

La familia PriA/HisA

PriA es la familia de enzimas de Actinobacteria homóloga a la familia HisA existente en Enterobacteria, Cyanobacteria, *Pseudomonas* y Archaea. Según las definiciones de este trabajo PriA es una familia promiscua ya que se ha encontrado que varios miembros de PriA *in vitro* tienen la capacidad de catalizar tanto la reacción correspondiente a su homólogo HisA en la síntesis de histidina como la isomerización que cataliza TrpF cuya reacción está involucrada en la producción de triptofano. Es decir, se ha mostrado mediante cinéticas enzimáticas y análisis de complementación genética que en varias Actinobacterias PriA participa a la vez en las rutas de síntesis de histidina y triptofano mediante la isomerización de un anillo de cinco carbonos Figure 1. Los primeros miembros caracterizados como promiscuos en esta familia fueron *Streptomyces coelicolor* y *Mycobacterium tuberculosis* [@baronagomez_occurrence_2003] en 2003. La mayoría de las actinobacterias han perdido el gen *trpF* aunque su actividad es esencial. Además conservan el resto del operón de triptofano, por lo que se cree que la promiscuidad de PriA está extendida en un gran subconjunto de Actinobacteria. En la ruta de histidina PriA isomeriza el sustrato ProFAR en PRFAR, realizando la función HisA. En la ruta de triptofano PriA lleva a cabo la isomerización de PRA en CdrP catalizando la reacción de TrpF en la ruta del triptofano. La vecindad genómica de PriA en *Streptomyces coelicolor* y en otras Actinobacterias contiene genes tanto del operón de histidina como del de triptofano.

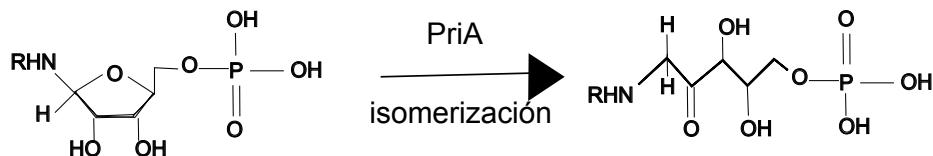


Figure 1: La reacción que cataliza PriA es una isomerización donde abre un anillo de cinco carbonos.

En este capítulo usamos como modelo de una familia promiscua la familia PriA/HisA porque además PriA ha mostrado un gradiente funcional en Actinobacteria. Esta variación divide a la familia en varias subfamilias según su capacidad catalítica en los sustratos ProFAR y PRA. La evolución de las subfamilias de PriA y sus funciones es compleja y muestra dinámicas distintas en cada linaje (ver introducción). Por ejemplo incluye a los de miembros de PriB, la subfamilia ubicada en el género *Streptomyces* con baja capacidad de catálisis para la función TrpF. Varios *Streptomyces* con un ortólogo en la familia PriB, se diferencian de otras Actinobacterias en que contienen en su genoma un gen *trpF* localizado fuera del contexto genómico inmediato de los operones de histidina y triptofano. Otra subfamilia de PriA es subHisA, que ha perdido totalmente la actividad TrpF, existen miembros de subHisA en *Corynebacteria* y en *Actinomycetes*. Finalmente en *Actinomycetes* también encontramos la subfamilia subTrpF que ha perdido la actividad de HisA.

En este capítulo se exploran cuatro aspectos de la familia PriA. *i)* La distribución y el contexto genómico de PriA en diversos linajes genómicos. *ii)* La información contenida a nivel de aminoácidos en variantes de PriA como medio de estudio de rutas evolutivas y su relación con la reconstrucción de su estructura tridimensional. *iii)* Las posibles afinidades de PriA por otros sustratos con métodos bioinformáticos, y finalmente *iv)* La validación experimental de la actividad de PriA con sus sustratos PRA, ProFAR o combinaciones de los dos.

La familia PriA cambió su función y sus patrones de promiscuidad en los cuatro linajes analizados

Las expansiones no son condición necesaria para la promiscuidad

Para explorar PriA en diversos linajes genómicos se utilizaron las herramientas EvoMining y CORASON descritas en los capítulos previos. Se investigaron las expansiones de la familia PriA en los linajes Actinobacteria, Cyanobacteria, *Pseudomonas* y Archaea. En la Figure 2 se muestra el número promedio de copias por genoma en los linajes genómicos seleccionados. En Actinobacteria, donde se sabe que PriA es promiscua no

se detectaron copias extra. Según EvoMining en Actinobacteria no hay expansiones, prácticamente todas las copias son reconocidas como de metabolismo central (rojo o morado) aunque algunas PriA además son marcadas por antiSMASH como parte de algún cluster biosintético (morado). En cambio en Cyanobacteria, *Pseudomonas* y Archaea la figura muestra en negro las copias extra de las que no se conoce su destino metabólico. El caso de Archaea es llamativo porque las copias de metabolismo central llegan en promedio hasta .5 copias por genoma, es decir muchos genomas de Archaea no cuentan con una copia de PriA, y en cambio, contrario a Actinobacteria, un 50% de la copias es marcado en negro, es decir varios de los genomas que tienen al menos una copia de PriA en realidad tienen dos copias. Esta figura muestra que en Actinobacteria PriA constituye un ejemplo de familia promiscua mayoritariamente distribuida con una sola copia por genoma por lo que es evidente que para que una familia sea promiscua no es imperativo tener copias extra con marcas de reclutamiento en metabolismo especializado. Aunque las copias extra suelen ser una indicación de promiscuidad, no son una condición necesaria. Tanto EvoMining como CORASON mostraron que existen excepciones de organismos donde PriA tiene doble copia, tanto en Actinobacteria como en otros linajes genómicos.

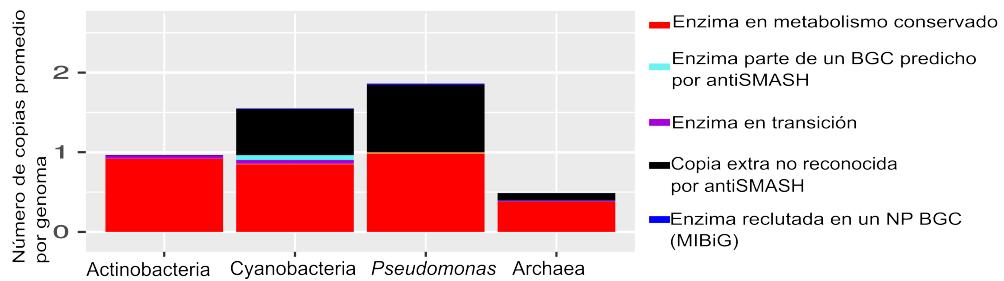


Figure 2: Número promedio de copias por genoma de PriA en Actinobacteria, Cyanobacteria, *Pseudomonas* y Archaea. Los colores muestran el destino metabólico asignado a cada copia según EvoMining. En rojo están los BBH a las enzimas semilla de metabolismo central. En morado las enzimas de metabolismo central también reconocidas por antiSMASH como parte de unBGC y en negro las copias sin un destino metabólico conocido.

Las expansiones condujeron a la creación de la subfamilia HisF

Después del conteo de número de copias promedio, se analizaron los árboles de PriA de EvoMining coloreados de acuerdo al número de copias, Figure 3En Actinobacteria la mayoría de las hojas son verdes mostrando que existe sólo una copia por genoma en ese organismo. Sin embargo, existen varias hojas de color amarillo, lo que indica que hay dos copias en algunos genomas. En Cyanobacteria, *Pseudomonas* y Archaea en contraste con Actinobacteria, se muestran una mezcla entre organismos que poseen una (verdes) o dos copias (amarillos) de PriA. Sin embargo al analizar detalladamente los árboles producidos por EvoMining en los distintos linajes, tanto en Cyanobacteria como en *Pseudomonas* la copia extra está en una rama divergente y muy poblada del árbol de esta Familia Expandida, lo que indica que esas copias extra son en realidad miembros de otra subfamilia enzimática que aun guarda cierta similitud de secuencia con PriA. Esta segunda copia está en su mayoría anotada por RAST como imidazol glicerol fosfato sintasa ciclase en ambos linajes. En Archaea sin embargo, diversas especies de la clase Methanomicrobia sí tienen dos copias cercanas a PriA.

Cada linaje tiene distintos destinos metabólicos de PriA

Después de explorar cuáles organismos tienen expansiones de PriA, analizamos el posible destino metabólico de las copias extra de la familia como se muestra en la Figure 4. El árbol de Actinobacteria está poblado de hojas rojas Figure 4, es decir de PriA dedicadas al metabolismo conservado, en este caso relacionado a las rutas de Histidina y Triptofano. Sin embargo hay algunas hojas grises, como es el caso de los dos *Serenicoccus*. Es posible que estas PriA puedan tener funciones alternativas o ser otros eventos con cambio de promiscuidad. Además, en Actinobacteria la PriA de *Janibacter hoilley*, la rama más larga del árbol, es muy divergente. Esto se debe a que existe una fusión de PriA con HisH. La fusión de genes bacterianos es un

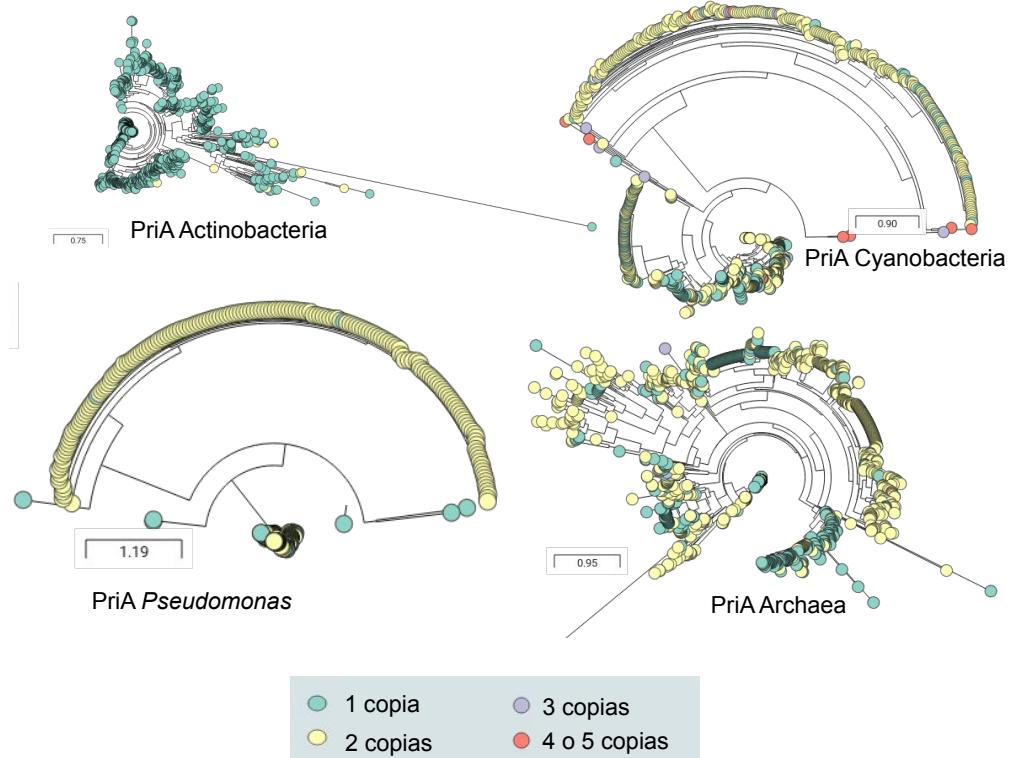


Figure 3: Número de copias de PriA en Actinobacteria, Cyanobacteria, Pseudomonas y Archaea. En Actinobacteria, donde se ha comprobado su promiscuidad enzimática la moda en el número de copias es uno (verde). En Cyanobacteria y Archaea se observa una mezcla de hojas verdes y hojas amarillas, donde las amarillas son nodos que pertenecen a organismos con dos copias. También en estos linajes algunos organismos poseen tres copias (morado) y cuatro o cinco (rojo). En Pseudomonas la mayoría de los organismos muestran dos copias

mecanismo común para formar proteínas multidominio [@pasek_gene_2006]. De hecho, dentro del operón de histidina existen fusiones reportadas HisNB y HisIE con proteínas bifuncionales en ciertas proteobacterias [@fani_origin_2005]. La fusión aquí presentada no parece ser un artefacto de anotación ya que hay otros genomas de *Janibacter* con secuencias de PriA más grandes que el promedio. Los árboles que se produjeron por EvoMining están disponibles para exploración interactiva en Microreact en los links de la Table 1:

En Cyanobacteria hay pocas predicciones de EvoMining, lo que se muestra en escasas hojas verdes Figure 4 que además no están localizadas cerca de su reclutamiento la HisA de saxitoxin, el BGC proveniente de Cyanobacteria. Varias de las copias extra de Cyanobacteria corresponden a HisF, la subunidad enzimática imidazol glicerol fosfato sintasa parte también del operón de histidina. HisA proviene de una duplicación de un ancestro común de la mitad de su tamaño y HisF proviene de la duplicación de HisA [@fani_evolution_1994,@fani_paralogous_1997]. Por ello puede considerarse que EvoMining en efecto está encontrando una expansión de HisA que consiste en la familia HisF en el linajes Cyanobacteria, lo mismo ocurre en *Pseudomonas*. En su momento HisF fue parte del *dispensable genome*, no todos los organismos poseían esta expansión, posteriormente, al parecer antes del último ancestro común de Archaea, Bacteria y Eukarya, la síntesis de histidina fue fijada [@fondi_evolution_2009]. Este es un caso donde lo que alguna vez fue *dispensable genome* / metabolismo secundario, se ha vuelto *shell genome* / metabolismo conservado y donde se muestra que en estos linajes Cyanobacteria y *Pseudomonas* EvoMining aún es capaz detectar esa marca de expansión. De hecho se ha reconstruido una HisF ancestral con el objetivo de probar si retenía la actividad de HisA, es decir si poseía promiscuidad enzimática [@merkl_reconstruction_2016]. LUCA-HisF mostró no ser promiscua para los sustratos de HisA ni tampoco para su similar el sustrato TrpF.

En *Pseudomonas* hay una gran población de predicciones de EvoMining, pero el árbol tiene similitudes con Cyanobacteria, la rama divergente corresponde en su mayoría a copias de HisF. De hecho, el reclutamiento que hace que el árbol de *Pseudomonas* tenga una rama verde está anotado funcionalmente no como una HisA, sino como una HisF parte de un BGC que produce un lipopolisacárido en la Proteobacteria *Legionella pneumophila*. En Archaea, la rama central con una mezcla de hojas grises y rojas Figure 4 contiene también copias de HisF provenientes de genomas de *Sulfolobus*. Al verificar en el árbol interactivo en Microreact Table 1 el número de copias presente en organismos con un homólogo clasificado como HisF se comprobó que los *Sulfolobus* sólo poseen una copia en este árbol. Sin embargo, tanto al explorar manualmente los genomas de *Sulfolobus*, como al revisar la literatura sobre metabolismo de histidina en Archaea se encontró que sí se han encontrado homólogos de PriA en genomas de *Sulfolobus* [@fondi_evolution_2009]. Por tanto este árbol sugiere que aunque existe tanto PriA como su expansión HisF en el género *Sulfolobus*, la copia HisF es la más parecida a la semilla con que se generó este árbol. Además la similitud de secuencia del homólogo de HisA no fue suficiente para recuperar copias de HisA en este experimento computacional. Algunas de las hojas grises del árbol sin destino metabólico conocido serán exploradas en la siguiente sección.

Table 1: Árboles EvoMining de PriA/HisA en MicroReact

Linaje	Link al árbol de EvoMining en Microreact
Actinobacteria	7g2IGfkv9
Cyanobacteria	qF6jWRMox
<i>Pseudomonas</i>	ydff6DWqs
Archaea	Ig-m9Cm6f

Homólogos de PriA han sido reclutados a clusters de metabolismo especializado

Se encontraron algunos homólogos reclutados a BGC conocidos, marcados en azul en el árbol, que son miembros de la familia expandida PriA/HisA en estos linajes genómicos. La información de dichos BGC está listada en la tabla section . Entre ellos se encuentran dos toxinas de Cyanobacteria [@moustafa_origin_2009], un lipopolisacárido producido por una Proteobacteria y un BGC productor de cloro-pentostatina producido en Actinobacteria [@gao_biosynthesis_2017].

Table: Reclutamientos de expansiones de PriA en MIBiG

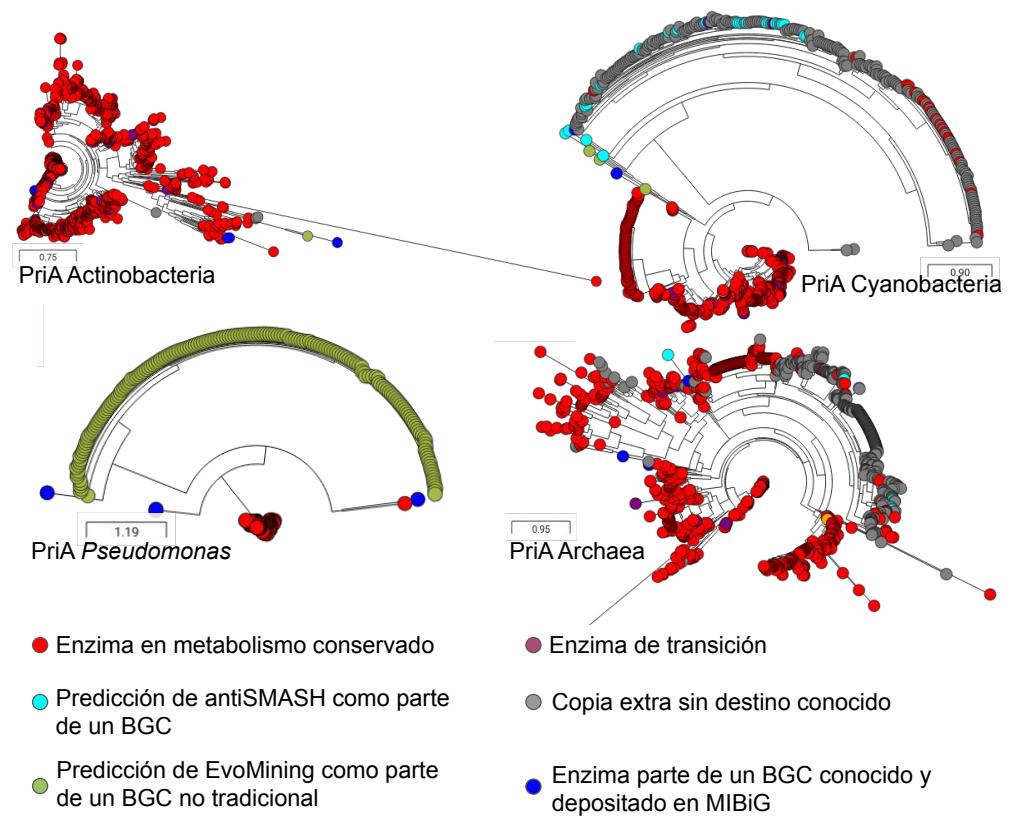


Figure 4: Árboles de destino metabólico de PriA en Actinobacteria, Cyanobacteria, Pseudomonas y Archaea según EvoMining

Compuesto	Organismo	BGC origen	Clase
Pentostatina	<i>Actinomadura</i> sp. ATCC 39365	Actinobacteria	Otros
Saxitoxina	<i>Cylindrospermopsis raciborskii</i> T3	Cyanobacteria	Alcaloide
Toxina	<i>Dolichospermum circinale</i> AWQC131C	Cyanobacteria	Otros
Lipopolisacárido	<i>Legionella pneumophila</i>	Proteobacteria	Sacárido

La pentostatina es un antibiótico nucleosídico derivado de adenosina, cuyo cluster es llamado *ada*. La PriA del cluster *ada* es llamada *adaK* y sí parece participar del cluster, ya que muestra una isomerización sobre un sustrato similar a los nativos de PriA, Figure 5, con un anillo de 5 carbonos, dos OH, un oxígeno y un grupo fosfato. Esta isomerización es muy parecida a la que realiza PriA sobre ProFAR y PRA. Esta PriA no es una copia extra, sino la copia única de este organismo. Este contexto genómico no se encuentra conservado en las copias vecinas en el árbol. Es relevante mencionar que la mutante de PriA no suprime la producción de este antibiótico en *Actinomadura* sp. ATCC 39365, por lo que los autores especulan que otra enzima podría estar realizando la isomerización redundantemente. *Actinomadura* no posee una copia de TrpF que sería el candidato inmediato.

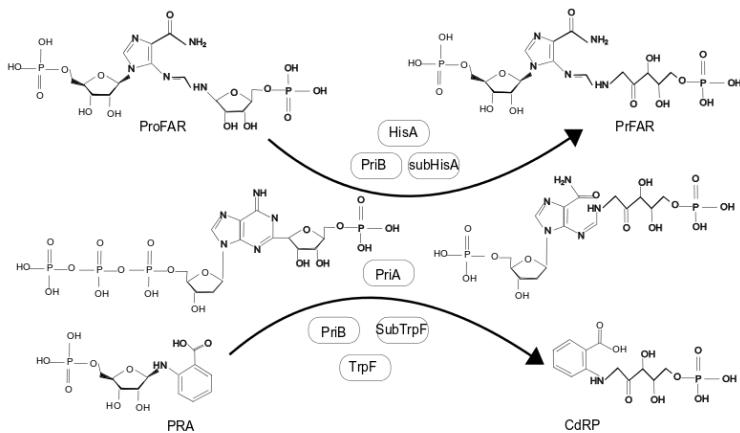


Figure 5: PriA participa en la síntesis del antibiótico ada en *Actinomadura*. Los sustratos nativos de PriA, ProFAR y PRA son isomerizados de manera muy similar a un paso en la ruta de síntesis de ada

Análisis de contextos genómicos de PriA/HisA en distintos linajes utilizando CORASON como herramienta de visualización.

Contextos genómicos de HisA en Actinobacteria

En Actinobacteria, se observó que todos los *Streptomyces* tienen el cluster de PriA parcialmente conservado con respecto al BGC de *Streptomyces coelicolor*. Ejemplos de ello son *S. roseous*, *S. sviceus*, *S. sp C* y *S. Mg1* donde genes tanto de la ruta de histidina como de triptofano rodean a PriA. Otros como *S rimosus*, *S HmicA12* y *S. sp CT34* tienen los genes de triptofano más alejados Figure 6 . Como ya se describió en la sección de EvoMining, el único organismo de este género con una copia extra de PriA es *Streptomyces* CT34. Esta copia parece deberse a transferencia horizontal dado que su mejor hit en NCBI proviene de una *Lentzea*. Aún así parece ser un homólogo lejano ya que tuvo 50% de identidad en 98% de cobertura con respecto a la copia de *Lentzea*. Otro caso interesante en Actinobacteria son las Actinomaduras, ya que CORASON muestra que el cluster *ada* no está conservado en ellas (datos no mostrados). Además, también en Actinobacteria CORASON muestra en *Sporichthya polymorpha* DSM 43042 una PriA precedida por una NRPS, una enzima por excelencia de productos naturales. En otros organismos antiSMASH predice que PriA forma parte de clusters putativos, por ejemplo en *Modestobacter marinus* NC_0179551, *Geodermatophilus obscurus*, y en *Streptacidiphilus jeojiense*. Los contornos de los genes reconocidos por antiSMASH como parte de un BGC son marcados en azul en las figuras de CORASON.

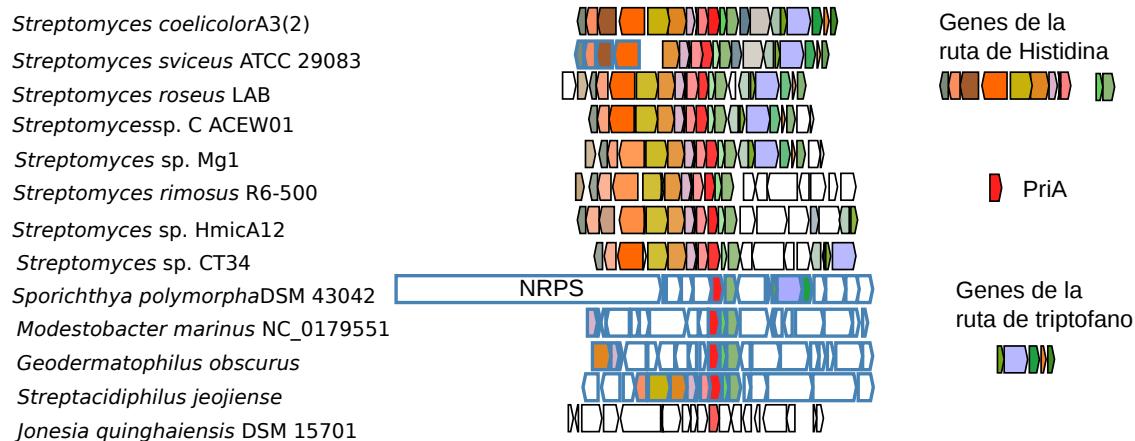


Figure 6: Contextos de PriA en Actinobacteria

Contextos genómicos de HisA en Archaea

En cuanto al contexto de HisA en Archaea la Figure 7 muestra que existen contextos como los de *Thermococcus kodakarensis* KOD1 y *Thermococcus sp JCM 11816* donde *hisA* está rodeada de genes tanto de histidina como de triptofano. Sin embargo esta configuración conjunta no es la generalidad. En Archaea, se conocen tres escenarios para la organización de los genes de síntesis de histidina. En el primero los genes de síntesis están dispersos en el genoma. El segundo escenario es la existencia de suboperones pequeños como *hisHAF* y finalmente el tercer escenario es el operón con la mayoría de los genes juntos aunque no siempre en el mismo orden [fondi_evolution_2009].

En efecto en Euryarchaeota están presentes los tres escenarios, en *Archaeoglobales* y *Halobacteriales* los genes del operón *his* están distribuidos en el genoma, como muestran los contextos de *hisA* en *Halorubrum tebenquichense* y *Halototius sp.* sin ningún otro gen del operón cerca. En *Thermosarcinales* donde ya se conocía la existencia del suboperón *hisGBA*, se observó que éste se encuentra conservado en la mayoría de los miembros del orden como ejemplifica el contexto correspondiente a *Methanosarcina lacustris*. En *Thermococcales* se conocía la configuración *hisGDBHAFIEC* que es confirmada por los tres *Thermococcus* de

la Figure 7, donde además se muestran dos contextos acompañados del su operón de triptofano *trpCDEGF*, resultando un contexto total de *hisGDBHAFIECXtrpCDEGF*. En el phylum Thaumarchaeota se habían estudiado sólo dos organismos y estos presentaban la configuración *hisGDCXBHAI*. En contraste en este trabajo encontramos en el orden *Nitrosopumilales* la configuración *hisDCXBHAFIE*, donde además la *hisX* es una fosfatasa homóloga a la *hisX* presente en *T. kodakarensis*. Este patrón es común en Thaumarchaeota, aunque también se muestra la existencia del suboperón *hisAF* en *Aigarchaeota* así como la configuración de genes dispersos en el genoma en *Thaumarchaeota archaeon JGI OTU-1*. En Crenarchaeota *hisCGABFDEHI* está presente según la anotación funcional de RAST, aunque no todos los genes aparecen en color en la figura debido a la baja identidad de secuencia que conservan con respecto al contexto de referencia de *T. kodakarensis*. En *Pyrobaculum neutrophilum* está parcialmente conservado el operón con la configuración *hisCEDFBBA*, nuevamente con algunos genes con baja identidad de secuencia respecto a *T. kodakarensis*. La configuración genes dispersos en el cromosoma en el phylum Crenarchaeota es mostrada en el contexto de *Ignicoccus islandicus*. Finalmente fue analizado el contexto de *hisA* en Bathyarchaeota un phylum donde no se contaba con información previa. En Bathyarchaeota se conserva como en Thaumarchaeota el ordenamiento *hisGDCXBHAI*, aunque existen otras configuraciones como muestra *Bathyarchaeota archaeon B25*.

Otra característica interesante de la visualización de los contextos de HisA en Archaea es que los tamaños de *hisA* son más variados que en Actinobacteria. Esta característica se debe tanto a variación intrínseca del tamaño de *hisA* como a diversas fusiones. Por ejemplo en *Halorubrum tebenquichense* DSM 14210 *hisA* se encuentra fusionada con *flaJ* una proteína relacionada con la construcción de flagelos en Archaea. En *Candidatus Bathyarchaeota archaeon B26-1* *hisA* se muestra fusionada con *hisF*.

En conclusión, no encontramos una configuración conservada que prevalezca en todo Archaea. Esto puede deberse a que Archaea es un dominio, no un phylum como Actinobacteria, y por tanto hay una mayor distancia entre los organismos que la conforman. Queda por

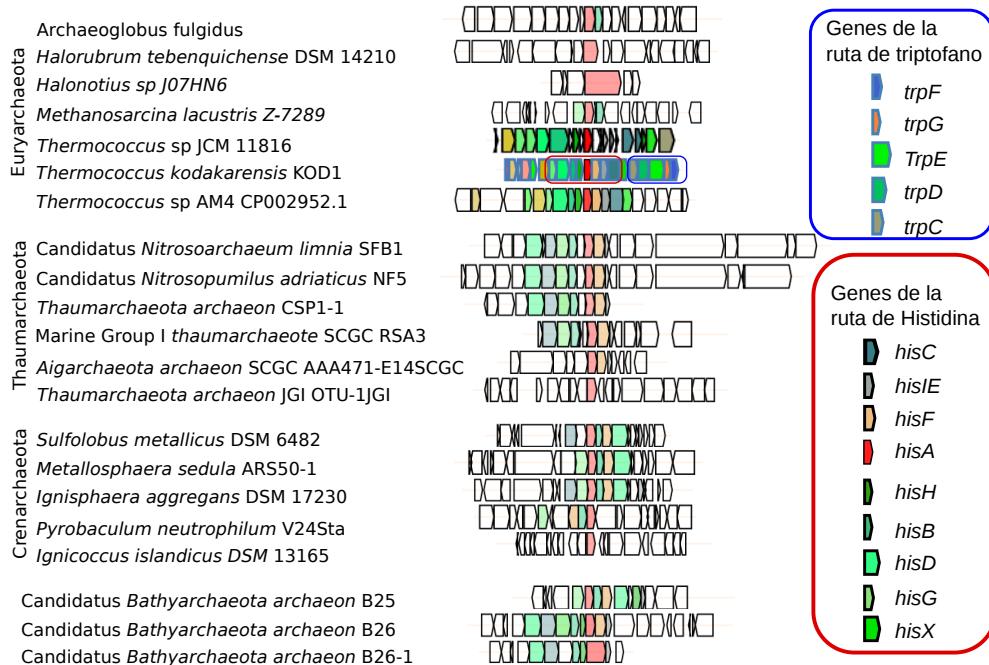


Figure 7: Contextos de PriA en Archaea

Contextos genómicos de HisA en saxitoxin

Como se mostró en la sección de análisis de expansiones de HisA con EvoMining en Cyanobacteria, *hisA* aparece dentro del cluster reportado como productor de saxitoxina en MIBiG. El contexto que rodea al

reclutamiento de *hisA* en el cluster de saxitoxina no está conservado en Cyanobacteria según la visualización de contextos producida por CORASON. Tomando como semilla la secuencia de HisA del BGC de saxitoxina y como referencia el cluster de saxitoxina, se muestra que si bien los reguladores del BGC saxitoxina sí están conservados en la vecindad genómica de HisA en otros organismos, este no es el caso para las enzimas biosintéticas del *cluster*. Además en el lado izquierdo de la Figure 8 se muestra que la HisA del BGC saxitoxina no está ubicada en una rama de PriA divergente, al contrario está embebida en la parte más conservada. Por estas razones es posible que PriA esté en la orilla del BGC de saxitoxina y más bien no participe en la síntesis de este compuesto o bien que mantenga su función primaria y a la vez tenga una función aún desconocida teniendo un papel en la regulación de la producción de saxitoxina.

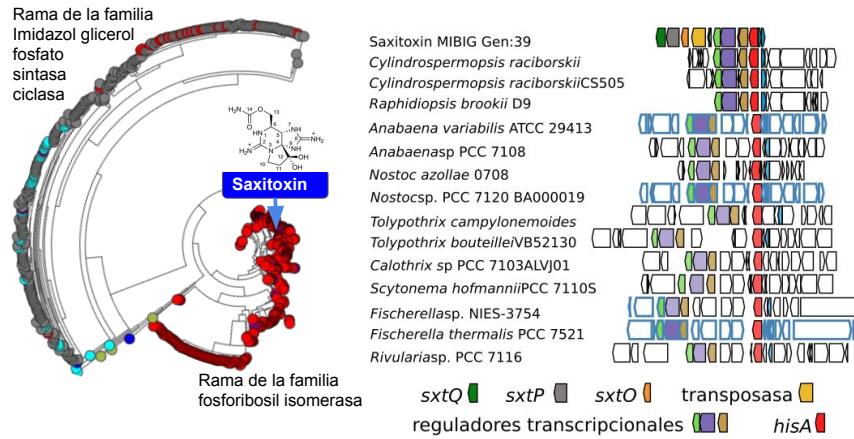


Figure 8: HisA en saxitoxin, un cluster de Cyanobacteria

Evolución molecular y estructural de PriA

En esta segunda sección concerniente a la familia PriA se busca información contenida en la secuencia de aminoácidos. En la primera parte se discute cómo en datos de evolución dirigida en el laboratorio no se encontró ninguna trayectoria en donde algún paso incrementara la actividad de PriA en sus dos sustratos nativos. En la segunda parte se muestra una reconstrucción de la estructura tridimensional de PriA basada en la covarianza de sus aminoácidos en secuencias del registro evolutivo.

Al transformar una subHisA en una PriA mediante mutaciones no se observó ninguna trayectoria creciente para ambos sustratos

En esta sección analizamos cómo cambia la capacidad catalítica de PriA sobre un sustrato mientras se varía la del otro. Para ello se utilizaron datos de mutantes de subHisA de *Corynebacterium diphtheriae*. Estas mediciones de cinéticas enzimáticas fueron obtenidos del trabajo de tesis de Lianet Noda [@noda_tesis_2012]. A partir de la secuencia original que se mostró es una subHisA, se realizaron mutantes con el objetivo de alcanzar la promiscuidad, es decir de convertir la enzima subHisA en una PriA. Se comenzó con diferentes mutantes puntuales adicionando una mutación cada vez, hasta llegar a una con 11 mutaciones. En esta colección de mutantes varias ganaron la función de PRA isomerasa, a distintos niveles. La que alcanzó mayor actividad PRA isomerasa fue la 9.3, una variante con nueve mutaciones. En estos datos quedaba pendiente la exploración de los caminos mutacionales, es decir cómo es el camino desde una mutante sencilla hasta una múltiple ¿cuántas rutas son posibles? ¿Existe alguna tendencia en ciertos momentos de la ruta sobre el incremento/decremento de alguna de las dos funciones?

Así pues se desarrolló un programa utilizando recursividad para reconstruir todas las rutas posibles. El total de rutas calculadas hasta la mutación 11 fue de 2928 caminos, las rutas posibles hasta la mutante 9.3 son 790. Estas rutas son mostradas en la Figure 9.

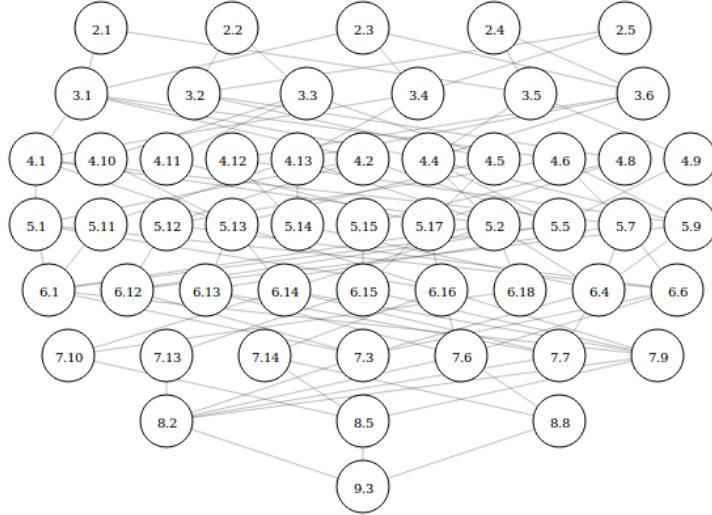


Figure 9: Rutas desde una subHisA hasta una variante con 9 mutaciones. En cada círculo el primer dígito indica el número de mutaciones

Al analizar todas las rutas que llevan a 9.3 se descubrió que no existe en ellas una trayectoria en la que todos los pasos incrementen la actividad de TrpF. Por otra parte, se observa que a pesar de que la primera mutación siempre decrece la actividad de HisA, en pasos subsecuentes se observa que esta actividad se incrementa, lo que muestra que hay epistasis positiva o amortiguadora. Esto es que los efectos negativos que de las mutaciones sobre SubHisA pueden ser menos negativos o hasta positivos cuando las mismas mutaciones ocurren en una proteína que ya tiene otras mutaciones. Como ejemplo, en la Figure 9 se muestran las rutas donde cada mutante mantiene un nivel mínimo de actividad de ProFAR isomerasa ($\frac{K_{cat}}{K_m} PriA_{ProFAR} \geq .004$). En azul sólido se ven los incrementos en PRA y en rojo sólido los incrementos en ProFAR. Las líneas punteadas indican que la actividad decreció en ese paso de la ruta. Entre una y cuatro mutaciones el azul sólido es predominante, es decir se incrementa la actividad para PRA, pero entre 4 y 5 mutaciones ningún paso incrementa la actividad de PRA y en cambio sí se incrementa la actividad para ProFAR, esta figura sugiere que al mejorar una actividad se compromete el mejoramiento de la otra. En este ejemplo, las mutaciones puntuales que llevan a una enzima monofuncional a adquirir promiscuidad no mantienen una tendencia no decreciente de principio a fin sobre ninguna ruta en ninguna de las dos reacciones isomerización de PRA e isomerización de ProFAR. Este tipo de trayectorias se conoce como no darwiniana ya que siempre existe algún paso donde decrece alguna de las actividades.

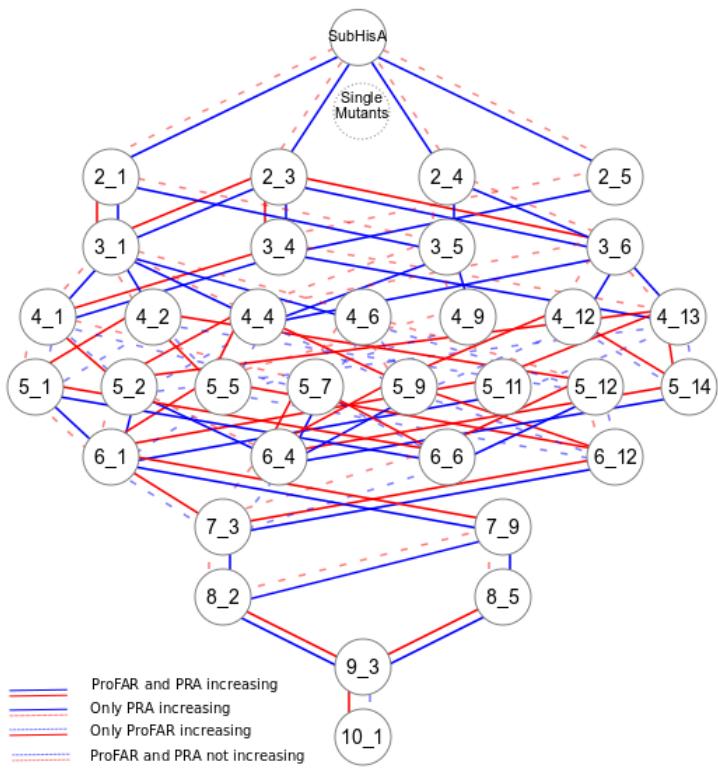


Figure 10: Rutas de evolución dirigida para ganar la función PRA. En la figura se muestra como la mayoría de los pasos que incrementan una función hacen decrecer la capacidad catalítica de la otra. El primer número de cada nodo indica cuántas mutaciones tiene esa variante y el segundo número es el identificador de esa combinación específica.

Los residuos que covariarían en el registro evolutivo de PriA permiten una reconstrucción aproximada de su estructura tridimensional

El estudio de la evolución de PriA en la sección anterior nos proporcionó el aprendizaje de que para adquirir una actividad el camino no es estrictamente creciente. En cambio, suele haber pasos donde alguna de las dos actividades baja. En esta sección utilizaremos el registro evolutivo resultado de millones de años, tomaremos miles de homólogos de PriA para inferir la estructura tridimensional de una secuencia. EVcouplings es un método que considera las secuencias génicas existentes como experimentos exitosos de la naturaleza. Con esta información obtiene la covariación entre pares de aminoácidos de las secuencias existentes en el registro evolutivo. Los pares fuertemente relacionados se denominan acoplamientos, estos acoplamientos a menudo están cerca físicamente en la estructura terciaria de la proteína. Se ha demostrado que muchas proteínas contienen suficientes acoplamientos distribuidos ampliamente en toda la secuencia, de forma que con ellos es posible la reconstrucción de su estructura tridimensional [@marks_protein_2011]. En esta sección se aplicará EVcouplings para reconstruir la estructura tridimensional de PriA.

Las diferencias a nivel estructural pueden amplificar la información proporcionada por variaciones a nivel de secuencia. Una métrica común de distancia entre dos estructuras de proteínas es el *root mean square deviation* (RMSD). Esta distancia se obtiene después de alinear las estructuras calculando el promedio de las distancias al cuadrado entre los átomos de carbono primarios [@kufareva_methods_2012]. La información de las covariación entre residuos es suficiente para que el método EVcouplings genere estructuras con un error de 2.7–4.8 Å respecto a la estructura cristalográfica conocida [@marks_protein_2011]. Por este motivo se decidió implementar EVcouplings y aplicarlo a la familia PriA. Este método es de difícil instalación ya que requiere varias dependencias. Por ello, desarrollé un contenedor Docker donde dependencias, software y base de datos quedan instalados. Este desarrollo fue incluido por los autores de EVcouplings como sugerencia de instalación. El contenedor docker implementa el *EVcouplings python framework* [@hopf_evcouplings_2019] que comprende cinco etapas para estudiar el análisis de coevolución de residuos de una familia de proteínas. Dichas etapas son i) Alineado, ii) análisis de acoplamiento, iii) plegamiento basado en acoplamientos iv) análisis de mutación y v) comparación con estructuras conocidas.

EVcouplings fue aplicado a la secuencia de PriA de *Streptomyces coelicolor* obtenida de la base de datos Uniprot con identificador HIS4_STRCO. Los parámetros de EVcouplings se dejaron con su configuración inicial incluyendo el umbral de recuperación para las secuencias utilizadas en el alineamiento. Con este alineamiento, utilizando la información de los acoplamientos de sus aminoácidos, se obtuvo el modelo *PriAEV* de la estructura tridimensional de PriA de *S. coelicolor*.

Ahora bien, 1VZW es una estructura de PriA de *S. coelicolor* obtenida mediante cristalográfia [@kuiper_twofold_2005]. En esta tesis calculé que el RMSD entre PriA (1VZW) y su homólogo lejano HisF (2A0N proveniente de *Thermotoga maritima* es de 7.34Å. Así mismo la distancia obtenida entre PriA (1VZW) y el modelo de TrpF 5LHE de *Thermococcus kodakaraensis* es de 12.088 Å. Por tanto cualquier modelo existoso de PriA debe tener un RMSD menor que estas distancias con respecto al modelo 1VZW. La distancia entre la estructura PriAEV y HisF es de 5.603 Å y el RMSD entre PriAEV y TrpF es de 12.554 Å. Además, la distancia entre estas 1VZW y PriAEV medida utilizando el RSMD (pymol 2.2.3) entre ellas es de 3.362 Å. Así pues PriAEV consigue diferenciarse de HisF y TrpF y estar a una distancia de 3.362 Å de 1VZW. Esta distancia es menor que que el 3.730 Å el RMSD entre la estructura de PriA 1VZW de *S. coelicolor* y la estructura de PriA *Streptomyces sviceus ATCC 29083* con identificador de PDB 4TX9. Además, las estructuras cristalográficas tienen una resolución de 1-3 Å como puede verse en la Table 4, por tanto un RMSD de 3 Å es aceptable para decir que las estructuras son similares. Sin embargo los RMSD entre otras estructuras cristalográficas de *Streptomyces coelicolor* son menores a 2Å, por ejemplo el (RMSD (2VEP, 1VZW) = 0.449Å, RMSD (1VZW, 2X30)= 1.929 Å, RMSD (1VZW,5DN1)=2.802 Å). Así pues EVcouplings obtuvo una estructura de PriA de *S. coelicolor* similar a la cristalográfica 1VZW, pero aparentemente no mejor que otras estructuras cristalográficas de *S. coelicolor*. La comparación de la estructura PriAEV lograda por EVcouplings con respecto a 1VZW es mostrada en la Figure 11.

La importancia de haber montado EVcouplings en el Laboratorio de Evolución de la diversidad metabólica es que ahora puede obtenerse un modelo tridimensional para cada secuencia de PriA disponible en horas, en contraste con otros métodos computacionales que pueden tardar días. Con estas secuencias se podría

corroborar si diferenciar a nivel familia mediante la información de estructuras tridimensionales es posible para PriA en Actinobacteria. Sin embargo, es posible que el RMSD de la estructura de toda la proteína no sea suficiente para diferenciar por familia, ya sea porque las diferencias estructurales no son suficientes, porque los métodos de construcción de las estructuras necesiten mejorarse, porque se deba refinar la selección de proteínas en el alineamiento diferenciando entre secuencias conocidas de PriA, subHisA, PriB y subTrpF, o bien porque deban considerarse para el RMSD solo las regiones de la estructura con más diferencias conocidas incluyendo todos los átomos de esas regiones en lugar de sólo los átomos de carbono alfa. . En la siguiente sección mediante una colaboración se aborda un poco la dinámica molecular de la familia PriA que es un paso más allá de la comparación estructural.

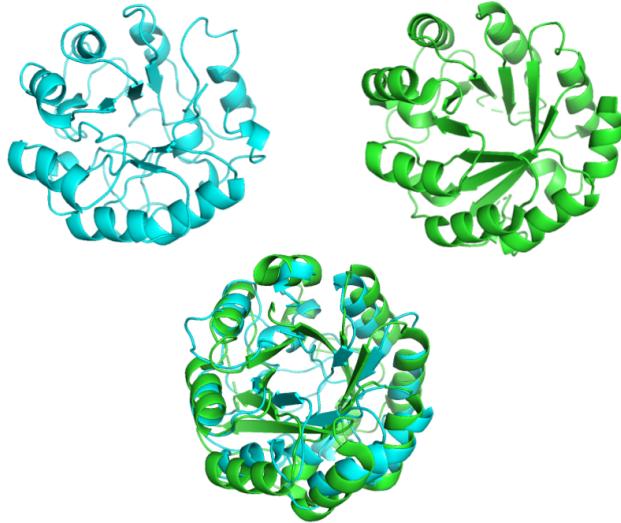


Figure 11: Comparación de estructura tridimensional de PriA generada por EVcouplings con una estructura cristalográfica de *Streptomyces coelicolor*. En azul se muestra la estructura generada considerando acoplamientos de los aminoácidos de PriA según el registro evolutivo. En verde está la estructura 1vzw obtenida experimentalmente. En la parte inferior de la figura se muestran las dos estructuras alineadas. Su RMSD es de 3.362 Å

Finalmente los aminoácidos utilizados en la evolución dirigida de la sección anterior fueron comparados con los provistos por EVcouplings como altamente partícipes en la covariación. Los 10 aminoácidos con más acoplamientos fueron 90L, 117V, 127V, 48W, 208I, 87D, 135T, 21V, 109E. El número indica la posición de la secuencia y la letra es el código de los aminoácidos. Por ejemplo 87D es la posición 87 el aminoácido ácido aspártico. Sólo el 21V es parte los aminoácidos mutados en el estudio previamente descrito. En la Table 2 se muestra en la primera columna los aminoácidos variados en el estudio de mutación dirigida en *Corynebacterium*, en la segunda columna el aminoácido correspondiente en la secuencia de *Streptomyces coelicolor* y finalmente su correspondiente acoplamiento más significativo.

Table 2: Árboles EvoMining de PriA en MicroReact

<i>Corynebacterium</i>	<i>Streptomyces</i>	Acoplamiento más relevante
D20V	21V	173T
L48I	49L	76I
F50L	51L	70V
M66I	67I	80L
T80S	81S	102N
A97C	98C	107A
D127A	128G	164V
A129D	130D	168I
T139L	136L	151T

<i>Corynebacterium</i>	<i>Streptomyces</i>	Acoplamiento más relevante
Y214L	211Y	-
E230A	227A	234E

Una forma gráfica de ver esta tabla es la Figure 12. Esta figura es simétrica porque los acoplamientos son simétricos. Los puntos negros indican acoplamientos entre pares de aminoácidos. Entre más cerca de la diagonal hay más acoplamientos porque son residuos cercanos en la secuencia lineal de la proteína. En las líneas naranjas muestro cómo el aminoácido 21D tiene un acoplamiento con el 173T, que está muy cerca del 175D que ha sido asociado con la actividad de isomerización de PRA, pero no de ProFAR. A futuro, para obtener resultados más precisos sobre la covariación de residuos en Actinobacteria, se debe proveer un alineamiento exclusivo de Actinobacteria.

Finalmente se investigó también la variación individual de los aminoácidos catalíticos D11, D130 [@due_bisubstrate_2011] y D175 reportados en *Mycobacterium tuberculosis*[@verduzco-castro_co-occurrence_2016]. Estas posiciones corresponden en *S coelicolor* a D11, D131 y D171. En este trabajo se encontró que existe un conjunto de secuencias donde estos residuos presentan variantes Figure 13. Los homólogos más relevantes de PriA/HisA abordados en este capítulo junto con las secuencias seleccionadas divergentes en los residuos mencionados fueron alineadas para mostrar la diversidad de las familias PriA/HisA a nivel de aminoácidos. Aunque existen variantes de 131D fuera del phylum Actinobacteria, como en *Pseudomonas*, Cyanobacteria, Enterobacteria, Proteobacteria y Chloroflexi también en Actinobacteria se encontraron 159 *Corynebacterium*, 49 *Streptomyces* y 2 *Actinokineospora* con un residuo diferente al aspártico en la posición 131. Están depositados en la base pública NCBI hasta Enero de 2019, 349 genomas públicos de *Streptomyces* 2019, así pues aproximadamente un 14% tiene una variante en la posición 131. Entre estos *Streptomyces* están el ya mencionado *Streptomyces* CT 34 que posee dos copias de PriA una de ellas muy divergente y el *Streptomyces rimosus* del que se habló en en capítulo anterior porque sintetiza rimosamidas. Es llamativo que las variantes de 131D se encuentran principalmente en *Corynebacterium* y en *Streptomyces* que corresponden a los géneros donde se han ubicado a la familia subHisA y a la familia PriB respectivamente. La variabilidad mostrada en estos géneros podría estar relacionada con la existencia de estas familias. Otros los *Streptomyces* con una variante en 131D son *S. oceanii*, *S. scabiei*, *S. fradiae*, *S. rimosus*, *S. sp. CT34* y *S. MUSC 14* y *S. D11H*. De ellos los tres últimos son además los únicos *Streptomyces* con la variante D11H que comparten con *Actinokineospora auranticolor*. De hecho las tres posiciones conservadas en la mayoría de las PriA presentan variantes en estos cuatro organismos, por lo que sería interesante caracterizarlos bioquímicamente. Entre los organismos con doble copia de PriA están *Sericicoccus marinus*, *Sericicoccus profundi* y *Ornithinimicrobium pekingense*, ambas copias fueron incluidas en el alineamiento. Estas seis secuencias de PriA tienen aspártico en las tres posiciones conservadas, las diferencias que muestran en otras posiciones indican que las copias primarias y las secundarias de cada organismo se agrupan entre sí. La secuencia de *Actinomadura sp. ATCC 39365* involucrada en el BGC ada es similar a la PriA de *S. coelicolor* por lo que posiblemente PriA de *S. coelicolor* también podría presentar esta actividad. Se incluyó en el análisis una secuencia reconstruida del ancestro común de HisA en Bacteria [@plach_long-term_2016], esta secuencia muestra el aspártico en las tres posiciones. En Cyanobacteria la HisA del BGC saxitoxina es prácticamente igual a la de *Cylindrospermopsis raciborskii* CS50 y ambas conservan el aspártico en las tres posiciones. En Archaea las PriA más grandes encontradas *Halorubrum tebenquichense* y *Halototius sp J07HN6* tienen la variante en D171E, que comparten con los *Streptomyces* y la *Actinokineospora* con variante D11H y con los *Streptomyces* con la variante D130E. Finalmente también en Archaea *Thermococcus sp JCM 11816* que comparte el contexto genómico con genes de triptófano presenta los tres residuos conservados.

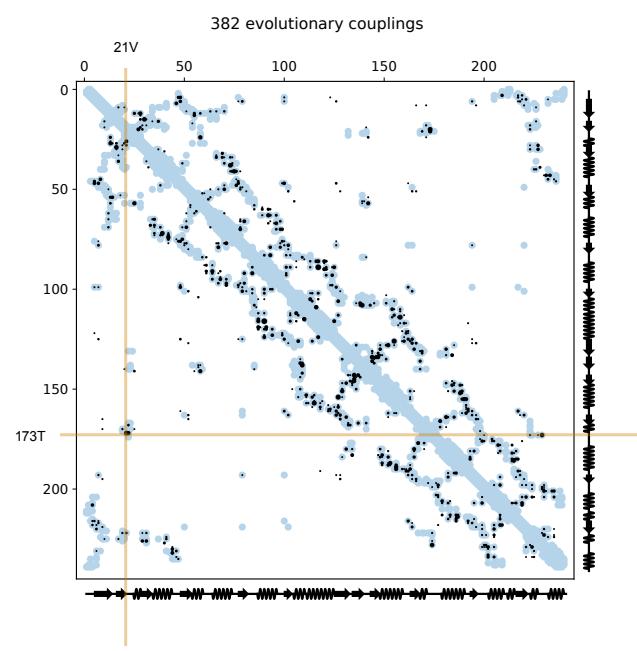


Figure 12: Visualización de los acoplamientos en la familia PriA

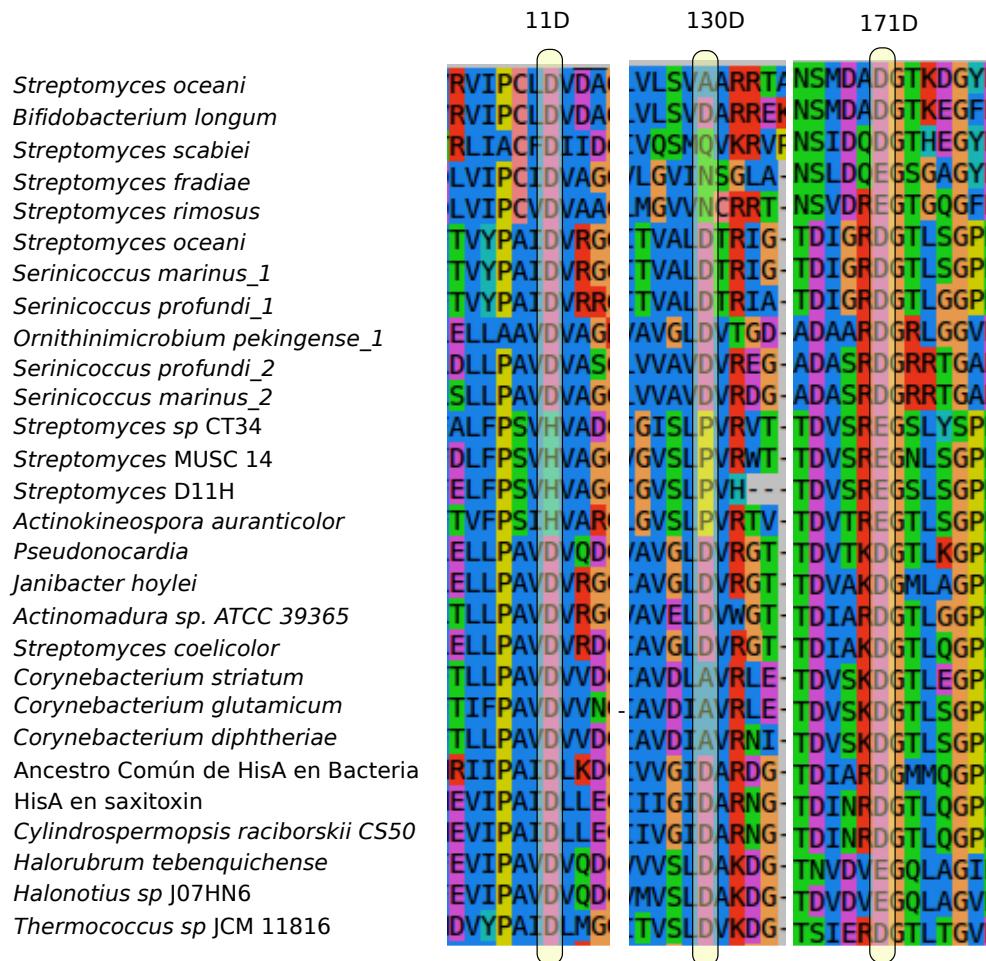
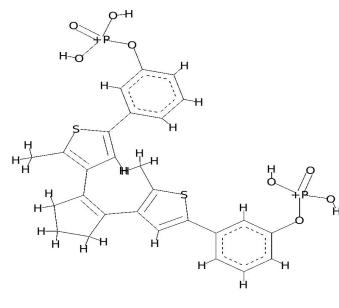


Figure 13: Miembros de PriA que poseen variantes los residuos catalíticos D11, D130

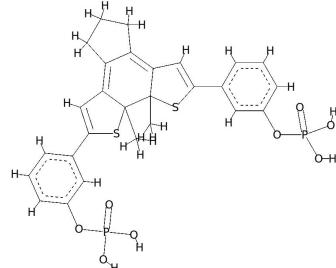
Afinidad de enzimas selectas por sustratos químicamente parecidos a PRA y PROFAR

Además de los sustratos conocidos ProFAR y PRA en los que PriA es capaz de realizar una isomerización, es posible que PriA pueda ser promiscua en otros sustratos. De hecho, como se vio en la sección de EvoMining de este capítulo PriA parece participar en la síntesis del antibiótico pentostaina (*ada* BGC). Tomando este ejemplo como inspiración, se buscaron sustratos químicamente parecidos a ProFAR y PRA. Esta sección buscará probar la afinidad de sustratos parecidos a los nativos de PriA para posteriormente probar alguno en copias selectas de PriA provenientes de diversos organismos. Veinte sustratos (S1, S2, ... S20) fueron recolectados tanto de la literatura [@adams_promiscuous_2014,@due_bisubstrate_2011,@reisinger_exploiting_2014,@verduzco-castro_co-occurrence_2016] como de predicciones químicoinformáticas [@jeffryes_mines_2015]. Estos sustratos son mostrados en Figure 14, Figure 15,Figure 16,Figure 17. Los sustratos nativos son PRA (S3) y el sustrato PROFAR es S7. Por otra parte, S13-S16 son sustratos activados por luz. Los sustratos S17 (PRAP) y S18 (Compuesto V) se encontraron en la literatura, mientras que S6 (GMP), S11 (GTP) y otros fueron sugeridos por el grupo del dr. Chris Henry debido a sus similitudes químicoinformáticas. En la siguiente sección, posteriormente a la selección de sustratos seleccionaremos las secuencias de PriA con las cuales se realizará el docking enzima-sustrato.



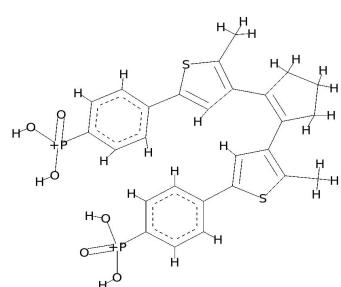
13

DTE-meta-phosphate(dte6_Open form)



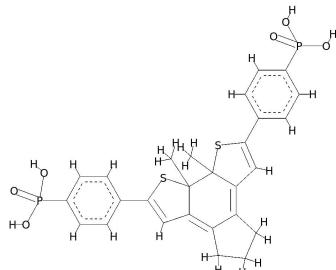
14

DTE-meta-phosphate(dte6_Closed form)



15

DTE-Para-Phosphonate(dte13_Closed form)



16

DTE-para-phosphonate(dte13_Closed form)

Figure 14: Substatos químicamente similares a los sustratos de PriA (parte 1)

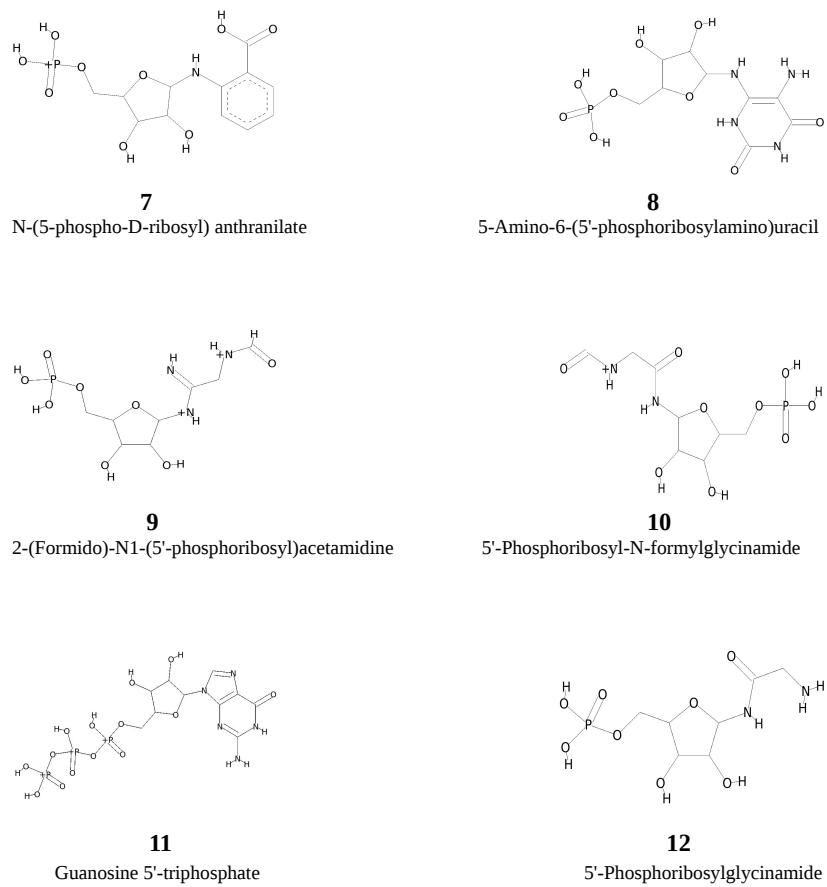
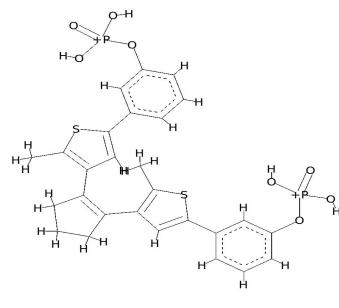
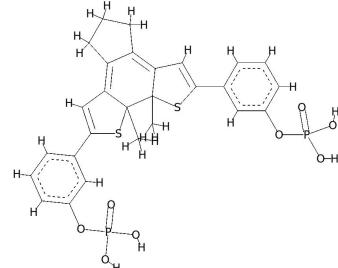


Figure 15: Substatos químicamente similares a los sustratos de PriA (parte 2)



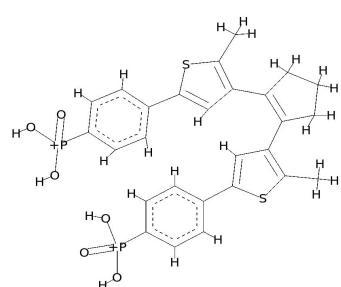
13

DTE-meta-phosphate(dte6_Open form)



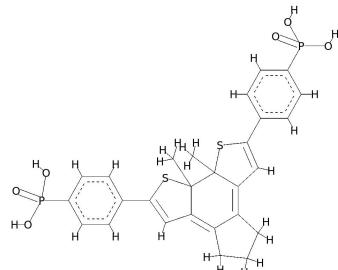
14

DTE-meta-phosphate(dte6_Closed form)



15

DTE-Para-Phosphonate(dte13_Closed form)



16

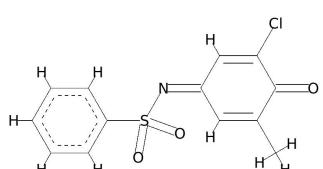
DTE-para-phosphonate(dte13_Closed form)

Figure 16: Substatos químicamente similares a los sustratos de PriA (parte 3)

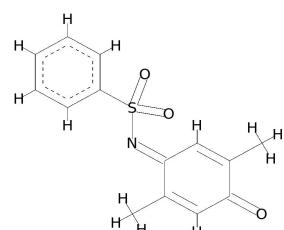


17
4N'-(5'-phosphoribosyl) 4-aminopyrrole
-2-carboxilate

18
2,5-di-amino-6-ribosylamino-4
(3H)-pyrimidinone 5'-phosphate



19
(E)-N-(3-chloro-5-methyl-4-oxocyclohexa-
2,5-dienylidene)benzenesulfonamide



20
2,5 dimethyl-N-(4-oxocyclohexa-
2,5-dienylidene)benzenesulfonamide

Figure 17: Substatos químicamente similares a los sustratos de PriA (parte 4)

Selección de secuencias de PriA o familias relacionadas para docking con sustratos similares a lo nativos

Una vez escogidos los sustratos quedaban por definir las secuencias con las que se llevaría a cabo el análisis bioinformático de acoplamiento enzima-sustrato, lo que hicimos en colaboración con el grupo del dr. Carrillo-Trip. Para ello seleccioné 39 secuencias de la familia PriA y sus subfamilias, entre ellas incluí varios *Streptomyces* para tener tanto secuencias pertenecientes a PriA y como a la subfamilia PriB. Estas secuencias seleccionadas de PriA / PriB están uniformemente distribuidas en *Streptomyces* de acuerdo a un árbol de especies de *Streptomyces* basado en la proteína RpoB. Además estos *Streptomyces* tienen variedad en cuanto a la presencia / ausencia de *trpF* en su genoma. Se incluyeron además otros homólogos de PriA que han sido caracterizados químicamente y finalmente se con el mismo criterio de caracterización bioquímica se agregaron secuencias de HisA de *Escherichia coli*, *Arthrobacter aurescens*, *Salmonella enterica* y *Acidimicrobium ferrooxidans* y secuencias TrpF provenientes *Jonesia denitrificans* y *Streptomyces sp Mg1* para ser utilizadas como controles.

Para realizar el docking se requieren estructuras cristalográficas. Cuando existían estructuras cristalográficas de la secuencia específica de PriA del organismo se utilizó dicha estructura. En caso contrario, el grupo del Dr Carrillo-Tripp generó estructuras homólogas utilizando la técnica de modelado por homología mediante el software Rosetta. Para cada secuencia decidí usar como plantilla la estructura tridimensional del homólogo de PriA más cercano que sí contara estructura con estructura cristalográfica. Los organismos con estructura cristalográfica de alguna familia relacionada a PriA están descritos en la Table 4. Para este análisis fueron las secuencias con estructura(s) cristalográfica(s) seleccionadas fueron: para la familia HisA la Enterobacteria *Salmonella enterica* (PDB:5AHE), para la familia subHisA la Actinobacteria *Actinomyces urogenitalis* (PDB:4X2R), la estructura que representa a subTrpF es la de *Arthrobacter aurescens* (PDB:4WD0). En cuanto a la familia PriA, varias estructuras cristalográficas de *Mycobacterium tuberculosis* (Mtub PDB:2Y88,2Y89,2Y85,3ZS4) y *Streptomyces coelicolor* (Scoe PDB:2VEP,2X30,1VZW) fueron incluidas. Las estructuras de PriB consideradas fueron las de *Streptomyces sviecius* (PDB:4U28,4TX9) y las de *Streptomyces sp Mg1* (4W9T, 4X9S). Finalmente la estructura cristalográfica de TrpF corresponde a *Jonesia denitrificans* (PDB:4WUI). Table 3.

Se incluyeron también enzimas TrpF provenientes de *Streptomyces Mg1*, *Jonesia denitrificans*. Las estructuras cristalográficas disponibles de PriB provienen de . . *Streptomyces globisporus*, *Actinomyces urogenitalis* (4X2R) y *Corynebacterium jeikeum*. De la familia subHisA se muestran *Corynebacterium efficiens* y . y finalmente, *Chlamydia trachomatis*, *Streptomyces sp. Mg1* y *Actinomyces odontolyticus*. Las representantes de la familia TrpF son las Actinobacterias *Jonesia denitrificans* y *Streptomyces sp Mg1*. A continuación se muestra una tabla con las estructuras cristalográficas disponibles de PriA y familias relacionadas.

Table 3: Secuencias selectas de PriA y familias relacionadas para el análisis de docking

Abreviatura	Organismo
Save	<i>Streptomyces avellaneus</i>
Scar	<i>Streptomyces carneus</i>
Spur	<i>Streptomyces purpeofuscus</i>
Smeg	<i>Streptomyces megasporus</i>
Sfrad	<i>Streptomyces fradiae</i>
Svar	<i>Streptomyces varsoviensis</i>
Satra	<i>Streptomyces atratus</i>
Srim	<i>Streptomyces rimosus</i> R6-500
SCT34	<i>Streptomyces sp. CT34</i>
S1813	<i>Streptomyces sp. NRRL S-1813</i>
Ssul	<i>Streptomyces sulphureus</i> DSM 40104
Scla	<i>Streptomyces clavuligerus</i> ATCC 27064
Stsu	<i>Streptomyces tsukubaensis</i> NRRL18488
Sbik	<i>Streptomyces bikiniensis</i>

Abreviatura	Organismo
Sven	<i>Streptomyces venezuelae</i> ATCC 10712
Sful	<i>Streptomyces fulvoviridis</i>
Scal	<i>Streptomyces californicus</i>
Sbaa	<i>Streptomyces baarnensis</i>
Salb	<i>Streptomyces albus</i> J1074
Siak	<i>Streptomyces iakyurus</i>
Sgha	<i>Streptomyces ghanaensis</i> ATCC 14672
Sbic	<i>Streptomyces bicolor</i>
Sipo	<i>Streptomyces ipomoeae</i> 91-03
Sbot	<i>Streptomyces bottropensis</i> ATCC 25435
SspC	<i>Streptomyces</i> sp. C
SMg1	<i>Streptomyces</i> sp. Mg1
Sxan	<i>Streptomyces xanthophaeus</i>
Skat	<i>Streptomyces katrae</i>
Slav	<i>Streptomyces lavendulae</i> subsp. <i>lavendulae</i>
Sery	<i>Streptomyces erythrochromogenes</i>
Sniv	<i>Streptomyces niveus</i> NCIMB 11891
Ssvi	<i>Streptomyces sviceus</i> ATCC 29083
Sfla	<i>Streptomyces flaveus</i>
Save	<i>Streptomyces avermitilis</i> MA-4680 = NBRC 14893
Scoe	<i>Streptomyces coelicolor</i> A3(2)
ArTC1	<i>Arthrobacter aurescens</i> TC1
Sglob	<i>Streptomyces globisporus</i> C-1027
Sgri	<i>Streptomyces griseolus</i>
S34	<i>Streptomyces</i> sp. CT34 (paralogo)

Table 4: Estructuras cristalográficas disponibles de PriA y familias relacionadas

Organismo	PDB	Familia	Resolución Å	Fecha
<i>Salmonella enterica</i>	5AHE	HisA	1.7	2015
<i>Salmonella enterica</i>	5AB3	HisA	1.8	2016
<i>Salmonella enterica</i>	5ABT	HisA	1.65	2016
<i>Salmonella enterica</i>	5AC7	HisA	1.9	2016
<i>Salmonella enterica</i>	5AC8	HisA	1.7	2016
<i>Salmonella enterica</i>	5AC6	HisA	1.99	2016
<i>Salmonella enterica</i>	5A5W	HisA	1.6	2015
<i>Salmonella enterica</i>	5AHF	HisA	2.2	2012
<i>Thermotoga maritima</i>	2W79	HisA	1.85	2008
<i>Thermotoga maritima</i>	1QO2	HisA	1.85	2000
<i>Streptomyces</i> sp. Mg1	4X9S	PriB	1.6	2014
<i>Streptomyces sviceus</i>	4TX9	PriB	1.6	2014
<i>Streptomyces sviceus</i>	4U28	PriB	1.33	2014
<i>Streptomyces</i> sp. Mg1	4W9T	PriB	1.57	2014
<i>Arthrobacter aurescens</i>	4WD0	subTrpF	1.5	2014
<i>Streptomyces coelicolor</i>	5DN1	PriA	1.95	2015
<i>Streptomyces coelicolor</i>	1VZW	PriA	1.8	2004
<i>Streptomyces coelicolor</i>	2VEP	PriA	1.8	2007
<i>Streptomyces coelicolor</i>	2X30	PriA	1.95	2010
<i>Mycobacterium tuberculosis</i>	2Y85	PriA	2.4	2011
<i>Mycobacterium tuberculosis</i>	2Y88	PriA	1.33	2011

Organismo	PDB	Familia	Resolución Å	Fecha
<i>Mycobacterium tuberculosis</i>	2Y89	PriA	2.5	2011
<i>Mycobacterium tuberculosis</i>	3ZS4	PriA	1.9	2012
<i>Actinomyces urogenitalis</i>	4X2R	SubHisA	1.05	2014
<i>Corynebacterium efficiens</i>	4AXK	SubHisA	2.25	2013
<i>Thermococcus kodakaraensis</i>	5LHE	TrpF	1.85	2016
<i>Thermococcus kodakaraensis</i>	5LHF	TrpF	1.75	2016
<i>Thermus thermophilus</i>	1V5X	TrpF	2	2003
<i>Thermotoga maritima</i>	1DL3	TrpF	2.7	1999
<i>Thermotoga maritima</i>	1LBM	TrpF	2.8	2002
<i>Thermotoga maritima</i>	1NSJ	TrpF	2	1996
<i>Jonesia denitrificans</i>	4WUI	TrpF	1.09	2014
<i>Pyrococcus furiosus</i>	4AAJ	TrpF	1.75	2012

El análisis de PriA a nivel estructural sugiere que GTP es el sustrato más afín

Con las estructuras de las enzimas seleccionadas de PriA se realizaron simulaciones de docking respecto a los veinte sustratos descritos. El procedimiento puede ser consultado en Docking Protocols. En la Figure 18 presento una visualización de los resultados obtenidos. En esta figura se marcan en color los organismos de los que se tiene una caracterización bioquímica, en azul están las secuencias correspondientes a TrpF, en rojo las de HisA, en rosa las secuencias de PriA y subHisA y en verde las de PriB y subTrpF. Con un asterisco fueron marcadas las secuencias de las que se cuenta con estructura cristalográfica. En el caso de las secuencias que no contaban con estructura cristalográfica, la estructura que sirvió como base está anotada después de la abreviatura del organismo. Las abreviaturas corresponden a los organismos de la Table 3. Los cuadrados grises son datos que faltaron de simular. Entre más azul es mayor la afinidad del sustrato por la secuencia, entre más rojo menor la afinidad de la enzima por este sustrato. Se requieren más controles para este experimento, pero una tendencia general es que la columna correspondiente a S11 (GTP) es el sustrato que tiene mayor afinidad. Por ello, aunque afinidad no tiene porque corresponder con capacidad catalítica, se decidió investigar la capacidad de secuencias de PriA para catalizar alguna reacción sobre el GTP, esa investigación es el objeto de la última sección de este capítulo.

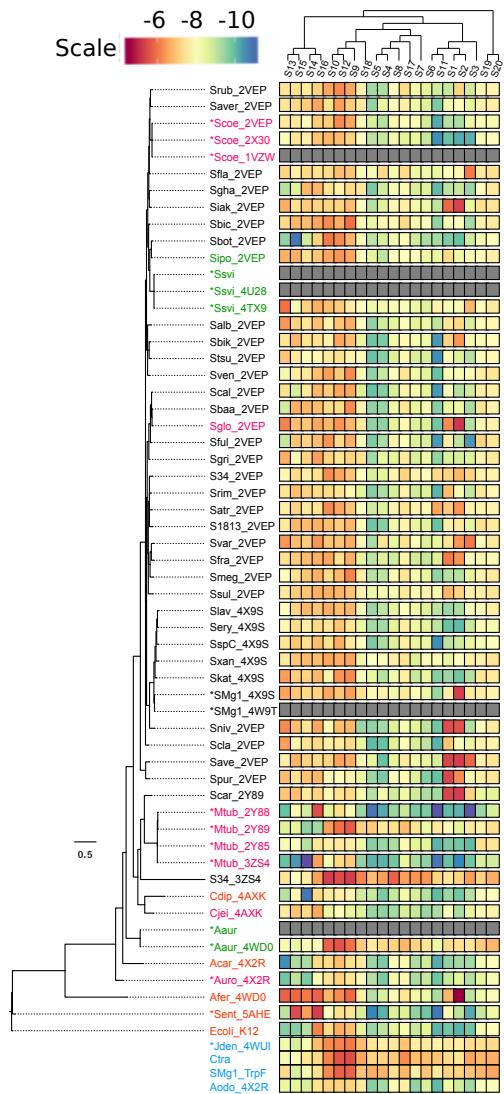


Figure 18: Heatplot del docking de enzimas relacionadas a PriA vs posibles substratos

PriA en cinéticas enzimáticas no tradicionales.

Después de la exploración genómica de PriA/HisA se trabajó en caracterizaciones experimentales no tradicionales de PriA. No se tuvo el tiempo ni la experiencia para tener réplicas de los resultados descritos a continuación, sin embargo se incluyen en esta tesis porque pueden ser un buen comienzo para luego retomar este trabajo. Los protocolos fueron cuidadosamente descritos y se encuentran incluidos en los anexos. Este capítulo se enfoca en el montaje de dos experimentos, el primero sugerido por la sección anterior: *i)* Cinéticas de PriA en GTP. Mientras que el segundo está motivado por la curiosidad sobre el cambio en la capacidad catalítica según se provea primero un sustrato u otro, este experimento consiste en la *ii)* Medición simultánea de la actividad de PriA sobre ProFAR y PRA.

Lo primero que se realizó fue la sobre-expresión heteróloga de PriA. La secuencia de PriA fue clonada en cepas de *E. coli* V68 donde se indujo la sobre expresión. A partir de este cultivo productor de PriA se purificó la proteína. Se corroboró el éxito de esta actividad mediante un gel de proteína en la que puede verse expresión en la barra correspondiente al tamaño de PriA Figure 19. La proteína purificada fue almacenada a -80° en esferas de proteína selladas con nitrógeno líquido para su mejor preservación, tal y como se describe en el anexo de los procedimientos. Con este *stock* de proteína se realizaron cinéticas enzimáticas *in vitro*.



Figure 19: Gel de proteína donde se muestra la banda correspondiente a PriA después de ser sobre-expresada y purificada

PriA puede metabolizar GTP

Como se sugirió en la sección anterior PriA puede tener actividad sobre GTP. Ahora bien, la actividad de PriA en ProFAR es medida en un ensayo estandarizado mediante la detección de cambio de fluorescencia. Para esta medición por fluorimetría se coloca la enzima en placas de 96 pozos (Nuc 96-Well Optical Bottom Plates), se agrega el sustrato y utilizando un lector de fluorescencia de placas (TECAN infinite M1000) se mide el cambio de fluorescencia en el tiempo. Una variación de este ensayo permitió obtener información sobre la actividad PriA sobre dGTP. Los parámetros utilizados para la medición fueron tomados del trabajo no publicado de Verduzco-Castro en este laboratorio (excitación a 255 nm y emisión a 334 nm). Para comenzar esta investigación se utilizó primero dGTP en lugar de GTP esto debido a que el dGTP es un sustrato siempre disponible en cualquier laboratorio ya que es necesario para la reacción de PCR. Así pues, ensayos preliminares de actividad de PriA sobre este dGTP se realizaron en la proteína purificada de PriA de *Streptomyces coelicolor*, activa en ambos sustratos nativos utilizando como control su mutante inactiva que sustituye en la posición 11 el ácido aspártico por una alanina (D11A). La Figure 20 muestra el resultado

de este experimento donde se aprecia un decrecimiento mayor en la fluorescencia del pozo que contiene la proteína activa respecto del pozo que contiene a la mutante.

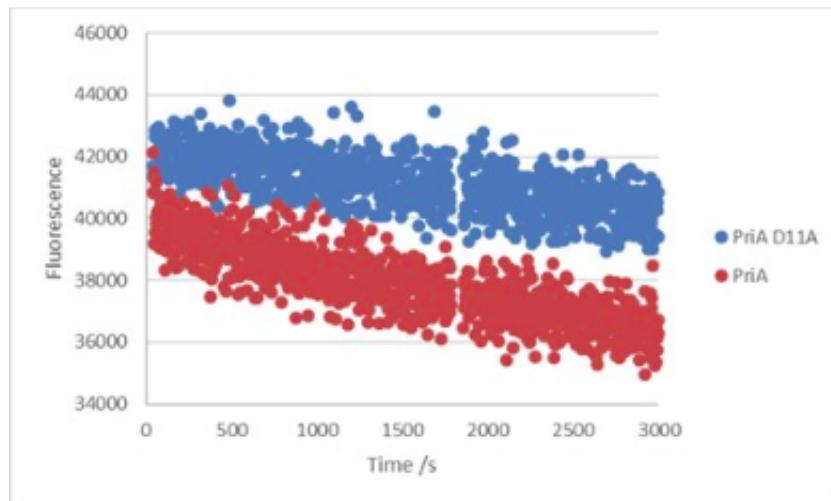


Figure 20: PriA de *S.coelicolor* y su mutante no funcional sobre dGTP

Este experimento de detección de actividad de dGTP fue ampliado incluyendo ahora proteínas provenientes de otros organismos. Para ello se utilizaron los stocks disponibles de PriA provenientes de *Streptomyces roseus*, *Thermomonospora curvata* y *Mycobacterium smegmatis*; los stocks de PriB provenientes de *Streptomyces sp C*, *Streptomyces sviceus* y de *Streptomyces sp Mg1* y finalmente el stocks de TrpF de *Jonesia denitrificans*. El ensayo descrito arriba fue realizado en estas enzimas y el resultado es mostrado en Figure 21. En la parte inferior derecha vemos que como debe ser el control negativo, el pozo con Buffer y sustrato no muestra ningún cambio en la fluorescencia. Sin embargo no es claro que exista una tendencia claramente decreciente en los pozos que contienen enzima. Posiblemente la pendiente que se observa en los datos de *Thermomonospora curvata*, indique cierta actividad sobre dGTP. Este comportamiento también puede deberse a que las proteínas llevaban tiempo almacenadas. Los experimentos deben ser repetidos con proteína recién purificada.

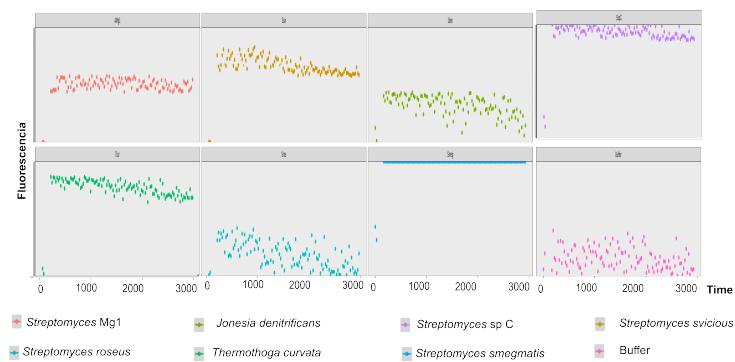


Figure 21: Actividad de PriA y familias relacionadas de distintos organismos sobre dGTP. Este experimento no muestra disminución de fluorescencia en las enzimas seleccionadas lo que implicaría que no existe actividad importante sobre dGTP, excepto posiblemente en el caso de *Thermomonospora curvata*

Finalmente se probó con los mismos parámetros la actividad de PriA de *Streptomyces coelicolor* sobre GTP. Se comprobó que la enzima estuviera activa mediante ensayos enzimáticos exitosos en sus otros dos sustratos. Sin embargo, no se pudo detectar cambio en la fluorescencia de GTP debido a la actividad de PriA. Es posible que la PriA de *S coelicolor* no pueda catalizar ninguna reacción sobre el GTP por lo que una variante que a futuro podría aportar a este experimento es utilizar análogos fluorescentes de GTP.

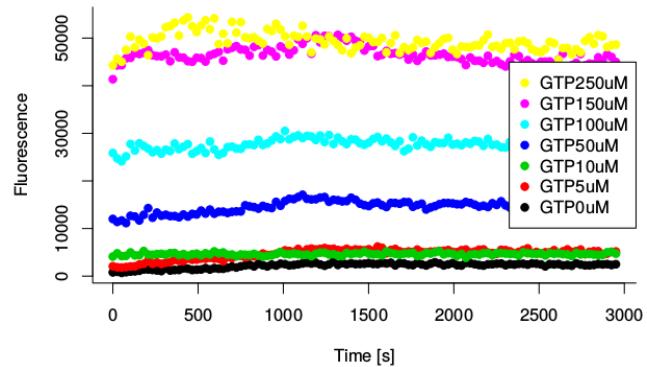


Figure 22: PriA no mostró actividad sobre GTP. El ensayo enzimático de PriA de *Streptomyces coelicolor* sobre GTP a distintas concentraciones no mostró cambios en la fluorescencia de GTP.

Cinéticas simultáneas para PRA y ProFAR

Otra pregunta que me surgió al investigar la promiscuidad enzimática es si esta es una propiedad de la población o de cada una de las enzimas existentes en la célula. En PriA esta pregunta puede traducirse a ¿cómo saber si cada molécula de PriA cataliza PRA y ProFAR *in vivo*? o si en cambio, existe una subpoblación que cataliza la isomerización de PRA y otra subpoblación de moléculas que cataliza isomerización de ProFAR. Una buena aproximación *in vitro* a este problema sería poder medir el comportamiento de una sola molécula en un medio con ambos sustratos. Una serie de estos experimentos mostraría si existen subpoblaciones dedicadas a la catálisis de uno u otro sustrato, o bien si todas las enzimas de una población son igualmente promiscuas. La existencia de subpoblaciones podría deberse a interacciones de la enzima con el sustrato, como por ejemplo que agregar PRA primero al medio sesgue el comportamiento de la enzima y la haga mejor para catalizar PRA que ProFAR o visceversa.

Para empezar a abordar este problema quisimos medir la actividad de una población de PriA sobre PRA y ProFAR. or separado, ambos ensayos están montados, pero nunca se ha intentado un ensayo simultáneo. Las constantes catalíticas de PriA para la isomerización de ProFAR se mide utilizando absorbancia mientras que el ensayo enzimático de isomerización de PRA se realiza mediante la medición de fluorescencia. En esta sección se empezó a desarrollar un método para medir simultáneamente las actividades de Pra y ProFAR. Primero intentamos medir ambas actividades secuencialmente pero monitorieando que la señal de la otra actividad no se viera afectada. Para esto agregamos la enzima entre los segundos 100 y 400 en los pozos que ya tienen todo para que se lleve a cabo la catálisis de PRA y observamos cómo la fluorescencia baja a todas las concentraciones mientras que la absorbancia se mantiene casi constante (FIGURA-N). Luego cerca del segundo 1500 agregué el sustrato ProFAR y se observa que la absorbancia cae mientras que la fluorescencia en los mismos pozos permanece casi constante. El siguiente paso fue medir simultáneamente la actividad de una población de PriA sobre PRA y ProFAR y aunque si lo medimos no pudimos optimizarlo lo suficiente para medir las constantes catalíticas (datos no mostrados). Las pruebas piloto sugieren que es posible cuantificar la catálisis de ProFAR y PRA simultáneamente y en el caso de los homólogos que también toman a GTP se podrían medir incluso las tres actividades catalíticas para identificar si hay interacciones entre los sustratos.

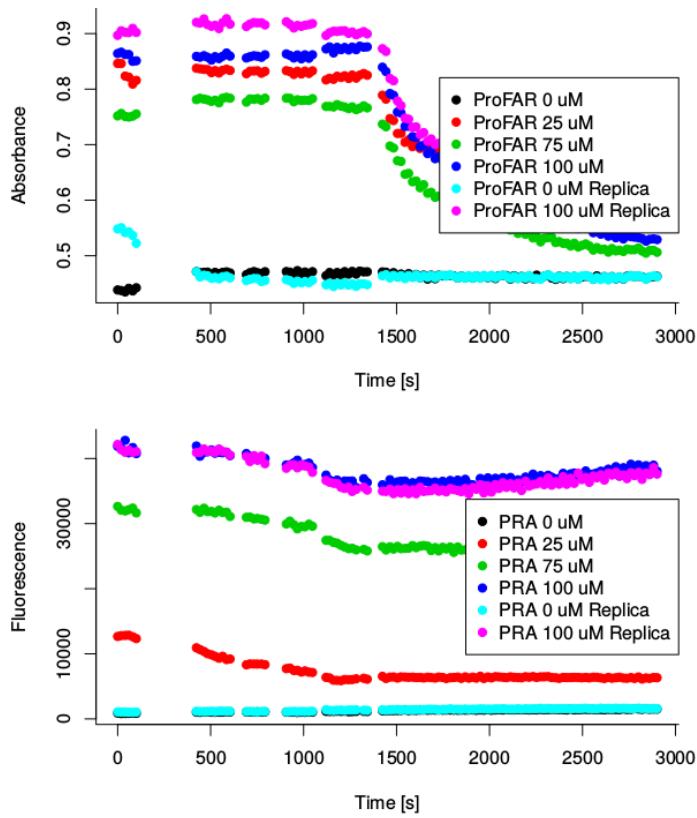


Figure 23: Se pueden medir las dos actividades de PriA simultáneamente. Se midió secuencialmente la actividad de ProFAR isomerasa y de PRA isomerasa en la misma población de enzimas PriA. Se cuantificó la absorbancia (arriba) y la fluorescencia (abajo) durante casi 3000 segundos. entre los segundos 100 y 400 se agregó la enzima a un buffer que ya tenía PRA y al segundo 1400 se agregó el sustrato ProFAR a la misma preparación.