineered imine reductases and methods for the reductive amination of ketone and amine compounds. (2014).
 35. Mihara, H. *et al.* N-methyl-L-amino acid dehydrogenase from *Pseudomonas putida*: A novel member of an unusual NAD(P)-dependent oxidoreductase superfamily. *FEBS J.* **272**, 1117–1123 (2005).
 36. Sg, S. G., Jeffrey, C. & Us, N. J. Engineered imine reductases and methods for the reductive amination of ketone and amine compounds. vol. 1 (2013).
 37. Ye, L. J. *et al.* Engineering of amine dehydrogenase for asymmetric reductive amination of ketone by evolving *Rhodococcus phenylalanine dehydrogenase*. *ACS Catal.* **5**, 1119–1122 (2015).
 38. Abrahamson, M. J., Vázquez-Figueroa, E., Woodall, N. B., Moore, J. C. & Bommaris, A. S. Development of an amine dehydrogenase for synthesis of chiral amines. *Angew. Chemie - Int. Ed.* **51**, 3969–3972 (2012).
 39. Mayol, O. *et al.* A family of native amine dehydrogenases for the asymmetric reductive amination of ketones. *Nat. Catal.* **2**, 324–333 (2019).
 40. Wetzl, D. *et al.* Asymmetric Reductive Amination of Ketones Catalyzed by Imine Reductases.

- ChemCatChem* **8**, 2023–2026 (2016).
41. Huber, T. *et al.* Direct reductive amination of ketones: Structure and activity of S-selective imine reductases from streptomyces. *ChemCatChem* **6**, 2248–2252 (2014).
 42. Scheller, P. N., Lenz, M., Hammer, S. C., Hauer, B. & Nestl, B. M. Imine Reductase-Catalyzed Intermolecular Reductive Amination of Aldehydes and Ketones. *ChemCatChem* **7**, 3239–3242 (2015).
 43. France, S. P. *et al.* Identification of Novel Bacterial Members of the Imine Reductase Enzyme Family that Perform Reductive Amination. *ChemCatChem* **10**, 510–514 (2018).
 44. Matzel, P., Gand, M. & Höhne, M. One-step asymmetric synthesis of (R)- and (S)-rasagiline by reductive amination applying imine reductases. *Green Chem.* **19**, 385–389 (2017).
 45. Aleku, G. A. *et al.* A reductive aminase from *Aspergillus oryzae*. *Nat. Chem.* **9**, 961–969 (2017).
 46. Sharma, M. *et al.* A Mechanism for Reductive Amination Catalyzed by Fungal Reductive Aminases. *ACS Catal.* **8**, 11534–11541 (2018).
 47. Aleku, G. A. *et al.* Stereoselectivity and Structural Characterization of an Imine Reductase (IRED) from *Amycolatopsis orientalis*. *ACS Catal.* **6**, 3880–3889 (2016).
 48. Man, H. *et al.* Structure, activity and stereoselectivity of NADPH-dependent oxidoreductases catalysing the S-selective reduction of the imine substrate 2-methylpyrroline. *ChemBioChem* **16**, 1052–1059 (2015).
 49. France, S. P. *et al.* Biocatalytic Routes to Enantiomerically Enriched Dibenz[c,e]azepines. *Angew. Chemie - Int. Ed.* **56**, 15589–15593 (2017).
 50. Brunhuber, N. M. W., Thoden, J. B., Blanchard, J. S. & Vanhooke, J. L. Rhodococcus L-phenylalanine dehydrogenase: Kinetics, mechanism, and structural basis for catalytic specificity. *Biochemistry* **39**, 9174–9187 (2000).
 51. Xiong, J. *Essential bioinformatics*. (Cambridge University Press, 2006).
 52. Luscombe, N. M., Greenbaum, D. & Gerstein, M. What is bioinformatics? A proposed definition and overview of the field. *Methods Inf. Med.* **40**, 346–358 (2001).
 53. Hagen, J. B. The origins of bioinformatics. *Nat. Rev. Genet.* **1**, 231–236 (2000).
 54. Borém, A. & Fritsche-Neto, R. Omics in Plant Breeding. *Omi. Plant Breed.* **9781118820**, 1–230 (2014).
 55. Hersh, R. & Dayhoff, M. O. Atlas of Protein Sequence and Structure, 1969 Volume 4. *Syst. Zool.* **19**, 192 (1970).
 56. Bernstein, F. C. *et al.* The protein data bank: A computer-based archival file for macromolecular structures. *Arch. Biochem. Biophys.* **185**, 584–591 (1978).
 57. Needleman, S. B. & Wunsch, C. D. A General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins. *Mol. Biol.* 453–463 (1989) doi:10.1016/b978-0-12-131200-8.50031-9.
 58. Chou, P. Y. & Fasman, G. D. Prediction of Protein Conformation. *Biochemistry* **13**, 222–245 (1974).
 59. Krieger, E., Nabuurs, B. S. & Vriend, G. Homology Modeling. 1–33 (1963).
 60. Polanski, J. *et al.* 4.14 Chemoinformatics. 459–506 (2009).
 61. Chaitanya, M. *et al.* Exploring the molecular basis for selective binding of *Mycobacterium tuberculosis* Asp kinase toward its natural substrates and feedback inhibitors: A docking and molecular dynamics study. *J. Mol. Model.* **16**, 1357–1367 (2010).

62. Shen, J. *et al.* Homology modeling, molecular docking, and molecular dynamics simulations elucidated α -fetoprotein binding modes. *BMC Bioinformatics* **14**, (2013).
63. Ogawa, J. & Shimizu, S. Industrial microbial enzymes: Their discovery by screening and use in large-scale production of useful chemicals in Japan. *Curr. Opin. Biotechnol.* **13**, 367–375 (2002).
64. Gupta, R., Rathi, P., Gupta, N. & Bradoo, S. Lipase assays for conventional and molecular screening: an overview. *Biotechnol. Appl. Biochem.* **37**, 63 (2003).
65. Towards, A. & Biological, S. *Modern Biocatalysis*. Royal Society of Chemistry (2008). doi:10.1002/9783527623839.
66. D'Haeseleer, P. What are DNA sequence motifs? *Nat. Biotechnol.* **24**, 423–425 (2006).
67. Rabiner, L. R. & Juang, B. H. An Introduction to Hidden Markov Models. *IEEE ASSP Mag.* **3**, 4–16 (1986).
68. Madeira, F. *et al.* The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).
69. Zvelebil, M. & Baum, J. O. *Understanding bioinformatics*. (2008).
70. Kobe, B. & Kemp, B. E. Active site-directed protein regulation. *Nature* **402**, 373–376 (1999).
71. Berman, H. M. *et al.* The protein data bank. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **58**, 899–907 (2002).
72. McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallogr. Sect. F Structural Biol. Commun.* **70**, 2–20 (2014).
73. Bornscheuer, U. T. & Höhne, M. Protein engineering - Methods and protocols. *Methods Mol. Biol.* **1685**, 1–347 (2018).
74. Sears, D. Protein structure prediction: Concepts and applications. *Biochem. Mol. Biol. Educ.* **35**, 79–80 (2007).
75. Tradigo, G., Rondinelli, F. & Pollastri, G. Algorithms for structure comparison and analysis: Docking. *Encycl. Bioinforma. Comput. Biol. ABC Bioinforma.* **1–3**, 77–80 (2018).
76. Vakser, I. A. Protein-protein docking: From interaction to interactome. *Biophys. J.* **107**, 1785–1793 (2014).
77. Rohs, R., Bloch, I., Sklenar, H. & Shakked, Z. Molecular flexibility in ab initio drug docking to DNA: Binding-site and binding-mode transitions in all-atom Monte Carlo simulations. *Nucleic Acids Res.* **33**, 7048–7057 (2005).
78. Yuriev, E., Agostino, M. & Ramsland, P. A. Challenges and advances in computational docking: 2009 in review. *J. Mol. Recognit.* **24**, 149–164 (2011).
79. Vila, M. A. *et al.* Computational insights into the oxidation of mono- and 1,4 disubstituted arenes by the Toluene Dioxygenase enzymatic complex. *J. Mol. Catal. B Enzym.* **133**, S410–S419 (2016).
80. Catucci, G. *et al.* Enzyme-substrate matching in biocatalysis: in silico studies to predict substrate preference of ten putative ene-reductases from *Mucor circinelloides* MUT44. *J. Mol. Catal. B Enzym.* **131**, 94–100 (2016).
81. Mitsukura, K. *et al.* A NADPH-dependent (S)-imine reductase (SIR) from *Streptomyces* sp. GF3546 for asymmetric synthesis of optically active amines: Purification, characterization, gene cloning, and expression. *Appl. Microbiol. Biotechnol.* **97**, 8079–8086 (2013).
82. Mitsukura, K. *et al.* Purification and characterization of a novel (R)-imine reductase from *Streptomyces* sp. GF3587. *Biosci. Biotechnol. Biochem.* **75**, 1778–1782 (2011).
83. BLAST[®]. <https://blast.ncbi.nlm.nih.gov/Blast.cgi>.

84. Fademrecht, S., Scheller, P. N., Nestl, B. M., Hauer, B. & Pleiss, J. Identification of imine reductase-specific sequence motifs. *Proteins Struct. Funct. Bioinforma.* **84**, 600–610 (2016).
85. Lu, S. Rapid Screening of Recombinant Plasmids. **235**, 169–174 (2003).
86. Geneious. <https://www.geneious.com/home/>.
87. Vector NTI Software. <https://www.thermofisher.com/uy/en/home/life-science/cloning/vector-nti-software.html>.
88. Velikogne, S., Resch, V., Dertnig, C., Schrittwieser, J. H. & Kroutil, W. Sequence-Based In-silico Discovery, Characterisation, and Biocatalytic Application of a Set of Imine Reductases. *ChemCatChem* **10**, 3236–3246 (2018).
89. Allouche, A. Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. *J. Comput. Chem.* **32**, 174–182 (2012).
90. GenScript. <https://www.genscript.com/>.
91. Kumar, S., Stecher, G., Li, M., Niyaz, C. & Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
92. Gattiker, A., Gasteiger, E. & Bairoch, A. ScanProsite: a reference implementation of a PROSITE scanning tool. *Appl. Bioinformatics* **1**, 107–108 (2002).
93. Apweiler, R. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **32**, 115D – 119 (2004).
94. Edgar, R. C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 1–19 (2004).
95. Unger, T., Jacobovitch, Y., Dantes, A., Bernheim, R. & Peleg, Y. Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. *J. Struct. Biol.* **172**, 34–44 (2010).
96. Van Den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* **67**, 67–74 (2006).
97. Freeze 'N Squeeze™ DNA Gel Extraction Spin Columns. <https://www.bio-rad.com/en-us/product/freeze-n-squeeze-dna-gel-extraction-spin-columns?ID=b6193ff4-ba82-4a64-8a32-f1507c09f308>.
98. Stoermer, M. Homology Models of Coronavirus 2019-nCoV 3CLpro Protease. *ChemRxiv* 1–11 (2020) doi:10.26434/CHEMRXIV.11637294.V3.