

La familia PriA

PriA es una familia de enzimas de Actinobacteria homóloga a la familia HisA en Enterobacteria. Según las definiciones de este trabajo PriA es una familia promiscua. Varios miembros de PriA han sido sujetos a estudios de cinéticas enzimáticas demostrándose su capacidad de catalizar tanto la reacción correspondiente a HisA como la de TrpF. Es decir PriA participa a la vez en las rutas de síntesis de Histidina y Triptofano, al menos en varias Actinobacterias. Los primeros miembros caracterizados como promiscuos en esta familia en el 2003 fueron *Streptomyces coelicolor* y *Mycobacterium tuberculosis* [@baronagomez_occurrence_2003]. La mayoría de las actinobacterias han perdido el gen trpF en la ruta del triptofano, por lo que se cree que la promiscuidad de PriA está extendida en un gran subconjunto de Actinobacteria. PriA en la ruta de histidina isomeriza el sustrato ProFAR en PRFAR, realizando la función HisA. En la ruta de triptofano lleva a cabo la isomerización de PRA en CdRP actuando análogamente a la función TrpF en la ruta del triptófano. En *Streptomyces coelicolor* varios genes de los operones de histidina y triptofano están en la vecindad genómica de PriA.

Además, en Actinobacteria PriA ha mostrado un gradiente funcional. Esta variación divide a la familia en varias subfamilias según su capacidad catalíticas en los sustratos ProFAR y PRA. En la section se muestra la constantes catalíticas de PriA estimadas en diferentes organismos para estos dos sustratos. Estos datos ejemplifican las subfamilias de PriA, en la tabla se encuentran ejemplos 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 *Actinomyces*. Finalmente en *Actinomyces* encontramos la subfamilia *subTrpF* que ha perdido la actividad de HisA.

Fuente	Familia	HisA <i>in vivo</i>	TrpF <i>in vivo</i>	$K_{cat}^{ProFAR}[M^{-1}s^{-1}]$	$K_m^{ProFAR}[\mu M]$	$\frac{K_{cat}}{K_m} ProFAR$	$K_{cat}^{PRA}[M^{-1}s^{-1}]$	$K_m^{PRA}[\mu M]$	$\frac{K_{cat}}{K_m} PRA$	Referencia
Afer	HisA	1.1 ± 0.2	0.05 ± 0.001	0.045	-10.1	-12.3	Henn [?]			
Ecoli	HisA	1.6	4.9	3.1	-9.9	-16	Henn-Sax et al. (2002)			

Table 1: Datos Cinéticos de subfamilias de PriA

Fuente	Familia <i>in vivo</i>	HisA							Referencia
		-	-	1.6	4.9	3.1	-	-	
<i>Escherichia coli</i>	HisA	-	-	1.6	4.9	3.1	-	-	0 Henn [@henn-sax_two_2002]
<i>Escherichia coli</i>	TrpF	-	-	-	-	0	12.2	34.5	2.82 Sterner [@sterner_phosphoribosyl_1]
<i>Mycobacterium smegmatis</i>	PriA	*	*	2.6 ± 0.5	0.85±0.04	0.33	7.9 ± 2.4	3.1± 0.43	0.39 Verduzco [@verduzco-castro_co-occurrence_2016]
<i>Mycobacterium tuberculosis</i>	PriA	-	-	19	0.23	12	21	3.6	0.17 Due [@due_bisubstrate_2011]
<i>Streptomyces globisporus</i>	PriA	*	*	4.2 ± 0.8	0.74 ± 0.03	0.18	11 ± 1.0	3.8 ± 0.2	0.34 Verduzco [@verduzco-castro_co-occurrence_2016]
<i>Streptomyces coelicolor</i>	PriA	-	-	3.6 ± 0.7	1.3 ± 0.2	0.4	5.0 ± 0.08	3.4 ± 0.09	0.7 Noda [@noda-garcia_identification_2010]
<i>Streptomyces ipomoeae</i>	PriB	*	*	3.8 ± 0.2	0.82 ± 0.02	0.21	60.8 ± 1.1	8.25 ± 0.4	0.14 Verduzco [@verduzco-castro_co-occurrence_2016]
<i>Streptomyces Mg1</i>	PriB	*	*	13.2 ± 3.4	0.92 ± 0.19	69	129.6 ± 34	0.29 ± 0.04	0.0022 Verduzco [@verduzco-castro_co-occurrence_2016]
<i>Streptomyces</i> sp. C	PriB	*	*	11.4 ± 3.4	2.53 ± 0.74	0.22	149. 9 ± 29	1.4 ± 0.12	9 Verduzco [@verduzco-castro_co-occurrence_2016]
<i>Streptomyces sviceus</i>	PriB	*	*	3.9 ± 0.89	0.69 ± 0.04	0.18	24.5 ± 4.0	1.6 ± 0.29	67 Verduzco [@verduzco-castro_co-occurrence_2016]
<i>Corynebacterium diphtheriae</i>	subHisA-	-	-	4.4 ± 0.5	2.6 ± 0.3	0.59			0 Noda [@noda-garcia_evolution_2013]
<i>Corynebacterium jeikeium</i>	PriA	-	-	2.3 ± 0.2	0.9 ± 0.08	0.39	5.1 ± 1.0	1.6 ± 0.16	0.31 Noda [@noda-garcia_evolution_2013]
<i>Corynebacterium striatum</i>	subHisA-	-	-	6.9 ± 0.7	2.1 ± 0.5	0.3			0 Noda [@noda-garcia_evolution_2013]
<i>Corynebacterium diphtheriae</i> L48I-F50L-T80S	subHisA			4.5 ± 1.5	0.6 ± 0.08	0.13	133 ± 10	0.05 ± 0.01	0.0004 Noda [@noda-garcia_evolution_2013]

Fuente		HisA								Referencia
		Familia	<i>in vivo</i>							
<i>Actinomyces urogenitalis</i> DSM 15434	PriB	*	*	2.1 ± 0.5	1.8 ± 0.2	0.9	26.3 ± 6.3	0.37 ± 0.09	14	Verduzco [@verduzco-castro_co-occurrence_2016]
<i>Actinomyces odontolyticus</i> ATCC 17982	subTrpF	*	-	-	-	0			0.02	Juarez [@juarez-vazquez_evolution_2017]
<i>Actinomyces oris</i> K20 BABV01	PriA					0.02			0.01	Juarez [@juarez-vazquez_evolution_2017]
<i>Actinomyces</i> sp. oral taxon 171 str. F0337	PriA					0.01			4	Juarez [@juarez-vazquez_evolution_2017]
<i>Actinomyces</i> sp. oral taxon 848 str. F0332	subTrpF	*	-	-	-	0			0.0001	Juarez [@juarez-vazquez_evolution_2017]
<i>Actinomyces urogenitalis</i> DSM 15434	PriA					0.01			0.02	Juarez [@juarez-vazquez_evolution_2017]
<i>Bifidobacterium adolescentis</i> L2-32	PriA	*	*			0.2			0.1	Juarez [@juarez-vazquez_evolution_2017]
<i>Bifidobacterium gallicum</i> DSM 20093	PriA	*	*			0.1			0.04	Juarez [@juarez-vazquez_evolution_2017]
<i>Bifidobacterium longum</i> ATCC 15697	PriA	*	*			0.1			0.3	Juarez [@juarez-vazquez_evolution_2017]
Camera CAM1	Metagenoma			1.7 ± 0.1	0.3 ± 0.03	0.2	40 ± 7	3.5 ± 0.04	0.09	Noda [@noda-garcia_insights_2015]
CAM1_A81G	Metagenoma			1.7 ± 0.2	0.1 ± 0.01	0.06	32.2 ± 1.7	1.9 ± 0.1	0.06	Noda [@noda-garcia_insights_2015]
CAM1_A81S	Metagenoma			4.0 ± 0.9	0.2 ± 0.03	0.04	23.5 ± 6.5	0.5 ± 0.1	0.02	Noda [@noda-garcia_insights_2015]
CAM2	Metagenoma		n.d.	n.d.	0	n.d.	n.d.	n.d.	0	Noda [@noda-garcia_insights_2015]
PriA_Ancestral	Ancestral			9.4±1.6	0.3±0.009	0.03	4.3±0.4	0.6±0.02	0.13	Verduzco, Noda*
PriA_SubHisA	Ancestral			3.7±1.01	0.5±0.03	0.1	-	-	0	Verduzco, Noda*
SubHisA_Ancestral	Ancestral			6.3±0.7	0.15±0.03	0.02	-	-	0	Verduzco, Noda*
SubHisA_PriA	Ancestral			27.7±3.4	0.05±0.005		167.82	0.03±0.002	0.0001	Verduzco, Noda*
<i>scabies</i>	0	0	163.6	0.1						Verduzco*
<i>Aviso</i>		46	1.37	36	3.4					Juárez*

- Datos no publicados del Laboratorio de Evolución de la diversidad metabólica.

En este capítulo en las siguientes secciones 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 PriA en sustratos diferentes de PRA y ProFAR o bien en combinaciones de los mismos.

PriA como modelo de familia enzimática donde las expansiones no son condición necesaria para la promiscuidad.

PriA en EvoMining

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 Actinobacteria, donde se sabe que PriA es promiscua no se detectaron copias extra. En la Figure 1 se muestra el número promedio de copias por genoma en los linajes genómicos seleccionados. 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. Esta observación demuestra 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.

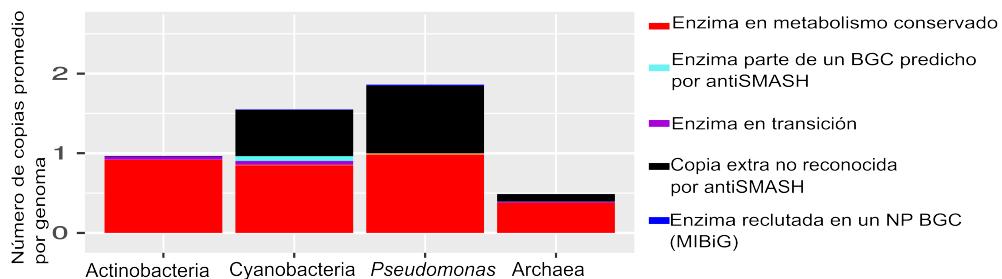


Figure 1: Número promedio de copias por genoma de PriA en Actinobacteria, Cyanobacteria, *Pseudomonas* y Archaea. Los colores muestran el destino metabólico asignado a cada copia según EvoMining. En rojo están los BBH a las enzimas semilla de metabolismo central. En morado las enzimas de metabolismo central también reconocidas por antiSMASH como parte de unBGC y en negro las copias sin un destino metabólico conocido.

Después del conteo de número de copias promedio, se analizaron los árboles de PriA de EvoMining, tanto los coloreados de acuerdo al número de copias, Figure 2, como los coloreados según el destino metabólico, Figure 3. En Actinobacteria la mayoría de las hojas son verdes reafirmando que existe sólo una copia por genoma en ese organismo. Sin embargo, existen varias hojas de color amarillo, indicando de dos copias en ese genoma. Las Actinobacterias con dos copias son *Ornithinimicrobium pekingense* DSM 21552, *Pseudonocardia* sp. P2, *Serinicoccus marinus* DSM 15273, *Serinicoccus profundi* MCCC 1A05965, *Sphaerotilus thermophilus* DSM 20745 y *Streptomyces* sp. CT34.

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, observamos que las copias extra tanto de Cyanobacteria como de *Pseudomonas* son en realidad miembros de otras familias enzimáticas, que guardan cierta similitud de secuencia con PriA. En Cyanobacteria y *Pseudomonas* la copia extra está en una rama divergente y muy poblada del árbol. En ambos casos esta segunda copia está en su mayoría anotada por RAST como imidazol glicerol fosfato sintasa ciclase. En Archaea sin embargo, diversas especies de la clase Methanomicrobia sí tienen dos copias de PriA.

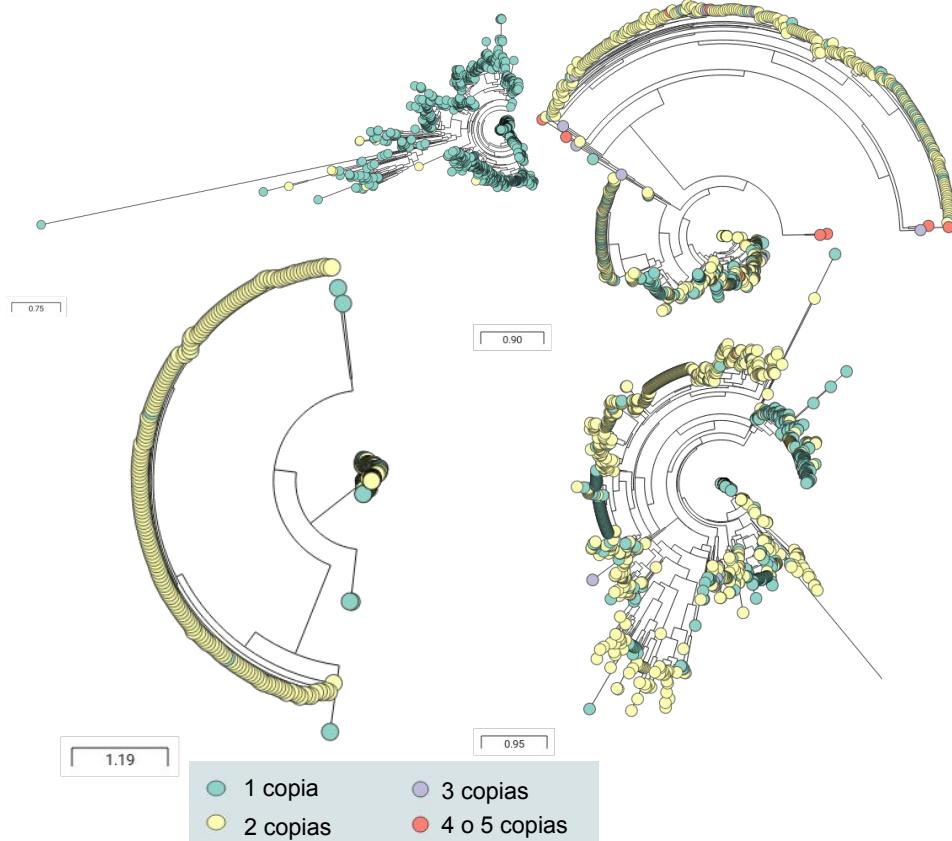


Figure 2:

Después de explorar cuáles organismos tienen expansiones de PriA, a continuación se muestra en la Figure 3 el posible destino metabólico de las copias extra de la familia. El árbol de Actinobacteria está poblado de hojas rojas, es decir de PriA dedicadas al metabolismo conservado, en este caso a las rutas de Histidina y Triptofano. Sin embargo hay algunas hojas grises, como es el caso de los dos *Serenicoccus* mencionados previamente. Es posible que estas PriA puedan tener funciones alternativas. 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, que no parece ser un artefacto de anotación ya que hay otros *Janibacter* con PriA ligeramente más grandes que el promedio.

En Cyanobacteria hay pocas hojas verdes (EvoMining predictions) y no están localizadas cerca de la HisA de saxitoxin, el BGC proveniente de Cyanobacteria. En *Pseudomonas* hay una gran población de predicciones de EvoMining, pero son falsos positivos, ya que estas no corresponden a una rama dedicada al metabolismo secundario, sino a la rama de la ciclase. En Archaea algunas de las hojas grises, es decir sin destino metabólico conocido serán exploradas en la siguiente sección.

Todos los reclutamientos (marcados en azul en el árbol) que tuvieron estas expansiones en estos linajes

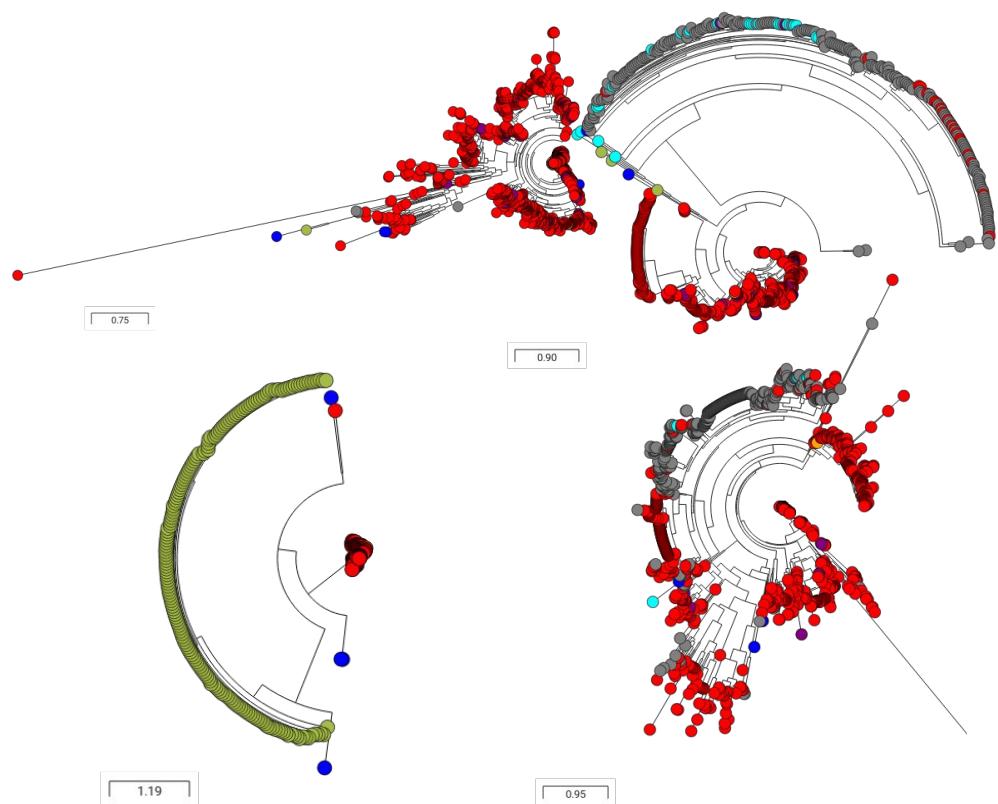


Figure 3: Árboles de destino metabólico de PriA en Actinobacteria, Cyanobacteria, Pseudomonas y Archaea según EvoMining

jes genómicos están listados en la tabla Table 2. 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 2: Reclutamientos de expansiones de PriA en MIBiG

Compuesto	Actinobacteria	Cyanobacteria	Pseudomonas	Archaea	BGC origen	Clase
saxitoxin	x	x	x	x	Cyanobacteria	Alcaloide
lipopolisacárido	x	x	x	x	Proteobacteria	Sacarido
toxin	x	x	x	x	Cyanobacteria	Otros
ada	x	x	x	x	Actinobacteria	Otros

La pentostatina es un atibiótico 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 4, 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.

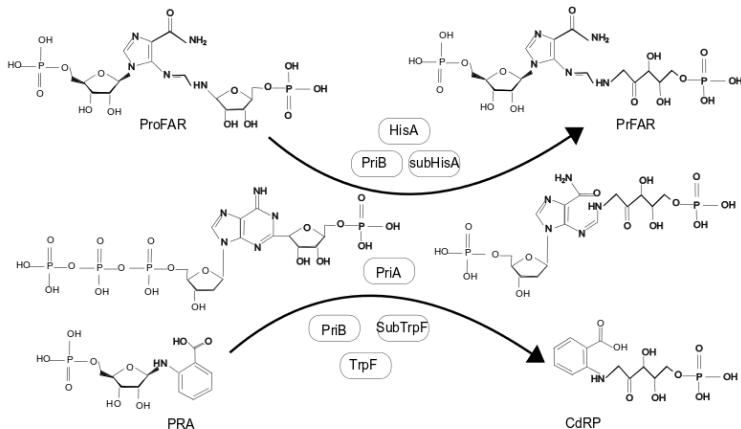


Figure 4: PriA participa en la síntesis del antibiótico ada en *Actinomadura*. Los sustratos nativos de PriA, ProFAR y PRA son isomerizados de manera muy similar a un paso en la ruta de síntesis de ada

Finalmente, los árboles que se produjeron por EvoMining están disponibles para exploración interactiva en Microreact en los links de la ??:

Table 3: Árboles EvoMining de PriA en MicroReact

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

Análisis de contextos genómicos de PriA en distintos linajes utilizando CORASON

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 5 . 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 jeoijense*. Los contornos de los genes reconocidos por antiSMASH como parte de un BGC son marcados en azul en las figuras de CORASON.

En cuanto a Archaea, los contextos mostrados en los páneles b y c de la Figure 5 muestran que si bien existen algunos contextos como el de ciertos *Thermococcus* donde PriA está rodeada de genes de histidina y triptofano, esta configuración no es la generalidad. Al contrario, parece que no hay 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.

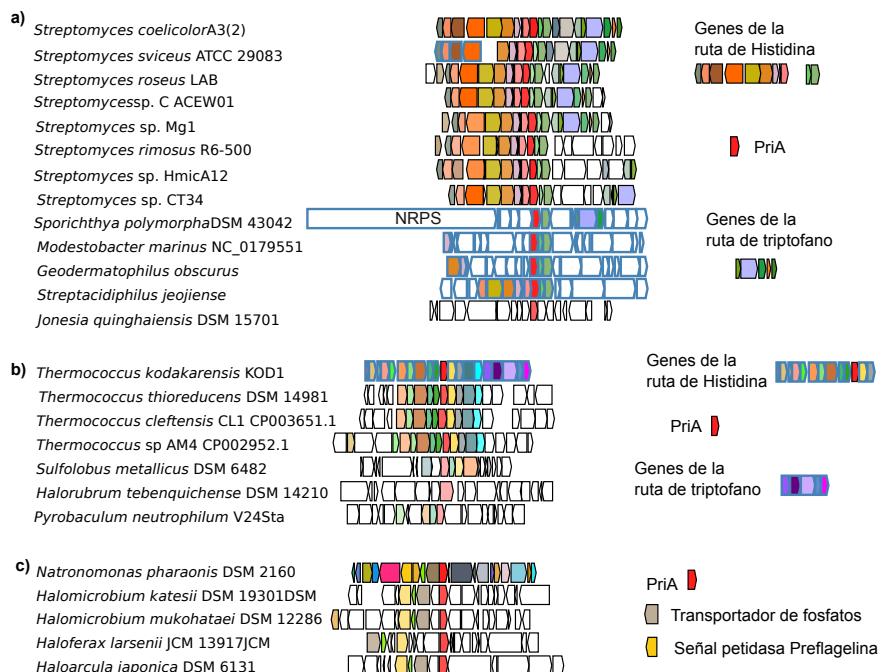


Figure 5: Contextos de PriA en Actinobacteria y Archaea

En cuanto a PriA en Cyanobacteria la visualización de contextos producida por CORASON tomando como referencia el cluster de saxitoxin, muestra que las enzimas biosintéticas de ese no están conservadas en otros organismos. Además en el lado izquierdo de la Figure 6 se muestra que la PriA del BGC saxitoxin no está ubicada en una rama de PriA divergente, al contrario está en la parte más conservada. Por estas razones es posible que PriA esté en la orilla del BGC de saxitosin y más bien no participe en la síntesis de este compuesto.

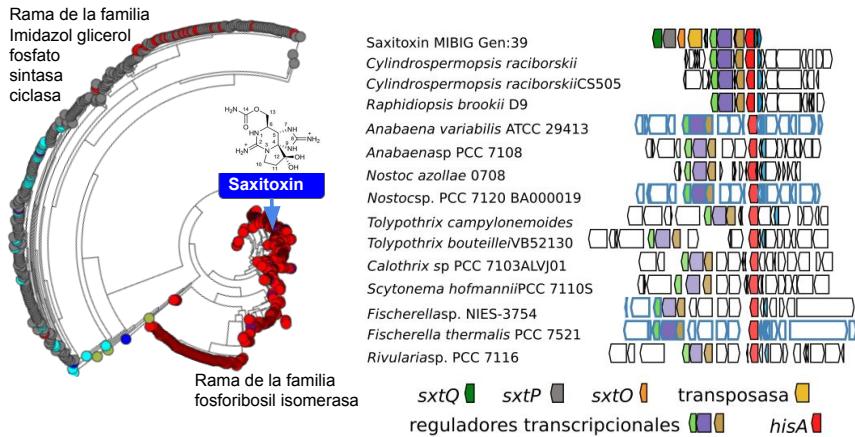


Figure 6: HisA en saxitoxin, un cluster de Cyanobacteria

El monitoreo en cambios de aminoácidos en PriA para estudiar rutas evolutivas y reconstrucciones de la estructura tridimensional.

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 laboatorio 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 (darwiniana)

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 las rutas, 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 script 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 7.

Al analizar los cambios sufridos en cada actividad en cada paso de cada ruta se descubrió que no existe en ellas una trayectoria no decreciente para ningún sustratos. Como ejemplo, en la Figure 7 se muestran las rutas donde cada mutante mantiene un nivel mínimo de actividad de ProFAR isomerasa ($\frac{K_{cat}}{K_m} \text{PriA}_{\text{ProFAR}} \geq .004$). 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

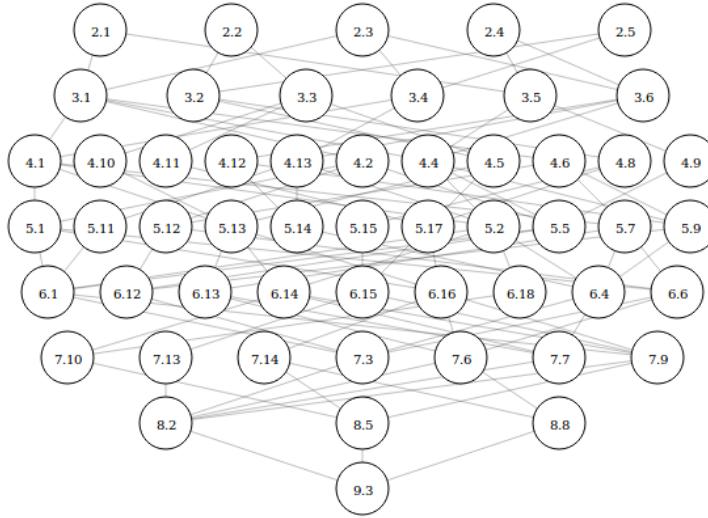


Figure 7: Rutas desde una subHisA hasta una variante con 9 mutaciones. En cada círculo el primer dígito indica el número de mutaciones

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.

Los residuos que covarián 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, si no que 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. Por este motivo se decidió implementar Evcouplings [@marks_protein_2011], para poder aplicarlo a la familia PriA. Este método es de difícil instalación ya que tiene muchas dependencias, por ello desarrollé un contenedor docker donde las dependencias y el software quedan instalados. Este desarrollo fue incluido por los desarrolladores originales como sugerencia de instalación.

El contenedor docker implementa Evcouplings python framework [@hopf_evcouplings_2019] que comprende cinco etapas para estudiar el análisis de coevolución de residuos de una familia de proteínas. Estas etapas son i) Alineado, ii) análisis de acoplamiento, iii) plegamiento basado en acoplamientos iv) análisis de mutación y

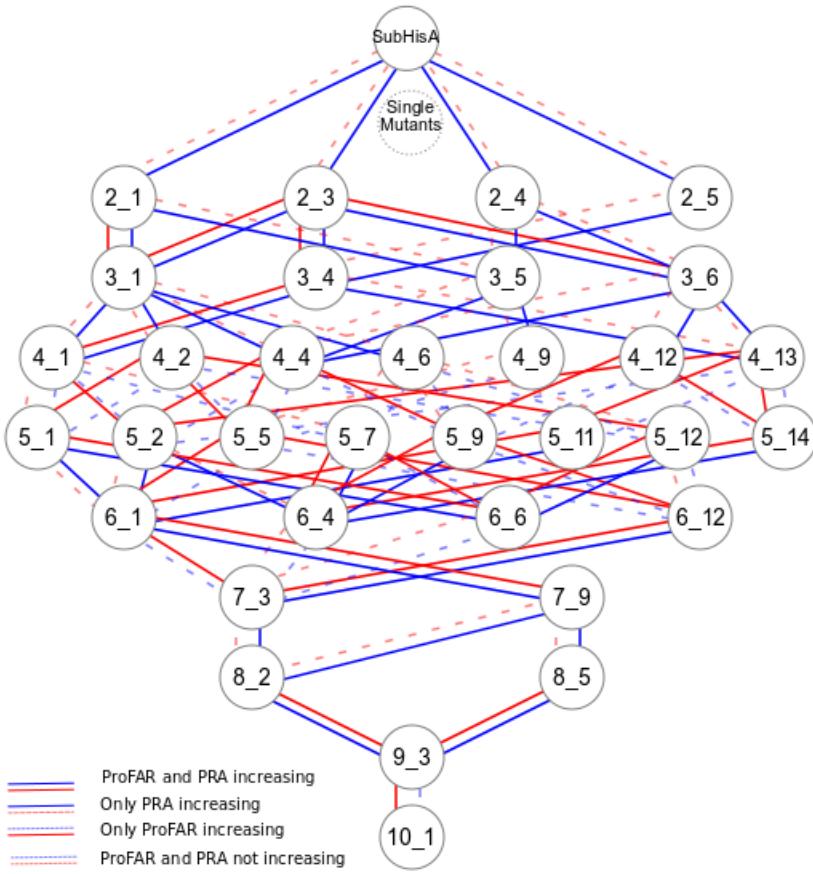


Figure 8:

v) comparación con estructuras conocidas.

EVcouplings fue aplicado a PriA de *Streptomyces coelicolor* con identificador de Uniprot HIS4_STRCO. Las secuencias para las alineaciones se recuperaron automáticamente de Uniprot, otros parámetros se dejaron con la configuración inicial del archivo de configuración. Como resultado se modeló la estructura de PriA con los acoplamientos de sus aminoácidos. Una comparación con una estructura cristalográfica es mostrada en la Figure 9

La estructura reconstruida es parecida, pero para obtener mejor definición es posible que se deba refinar la selección de las secuencias del alineamiento, diferenciando entre secuencias conocidas de PriA, subHisA, PriB y subTrpF.

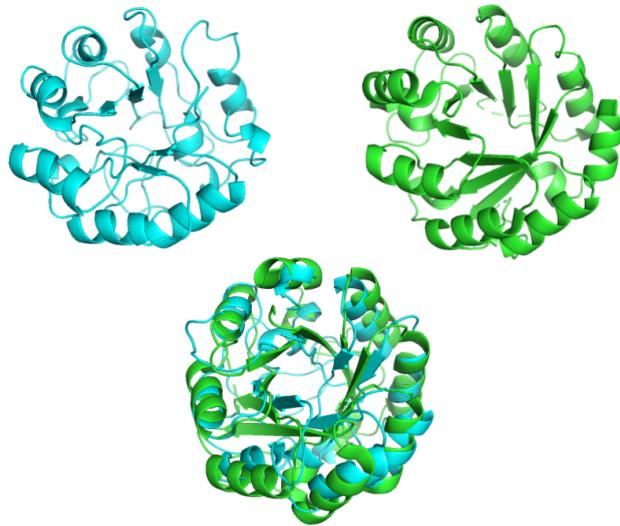


Figure 9:

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. En la siguiente tabla se muestra en la primera columna los aminoácidos variados en la sección anterior, y en la segunda columna el aminoácido correspondiente en la secuencia de *Streptomyces coelicolor*.

<i>Corynebacterium</i>	<i>Streptomyces</i>
D20V	21V
L48I	49L
F50L	51L
M66I	67I
T80S	81S
A97C	98C
D127A	128G
A129D	130D
T139L	136L
Y214L	211Y
E230A	227A

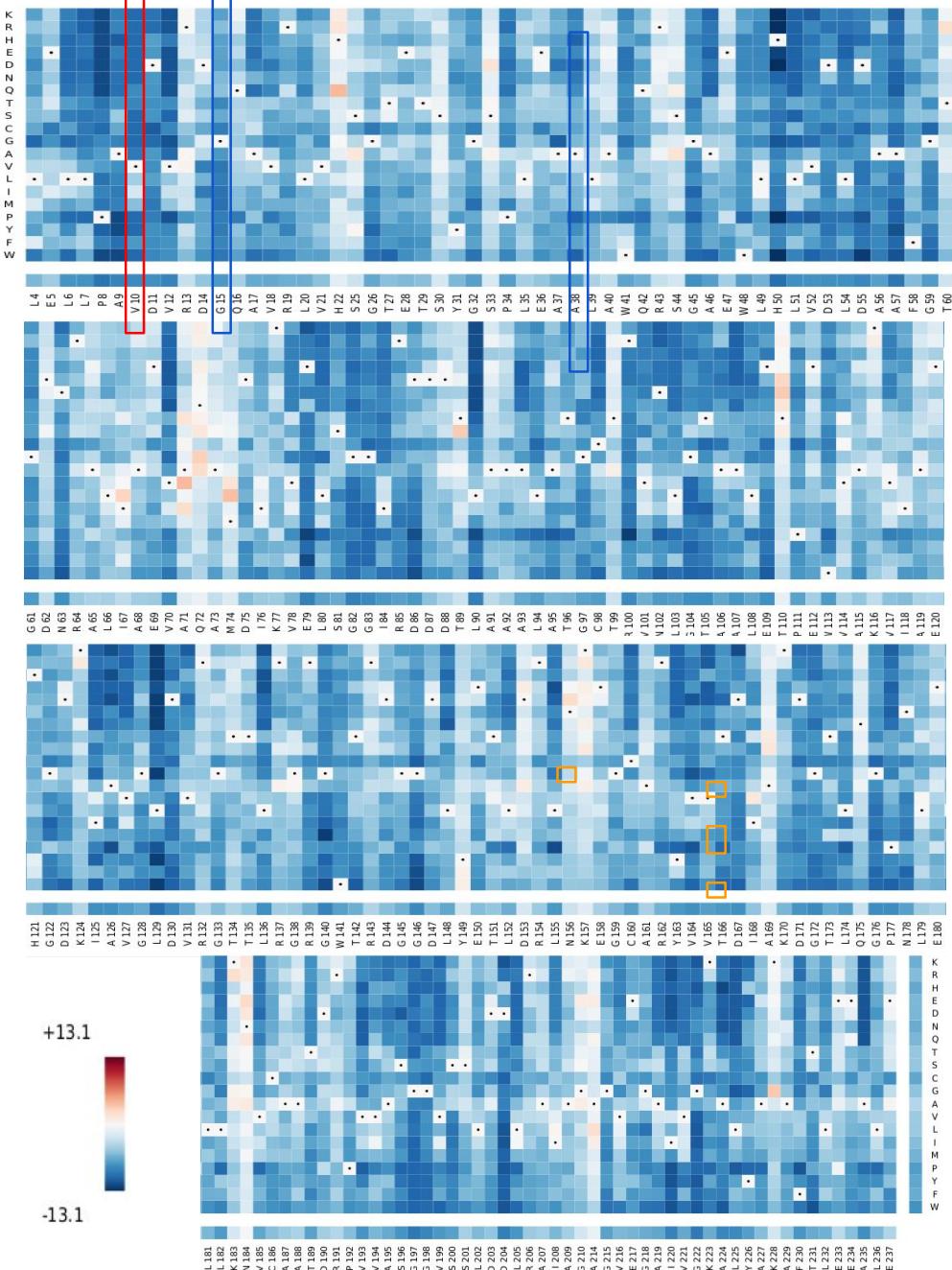
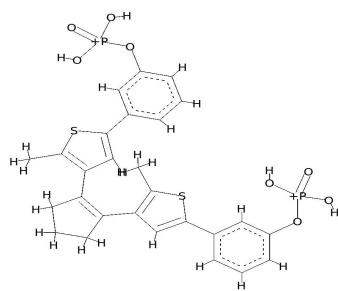


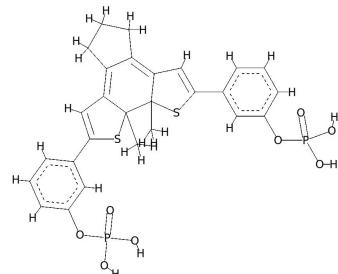
Figure 10:

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

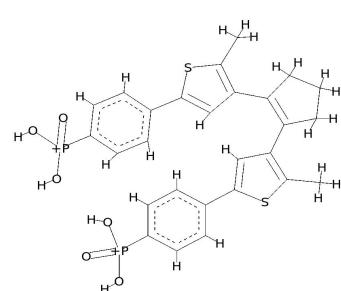
Además de los sustratos conocidos ProFAR y PRA en los que PriA es capaz de realizar una isomerización, es posible que PriA pueda ser promiscua en otros sustratos tal y como ya se mencionó para el antibiótico *ada* en 1. De hecho como se vio en la sección de EvoMining de este capítulo, PriA participa en la síntesis del antibiótico pentostaina. Tomando este ejemplo como inspiración, se colectaron en la literatura sustratos parecidos químicamente a ProFAR y PRA. En esta sección se buscará mostrar sustratos parecidos a los nativos con la idea de posteriormente probarlo en copias seleccionadas de PriA provenientes de diversos organismos.



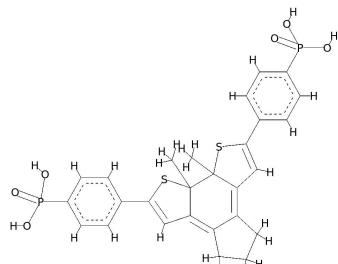
13
DTE-meta-phosphate(dte6_Open form)



14
DTE-meta-phosphate(dte6_Closed form)



15
DTE-Para-Phosphonate(dte13_closed form)



16
DTE-para-phosphonate(dte13_closed form)

Figure 11: Substatos químicamente similares a los sustratos de PriA (parte 1)

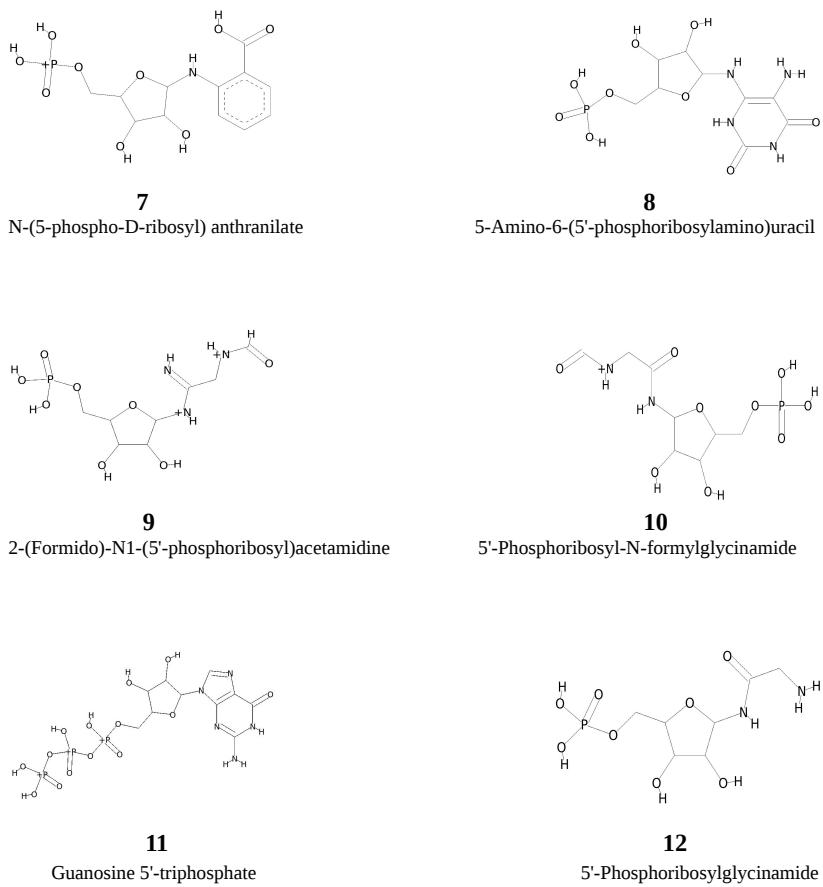
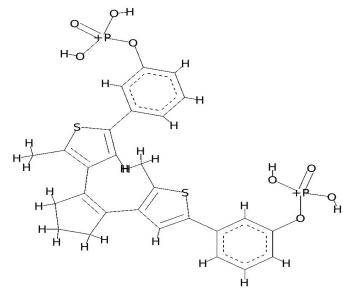
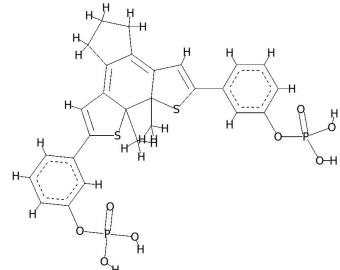


Figure 12: Substatos químicamente similares a los sustratos de PriA (parte 2)



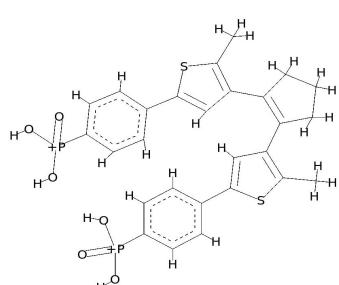
13

DTE-meta-phosphate(dte6_Open form)



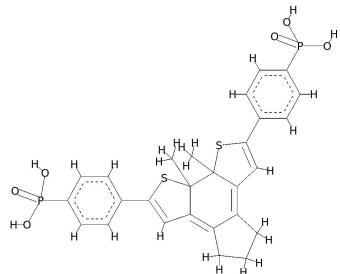
14

DTE-meta-phosphate(dte6_Closed form)



15

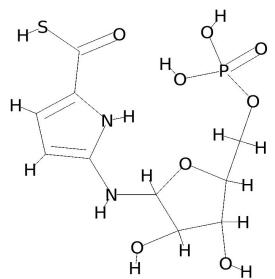
DTE-Para-Phosphonate(dte13_Closed form)



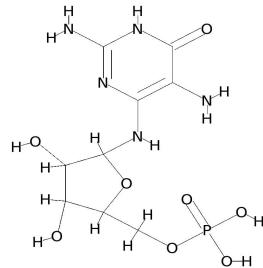
16

DTE-para-phosphonate(dte13_Closed form)

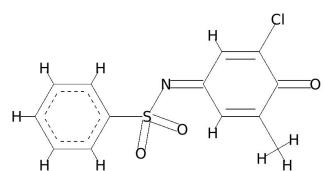
Figure 13: Substatos químicamente similares a los sustratos de PriA (parte 3)



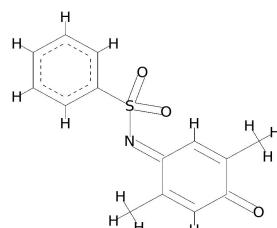
17
4N'-(5'-phosphoribosyl) 4-aminopyrrole
-2-carboxilate



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

Dado que en *Streptomyces* se han encontrado representantes de las familias PriA y PriB, se seleccionaron 39 secuencias homólogas para realizar el análisis del sustrato de la enzima de acoplamiento. Estas secuencias seleccionadas de PriA / Prib pertenecen a *Streptomyces* uniformemente distribuidas en un árbol de especies RpoB con diferentes condiciones sobre la presencia / ausencia de TrpF. En este estudio se incluyeron otros homólogos de PriA Actinobacterial caracterizados químicamente, y finalmente se agregaron HisA de *Escherichia coli*, *Arthrobacter Aurescens*, *Salmonella enterica* y *Acidimicrobium ferrooxidans* y Actinobacterial TrpF como controles.

Cuando existían estructuras de cristal, de lo contrario, se generaban estructuras homólogas utilizando como plantilla la enzima más cercana disponible con estructura de cristal.

Controls

HisA Enterobacteria enzymes from *Salmonella enterica* (PDB:5AHE), *Escherichia coli K12* *Acidimicrobium ferrooxydans* (PDB:4WD0)

TrpF Actinobacteria *Jonesia denitrificans* and *Streptomyces sp Mg1* sequences.

Chemically characterized Actinobacterial enzymes

PriA

Mycobacterium tuberculosis (Mtub PDB:2Y88,2Y89,2Y85,3ZS4) *Streptomyces coelicolor* (Scoe PDB:2VEP,2X30,1VZW), *Streptomyces globisporus*, *Actinomyces urogenitalis* 4X2R *Corynebacterium jeikeum*

subHisA

Corynebacterium diphtheriae *Actinomyces car* (PDB:4X2R)

subTrpF

Arthrobacter aurescens (PDB:4WD0)

PriB

Streptomyces ipomoeae, *Streptomyces sviceus* (PDB:4U28,4TX9)

TrpF controls *Jonesia denitrificans* (PDB:4WUI) *Chlamidya trachomatis*, *Streptomyces sp. Mg1* TrpF and *Actinomyces odontolyticus* were included

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

```
table <- read.csv("chapter4/EstructurasPDB", row.names = 1,sep="\t")
kable(table, caption = "Enzyme PDB \\label{tab:Enzyme PDB}",caption.short = "Enzyme PDB ")
```

Table 5: Enzyme PDB

	Organismo	Family	Observations	Resolution	Year
4WUI	Jonesia denitrificans	TrpF		1.09	2014
4X9S	Streptomyces sp. MG1	PriB		1.60	2014
5DN1	Streptomyces coelicolor	PriA		1.95	2015
1DL3	Thermotoga maritima	TrpF		2.70	1999
1LBM	Thermotoga maritima	TrpF	RCDRP	2.80	2002
1NSJ	Thermotoga maritima	TrpF		2.00	1996
1V5X	Thermus thermophilus	TrpF		2.00	2003
1VZW	Streptomyces coelicolor	PriA		1.80	2004
2VEP	Streptomyces coelicolor	PriA		1.80	2007
2X30	Streptomyces coelicolor	PriA	R139N	1.95	2010
2Y85	Mycobacterium tuberculosis	PriA	RCDRP	2.40	2011
2Y88	Mycobacterium tuberculosis	PriA	D11N PRFAR	1.33	2011
2Y89	Mycobacterium tuberculosis	PriA	D11N	2.50	2011
3ZS4	Mycobacterium tuberculosis	PriA	PRFAR	1.90	2012
4AAJ	Pyrococcus furiosus	TrpF		1.75	2012
4TX9	Streptomyces sviceus	PriB	ProFAR	1.60	2014
4U28	Streptomyces sviceus	PriB		1.33	2014
4W9T	Streptomyces sp. Mg1	PriB		1.57	2014
4WD0	Arthrobacter aurescens	PriB		1.50	2014
4X2R	Actinomyces urogenitalis			1.05	2014
4AXK	Corynebacterium efficiens	SubHisA		2.25	2013
5AHE	Salmonella enterica	HisA		1.70	2015
5AB3	Salmonella enterica	HisA	D7N, D10G, dup13-15, Q24L, G102A	1.80	2016
5ABT	Salmonella enterica	HisA	D7N, G102A, V106M, D176A	1.65	2016
5AC7	Salmonella enterica	HisA	D7N, D10G, dup13-15	1.90	2016
5AC8	Salmonella enterica	HisA	D10G, dup13-15, G102A	1.70	2016
5AC6	Salmonella enterica	HisA	D7N, D10G, dup13-15, Q24L, G102A	1.99	2016
5A5W	Salmonella enterica	HisA	HisA D7N D176A with ProFAR	NA	2015
5AHF	Salmonella enterica	HisA	HisA D7N with ProFAR	NA	NA
4GJ1	Campylobacter jejuni	HisA		2.15	2012
2W79	Thermotoga maritima	HisA		1.85	2008
1QO2	Thermotoga maritima	HisA		1.85	2000
5LHE	Thermococcus kodakaraensis	TrpF		1.85	2016
5LHF	Thermococcus kodakaraensis	TrpF		1.75	2016

Sustratos similares a PRA y PROFAR sugeridos por distancias de Tanimoto o probados previamente

S1, S2, ..., S20 sustratos fueron recolectados de la literatura y las predicciones de la quimioinformática. S3 PRA y S7 PROFAR son sustratos nativos, S13-S16 son sustratos activados por la luz, S17 PRAP, S18 Compuesto V, se encontraron en la literatura, S6 GMP, S11 GTP y otros fueron sugeridos por chemoinformatics.

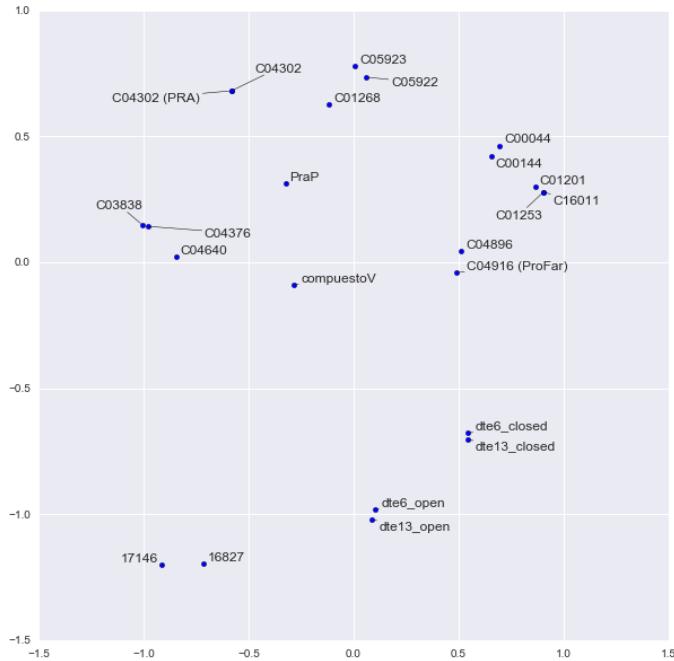


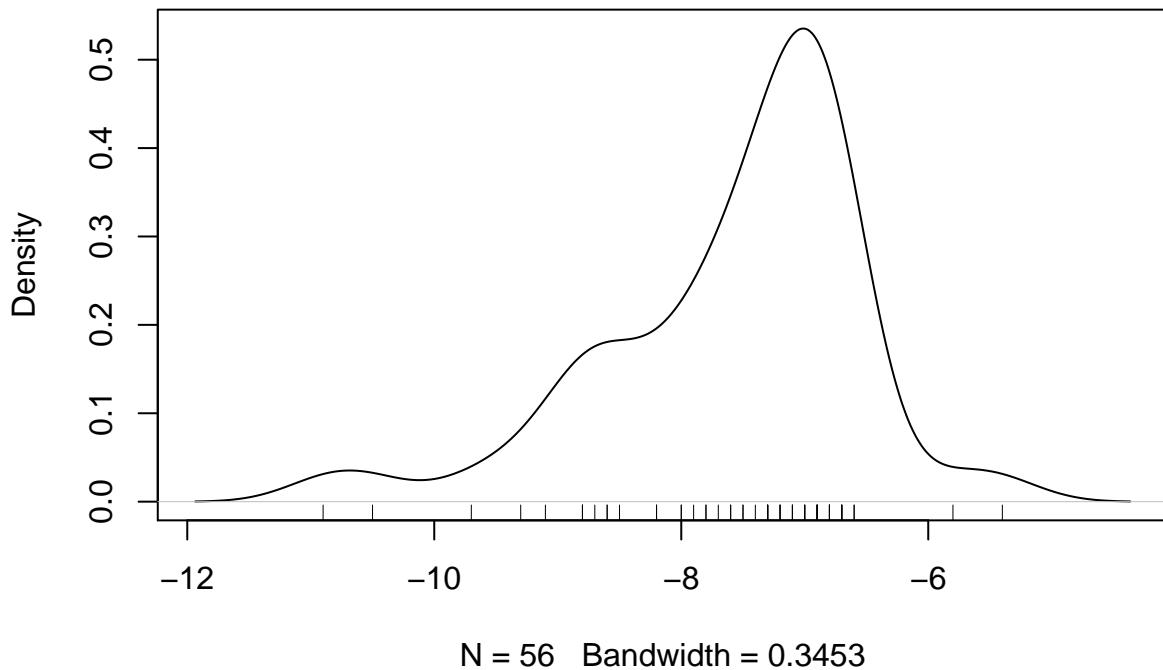
Figure 15:

On next figure we can see their chemical structures.

```
### Docking between PriA enzymes and selected substrates  
Para diversas enzimas PriA de Streptomyces se realizaron simulaciones de docking. Se incluyeron también como controles enzimas TrpF provenientes de Streptomyces Mg1, Jonesia denitrificans. Los procedimientos pueden ser consultados en Docking Protocols
```

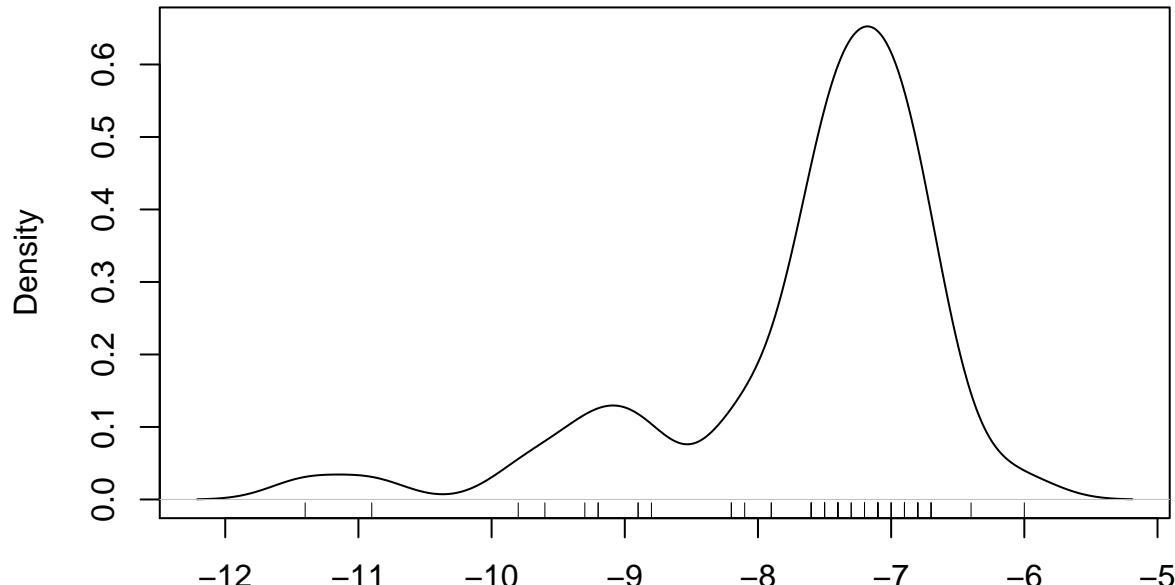
```
## Called from: eval(expr, envir, enclos) ## debug en <text>#4: plot(density(docking[,  
i], na.rm = T))
```

```
density.default(x = docking[, i], na.rm = T)
```



```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en  
<text>#4: plot(density(docking[, i], na.rm = T))
```

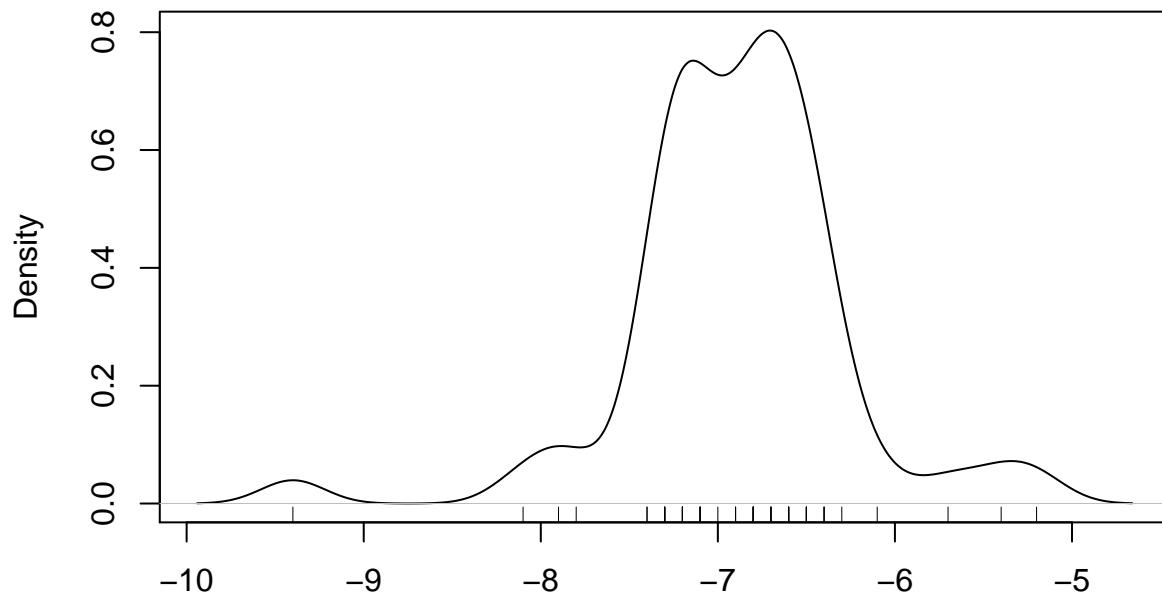
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.2702

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

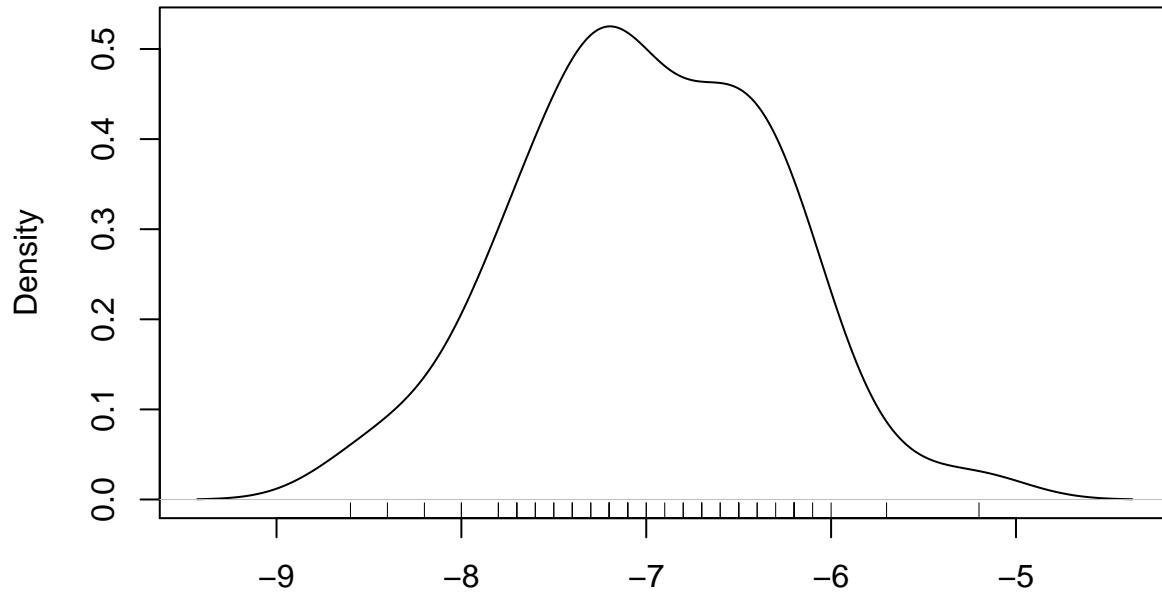
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.1802

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

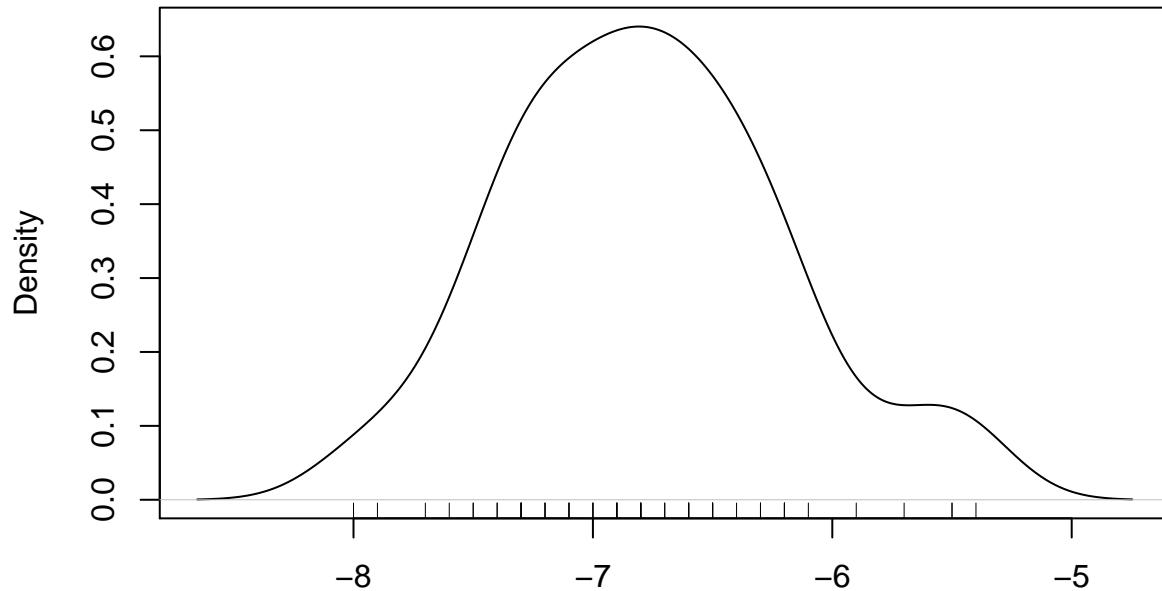
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.2764

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

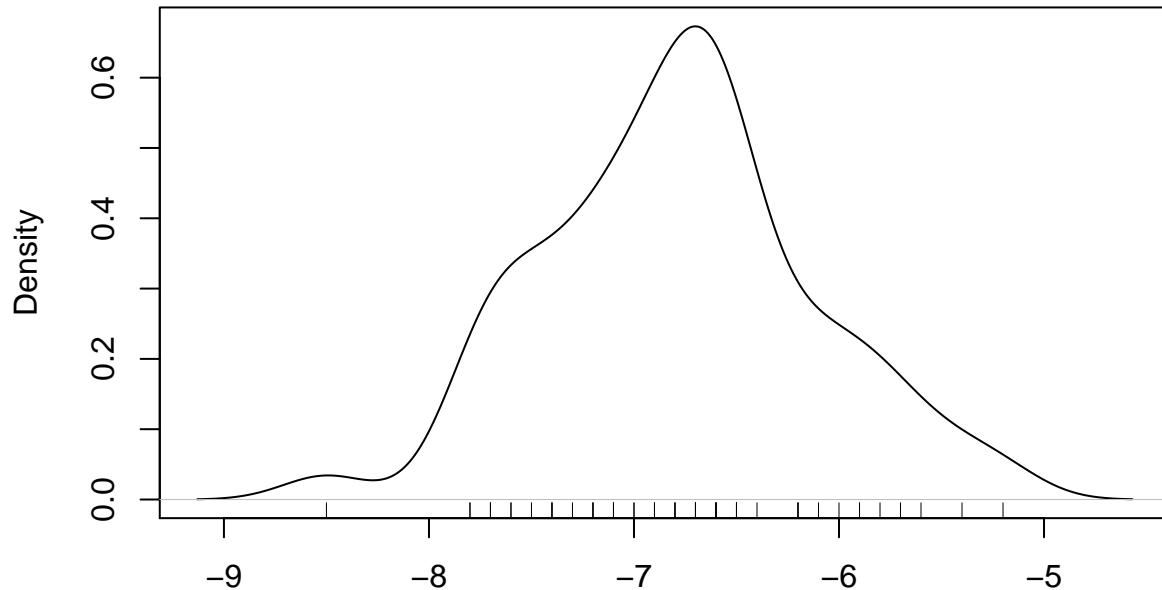
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.2177

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

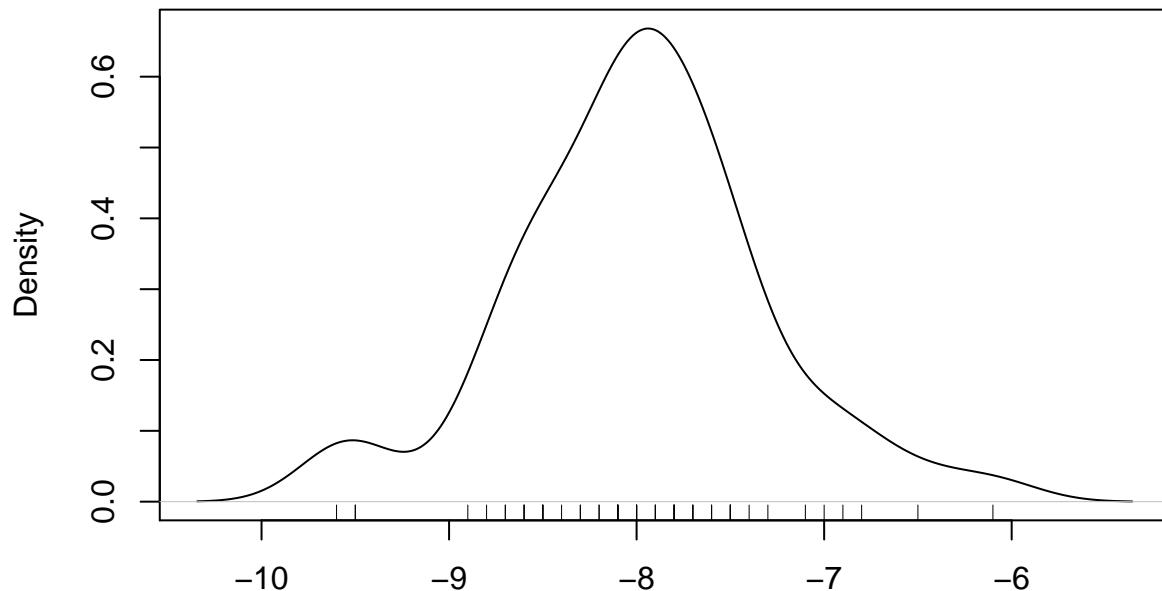
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.2102

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

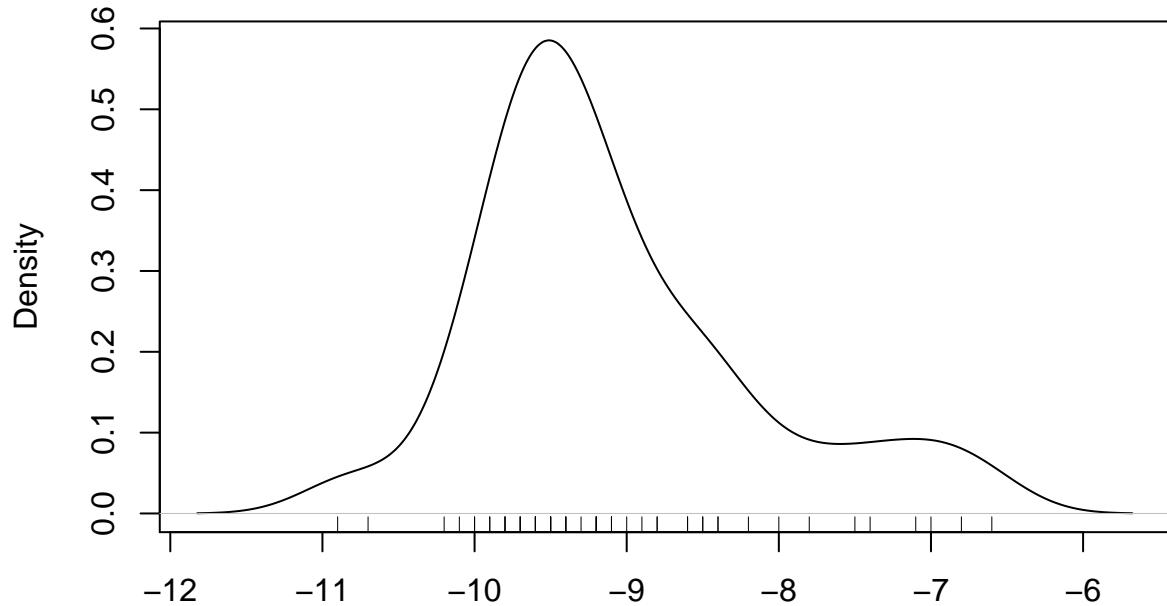
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.2477

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

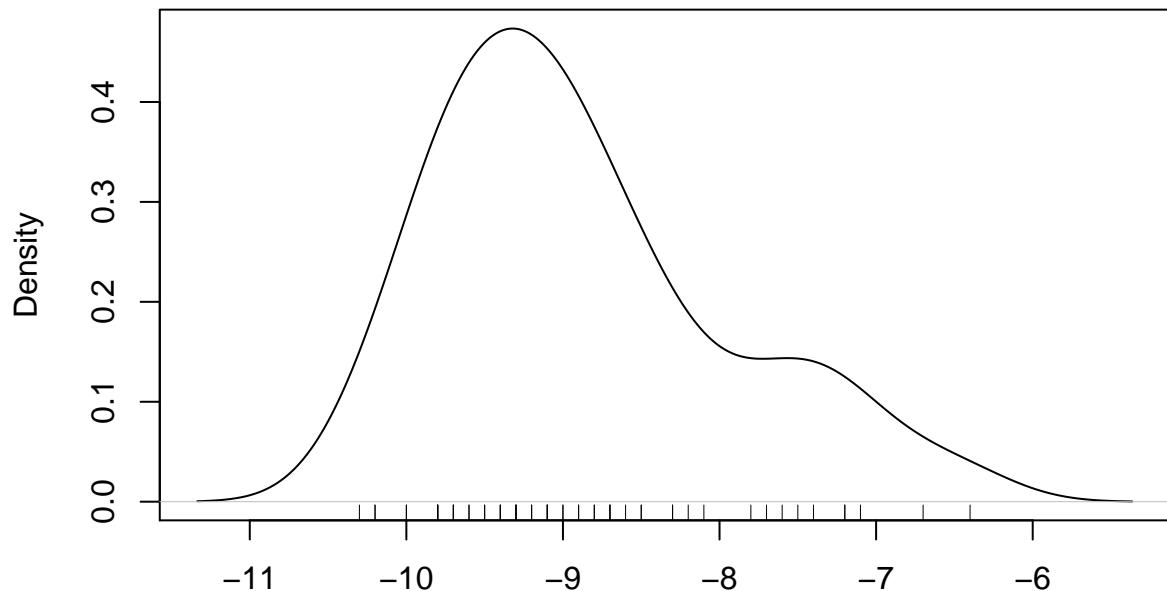
density.default(x = docking[, i], na.rm = T)



N = 56 Bandwidth = 0.3078

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

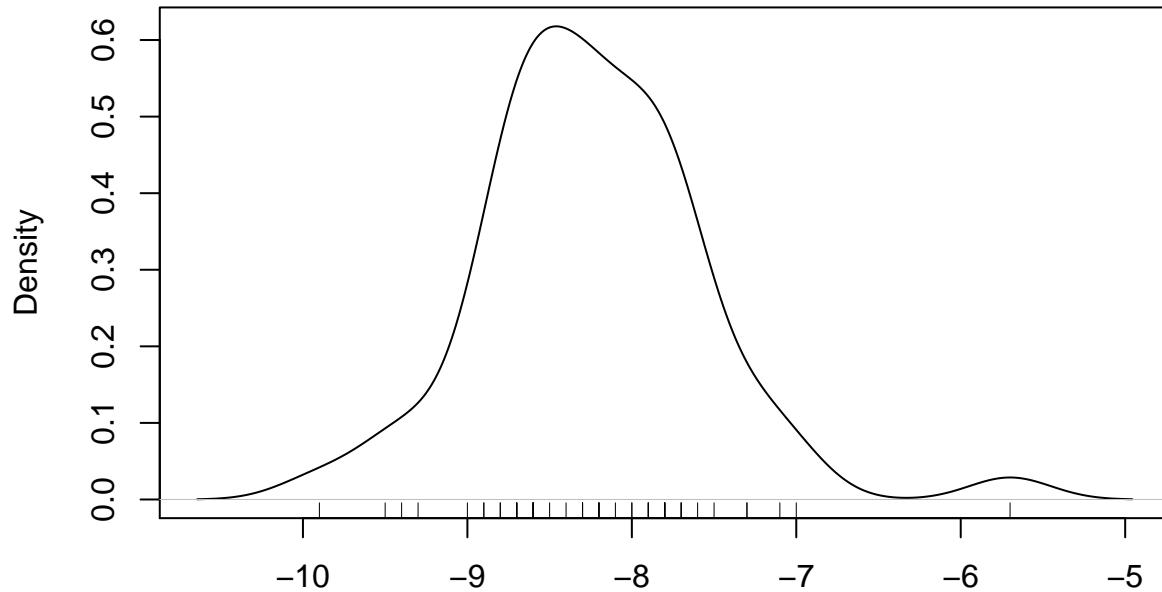
density.default(x = docking[, i], na.rm = T)



N = 56 Bandwidth = 0.3453

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

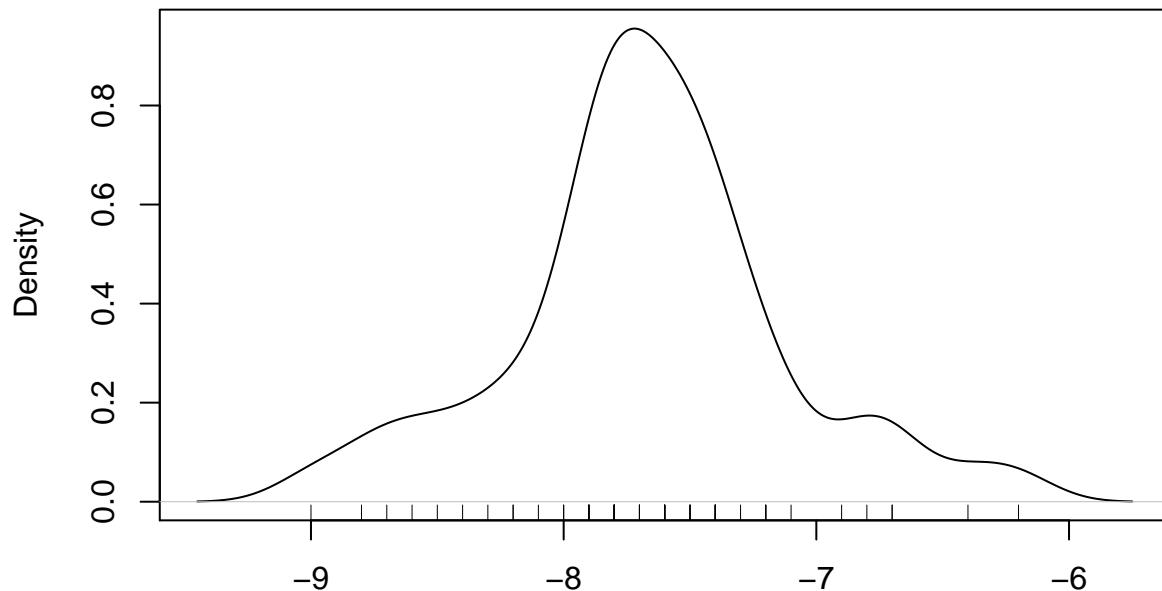
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.2477

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

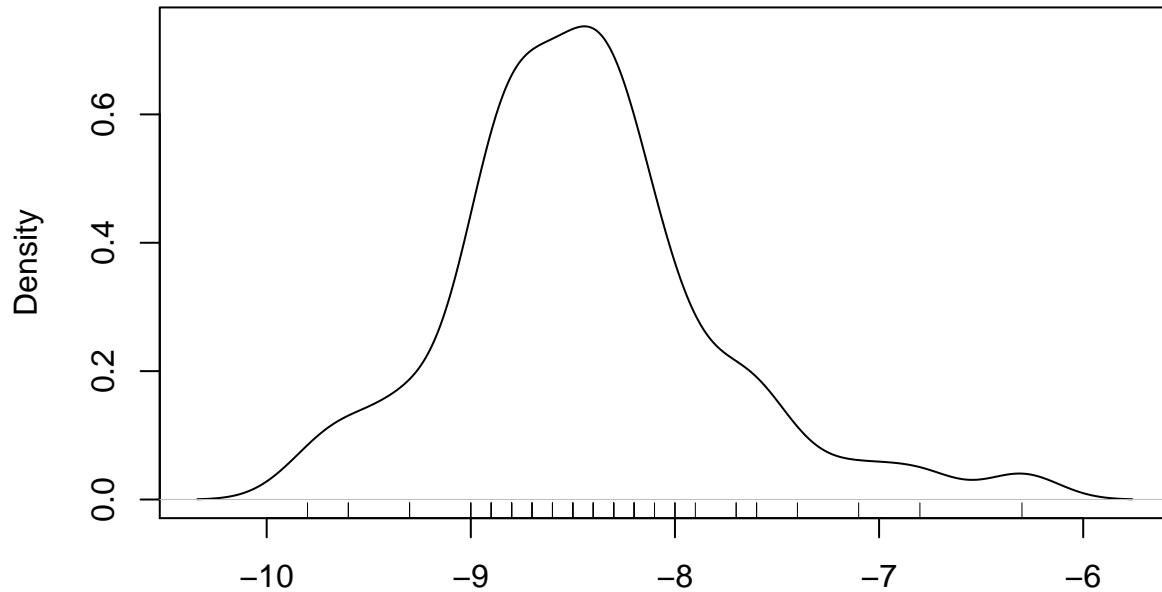
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.1501

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

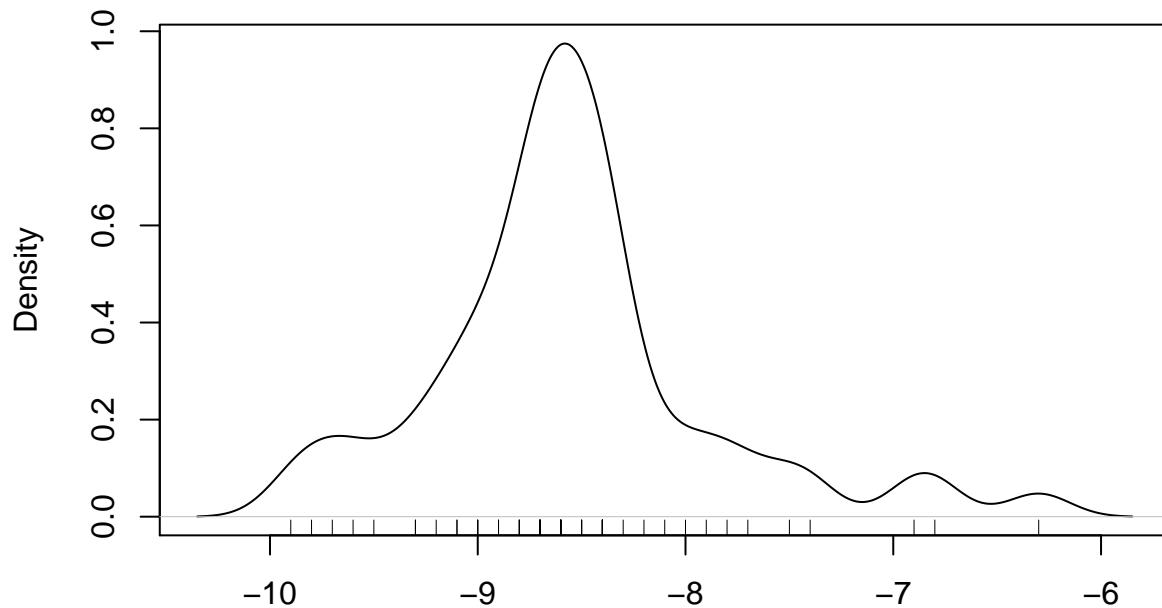
density.default(x = docking[, i], na.rm = T)



N = 56 Bandwidth = 0.1802

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

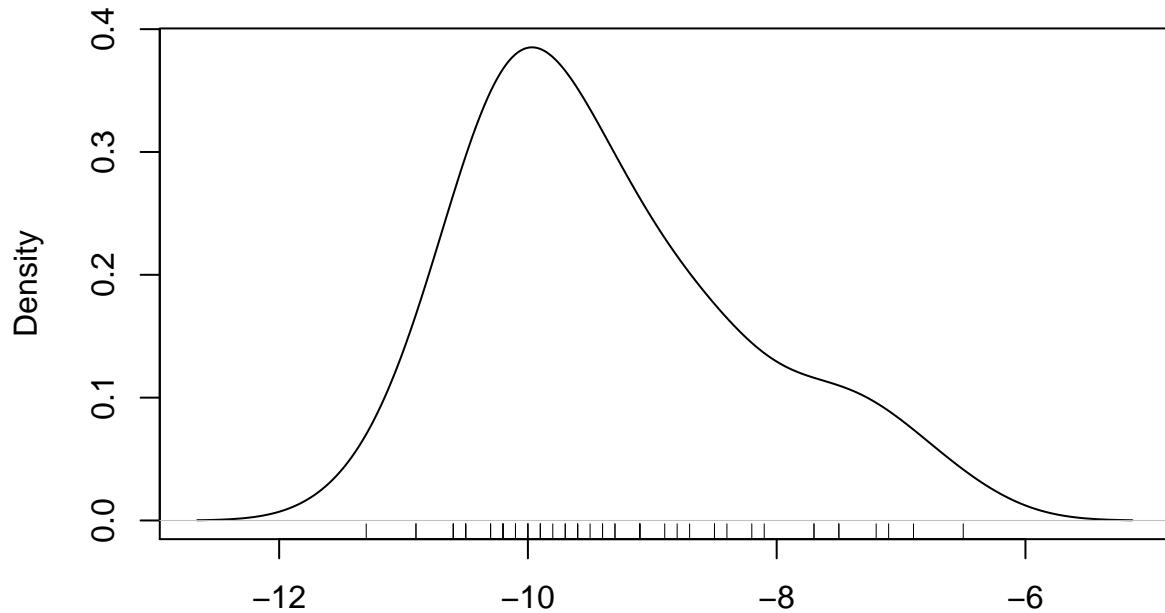
density.default(x = docking[, i], na.rm = T)



N = 56 Bandwidth = 0.1501

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

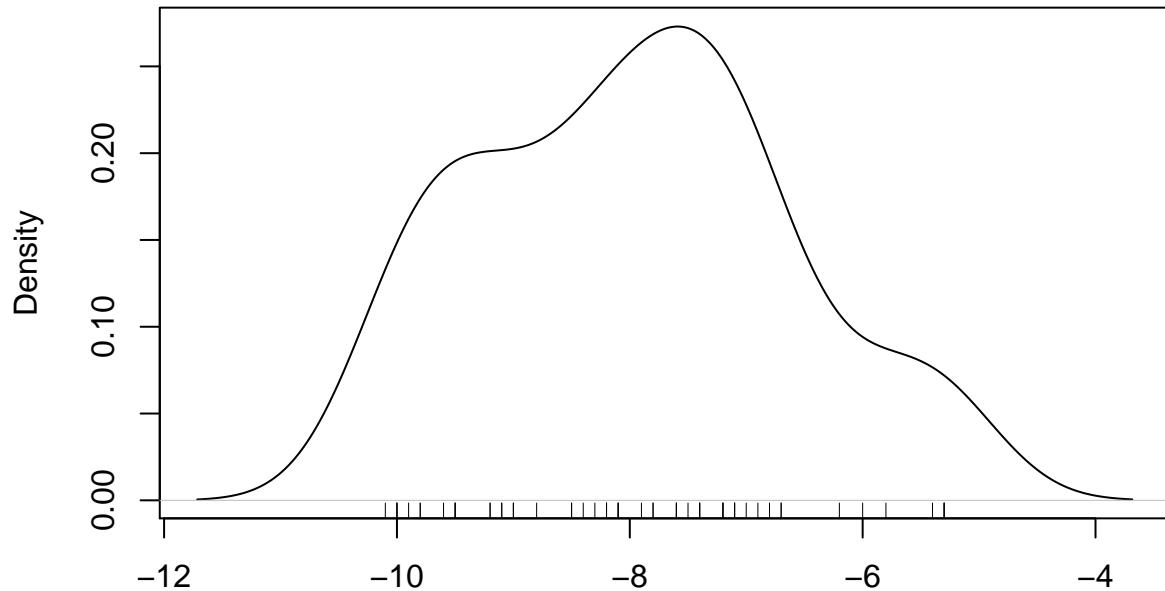
density.default(x = docking[, i], na.rm = T)



N = 56 Bandwidth = 0.453

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

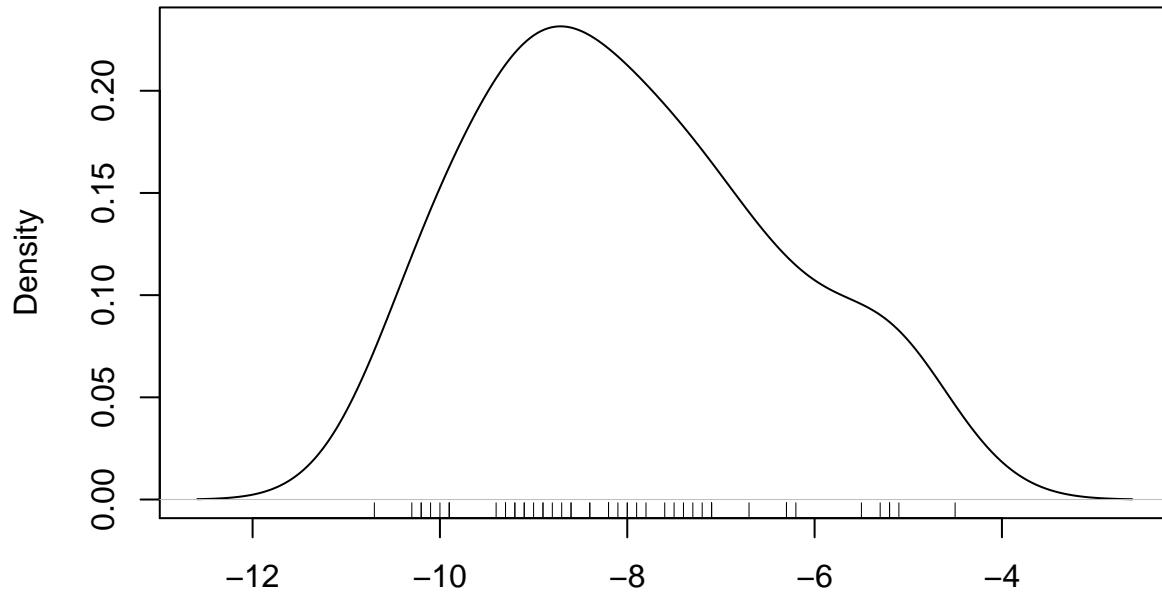
density.default(x = docking[, i], na.rm = T)



N = 56 Bandwidth = 0.5386

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

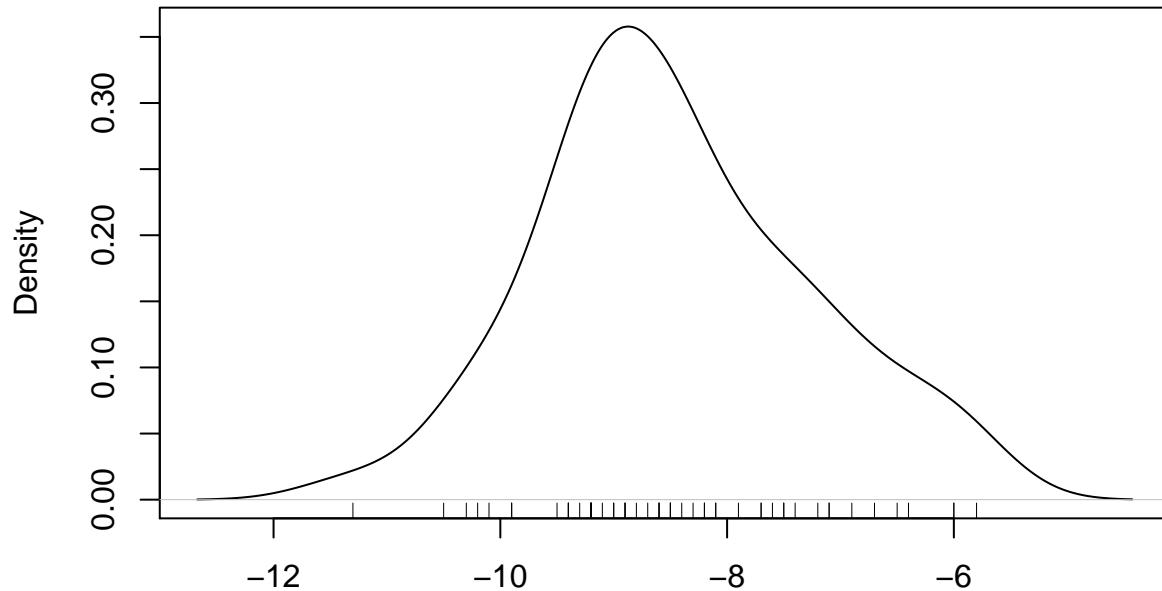
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.6305

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

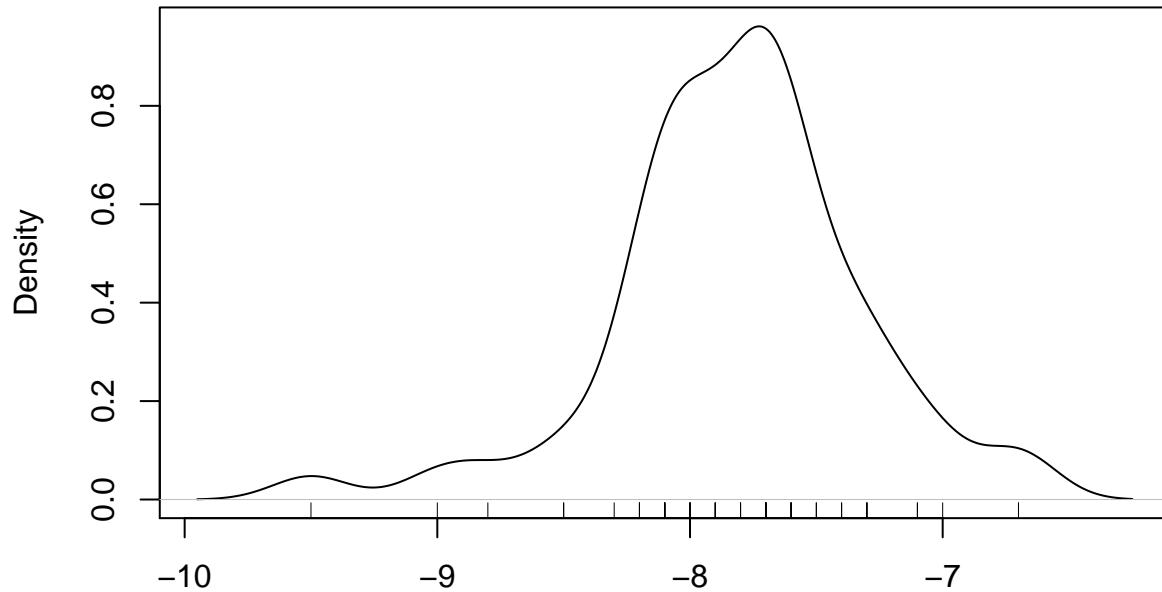
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.4579

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

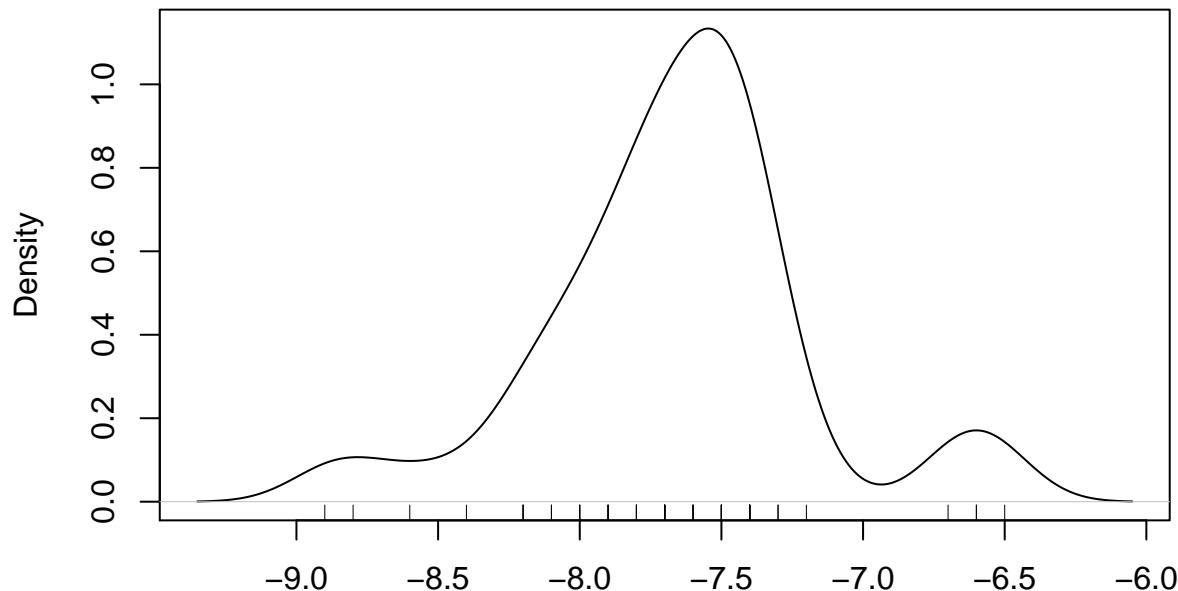
```
density.default(x = docking[, i], na.rm = T)
```



N = 56 Bandwidth = 0.1501

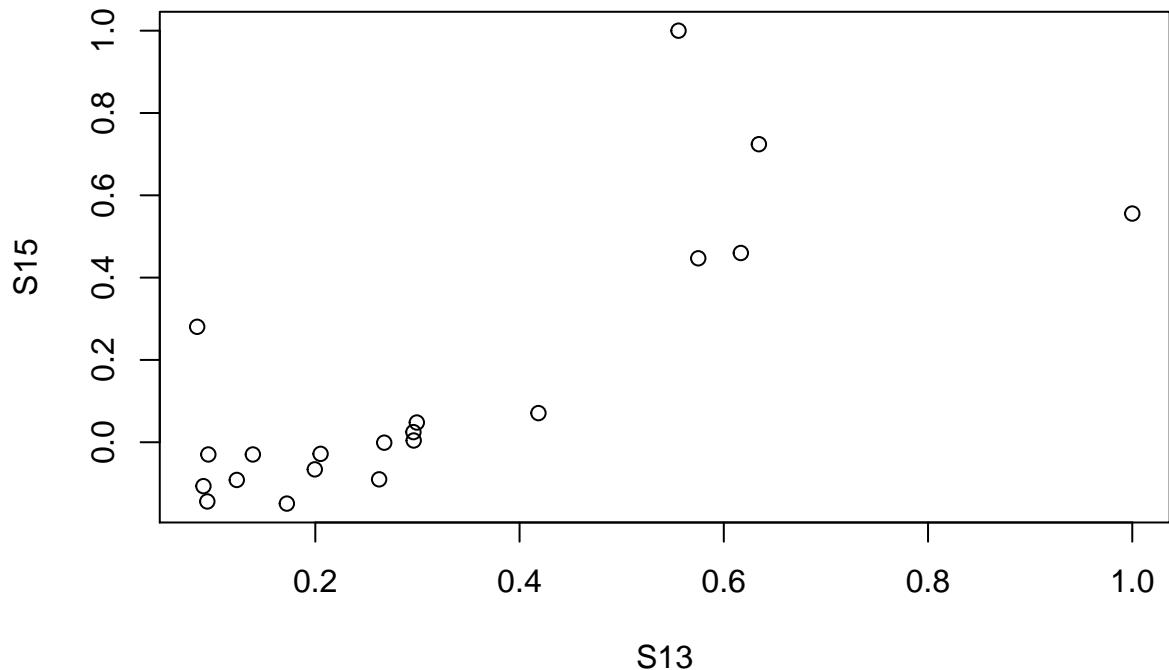
```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser() ## debug en
<text>#4: plot(density(docking[, i], na.rm = T))
```

```
density.default(x = docking[, i], na.rm = T)
```

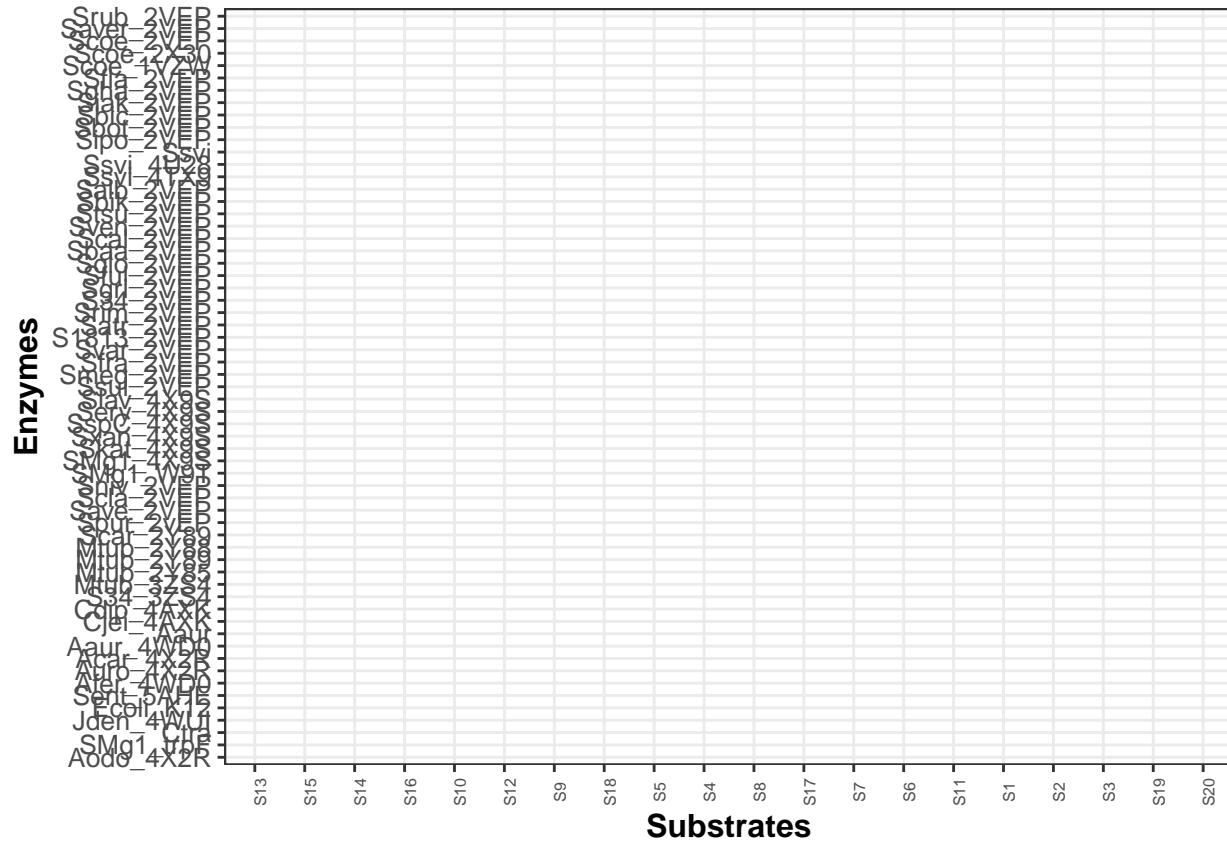


N = 56 Bandwidth = 0.1501

```
## debug en <text>#4: rug(docking[, i]) ## debug en <text>#4: browser()
r plot(cor(docking[-c(5,12,13,38,51),-1]))
```



```
r ## Leer sobre la incertidumbre del 2 y explicarla      ## Y leer el paper de Julian y
el de mauricio sobre reportes de docking -->
ggplot(docking.m, aes(x=variable, y=Enzima)) + labs(x = "Substrates", y = "Enzymes",text = element_text
```



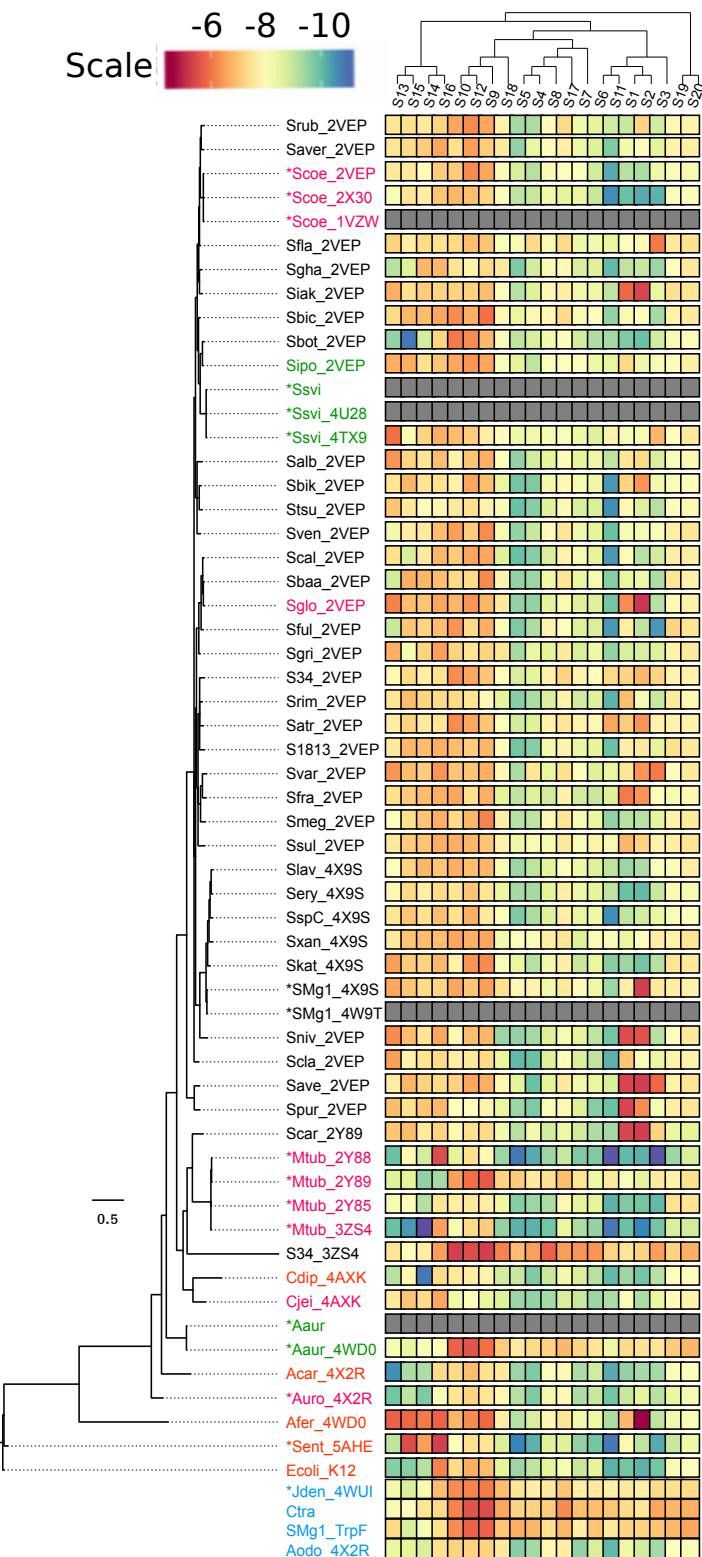


Figure 16:
33

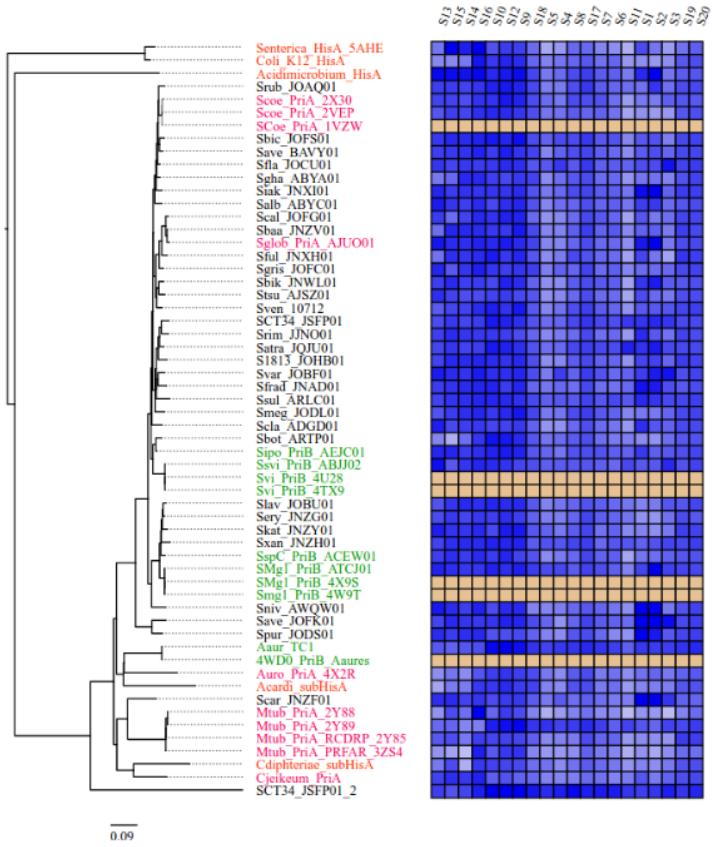
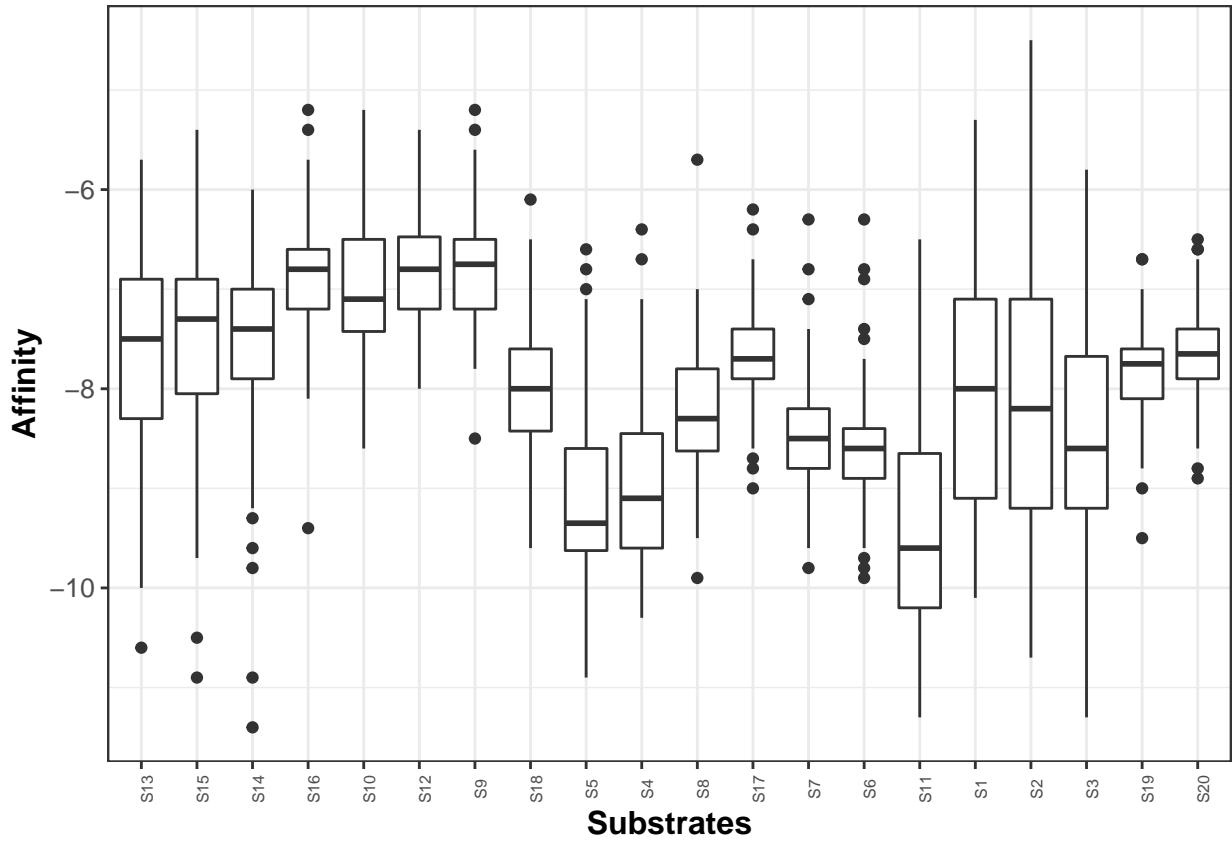


Figure 17:

GTP es el sustrato por el que PriA muestra mayor afinidad

```
## boxplot de los sustratos
ggplot(docking.m, aes(x=variable, y=value)) + labs(x = "Substrates", y = "Affinity",text = element_text)

## Warning: Removed 100 rows containing non-finite values (stat_boxplot).
```



Molecular dynamics vs experimental data

PriA activity on S3 (PRA)

organism	Family	K_M	k_{cat}	$\frac{k_{cat}}{K_M}$	Pre MD	Pos MD	Reference
Afer	HisA	1.1 ± 0.2	0.05 ± 0.001	0.045	-10.1	-12.3	Noda-García L et al 2015
Ecoli	HisA	1.6	4.9	3.1	-9.9	-16	Henn-Sax et al. (2002)
Sent	HisA	17.0 ± 0.1	7.8 ± 2.4	4.5×10^5	-10.3	-20.1	Söderholm A et al (2015)
Aaur	PriB	2.1 ± 0.5	1.8 ± 0.2	0.9	-7.4		verduzco-castro 2016
Sipo	PriB	3.8 ± 0.2	0.82 ± 0.02	0.21	-8.2	-14.7	verduzco-castro 2016
SspC	PriB	11.4 ± 3.4	2.53 ± 0.74	0.22	-8.5	-12.7	verduzco-castro 2016
SMg1	PriB	13.2 ± 3.4	0.92 ± 0.19	0.069	-8	-15.2	verduzco-castro 2016
Ssvi	PriB	3.9 ± 0.89	0.69 ± 0.04	0.18	-8.2	-16.7	verduzco-castro 2016
ScOE	PriA	3.6 ± 0.7	1.3 ± 0.2	0.4	-8.4	-15	Noda-García et al (2010)
Sglob	PriA	4.2 ± 0.8	0.74 ± 0.03	0.18	-9.2	-16.7	verduzco-castro
Mtub 2Y85	priA	190.23	$0.012 - 9.7$				Due et al 2011
Mtub 3ZS4	priA	?	-9.9				Due et al 2011 (To be published)
Auro	priA	4.0 ± 0.9	0.2 ± 0.03	0.04	-9.2		Vazquez-Juarez (2016)
Cjei	PriA	2.3 ± 0.2	0.9 ± 0.08	0.39	-8.5		Noda-García et al (2013)
Cdip	subHisA	4.4 ± 0.5	2.6 ± 0.3	0.59	-9.2		Noda-García et al (2013)
SMg1 TrpF	TrpF3	-	-	-	-6.9	-9.6	verduzco-castro 2016
Jden	TrpF3	-	-	-7.2	-9.4	16.8 ± 3.3	Verduzco-Castro E et al 2016
Acar	SubHisA	0.02					
Aodo	SubTrpF	-	-	-			

PriA activity on S7 (PROFAR)

\$\$

organism	Family	K_M	k_{cat}	$\frac{k_{cat}}{K_M}$	Pre MD	Pos MD	Reference
Afer	HisA	-	-	-	-9.2	-9	Noda-García L et al. (2015)
Ecoli	HisA	-	-	-	-9	-11.1	Henn-Sax et al. (2002)
Sent	HisA	-	-	-	-9.6	-10.2	Söderholm A et al (2015)
Aaur	PriB	26.3 ± 6.3	0.37 ± 0.09	0.014	-7.1	-	verduzco-castro 2016
Sipo	PriB	60.8 ± 1.1	8.25 ± 0.4	0.14	-8	-8.5	verduzco-castro 2016
SspC	PriB	149.9 ± 29	1.4 ± 0.12	0.009	-8.5	-10.8	verduzco-castro 2016
SMg1	PriB	129.6 ± 34	0.29 ± 0.04	0.0022	-7.5	-11	verduzco-castro 2016
Ssvi	PriB	24.5 ± 4.0	1.6 ± 0.29	0.067	-8	-9.7	verduzco-castro 2016
Scoe	PriA	5.0 ± 0.08	3.4 ± 0.09	0.7	-8	-9.4	Noda-García et al (2010)
Sglob	PriA	11 ± 1.0	3.8 ± 0.2	0.34	-8.7	-9.4	verduzco-castro 2016
Mtub2Y85	priA	21	3.6	0.17	-8.6		Due et al 2011
Mtub3ZS4	priA				-9.3		Due et al 2011 (To be published)
Auro	priA	23 ± 6.5	0.5 ± 0.05	0.02	-9.3		Vazquez-Juarez (2016)
Cjei	PriA	5.1 ± 1.0	1.6 ± 0.16	0.31	-9		Noda-García et al (2013)
Cdip	subHisA	-	-	-	-8.8		Noda-García et al (2013)
SMg1 TrpF	TrpF3	8.4 ± 1.7	10.5 ± 2.4	1.25	-7.6	-9	verduzco-castro
Jden	TrpF3	16.8 ± 3.3	27 ± 1.6	1.6	-7.6	-7.7	verduzco-castro
Acar	SubHisA	Na	Na	0.02	Na	Na	Na
Aodo	SubTrpF	-	-	-	-	Na	Na

\$\$

Con actividad de FolE i.e activa para el compuesto V Adams et al (2014) Genome size vs Total antismash cluster coloured by order

```
#sessionInfo()
```

PriA en cinéticas enzimáticas no tradicionales.

Además de la exploración genómica de PriA se realizaron caracterizaciones experimentales. *i)* Se realizaron cinéticas de PriA en GTP . *ii)* Se avanzó en medir simultáneamente la actividad de PriA sobre ProFAR y PRA.

Es posible que PriA tenga actividad en GTP

Activity was measured fluorometrically in 96-well plates (Nuc 96-Well Optical Botto Plates) in a TECAN infinite M1000 plate reader (excitation at 286 nm and emission at 386 nm)

Preliminary activity assays were performed on an active PriA from *Streptomyces coelicolor* and an inactive mutant D11A.

Enzymes were cloned on coli V68 strains, overexpression were induced and protein were purified.

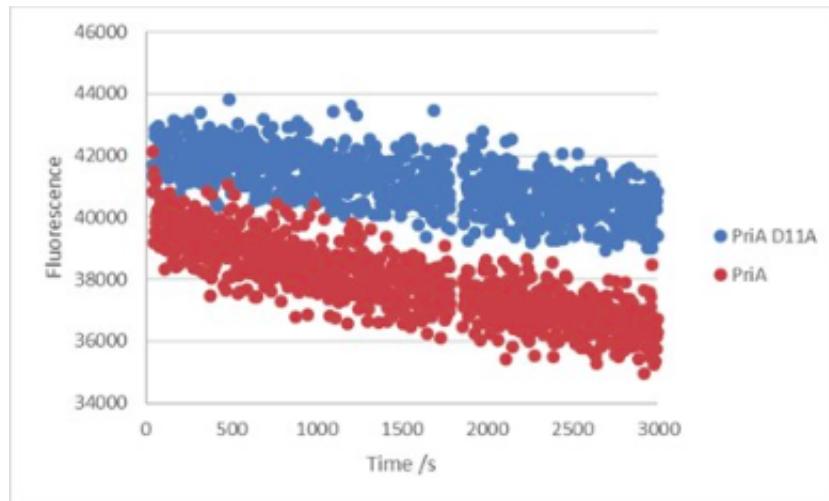


Figure 18:

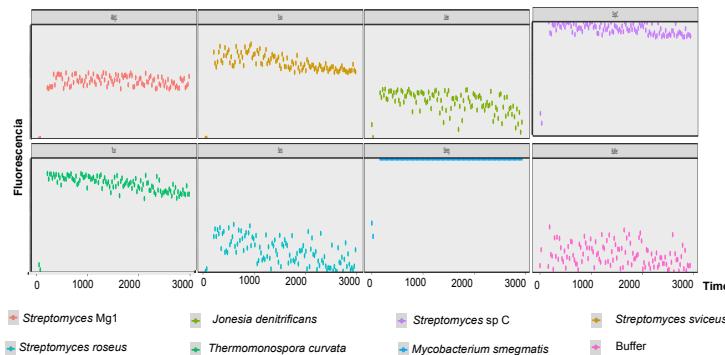


Figure 19:

las dos enzimas con actividad *Thermomonospora curvata* es una Actinobacteria termofílica de la familia *Thermomonosporaceae*, puede ser encontrada en compostas ya que participa en la degradación de celulosa [chertkov_complete_2011].

Jonesia trpF no mostró actividad *Jonesia denitrificans* está clasificada como un organismo patogénico para animales, su genoma fue originalmente aislado de from cooked ox blood [pukall_complete_2009].

```
docking <- read.csv("chapter4/SmallHeat.data", header=TRUE, sep="\t")
kable(docking,
      caption = "Enzymes docking \\label{tab:docking}",
      caption.short = "Enzymes docking ")
```

Table 6: Enzymes docking

Enzima	S13	S15	S14	S16	S10	S12	S9	S18	S5	S4	S8	S17	S7	S6	S11	S1
Srub_2VEP	-7.4	-7.3	-7.5	-7.1	-6.5	-6.2	-6.5	-7.7	-9.4	-9.3	-7.9	-7.2	-8.3	-8.6	-8.9	-9.0
Saver_2VEP	-7.4	-7.2	-7.0	-6.5	-7.3	-6.4	-7.0	-7.5	-9.6	-8.5	-7.9	-7.6	-8.4	-8.7	-9.8	-8.3
Scoe_2VEP	-7.5	-7.5	-7.9	-7.0	-7.0	-6.2	-6.5	-7.8	-8.8	-9.2	-7.8	-7.9	-8.0	-8.9	-10.3	-9.2
Scoe_2X30	-8.1	-7.4	-7.6	-6.9	-6.7	-6.8	-7.1	-7.9	-9.1	-9.0	-8.3	-8.6	-8.5	-9.0	-10.6	-10.0
Scoe_1VZW	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sfla_2VEP	-7.1	-7.6	-7.2	-7.3	-7.2	-6.8	-6.9	-8.1	-8.2	-7.2	-8.2	-7.2	-8.4	-8.3	-8.5	-7.9
Sgha_2VEP	-9.2	-8.7	-6.7	-6.7	-7.4	-7.7	-7.2	-7.5	-9.8	-8.9	-8.2	-7.8	-8.8	-8.7	-10.1	-9.1
Siak_2VEP	-6.6	-7.3	-7.0	-7.1	-7.1	-7.1	-6.8	-8.0	-9.2	-8.7	-7.8	-7.6	-8.3	-8.4	-9.1	-5.8
Sbic_2VEP	-7.2	-6.7	-6.8	-6.5	-6.2	-6.6	-5.9	-7.8	-8.5	-7.8	-7.8	-7.2	-8.2	-8.0	-9.6	-8.2
Sbot_2VEP	-9.6	-10.9	-8.9	-7.1	-6.0	-6.2	-6.7	-8.1	-9.1	-8.8	-8.3	-7.9	-8.9	-9.3	-9.4	-9.8
Sipo_2VEP	-6.6	-6.6	-7.3	-6.9	-6.5	-6.3	-6.5	-8.1	-8.6	-9.1	-8.0	-7.9	-8.0	-8.4	-8.5	-7.2
Ssvi	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Ssvi_4U28	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Ssvi_4TX9	-5.9	-8.2	-7.1	-6.6	-6.8	-7.0	-7.6	-7.9	-8.4	-8.3	-8.2	-8.1	-8.4	-8.7	-7.7	-8.2
Salb_2VEP	-6.3	-7.1	-7.4	-6.8	-7.7	-6.9	-6.6	-8.0	-9.6	-8.9	-8.6	-7.9	-8.6	-8.4	-9.1	-7.4
Sbik_2VEP	-7.4	-6.7	-7.4	-7.3	-7.7	-6.5	-6.8	-8.5	-9.9	-9.8	-8.3	-7.8	-8.2	-8.4	-10.5	-7.1
Stsu_2VEP	-6.9	-7.5	-7.9	-7.8	-7.5	-7.3	-7.7	-8.1	-10.0	-9.7	-8.7	-7.8	-8.8	-8.8	-10.6	-7.8
Sven_2VEP	-8.3	-7.6	-7.5	-6.8	-6.4	-7.0	-6.2	-8.2	-9.6	-9.0	-7.9	-7.4	-8.3	-8.6	-10.0	-8.2
Scal_2VEP	-7.3	-8.8	-7.5	-6.7	-7.0	-6.5	-6.5	-8.7	-10.0	-9.7	-8.8	-7.7	-8.2	-8.6	-10.5	-8.1
Sbaa_2VEP	-8.9	-6.6	-6.7	-6.8	-7.2	-7.2	-6.2	-7.7	-9.6	-9.5	-8.4	-7.4	-8.5	-8.2	-9.7	-8.5
Sglo_2VEP	-6.1	-6.8	-6.9	-6.5	-6.7	-6.3	-6.7	-7.5	-9.6	-9.6	-8.6	-7.8	-8.7	-8.7	-9.9	-6.2
Sful_2VEP	-9.0	-6.9	-7.1	-6.8	-6.3	-7.2	-6.6	-8.1	-9.8	-9.3	-7.7	-8.0	-8.7	-8.8	-10.5	-7.6
Sgri_2VEP	-6.6	-8.2	-7.1	-6.4	-7.1	-7.4	-7.1	-7.4	-9.4	-8.5	-8.6	-7.5	-8.8	-8.4	-9.6	-8.8
S34_2VEP	-7.8	-7.2	-7.6	-7.3	-6.2	-6.5	-6.9	-8.0	-8.8	-8.6	-7.7	-7.1	-7.7	-7.9	-7.2	-7.1
Srim_2VEP	-7.4	-6.8	-7.1	-7.2	-7.2	-7.4	-7.8	-8.6	-9.7	-9.3	-8.8	-7.7	-8.9	-8.7	-10.2	-6.8
Satr_2VEP	-7.6	-7.1	-7.5	-7.4	-6.2	-6.5	-6.9	-8.0	-8.9	-8.7	-7.7	-7.4	-7.7	-7.8	-6.5	-7.0
S1813_2VEP	-7.5	-6.7	-6.8	-6.6	-6.8	-7.0	-6.7	-7.9	-9.9	-9.7	-8.3	-7.7	-8.5	-8.6	-10.0	-7.6
Svar_2VEP	-6.3	-6.8	-7.0	-6.3	-6.9	-6.9	-6.6	-7.8	-9.3	-7.4	-8.5	-7.3	-8.0	-8.7	-8.5	-7.4
Sfra_2VEP	-7.3	-7.0	-6.8	-6.6	-6.5	-7.3	-6.6	-8.6	-9.1	-8.7	-8.9	-7.7	-8.9	-9.0	-8.7	-6.0
Smeg_2VEP	-8.0	-7.3	-6.9	-6.6	-7.2	-6.7	-6.1	-7.6	-9.2	-9.1	-7.8	-7.4	-8.2	-8.5	-9.6	-9.2
Ssul_2VEP	-7.5	-7.0	-6.9	-6.6	-6.6	-6.7	-6.9	-7.7	-8.4	-8.2	-8.2	-7.5	-8.1	-7.7	-8.2	-6.7
Slav_4X9S	-8.0	-7.1	-6.7	-6.9	-6.6	-6.8	-6.7	-8.2	-9.5	-9.1	-8.6	-8.0	-8.4	-8.7	-9.1	-9.5
Sery_4X9S	-7.8	-7.2	-7.6	-7.2	-7.4	-6.9	-7.4	-8.5	-9.4	-9.4	-8.6	-7.6	-8.4	-8.6	-9.1	-10.0
SspC_4X9S	-7.4	-6.9	-7.3	-6.8	-7.3	-6.6	-7.5	-8.0	-9.8	-9.3	-8.7	-7.6	-8.5	-8.4	-10.6	-9.1
Sxan_4X9S	-7.5	-6.9	-6.8	-6.8	-6.5	-6.6	-6.5	-8.3	-8.0	-8.1	-7.6	-7.4	-8.7	-8.1	-8.5	-8.5
Skat_4X9S	-6.4	-7.0	-7.1	-6.7	-7.7	-6.3	-6.2	-8.1	-8.5	-9.2	-8.3	-7.6	-9.0	-8.5	-9.7	-9.5
SMg1_4X9S	-6.5	-6.9	-7.2	-7.1	-6.5	-6.9	-6.4	-7.3	-7.8	-7.7	-8.3	-7.5	-7.9	-8.4	-9.5	-7.6
SMg1_W9T	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Sniv_2VEP	-6.2	-6.8	-7.4	-6.5	-7.8	-6.8	-6.7	-9.5	-9.6	-9.4	-8.7	-8.2	-8.6	-9.1	-9.9	-5.4
Scla_2VEP	-6.4	-7.7	-7.4	-7.2	-7.6	-7.5	-7.2	-8.5	-10.1	-10.0	-8.7	-7.9	-8.8	-9.1	-10.2	-6.9

Enzima	S13	S15	S14	S16	S10	S12	S9	S18	S5	S4	S8	S17	S7	S6	S11	S1
Save_2VEP	-7.6	-6.7	-7.1	-7.2	-7.1	-6.5	-6.6	-7.8	-8.6	-10.0	-8.6	-7.5	-8.3	-8.5	-8.8	-5.3
Spur_2vEP	-6.8	-7.3	-6.9	-6.7	-8.0	-7.9	-7.7	-8.5	-9.5	-9.8	-8.1	-7.8	-8.7	-9.7	-10.0	-5.3
Scar_2Y89	-6.9	-6.8	-7.5	-7.1	-7.8	-7.3	-7.2	-8.3	-9.2	-8.3	-8.9	-8.4	-8.9	-9.3	-9.4	-5.3
Mtub_2Y88	-10.0	-7.8	-8.9	-5.4	-8.6	-7.3	-7.8	-9.5	-10.9	-10.3	-9.5	-9.0	-9.8	-9.8	-11.3	-10.1
Mtub_2Y89	-8.7	-8.6	-9.6	-9.4	-6.4	-5.9	-5.6	-7.1	-7.0	-7.5	-7.3	-6.8	-7.4	-8.4	-7.5	-8.1
Mtub_2Y85	-8.2	-7.9	-9.2	-7.4	-7.6	-7.5	-7.6	-8.4	-9.7	-9.5	-9.3	-7.8	-8.6	-8.6	-10.2	-9.8
Mtub_3ZS4	-10.0	-10.5	-11.4	-6.4	-8.2	-7.2	-7.0	-9.6	-10.2	-10.2	-9.9	-8.5	-9.3	-9.6	-10.9	-10.0
S34_3ZS4	-7.4	-8.0	-7.6	-6.4	-5.2	-5.4	-5.2	-6.1	-6.8	-6.4	-5.7	-6.4	-6.3	-6.3	-7.1	-7.4
Cdip_4AXK	-9.2	-7.8	-10.9	-7.2	-7.5	-7.6	-7.7	-8.9	-9.0	-9.8	-9.0	-8.3	-8.8	-9.2	-10.1	-9.5
Cjei_4AXK	-7.5	-6.9	-7.2	-6.5	-8.4	-8.0	-8.5	-8.7	-9.5	-9.6	-9.4	-8.8	-9.0	-9.5	-9.3	-8.1
Aaur	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Aaur_4WD0	-8.2	-8.5	-8.1	-7.9	-5.7	-5.5	-5.9	-7.0	-7.5	-7.2	-7.1	-6.7	-7.1	-7.4	-8.4	-7.2
Acar_4X2R	-10.6	-9.3	-9.3	-7.2	-7.2	-6.8	-7.3	-7.5	-9.5	-9.8	-8.0	-7.8	-9.3	-8.9	-10.3	-9.6
Auro_4X2R	-9.9	-9.1	-9.8	-8.1	-7.4	-7.1	-7.4	-7.8	-9.2	-9.8	-8.3	-7.8	-9.3	-9.0	-9.9	-9.0
Afer_4WD0	-5.7	-5.8	-6.0	-5.7	-6.7	-6.2	-5.8	-7.6	-9.2	-8.8	-7.8	-7.4	-8.3	-8.4	-9.3	-6.7
Sent_5AHE	-9.1	-5.4	-6.4	-5.2	-8.0	-7.3	-7.4	-8.8	-10.7	-10.2	-8.7	-8.7	-9.6	-9.9	-10.9	-7.8
Ecoli_K12	-9.7	-9.7	-9.2	-6.1	-7.2	-6.6	-6.8	-8.6	-9.5	-9.1	-8.6	-8.2	-9.0	-8.6	-10.2	-9.9
Jden_4WUI	-8.3	-8.8	-8.2	-6.8	-6.2	-6.1	-6.0	-6.8	-7.4	-7.6	-7.5	-6.9	-7.6	-7.5	-7.7	-7.5
Ctra	-8.2	-8.0	-7.4	-7.2	-6.1	-5.5	-5.4	-6.5	-7.1	-7.1	-7.0	-6.2	-6.8	-6.8	-6.9	-7.2
SMg1_trpF	-7.2	-8.8	-8.2	-7.3	-6.2	-5.7	-5.7	-6.9	-6.6	-6.7	-7.3	-6.7	-7.6	-6.9	-7.5	-7.1
Aodo_4X2R	-8.5	-8.6	-8.8	-7.3	-7.1	-7.1	-7.1	-7.5	-9.7	-9.4	-7.8	-7.6	-9.6	-8.8	-10.1	-8.4

Enzima	S13	S15	S14	S16	S10	S12	S9	S18	S5	S4	S8	S17	S7	S6	S11	S1
--------	-----	-----	-----	-----	-----	-----	----	-----	----	----	----	-----	----	----	-----	----



Figure 20:

Cinéticas simultáneas

tle: "TrpF_kinetics"
 thor: "NellySelem"
 te: "May 18, 2017"
 tput: pdf_document

```
#First I saved file as scv tablar separated
#perl -p -i -e 's/,/\t/g' Pra_scoe1.csv
#An cut it to obtain just data
#tail -n +39 Pra_scoe1.csv | head -n-3 > Pra_scoe1.data
#perl -p -i -e 's/^\sNr\|.\|\s\[\w*\]\|/\.\s\[\.*\]\//g' Pra_scoe1.data


```

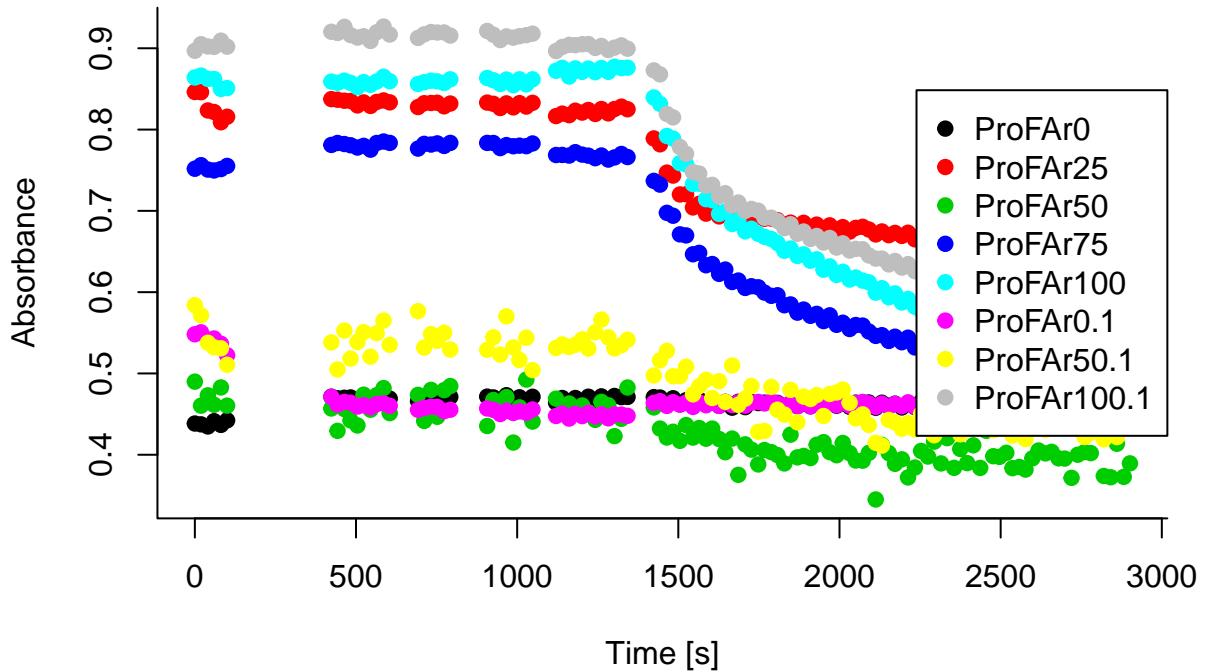
```

## [1] 0.00 19.15 40.50 59.65 81.00 100.15 423.50 442.65
## [9] 464.00 483.15 504.50 523.65 545.00 564.15 585.50 604.65
## [17] 692.40 711.55 732.90 752.05 773.40 792.55 907.00 926.15
## [25] 947.50 966.65 988.00 1007.15 1028.50 1047.65 1120.70 1139.85
## [33] 1161.20 1180.35 1201.70 1220.85 1242.20 1261.35 NA NA
## [41] 1282.80 1301.95 1323.30 1342.45 1424.00 1443.15 1464.50 1483.65
## [49] 1505.00 1524.15 1545.50 1564.65 1586.00 1605.15 1626.50 1645.65
## [57] 1667.00 1686.15 1707.50 1726.65 1748.00 1767.15 1788.50 1807.65
## [65] 1829.00 1848.15 1869.50 1888.65 1910.00 1929.15 1950.50 1969.65
## [73] 1991.00 2010.15 2031.40 2050.55 2071.90 2091.05 2112.50 2131.65
## [81] 2153.00 2172.15 2193.50 2212.65 2234.00 2253.15 2274.50 2293.65
## [89] 2315.00 2334.15 2355.50 2374.65 2396.00 2415.15 2436.50 2455.65
## [97] 2477.00 2496.15 2517.50 2536.65 2558.00 2577.15 2598.50 2617.65
## [105] 2639.00 2658.15 2679.50 2698.65 2720.00 2739.15 2760.50 2779.65
## [113] 2801.00 2820.15 2841.50 2860.65 2882.00 2901.15

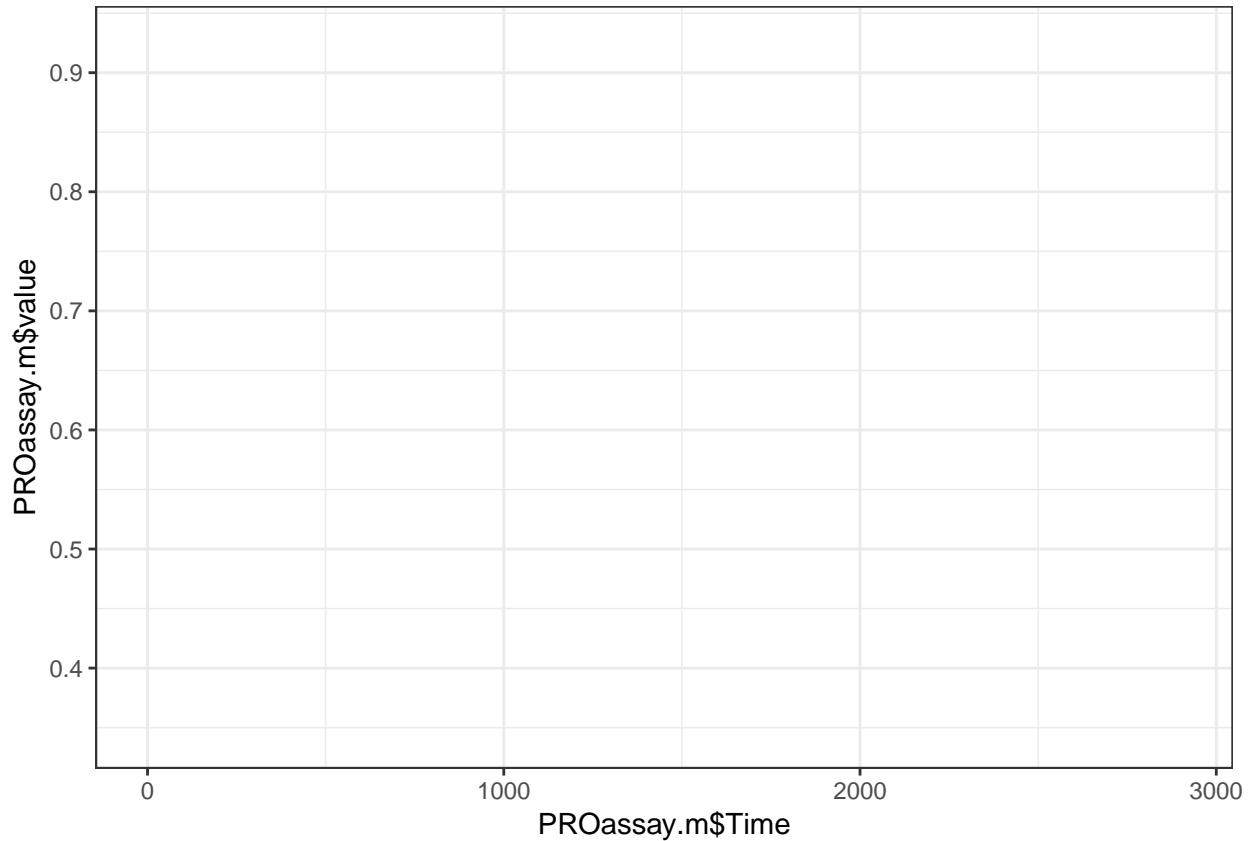
tablePRO<-tablePRO[grep("Pro|Time", colnames(tablePRO))]
#tablePRA<-tablePRA[grep("Pra/Time", colnames(tablePRA))]
```

```
PROassay.m <- melt(tablePRO,id="Time")
```

```
plot(PROassay.m$Time, PROassay.m$value, col=PROassay.m$variable, xlab="Time [s]",ylab="Absorbance", pch=19, bg="white", cex=1)
par(xpd = TRUE)
legend("right", legend = (unique(PROassay.m$variable)), col = (unique(PROassay.m$variable)),pch=19, bg="white")
```

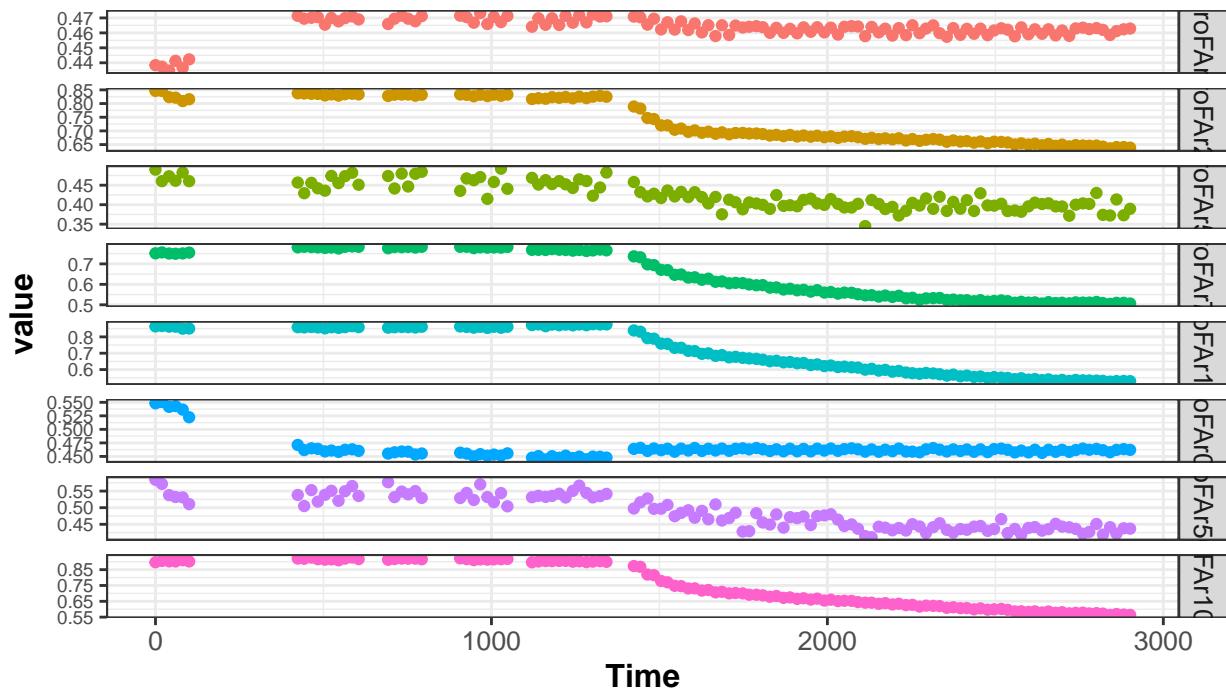


```
ggplot(PROassay.m, aes(x = PROassay.m$Time, y = PROassay.m$value),color="variable") + theme_bw()
```



```
qplot(Time,value, data = PROassay.m, colour=variable)+theme_bw() + theme(legend.position = "bottom", legend
```

Warning: Removed 16 rows containing missing values (geom_point).



variable	● ProFAR0	● ProFAR50	● ProFAR100	● ProFAR50.1
	● ProFAR25	● ProFAR75	● ProFAR0.1	● ProFAR100.1

```
## visualising full data
tablePRA$Temp<-NULL
tablePRA$Cycle<-NULL
```

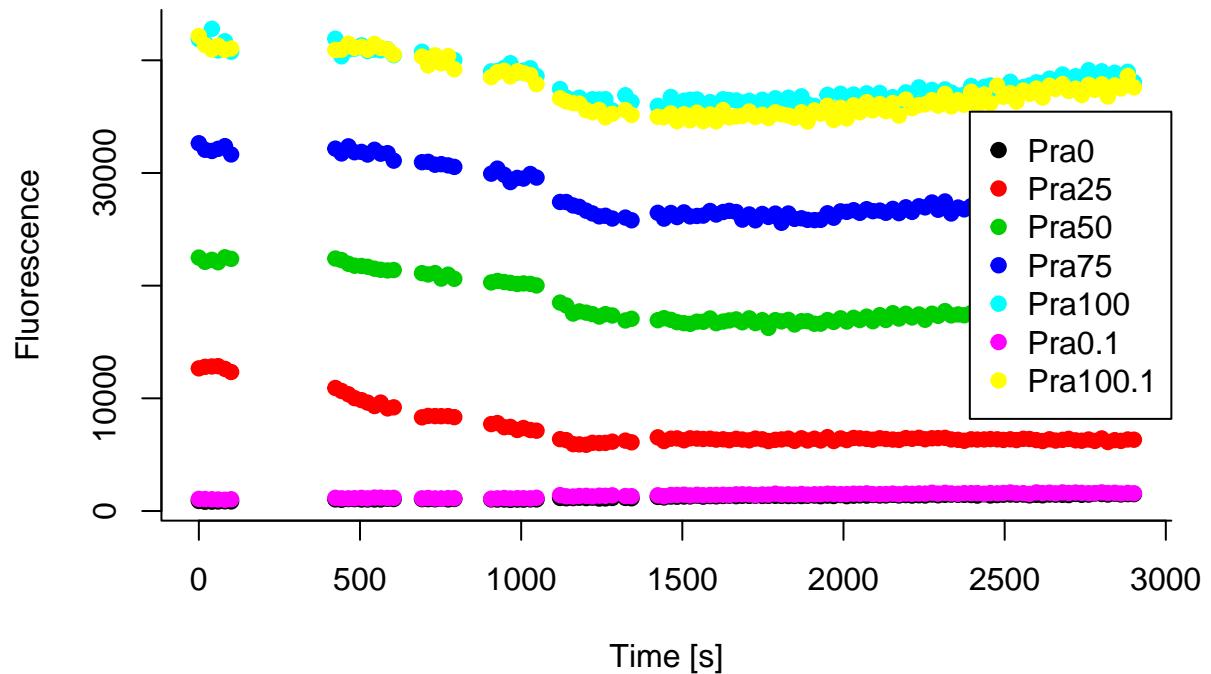
```
a<-tablePRA$Time[c(TRUE, FALSE)]
b<-tablePRA$Time[c(TRUE, FALSE)]+19.15
```

```
tablePRA$Time[c(TRUE, FALSE)] <- a
tablePRA$Time[c(FALSE, TRUE)] <- b
tablePRA$Time
```

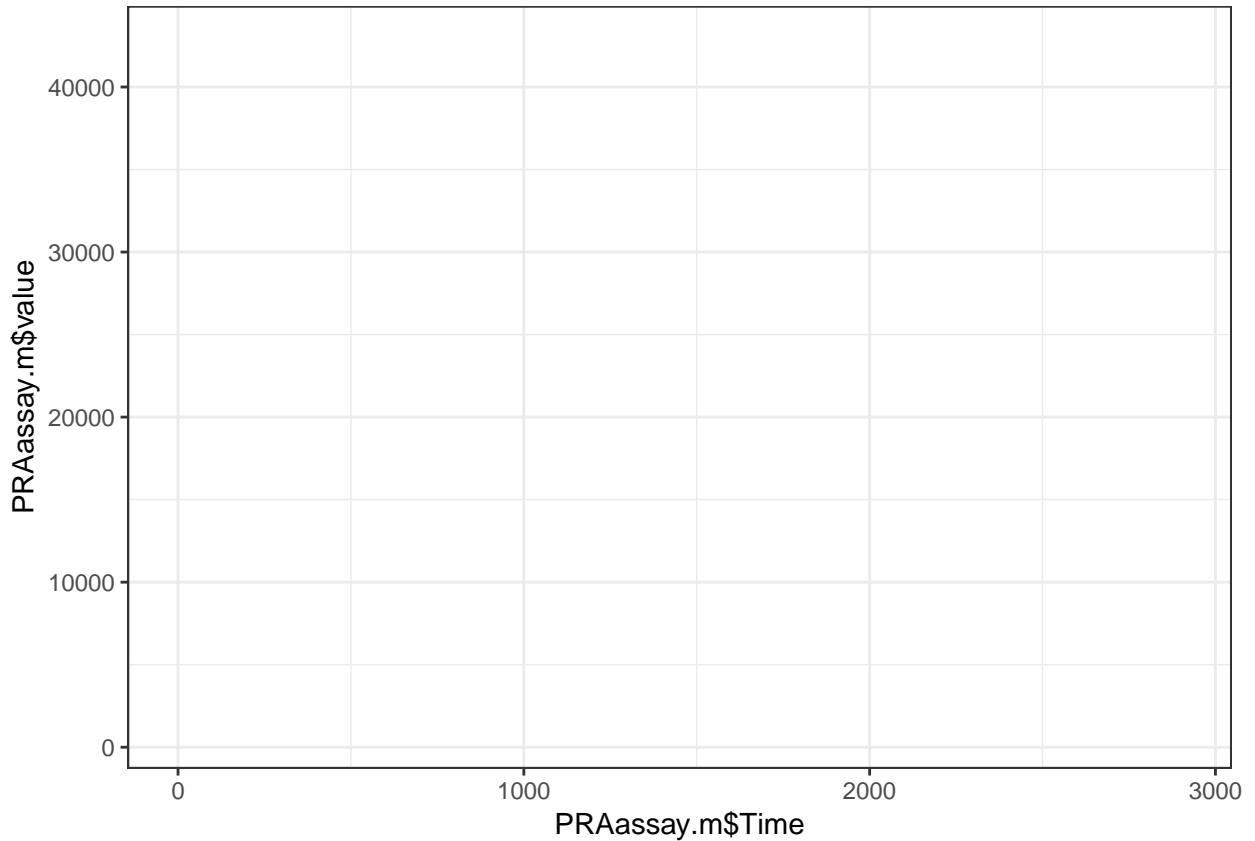
```
## [1] 0.00 19.15 40.50 59.65 81.00 100.15 423.50 442.65
## [9] 464.00 483.15 504.50 523.65 545.00 564.15 585.50 604.65
## [17] 692.30 711.45 732.80 751.95 773.40 792.55 907.00 926.15
## [25] 947.50 966.65 988.00 1007.15 1028.50 1047.65 1120.70 1139.85
## [33] 1161.20 1180.35 1201.70 1220.85 1242.20 1261.35 1282.80 1301.95
## [41] 1323.30 1342.45 1424.00 1443.15 1464.50 1483.65 1505.00 1524.15
## [49] 1545.50 1564.65 1586.00 1605.15 1626.50 1645.65 1667.00 1686.15
## [57] 1707.50 1726.65 1748.00 1767.15 1788.50 1807.65 1829.00 1848.15
## [65] 1869.50 1888.65 1910.00 1929.15 1950.40 1969.55 1991.00 2010.15
## [73] 2031.50 2050.65 2071.90 2091.05 2112.50 2131.65 2153.00 2172.15
## [81] 2193.50 2212.65 2234.00 2253.15 2274.50 2293.65 2315.00 2334.15
## [89] 2355.50 2374.65 2396.00 2415.15 2436.50 2455.65 2477.00 2496.15
## [97] 2517.50 2536.65 2558.00 2577.15 2598.50 2617.65 2639.00 2658.15
## [105] 2679.50 2698.65 2720.00 2739.15 2760.50 2779.65 2801.00 2820.15
## [113] 2841.50 2860.65 2882.00 2901.15
```

```
tablePRA<-tablePRA[grep("Pra|Time", colnames(tablePRA))]
```

```
PRAassay.m <- melt(tablePRA,id="Time")
plot(PRAassay.m$Time, PRAassay.m$value, col=PRAassay.m$variable, xlab="Time [s]",ylab="Fluorescence",
par(xpd = TRUE)
legend("right", legend = (unique(PRAassay.m$variable)), col = (unique(PRAassay.m$variable)),pch=19,bg="white")
```

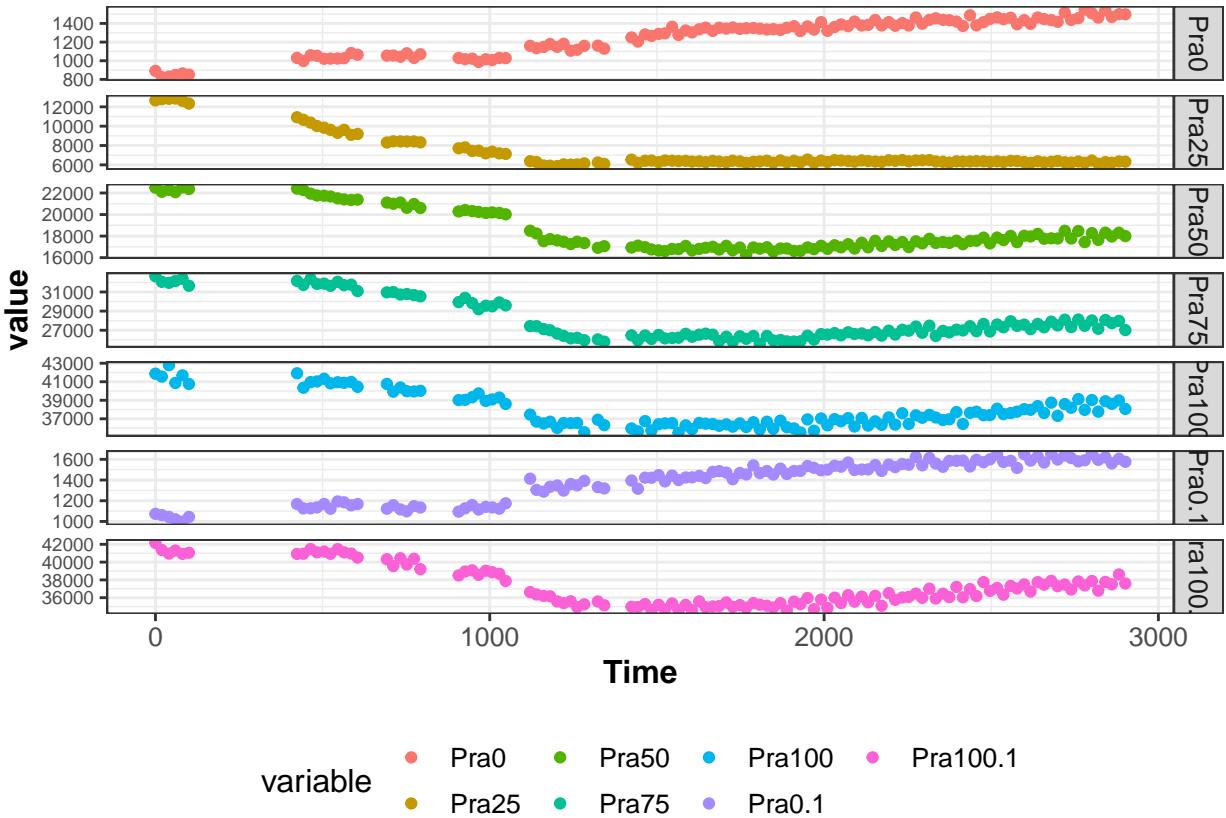


```
ggplot(PRAassay.m, aes(x = PRAassay.m$Time, y = PRAassay.m$value),color="variable") + theme_bw()
```



```
qplot(Time,value, data = PRAassay.m, colour=variable)+theme_bw() + theme(legend.position = "bottom", legend
```

Warning: Removed 7 rows containing missing values (geom_point).



```

##pendientes (slopes) of linear part of the curve
# time when TrpF is added
cuttime<-344.2
maxtime=620

PRAassay2Time<-tablePRA[which(tablePRA$Time >= cuttime & tablePRA$Time <= maxtime),]

## removi columnas que se salian del margen de medicion
#PRAassay2Time<-PRAassay2Time[, !(colnames(PRAassay2Time) %in% c("C62uM"))]

# vector with slopes for each dataset
V0 <- apply(PRAassay2Time, 2, function(x) coefficients(lm(x ~ PRAassay2Time$Time, na.action=na.omit))[2])

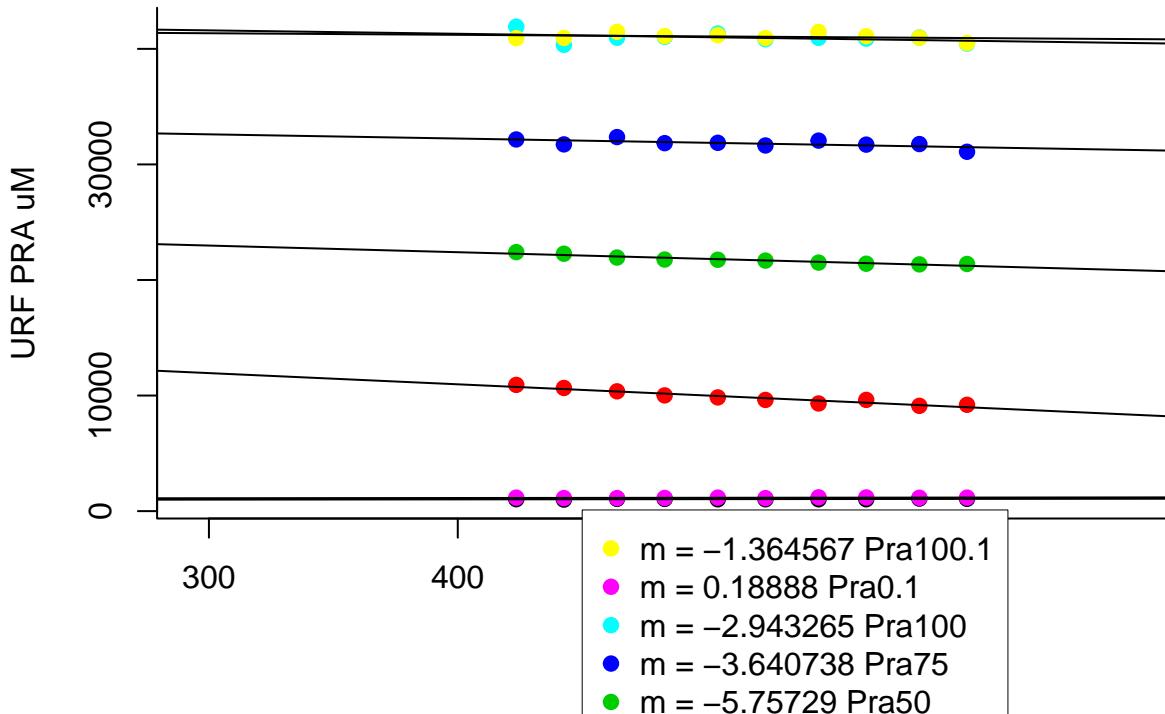
PRAassay2Time.m <- melt(PRAassay2Time,id="Time")
plot(PRAassay2Time.m$Time, PRAassay2Time.m$value, col=PRAassay2Time.m$variable,xlab="Time", ylab="URF P")

V1 <- apply(PRAassay2Time[,2:ncol(PRAassay2Time)], 2, function(x) coefficients(lm(x ~ PRAassay2Time$Time, na.action=na.omit))[2])
apply(V1, 2, function(x) {abline(x, col = PRAassay2Time.m$variable)})

## NULL

par(xpd = TRUE)
slopeText2="m ="
slopeText<-paste(slopeText2,rev(round(V0[-1],digits=6)),rev(names(V0[-1])))
legend(450,100, legend = slopeText, col = rev(unique(PRAassay2Time.m$variable)), pch=19,box.lwd=0, bg="white")

```



```

## get relevant velocities and get correspondent concentrations
## V0 or slopes are on [PRA]uM/s
# slopes stores the slope, with TRUE , FALSE the desired enzyme can be selected, One more than one is t
slopes=(round(V0[-1],digits=6))[c(TRUE)]
concentration<-as.matrix(PRAassay2Time[1,names(slopes)])
#slopes
#concentration
slopes=as.matrix((round(V0[-1],digits=6))[c(TRUE)])
row.names(slopes) <- NULL
slopes<-c(slopes)# plot on r base

slopes

## [1] 0.196328 -9.695510 -5.757290 -3.640738 -2.943265 0.188880 -1.364567
# in concentration I stored the initial substrate concentration
row.names(concentration) <- NULL
concentration<-c(concentration)
slopes

## [1] 0.196328 -9.695510 -5.757290 -3.640738 -2.943265 0.188880 -1.364567
concentration

## [1] 1031 10920 22410 32161 41918 1168 40922
## interpolar michaelis-menden

#https://rpubs.com/RomanL/6752
#https://davetang.org/muse/2013/05/17/fitting-a-michaelis-mentens-curve-using/
#S <-concentration
## from highest to lower concentration
S_ProFAR <-c(0,25,50,75,100,0,50,100)
v<-slopes

```

```

#v<-slopes[-length(slopes)]
#v <-slopes[c(TRUE, FALSE)]
v<--1*v
S_ProFAR
v
mm <- data.frame(S_ProFAR,v)
model.drm <- drm(v ~ S, data = mm, fct = MM.2())
summary(model.drm)
## first value equal km
## second value = vm
Km=coefficients(model.drm) [1]
Vmax=2*coefficients(model.drm) [2]
Enzyme=2.5 #2.5uM
Kcat=Vmax*Enzyme
Km
Vmax
Kcat
mml <- data.frame(S = seq(0, max(mm$S), length.out = 100))
mml$v <- predict(model.drm, newdata = mml)

## plot on r base
plot(mm,log=' ',xlim=c(0,max(mm$S)), ylim=c(0,max(mm$v)), xlab="Reads", ylab="Transcripts")

##plot on ggplot
ggplot(mm, aes(x = S, y = v)) + theme_bw() + xlab("Concentration [uM]") + ylab("Speed [d[PRA]uM/s]")
ggsave("mm.pdf", width = 6, height = 4)

```